N° 517-518-519 - 2020

Economie ET Statistique

Economics AND Statistics

Numéro spécial
Au-delà et autour du PIB : questions à la comptabilité nationale

Special Issue
Beyond and Around GDP: Questions to National Accounting

Insee

Economie ET Statistique

Economics AND Statistics

Directeur de la publication / Director of Publication:
Jean-Luc TAVERNIER
Rédactrice en chef / Editor in Chief:
Sophie PONTHIEUX
Responsable éditorial / Editorial Manager: Pascal GODEFROY
Assistant éditorial / Editorial Assistant: Étienne de LATUDE
Traductions / Translations: RWS Language Solutions
Chiltern Park, Chalfont St. Peter, Bucks, SL9 9FG Royaume-Uni
Maquette PAO et impression / CAP and printing: JOUVE
1, rue du Docteur-Sauvé, BP3, 53101 Mayenne

La revue est en accès libre sur le site **www.insee.fr**. Il est possible de s'abonner aux avis de parution sur le site. La revue peut être achetée sur le site **www.insee.fr**, rubrique « Services / Acheter nos publications ». La revue est également en vente dans 200 librairies à Paris et en province.

The journal is available in open access on the Insee website **www.insee.fr.** Publication alerts can be subscribed on-line. The printed version of the journal (in French) can be purchased on the Insee website **www.insee.fr**.

88, avenue Verdier - CS 70058 - 92541 Montrouge Cedex
Tél : +33 (0)1 87 69 50 00

N° 517-518-519 - 2020

Economie ET Statistique

Economics AND Statistics

Economie et Statistique / Economics and Statistics

Issue 517-518-519 – 2020

BEYOND AND AROUND GDP: QUESTIONS TO NATIONAL ACCOUNTING

Preface – National Accounting: Old Questions Revisited, Plus Some New Ones

Diane Coyle*

Whether or not the national accounts were 'one of the great inventions' of the past century (Landefeld, 2000), they have certainly been one of the most influential. They are the lens through which we have viewed economic progress, focused ultimately on the growth in real GDP. As a result, governments promising to deliver progress to their citizens have been evaluated by that metric, and made policy decisions in order to deliver it. This issue of *Economie et Statistique / Economics and Statistics* provides a timely overview of key critiques – some old and some new – of the national accounts.

It is timely not only because the process of preparing the revision of the System of National Accounts (SNA) by 2025 is under way, but also because the world economy has experienced its biggest decline in GDP in our memories due to the global pandemic. This is the moment to evaluate how well the national accounts still serve as what Hicks (1942) described as the 'social framework'. The questions raised in this issue, all priorities in the current process of SNA revision, concern both old critiques of the national accounts for omitting important aspects of economic welfare (distribution, non-monetary production, sustainability), and also newer critiques concerning the treatment of globalisation and digitalisation within the existing framework.

To begin with the long-standing questions, **Didier Blanchet and Marc Fleurbaey** in their paper put it succinctly: although the measure of real incomes provided by the national accounts captures an essential aspect of economic welfare, it omits non-monetary components, and can tell us nothing about long-term sustainability. These two problems differ in character.

When it comes to non-monetary contributors, there have been proposals for a number of alternative composite indicators, such as metrics of environmental damage or income inequality. The well-known Human Development Index is one example. These tend to fall foul of their arbitrary choice of components and weights. There is now a substantial literature on the measurement of subjective well-being, but it presents some unresolved challenges and ambiguities. The authors therefore favour a third approach, calculation of equivalent incomes as an inclusive measure taking into account differing non-monetised circumstances such as health or employment status. The relevance of direct measures of subjective well-being is also discussed by **Jérôme Accardo**, in a paper on the contributions of social statistics to the 'beyond GDP' literature: evaluations of non-market household production, disaggregation of accounts across categories of households for a better assessment of growth inclusiveness, etc.

Inclusivity is also the focus of the paper by **Facundo Alvaredo, Lucas Chancel, Thomas Piketty, Emmanuel Saez and Gabriel Zucman,** which describes their Distributional National Accounts (DINA), a method of assigning income and wealth to different groups

* *Bennett Institute, University of Cambridge (dc700@cam.ac.uk)*

Citation: Coyle, D. (2020). Preface – National Accounting: Old Questions Revisited, Plus Some New Ones. *Economie et Statistique / Economics and Statistics*, 517-518-519, 5–7. https://doi.org/10.24187/ecostat.2020.517t.2015

in a manner fully consistent with the SNA methodology. Their results show substantial increases in inequality – including a global increase since 1980 despite the growth of China and India – but with substantial differences across countries. Given the way the pandemic is amplifying existing inequalities, it seems likely distribution will remain in sharp focus, in a welcome return to the tradition of early national accountants including Simon Kuznets.

When it comes to sustainability, and the position of future generations, the statistical and conceptual challenges are greater. As Blanchet and Fleurbaey note, an assessment of sustainability necessarily involves forecasts to value stocks of wealth, which they suggest falls outside the remit of statistics production. Two aspects of sustainability seem critical to understanding economic welfare, however. One is human capital, which World Bank work estimates to be empirically the most important component of comprehensive wealth (World Bank, 2018). The other is natural capital and in particular climate. In his paper, **Nicolas Canry** discusses integrating measurement of human capital in the national accounts, as a component of investment rather than consumption, while **Jean-Marc Germain and Thomas Lellouch** discuss of ways towards an environmental economic accounting that would include the climate debt. Given recent progress in developing the Standard for Environmental Economic Accounts (UN) and its application in some countries (e.g. the Office for National Statistics in the UK), as well as broader interest in the measurement of produced and non-produced capitals (Zenghelis *et al*., 2020), statistics for sustainability seem sure to make progress.

Turning to the newer challenges, the need to understand the interaction between the phenomena of globalisation and the national accounts has become pressing as supply chains and the role of large multinationals comes into sharper focus. **Marie-Baïanne Khder, Jérémi Montornès and Nicolas Ragache** examine how the notorious upward revision of Ireland's 2015 annual GDP growth in Ireland to 26% has reflected tax-related relocation of intangible assets by multinationals. **Niamh Holton, Margaret Kinsella, Oisín Mangan, Shaun McLaughlin and Patrick Quill** explore inconsistencies between the national accounts and balance of payments data for Ireland, related to methodological differences in the measurement of R&D and intangible intellectual property assets. The approach to measurement needs to be guided by the questions being addressed. Is the aim to have a picture of domestic economic activity, in which case the impact of multinationals needs to be captured in separate indicators, or to understand the impact of, say, tax policies on international integration?

However, as **Didier Blanchet** points out in his paper, the importance of multinationals operating extended supply chains across borders means it is increasingly difficult to define a 'domestic' economy, particularly given the ease with which intangible factors and assets can be relocated. Globalisation here overlaps with digitalization, which has seen the growing importance of intangibles including data, and of large multinationals providing digital services – some 'free' to users – in many countries.

Derek Burnell and Amani Elnasri argue that digitalization does not pose conceptual challenges to the national accounts, although it does require careful attention to data collection and more co-operation between national statistical agencies. They estimate that measurement issues related to digitalization do not help explain much of the slowdown in Australia's multifactor productivity growth (in line with findings by Syverson, 2017 and Byrne *et al.*, 2016, for the US) – although they note that free digital goods may offer consumer surplus benefits that lie outside the SNA production boundary.

However, the location of the production boundary is precisely one of the questions raised by considering whether the national accounts serve as a useful framework for assessing economic progress. If so much activity affecting people's economic welfare lies outside the production boundary, is the boundary usefully located?

The other question is the distinction between nominal magnitudes and 'real' or 'volume' measures, which raises some longstanding issues about the extent to which GDP and the national accounts can capture changes in economic welfare (Coyle, 2020). Here, **Lorraine Aeberhardt, Florian Hatier, Marie Leclair, Benoît Pentinat and Jean-Denis Zafar** explore the price/quantity split for areas affected by digitalization such as communications, free digital services and online products, concluding that in the French case measurement errors of this type are not enough to account for the productivity slowdown. **Mo Abdirahman, Diane Coyle, Richard Heys and Will Stewart** consider the specific example of prices for telecommunications services in the UK, where potential price profiles range from almost flat in the official index to a 90% decline in a unit value index, with corresponding implications for that sector's real growth and productivity. Finally, drawing together these conundrums about price/quantity splits and the production boundary, **Alexandre Bourgeois** considers different possible treatments of free digital services, concluding that the various approaches address somewhat different questions.

There are indeed different questions we try to answer using the national accounts, from narrower questions about production to fundamental questions about economic welfare, and it can sometimes seem that this is the wrong tool for the job. However, GDP is so universally used as shorthand for progress – and rightly so at times of recession – that the national accounts framework has to respond to the fundamental questions addressed in this issue about the extent to which it can capture economic welfare. If all of the changes that are so clearly affecting our lives, from climate change to digitalization, lie outside the limits of the framework, the framework will no longer be serving its purpose. □

References

Byrne, D. M., Fernald, J. G. & Reinsdorf, M. B. (2016). Does the United States have a Productivity Slowdown or a Measurement Problem. *Finance and Economics Discussion Series* 2016-17. Board of Governors of the Federal Reserve System (U.S.).
https://ideas.repec.org/p/fip/fedgfe/2016-17.html

Coyle, D. (2020). Productivity Measurement: New goods, variety and quality. *Brookings Institute*, forthcoming.

Hicks, J. (1942). *The Social Framework*. Oxford: Clarendon Press (1st ed).

Landefeld, S. (2000). GDP: One of the Great Inventions of the 20th Century, BEA Survey of Current Business, January, pp. 6–14.
https://apps.bea.gov/scb/pdf/BEAWIDE/2000/0100od.pdf

Lange G. M., Wodon, Q. & Carey, K. eds (2018). The Changing Wealth of Nations. World Bank.
https://openknowledge.worldbank.org/bitstream/handle/10986/29001/9781464810466.pdf

Syverson, C. (2017). Challenges to Mismeasurement Explanations for the US Productivity Slowdown. *Journal of Economic Perspectives*, 31(2), 165–186.
https://doi.org/10.1257/jep.31.2.165

Zenghelis, D., Agarwala, M., Coyle, D., Felici, M., Lu, S. & Wdowin, J. (2020). Valuing Wealth, Building Prosperity. Working paper.
https://www.bennettinstitute.cam.ac.uk/publications/valuing-wealth-building-prosperity/

Building Indicators for Inclusive Growth and its Sustainability: What Can the National Accounts Offer and How Can They Be Supplemented?

Didier Blanchet* and Marc Fleurbaey**

Abstract – How can the national accounts be linked to the objective of obtaining an inclusive measurement of growth, integrating distributional issues and all determining factors of well-being, in both the short and long term? The accounts offer measurements of real income that have undeniable connections with the quantification of current well-being, but they ignore the non-monetary factors of such well-being and they do not allow for evaluation of its sustainability. We present a way of dealing with the first limitation, the notion of equivalent income. It fits well with the accounts approach, it has relatively strong normative justifications and it lends itself well to the micro-macro bridging exercises needed to evaluate inequalities. Creating overall measurements of sustainability seems much more problematic, as it is impossible to do so without projection models that go far beyond the framework of current statistical output.

JEL Classification: E01, I31, Q51
Keywords: national accounts, inclusive growth, equivalent income, sustainability

** Insee (didier.blanchet@insee.fr); ** CNRS - École d'économie de Paris (marc.fleurbaey@psemail.eu)*

This text is based on work by Fleurbaey & Blanchet (2013), to which readers should refer for more thorough analyses. An initial version of it was presented at the 16th symposium of the French National Accounting Association (Association de comptabilité nationale - ACN) in June 2017. The authors would like to thank the participants of that symposium and the three reviewers for their comments. The authors are solely responsible for any errors or omissions.

Received May 5, 2018, accepted after revisions March 7, 2019.
Translated from: "Construire des indicateurs de la croissance inclusive et de sa soutenabilité : que peuvent offrir les comptes nationaux et comment les compléter ?"
Citation: Blanchet, D. & Fleurbaey, M. (2020). Building Indicators for Inclusive Growth and its Sustainability: What Can the National Accounts Offer and How Can They Be Supplemented? *Economie et Statistique / Economics and Statistics*, 517-518-519, 9–24. https://doi.org/10.24187/ecostat.2020.517t.2020

The term inclusive growth has gradually spread to refer to growth that is not limited to that of the usual monetary aggregates (OECD, 2014). It covers two ideas, the idea of growth that includes all the determining factors of quality of life for individuals and the idea of growth that benefits everyone without leaving anyone by the wayside. The notion of inclusion can also extend to future generations, since the increase in current well-being should not come at the expense of their future well-being. This brings us closer to another traditional notion, that of sustainable development, as defined in the Brundtland report, with its three economic, social and environmental pillars (World Commission for the Environment, 1987).

Whichever of these concepts is used, they both raise the same issues for the national accounts. There is extensive literature on them, which was used as the basis for the Stiglitz-Sen-Fitoussi report published around ten years ago (Stiglitz *et al.* 2009). Recent reviews of this literature are provided by Coyle (2014), Gadrey & Jany-Catrice (2016) and Laurent & Le Cacheux (2016). GDP is the flagship indicator of the national accounts and does not claim to measure inclusive growth or sustainable development. From their introduction, the national accounts chose to focus on measuring production or activity rather than well-being, restricted to productive activities that result in monetary flows. Since GDP is an aggregated measurement, it also provides no information on the individual distribution of these monetary flows: the institutional sector accounts merely disaggregate those flows between the main stakeholder categories, which are corporations, households and government departments. Finally, GDP is a measurement of current activity, without any information on the sustainability of this level of activity or the well-being derived from it. The recurring issue, therefore, is determining how to supplement it. Can we use other indicators that are already available in the national accounts? Can we proceed by expanding their conceptual framework or should we work completely outside that framework, leaving national accountants to focus on their comparative advantage, i.e. the structured and as exhaustive as possible view of all monetary flows between economic agents?

The aim of this article is not to propose a systematic review of all these tracks and the way in which all or part of them are already implemented. An example of a much more in-depth survey is provided by Jorgenson (2018). The aim is only to present their analytical background, in as concise and educational a manner as possible, by returning to what we are ultimately attempting to measure, the degree to which the existing indicators do or do not measure it and what methodological problems we face if we want to go beyond their limits.

The paper is organised into four sections. The first section is positioned within the stylised framework of a single-good economy: it will allow several introductory clarifications on the link between measurements of current well-being, of its sustainability and the main concepts of national accounting: gross or net production, consumption, gross savings or savings net of capital depreciation. The difference between well-being, on the one hand, and production or consumption, on the other, will appear here quite easily: the first is not necessarily proportional to the other two, even if it depends positively on them. This is even more the case in respect of sustainable well-being.

This basic framework seems to validate the idea of a simple frontier to be drawn between national accounts centred on the measurement of part of the resources of well-being, and the actual assessment of such well-being. However, taking into account the multiplicity of goods and services blurs this border. Aggregating quantities of heterogeneous goods and services requires the choice of a common metric, and it is difficult to see how this can be done without reference to the relative well-being or utility derived from these goods and services. Aggregation based on prices is admitted only because prices are considered acceptable proxies for these relative utilities and, indeed, much effort is put into making this approximation as relevant as possible: investing in hedonic price calculations or trying to approximate the concept of "constant utility" price indices are well-known examples of this. What maintains the gap with the true notion of current well-being is the fact that these volume-price splits are at best a reference to an ordinal notion of well-being and, above all, the failure to take into account things that have no price as they are too far from the market for a monetary value to be easily imputed.

The rest of the article will be structured around this dual issue of aggregation by prices and determining the value of things that have no price. The two middle sections will focus solely on issues relating to current production, consumption and well-being. The first of

these two sections will examine in what sense volume-price splitting techniques can be said to go towards an assessment of the well-being content of production or consumption. The second one explores a possible avenue for incorporating the non-monetary determining factors of well-being: the pseudo-monetary approach based on the calculation of so-called "equivalent" incomes.

The final section revisits the issue of sustainability, in a more succinct manner, but still with the same problem of aggregation. The problem is to determine how to assess sustainability when it is not reduced to the preservation of a single transferable good but depends on a multitude of assets, whether produced or natural. The obstacles here are far greater than when measuring current well-being. The primary reason for this is the forward-looking nature of the question being asked, which forces us to question the future well-being content of these assets. Finding keys to aggregation requires making long-term projections about a future unknown by nature, which leads far beyond the standard statistics centred on the exploitation of directly observable data.

1. Production, Well-Being and Sustainability in a Stylised Single-Good Economy: Which Indicators Should Be Given Preference?

Let us start with the most rudimentary framework possible, that of an economy based on a general purpose single good, both a consumption good and a production good. This will help to set some base ideas on the link between measurements of production, well-being and sustainability, temporarily leaving aside the issue of the diversity of the goods produced and the existence of non-monetary components of well-being or those that cannot be monetised directly.

Adopting usual notations, K is the physical quantity of capital available on a given date, and L the amount of work provided by the agent(s). Combining them makes it possible to produce a quantity $Y = F(K, L)$ of the single good. Part of this production, C, will be consumed, while the other part will be saved. The savings rate is noted σ. Savings $S = \sigma Y = Y - C$, will first be used to offset capital depreciation δK, and, where sufficient, to increase the stock thereof, by the amount $\sigma Y - \delta K$.

Within such an economy, GDP is a measure of Y but not of current well-being in the cardinal sense of the term. Within this stylised framework, cardinal well-being is generally represented as a function $U(C)$, as the unconsumed part of the production does not generate current well-being. The most that can be done to reconcile the concepts of national accounting and the measurement of well-being is to consider C as one among all the possible parameterisations of this function $U(C)$, compatible with ordinal preferences, but whose limit is to ignore that perceived cardinal utility may not grow linearly with material consumption.

Let us now move onto the measurement of sustainability. As it does not measure well-being, Y measures the sustainability of that well-being even less. Formally, current well-being is said to be sustainable if its level can be reproduced to infinity, i.e. if, from the current state, there is at least one feasible trajectory ensuring, on any given date, a level of well-being that is never lower than that of the current period. Within the very simple economy considered here, there is an obvious criterion of sustainability, which is having a net savings rate above zero. If that is the case, there will be $K(t+1) \geq K(t)$, it will again be possible to consume C while leaving an amount of capital $K(t+2) \geq K(t+1)$ making it possible to do the same on the date $t+2$ and so on. Net savings or the variation in "wealth" K are thus the right concepts to measure the sustainability of this very simple economy.

At the same time, this framework allows us to understand the limitations of another quantity evaluated by national accounts, net production $Y_{net} = Y - \delta K$, which has sometimes been presented as an alternative to GDP as a measure of both well-being and its sustainability (Weitzman, 1976). This net GDP is effectively related to both of these notions. Based on the foregoing, Y_{net} measures the maximum level of sustainable consumption, since consuming a maximum of $Y - \delta K$ makes it possible to generate savings at least equal to δK, which exactly offset the capital depreciation. However, this is where the contribution of net GDP stops. The observation of Y_{net} taken in isolation is not sufficient to say whether we are on a sustainable trajectory or not. What is needed is to know whether actual consumption is higher or lower than this threshold Y_{net}. The correct sustainability indicator is still the net savings rate. It is the net savings rate and not Y_{net} that tells us whether or not there is overconsumption of what is produced as at the current date: this applies to

the usual asset produced and will also apply in our final section to environmental assets. The limitation of the indicator Y_{net} stems from the fact that this measure of current well-being and the measure of sustainability require at least two figures: by claiming to summarise both notions using a single figure, net production can measure neither of them.[1]

Now a few words on taking inequalities into account. Although highly simplistic in its description of the world of goods, the analytical framework of this section does not preclude taking into account a form of heterogeneity of individual situations. Indeed, many inequality analyses are implicitly placed within this single-good framework or, more specifically, they accept the homogenisation of the world of goods and services implicit in all monetary statistics (Alvaredo *et al.*, in this issue). Within this framework, for example, it is possible to replace the measurement of average income or average consumption with generalised averages, in the form proposed by Atkinson (1970):

$$W_{1-m} = \left(\frac{1}{n} \sum_{i=1}^{n} c_i^{1-m} \right)^{1/(1-m)}$$

in which m is the inequality aversion parameter: the scenario $m = 0$ returns the usual average, thus a total absence of taking inequality into account and the focus shifts towards increasingly disadvantaged individuals as the parameter m increases.

The issue of inequality can also play a role in an expanded definition of sustainability. If the collective well-being function involves inequality, the sustainability of collective well-being implies control of inequality dynamics. In this case, the preservation of the stock of capital K is only a necessary condition for sustainability: it must be accompanied by allocation mechanisms making it possible to ensure that the future benefit of this stock of capital is not increasingly appropriated by a part of the population. However, this means that the question of sustainability can no longer be answered by observing the net savings rate alone; it is necessary to add to it modelling of inequality dynamics. This last point gives a first taste of what will be the main message of the final section: except in hyper-simplified scenarios, evaluating sustainability cannot be limited to a simple instant accounting of flows and stocks; it is the full dynamics of the system that must be modelled. National accounts can provide a part of the data needed for this modelling, but they alone cannot deliver sustainability and non-sustainability messages.

To sum up, within the very basic framework from which we started, there is a range of indicators that fit together or complement each other in a fairly evident way: stock of capital, consumption, and net and gross production and savings, etc. What GDP measures is production Y, which is obviously not sufficient to fully describe the state of the economy. Net GDP provides interesting additional information, assuming a sufficiently precise measurement of the depreciation of capital, which is not an easy task, but that is not sufficient either. The reason is that, attempting to measure both the standard of living and its sustainability, it measures neither; measuring two distinct phenomena requires a pair of indicators. A good option would be the pair that combines current consumption and the net savings rate. It provides the same information as the net GDP/net savings pair, but in a form more directly oriented towards the joint measurement of current well-being and its sustainability. However, this solution remains unsatisfactory because consumption is merely a fairly poor proxy for well-being. This proxy ignores the fact that the relationship between consumption and cardinal well-being is not necessarily linear. All the physical accounting that is proposed here does not therefore answer the question of the utility that is really derived from the different quantities that are measured. This problem of the non-observability of well-being will arise with greater relevance when comparing two economies in which the preferences of agents are not necessarily the same and/or because well-being also depends on factors that are not produced and are therefore not measured by C or Y.

The situation will be even more complex once out of this framework of a single all-purpose good, but with the paradoxical effect of forcing a partial overlap between measures of production and well-being, as both measures face a common problem of the relative valuation of different goods and services. The diversity of goods and services will only add to the complexity of both the measurement of the current situation and that of sustainability. In the following two sections,

1. For the record, this point has been clearly identified and addressed in one of the founding texts of the "beyond GDP" literature, that of Nordhaus & Tobin, who proposed two versions of their "measure of economic well-being", MEW-A ("actual") measuring current well-being, and MEW-S ("sustainable") measuring sustainable well-being. It is from comparing the two that a message on the sustainability of the current living standards could be drawn (Nordhaus & Tobin, 1974).

we will leave aside the prospective question of sustainability and focus on the issues of current production and well-being.

2. Production, Income and Well-Being in the Presence of Multiple Goods: What Do Standard Indicators Say and Do Not Say?

The single good approach is obviously only a heuristic convenience. To what extent does the diversity of goods complicate the reading of the aggregates produced by the accounts? This issue has been debated since the 1940s (Hicks, 1940) and this debate played a major role in the shift away from the objective of measuring well-being[2]; however, it did not prevent the problem from repeatedly coming back. A new illustration of this is currently provided by the debate on the mismeasurement of growth, i.e. the capacity of GDP to measure the contribution of new forms of innovation made possible, *inter alia*, by the development of the digital economy.[3] Participants in this debate generally acknowledge that the purpose of national accounts is not to measure well-being. However, even when focusing on the volume of production, we inevitably end up looking for a common metric that can make it possible to aggregate volumes of production of all goods and services, both old and new, and we do not see what other theoretical metric to refer to than the utility that is derived from each of them. Aiming to disconnect completely the measurement of GDP from any reference to the notion of utility or well-being is a position that appears difficult to maintain. The way economists approach the subject inevitably requires using utility functions and other concepts provided by consumer theory.

To keep this paper as short as possible, we will not go into further detail on the issue of the renewal of goods, which is covered in another contribution to this issue, that of Aeberhardt *et al*.. We will restrict ourselves here to the simpler scenario in which the list of goods is fixed, even limiting ourselves to a scenario in which there are only two goods. This section and the next will also ignore the inter-temporal dimension: everything that is produced is immediately consumed, which will allow us to speak indifferently of production, income or consumption, to focus on how these notions both differ from and are connected with that of well-being. This framework will make it possible to show how indicators of volumes and well-being can diverge even more than in the single-good framework, with a risk of conflicting messages on the direction of developments. However, we will also see how methods aimed at avoiding this risk reintroduce some link between measures of volume and utility, in the ordinal sense of the term.

Thus x_1 and x_2 are the quantities of these two goods produced and consumed and we use x to write the pair $(x_1;x_2)$. It is assumed in this section that these are two market goods. The problem with aggregation is determining how to summarise the change in the quantities of these two produced and consumed goods using a single figure. Let us imagine, for example, a reference basket $x=(1;1)$ and another economy or the same economy at another time, with the basket $x'=(2;2)$. In this first example, it is commonplace to claim that both production and consumption are doubled when moving from one situation to the other, though it is not possible to be as certain with regard to well-being. However, what can be said about the magnitude of the increase if production or consumption changes from $x=(1;1)$ to $x'=(1.5;2)$, and what can be said about the direction of this development in the ambiguous scenario in which it changes to $x'=(0.5;2)$?

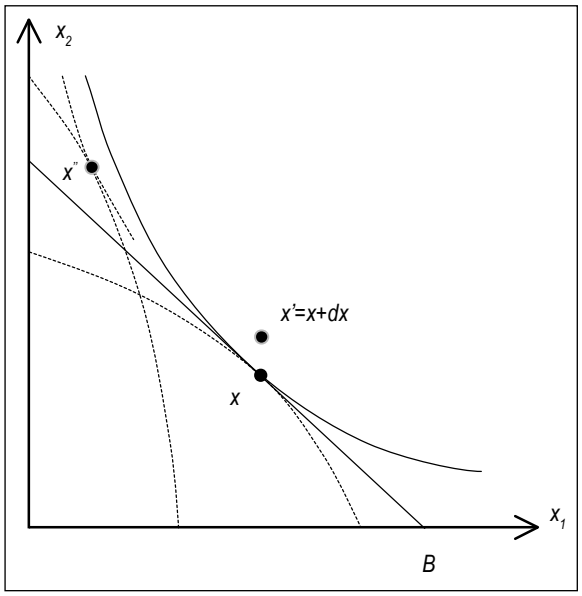
Faced with this question, the pragmatic response is to rely on the prices observed in the reference situation, i.e. $p=(p_1;p_2)$. Production or consumption x' are said to be higher (or lower) than production x if the aggregate at base prices $px'=p_1x'_1+p_2x'_2$ is higher (or lower) than the initial aggregate $px=p_1x_1+p_2x_2$, i.e. if the Laspeyres Index px'/px is higher (or lower) than one.

Such a calculation will not solve the issue of measuring cardinal well-being any more than the one-dimensional quantity x of the previous section, but the problem may go even further, as an error risk will also arise concerning the sign of its variation, which is therefore an error concerning the ordinal message.

Figure I sets out the problem assuming that the initial level x corresponds to a market equilibrium that maximises utility $U(x_1;x_2)$ (a concave indifference curve) under the production frontier represented by the convex curve at the bottom. The straight line B describes the

2. *See also the survey by Sen (1979) from the late 1970s and Vanoli's (2002) developments on this subject.*
3. *For an overview of this debate, see Blanchet* et al. *(2018).*

Figure I – **Well-being and volume indices at constant prices**



Reading Note: The production frontiers correspond to the convex curves. The indifference curves are the concave curves. From the initial equilibrium *x*, point *x'* corresponds to an improvement both in well-being and of the volume index at initial equilibrium prices, *px'/px*, as it is above budget constraint *B* of this initial equilibrium. This is not the case for point *x''*, which is positioned between the straight line *B* and the indifference curve passing through *x*.

budget constraint under which the maximisation of U is carried out. It is tangential to the two curves with a slope $-p_1/p_2$ corresponding to the equilibrium price system. The movements of x verifying $px'/px > 1$ are the set of movements that pass above the straight line B. As long as these movements are marginal, the tangency to the indifference curve at point x ensures that there will also be an increase in U: this is indeed the case in respect of point x'. In this case, the increase in the Laspeyres Index for production reflects an increase in well-being. Only the quantitative problem of the previous section remains: we know that well-being rises, but we cannot say by how much since we do not know to which quantitative levels of U the indifference curves passing through x and x' correspond.

However, this qualitative message ceases to be correct in the other reference scenario illustrated by point x''. In this case, the movement is not marginal. This point is always located above the straight line B, therefore giving $px''/px > 1$; however, it is on a lower indifference curve than that of the initial point x. Thus, there is an increase in the aggregate and a decrease in well-being. A Paasche Index based on prices associated with the state x'' would avoid this problem, as we have $p''x'' < p''x$, but this does not help since we do not know *a priori* which index to use when they send contradictory messages. This property applies to all points positioned between the budget straight line B and the indifference curve passing through x. The scenario involving marginal variations $x + dx$ only made it possible to avoid this problem because of the tangency between B and this indifference curve.

The same problem arises again when measuring production. We could have imagined that sending an incorrect message about well-being does not prevent having a correct message about production. However, this is not the case. Point x'' is indeed positioned above the initial production frontier, which would argue that production has increased. However, if this point also corresponds to a market equilibrium, it must result from a new production frontier of the type shown in the dotted line crossing the first one; therefore, it cannot be said that scenario x'' corresponds to an economy that is more productive than the one with the equilibrium point x. It is even rather tempting to say that this production is lower, since it can only provide a lower level of utility. This example illustrates the false simplicity of the notion of production: in the ambiguous scenario in which production decreases for one good and increases for the other, it is impossible to say whether total production is increasing or decreasing without reference to the way in which the consumer values such productions, therefore to the relative utilites. To ensure it is relevant, the measurement

of production has to rely on a well-being related metric.

Two responses to this problem will in fact go in the direction of a partial reconciliation between the measurement of quantities and the measurement of well-being, confirming the difficulty of completely disconnecting the two notions. The first is the replacement of the calculation using base prices with a calculation of volumes using chained prices from the previous year: the idea of chained prices is to decompose the non-marginal move from one basket to another as to a sequence of small variations of type $(p,x) \rightarrow (p',x')$, for which there is at least the assurance of having good qualitative information on the evolution of the well-being content of what is produced. The other is deflation using constant utility price indices, of which the use of volumes at the previous year's chained prices can be presented as an approximation. A constant utility price index indicates by how much income must change to maintain a reference level of utility in the presence of a price variation: it therefore measures the change in the price to be paid to obtain a given level of utility. Although this is only a theoretical reference that the practical indices can, at best, approximate, it is the most appropriate one for a good conceptualisation of what these price indices seek to measure (Triplett, 2001) and the term constant utility is obviously very illustrative of the link with the well-being approach. This link is also apparent in the use of hedonic price indices, another of the techniques used to improve volume-price splits.

Online Appendix C1[4] indicates more precisely how chained prices or constant utility price indices tend towards the estimation of a notion that is linked to the notion of consumer well-being or utility, and it also makes the link with the so-called equivalent income approach, which is the one that will be used hereinafter as one of the avenues allowing the non-monetary determining factors of well-being to be taken into account: equivalent income measures the minimum budget required to reach the utility level of the basket of interest under a price system chosen as the reference price system. The three approaches converge towards the same notion of volume or real income when assuming homothetic preferences that can be represented by a function $U(C_1,C_2) = F(G(C_1,C_2))$ with F a monotonic function and G a homogeneous function of degree 1. In this scenario, the volume-price splits make it possible to estimate the function G, in other words, broadly speaking, what C was for $U(C)$ in the single-good scenario, but incorporating a significant portion of the properties of $U(C_1,C_2)$, those reflecting the relative marginal preferences for different goods. Samuelson & Swamy (1974) speak of a "cardinal indicator of ordinal utility", a term subsequently used by Sen (1979), while stressing its ambiguities. Such ambiguities can be avoided by reserving the term "well-being" for U and keeping the more traditional terms of "volume of consumption" or "standard of living" for G. Nevertheless, this function G incorporates a significant amount of information on the function U, all that which concerns the extent to which substitution is possible between the goods, and we will see hereinafter that there are ethical arguments for considering that it is on the basis of this concept of standard of living that inter-personal comparison should be made, rather than on the basis of cardinal well-being.

Moreover, the volume-price splits are not the only area in which the national accountants' initial concept of market production is led to incorporate elements borrowed from the question of well-being, even without formally using the equivalent income approach, which we will discuss hereinafter. Having expanded the initial scope of market GDP to include public administration production basically stems from the idea that such production has to be counted as contributing to the well-being of the population: it would be extremely embarrassing to have a GDP that signals that living conditions are worse where a greater number of services are provided collectively and funded through taxation. Another specific form of production considered by the national accounts is that of housing services that homeowners households are deemed to provide to themselves: classifying these services as production is very conventional, the real reason for their inclusion in the accounts being to prevent GDP from showing a lower standard of living, and thus lower well-being, in countries with a higher proportion of homeowners.

The positioning of national accounts on the measurement of well-being is thus less clear-cut than suggested when it is claimed that the objective of GDP is "only" to measure production. The reason is that it is difficult to develop a measure of production without any reference to the well-being content of what is produced. National accountants, even though they defend

4. *See the link to the Online appendices at the end of the article.*

themselves from doing so, put a lot of effort into avoiding an excessively wide gap between what they measure and a certain notion of well-being.

However, these efforts are only able to achieve half the job. The example of self-produced housing services is one where a monetary equivalent can be easily imputed, based on observation of the rental market. But what is the best way to proceed in the total absence of such references? Should we abandon any idea of monetisation, which takes us in a very different direction from the national accounting approach, or can we resort to indirect monetisation, and with what theoretical or normative justifications?

3. The Non-Monetary Components of Well-Being: How Should They Be Aggregated?

So far, several types of solutions have been adopted or proposed in response to this issue of aggregation of the monetary and non-monetary components of well-being. The first is rather a non-response or, more precisely, it consists in acknowledging the impossibility of a shared response. This is the dashboard approach, which entails multiplying indicators that shed light on the different aspects of well-being. In a sense, these dashboards are unavoidable. At some point, it is necessary to go back to area-by-area investigations, and we will ultimately come to the conclusion that aggregation comes up against insurmountable limitations: it is necessary to learn to give up aggregating what cannot be aggregated. However, the problem with dashboards is their opposite tendency to provide too much information, in a manner that proves difficult to order and summarize, with the emblematic example being the sustainable development indicators adopted by the United Nations to monitor the 17 goals of its 2030 agenda (Cling *et al.*, 2019). Structured summary information is also required.

There are two other ways of obtaining aggregate indicators: first, the calculation of composite indicators such as the Human Development Index (HDI), which uses a statistical rule to aggregate GDP per capita, life expectancy and educational attainment, seen as the three essential components of well-being, and second, the measurement of subjective well-being.

The problem with the composite indicator approach is that it introduces a high degree of arbitrariness in the way it combines its arguments, and it can prove very problematic. Ravallion (2013), for example, details the implicit trade-offs that the HDI makes between GDP per capita and life expectancy depending on the level of development of the countries, showing how they can be deemed questionable.

The subjective approach has the advantage of bringing us closer to the notion of cardinal well-being, but at the risk of many of the biases (see Accardo, in this issue). In brief, its advantage is the fact that it is based on individual quantitative information that is fairly easy to collect and can be directly manipulated to produce aggregate indices. This information is deemed to respect the individual preferences of respondents, rather than the arbitrary weights used in composite indices and, in principle, it is these respondents who are best placed to know what is important to them. However, the problem is the lack of visibility regarding how individuals express their satisfaction in the scoring grid proposed to them by the interviewer: two individuals with similar living conditions may score their living conditions very differently. It can be said that the subjective nature of the measurement is both its advantage and its limitation: it is interesting to know how people evaluate their lives, but this does not necessarily provide a valid benchmark for interpersonal comparisons or, to a greater extent, for comparisons of living standards between countries and over time.

Therefore, we will focus here on a third approach, the pseudo-monetary approach based on the calculation of so-called "equivalent" incomes. It uses the ordinal representation of well-being, with good normative justifications, and it is the one that best fits with the national accounts' general framework. These properties do not necessarily justify giving this approach exclusive preference, but they nevertheless invite further exploration.

This notion of equivalent income, like the subjective approach, respects individual preferences, using weights for the determining factors of well-being that are consistent with these preferences. What this approach will have in common with the national accounts is that it expresses results in monetary units. Of course, there should be no misunderstanding about the meaning of this choice of unit: it is only a measurement benchmark, which does not mean that all the items considered can and must be produced and exchanged on the market. The approach includes both the case of market

goods and non-market goods that are to remain as such.

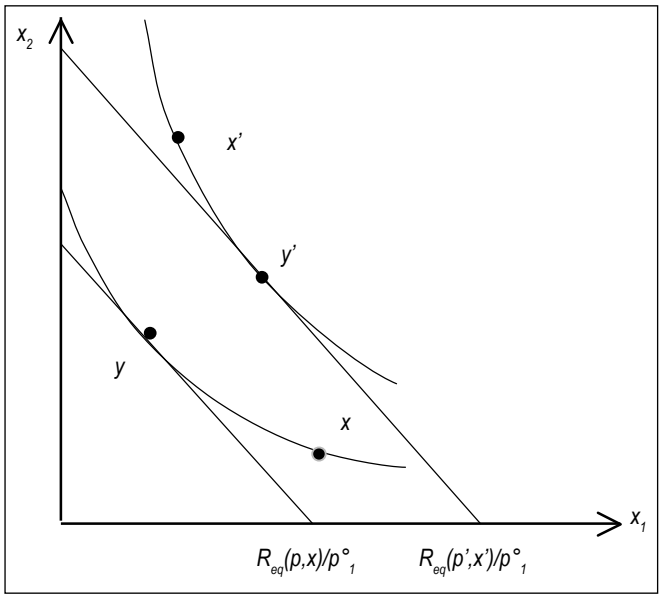
In the first case, in which there are only market goods, the equivalent income method consists in establishing a reference price system $p°$ and valuing the pairs (p,x) located on the different indifference curves by the amount $R_{eq}(p,x) = min(p°y; U(y) \geq U(x))$, which gives the minimum level of income required, under the reference price system $p°$, to achieve a level of utility at least equal to that obtained by the basket x under the price system p, taking into account the possibilities of substitution between goods along the indifference curve passing through x.

Figure II shows how this method allows a unique scalar to be associated with each indifference curve, with equivalent incomes normalised by using good 1 as the numeraire. In this example, the same levels of "utility" are associated respectively with the pairs (x,y) and (x',y') and these quantities make it possible to create a hierarchy for the two baskets x and x' for which a classification based on quantities alone would have been impossible since we have $x_1 > x'_1$ and $x_2 < x'_2$.

From there, it is easy to extend the same idea in case the two states to be compared differ not only in the price system and the resulting consumption of market goods, but also in the levels of a certain number of non-monetary factors of well-being. We will use $l = (x,e)$ to denote the extended consumption or production basket merging the goods x with a market price p and the vector e of the non-monetary determining factors of well-being, and we adopt the reference values $(p°,e°)$ for both p and e. R remains the monetary income px in the observed state (p,e). We term equivalent income equivalent the monetary income $R_{eq}(x,p,e,p°,e°)$ necessary to achieve the same level of utility under the reference conditions $(p°,e°)$ as under the observed configuration (x,p,e). The adoption of a unique reference vector $(p°,e°)$ makes it possible to compare the levels of well-being of any individuals whose situations differ in terms of consumption x and the levels of non-monetary factors, and who are not necessarily exposed to the same price system.

This method is already implemented in at least one relatively common production of public statistical systems, the computation of equivalence scales allowing the comparison of living standards across households of different compositions: the household structure is indeed a non-monetary parameter of the standard of living, for which a monetary equivalent is proposed by evaluating how much the monetary income of the household must be increased in order to preserve the levels of utility or well-being of its

Figure II – **Equivalent income with two market goods**



Reading Note: We want to compare the two baskets x and x' observed under the price systems p and p'. This is done by evaluating the two baskets y and y' providing the same levels of utility under a common reference price system $p°$. The equivalent incomes are the associated monetary incomes. Their levels can be read on the horizontal axis, after division by the price $p°_1$.

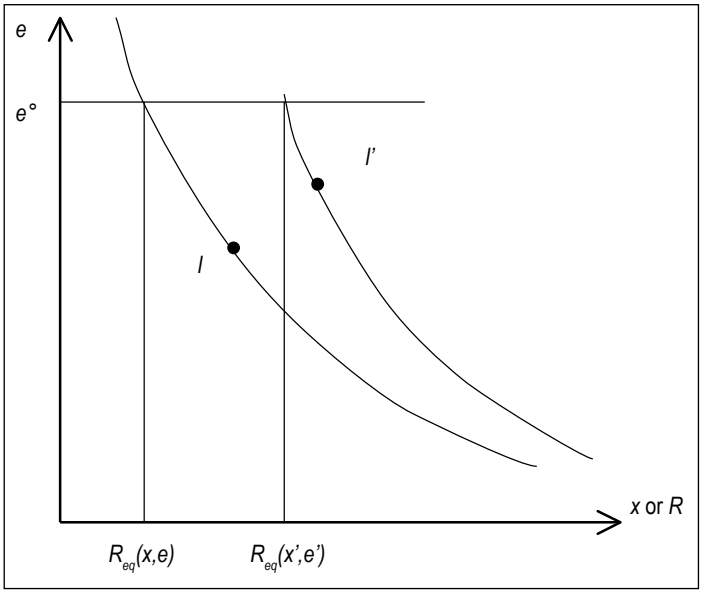
members when its size increases. The general principle is shown by Figure III in the scenario in which both x and e are one-dimensional, with the good x used as the numeraire, which makes it possible to equalize x and R. To make the link with the HDI, let us assume that e is the health status rather than household size and that the reference state used is good health. In other words, we will try to calculate the loss of income that, for an individual in good health, leads to the same drop in well-being as being in poor rather than good health. The equivalent incomes of two individuals in the situations $l = (x, e)$ and $l' = (x', e')$ can be read directly as the abscissa of the intersection points between their indifference curves and the horizontal line of level $e°$. In the example shown in the graph, the individual l' combines better health and higher income. Her equivalent income takes both of these factors into account.

We can see how this approach differs from both the subjective approach and the composite indicator approach. An indicator of subjective well-being will eventually reveal that the individual l is happier than the first one, if she is naturally undemanding and/or used to her lot in life. The equivalent income approach chooses to ignore this issue of character. But it takes into account the way in which this second individual weights material goods and health in the evaluation of her well-being, as opposed to the a priori weighting that would be assigned to them by a composite index, and it takes into account the way in which the individual's preferences would cause them to modify their basket of goods in response to a change in the price system with which they are confronted, if x is multidimensional.

We are thus on the middle path between the ignorance of individual preferences, which characterises composite indicators, and the complete taking into account of the satisfaction reported in the subjective approach. What the approach does take into account is a sub-set of the utility function's characteristics, those determined by ordinal preferences. In contrast, it neutralises everything that shifts the focus from ordinal preferences to cardinal well-being, including the fact that a basket of goods that is double the amount of another does not necessarily provide twice as much utility. We again find the distinction presented above between the notion of standard of living and its translation into perceived cardinal well-being.

Working on preferences corrected like this can be ethically defensible. The assessment of resource allocation needs not take into account the fact that individuals may have more or less demanding natures, except to recognise that policies should seek to systematically compensate individuals who are more dissatisfied than others

Figure III – **Equivalent income when well-being depends on a market good and a non-market factor**



Reading Note: Well-being depends on a market good x that is used as the numeraire (hence the integration of x and monetary income R) and on a non-market factor e. We want to compare the combinations $l=(x,e)$ and $l'=(x',e')$. We use a reference level $e°$ for e. The equivalent incomes $R_{eq}(x,e)$ and $R_{eq}(x',e')$ are those providing the same utility levels as l and l' for e and e' reduced to the common value $e°$.

by disposition. As regards the phenomenon of globally decreasing marginal utility, the idea is not to ignore it completely but rather to reintroduce it in a second step, when moving to social utility functions of the generalised average type in which this decreasing marginal utility allows us to account for the phenomenon of aversion to inequality. In addition, the way in which inequality is taken into account is preferable to approaches that would measure inequality on the monetary and the various non-monetary axes separately and would aggregate the various inequality indices thus obtained. When there are individual accumulations of handicaps on these various axes, the correct approach is to first assess the impact of these accumulations at the individual level, otherwise the overall impact of these different dimensions of inequality is minimised.

Of course, this approach itself raises certain questions, first of which is that of its implementation. Several pieces of work have attempted to apply it to a more or less wide range of non-monetary dimensions of well-being, including work by Fleurbaey & Gaulier, (2009), Murtin *et al*. (2015), Boarini *et al.* (2015, 2016), Decancq *et al*. (2015), Decancq & Schokkaert (2016) and Jones & Klenow (2016). Three types of techniques are possible in principle: (a) relying on calibrations of preferences, as revealed by behaviours, (b) the use of contingent evaluation techniques, i.e. direct questioning of individuals' willingness to pay for or receive given changes in their situations or environment, and lastly, (c) relying on subjective satisfaction data. We will focus on the latter in particular, because of its link to what was presented earlier. The idea is to obtain estimates of the degree to which individuals are willing to make trade-offs between material factors and other aspects of living conditions, by empirically analysing how they each affect subjective well-being, which is possible with surveys that combine a direct measurement of perceived well-being and objective components. Typically, if we have a measurement S of perceived well-being, regressing S on quantities x and y will provide coefficients the ratio of which can be interpreted as a measurement of substitutability between x and y. This is, of course, assuming that the various factors that bias the measurement of S do not hinder the estimation of these different coefficients: this will be the case if it is assumed that the noise that affects the measurement of S correlates neither with x nor y.

The other main question is that of the link between the practical and the ethical. The fact that the method requires the selection of reference values (p°, y°) means that it is necessary to establish principles on which to base that selection. These principles are fairly easy to establish when the non-monetary factor to be taken into account has unambiguous monotonic effects on well-being: either the highest or the lowest value of this factor is used, for example, the state of good health, which amounts to giving a monetary equivalent for the "disutility" associated with different levels of poor health. The choice is more complicated for a variable combining utility and disutility, such as working time. Here, the problem is determining the degree to which the individual would accept a decrease in their income decrease (resp. would like it to increase) to move from their actual working time d to a reference working time d°. However, total idleness $d^\circ = 0$ is not a more attractive reference than the maximum possible working time, as having productive work is also a factor of well-being. The result of the calculation may be sensitive to the selection of this reference working time and there is not always an obvious standard for setting it.[5]

4. Measuring Sustainability

The problems of measuring current well-being having been clarified, if not fully resolved, let us briefly revisit the question of its sustainability, outlined in the first section. How would it look if we were able to move towards a shared measurement of this current well-being? The first section gave the answer within the basic framework of a single-good economy. Assessing the sustainability of consumption and thus of current well-being was reduced to the calculation of a net savings rate, with the sole difficulty of knowing the rate of capital depreciation.

The important element of this first result, which may seem obvious but has not always been so in the search for indicators of sustainable growth, is re-emphasised here: the clear separation that is thus made between the measurement of sustainability and of current well-being. It is opposed to the idea that the measurement of sustainability could be reduced to calculations of green GDP and also to some attempts to calculate composite indices of sustainable development that mix the measurement of current progress and of its sustainability. It is fairly easy to understand that

5. *On this issue, see Appendix 1 in Fleurbaey & Blanchet (2013).*

by attempting to measure two different things using a single figure, neither of those things is measured. As its name suggests, current well-being is relating to the current situation. Sustainability is an issue relating to the possible prospects of the evolution of this well-being. Attempts to combine the two pieces of information in a single figure can only be explained by the obsession with establishing an international ranking of good or bad sustainable development practices, but it is clearly inappropriate. The one-dimensional approach inevitably leads to the risk of classifying countries with a high level of well-being but little concern for their future or that of the planet as a whole in the same manner as more sober countries that are more capable of sustaining their current way of life.

Presented in this manner, the problem of measuring sustainability thus consists of calculating one or more indicators that can alternatively be presented as indicators of net savings, net investment, over-consumption/ over-exploitation of resources, or even as a variation of an expanded notion of capital. This approach was initiated at the World Bank (Hamilton & Clemens, 1999; Lange *et al.* 2018), illustrated by Arrow *et al.* (2004), then taken up and expanded upon since 2012 under the name "inclusive wealth" as part of the United Nations Environment Program (2018).

All these terms do not necessarily have the same connotation: the term net investment is more reminiscent of the idea of renewing productive capital in the traditional meaning of national accounts, the terms over-consumption or over-exploitation are more reminiscent of the idea of the over-use of natural capital, which more clearly brings out the relationship with the notion of ecological footprint or its particular variations, such as the carbon footprint. Formally, all these terms refer to a common problem that is an extension of that in the first section: the fact that, in practice, sustainability does not boil down to the preservation of a unique productive resource, measured by K, it will depend on the evolution of a very wide range of assets K_i. It will simultaneously include the various components of natural capital, human capital, physical productive capital, financial capital and various forms of intangible capital, the list of which can be very wide ranging: one spontaneously has in mind the stock of knowledge and expertise, but the sustainability of our way of life also depends on the durability of several other intangible elements, such as the quality of institutions or social relationships.

What we are facing is therefore the same type of problem as that faced in the two previous sections, that of the heterogeneity of the "goods" or, more generally, of the items to be taken into account in the assessment. How can we hope to reduce a multiplicity of factors to a single figure for sustainability? There is little doubt that the problem will be at least as difficult as in the case of measuring current well-being. It is in fact much more difficult, as illustrated by the recurrent debate between "weak" and "strong" visions of sustainability, with the former interested in the expanded version of the list of K_i and considering that an increase in some of them can perfectly well compensate for the disappearance of others, while the latter focused instead on a subset of environmental assets deemed critical, refusing to consider the possibility of substituting them for non-environmental assets, with therefore very different sets of explicit or implicit weightings for each of these two approaches.

What should be done in this context? Market prices cannot be used as references. They can be accepted as proxies for the relative values to be assigned to different goods and services, when the issue is only that of measuring current well-being. Clearly, they can no longer fulfil this role in assessing sustainability, if only because some of the assets of interest cannot be assigned a market value. This leads back to an imputation problem, though one much more complex than imputing monetary equivalents to the non-monetary components of current well-being.

The theoretical answer to this question is detailed in Online Appendix C2. First, it involves monitoring "physical" measures of the various sub-components K_i of the "expanded" capital, as at date t, each of these items shows a net variation dK_i. In the case of exhaustible natural resources, this net variation will automatically be negative. In the case of renewable natural resources, this change will compare the drain on these resources and their regeneration, whether spontaneous or amplified by voluntary environmental restoration policies. In the case of productive physical capital, accurate measures of its volume and depreciation are needed. Concerning other categories of assets, it is for human capital that the exercise seems the least inaccessible, with the possibility of valuing human capital by the future income flows it is

likely to generate (on this point, see Canry in this issue). In contrast, huge difficulties can be foreseen in the case of intangible capital.

However, let us assume that this first step of calculating the dK_i is dealt with. In order to weight them, the contributions of each of them to the flow of future well-being must then be evaluated. The marginal value of an asset should be assessed by quantifying what its variations add or take away from the future flow of well-being. This can only be done by modelling these trajectories of well-being and the way in which they would be affected by more or less consumption of the asset under consideration at the current date, based on a comprehensive and integrated simulation of the economic, social and environmental dynamics, starting from specified initial conditions. This therefore requires much more than a set of separate evaluations of different assets. What we need to know is how the variation of one affects the dynamics of all the others. For example, in the case of climatic "capital", what is theoretically needed is a comprehensive cost-benefit evaluation of the long-term effects of current greenhouse gas emissions (on this point, see Germain & Lellouch in this issue). If this is achieved, it can become possible to overcome the conflict between the notions of strong and weak sustainability. An indicator that is considered "weak" and aggregates the variations of the different assets on a linear basis remains quite capable of addressing the issue of strong sustainability if the reaching of critical thresholds is reflected in the form of very high values imputed for the most affected natural assets, making any compensation by the accumulation of non-natural assets impossible (Fleurbaey & Blanchet, 2013).

The work conducted in this area is increasingly pointing towards the finding of a lack of sustainability for a significant number of countries. For example, the 2018 edition of the Inclusive Wealth Report covers 140 countries monitored since 1992 and shows a decline in natural capital in 127 of them, with a decline in overall inclusive wealth in 44 of them (United Nations, 2018). However, in spite of the efforts that have been developed, this approach continues to expose itself to criticism of insufficiently taking into account the environmental constraints (Roman & Thiry, 2016).

In addition, though presented as logical extensions of the standard national accounts analytical framework, such "expanded" accounting approaches are clearly outside the scope of normal statistical output. They cannot be based solely on the observation of current data, they force the confrontation of assumptions about what these dynamics are thought to be, leading at best to evaluation brackets. What the statistical system in general, and the national accounts in particular, can do is to feed base data into these exercises, i.e. evaluations of some of the K_i and of their variations, partially aggregated where possible, but without being able to pretend to step out of this role of supplier of base data.

* *
*

Let us recap the main lessons to be taken from this article. Much emphasis has been placed on what differentiates GDP and other national accounts indicators from the measurement of well-being, but also on what makes them related. The aim of measuring well-being was quickly abandoned when the tools of national accounting were introduced, with the kind of argument discussed in the second section: even when all goods have correctly measured prices, a volume index evaluated at constant prices provides, at best, an information on the direction in which well-being is changing, not on the intensity of its change, and good information on the direction of its changes is only guaranteed for marginal changes in quantities. It is this finding that has led national accountants to put forward a more modest and pragmatic conception of their tool: national accounts as an overall picture of the monetary and physical flows between economic agents, and GDP as the main summary of this overall picture, representative of the economic activity and overall income, particularly for the requirements of cyclical macroeconomic regulation and the steering of public finances.

There are reasons for this position, but it is also ambiguous, as the objective of measuring well-being remains indirectly present in many of the choices that have been made as the conceptual framework of the accounts has been enriched: attempting to quantify public sector production as well as possible, integrating into GDP a formal production of housing services that homeowners provide to themselves and improving the volume-price splits using hedonic pricing methods or by trying to get as close as possible to the notion of constant utility price indices are all ways of preventing GDP from deviating too far from what can be spontaneously expected from a well-being index. We

do not want GDP to show as less well-off the inhabitants of countries where a large number of services are provided outside the market, or countries where the vast majority of the inhabitants are homeowners. Furthermore, we do want GDP, through its deflator, to give the best possible account of the contribution to well-being of the decreasing costs of many products, or of the replacement of existing products by cheaper and/or higher performing ones. It is difficult to sustain such efforts while simultaneously claiming to be free from any concern for measuring well-being. GDP is not well-being, but it cannot be conceptualized independently from it (Schreyer, 2016); it would be counterproductive to ignore or downplay this link, both for users of the accounts and for national accountants themselves.

Nevertheless, this link well-being remains only very partial and the central framework of the accounts is not the right place for going beyond this partial character. There would undoubtedly be more to lose than to gain because, by aiming to take on too much, the main contribution of this central framework would be lost: its function as an information system on all current monetary flows between agents and on the monetary value of the assets they hold.

These shortcomings of the national accounts should rather be addressed outside their central framework. The composite indicator approach and the subjective approach have been mentioned without being expanded on in this article. They lack a normative basis and can pose significant problems for interpretation. In contrast, the equivalent income approach is positioned as a direct extension of the central accounts framework. It is in line with national accounts as far as market contributions to well-being are concerned, proposing a measure of the standard of living that takes into account what is known about the ordinal preferences of economic agents, revealed by their behaviour, that can be extended to the other components of this well-being. By also being applicable at the individual level, it equally allows a systematic approach to the issue of inequalities, more adequate than the composite index approach – which works directly on aggregate indicators – and more adequate than the subjective approach – which risks providing a very reduced view of real inequalities, due to the ability of individuals to adapt to their living conditions. All of these characteristics make it an avenue to consider in order to make the measurement of growth more "inclusive" in nature.

Saying this does not mean that we ignore the limitations of this approach. We have mentioned its dependence on the selection of reference standards, for both the prices and the non-monetary characteristics of individuals, as well as the implementation difficulty – the need to indirectly reconstruct monetary valuations. It also seems difficult to apply it to more than a small number of non-monetary components of well-being. Therefore, we still remain far from the level of granularity of the work carried out to construct the central framework of national accounts. Finally, we should also mention an obstacle that may be a major hindrance to communication, the difficulty of getting the public to accept the neutrality of the monetary metric, as the trend of resistance to GDP is also fuelled a great deal by a rejection of this metric, associated with the idea of generalised commodification of all aspects of existence. There is certainly a strong argument against this rejection, which is that implicit forms of monetary valuation are actually implemented in any approach to constructing an aggregate index. Aggregating automatically means assigning relative values to the things being aggregated, doing so using a monetary account unit is just one choice of account unit among others, but this argument is not necessarily easy to get across.

What all this could argue for is therefore a rather eclectic and tailor-made approach. The conclusions of the Stiglitz Report in 2009 were already pointing towards eclecticism and it is also a characteristic of the recent follow-up to that report under the aegis of the OECD (Stiglitz *et al.*, 2018). However, eclecticism does not exclude a certain form of structure. The plan could be to have (a) a main core of accounts focused on their core business, while at the same time ensuring that they provide the most ready-to-use components for measuring well-being, notably through the refinement of volume-price splits, and (b) a satellite account focused on the measurement of household well-being, with an important place being naturally reserved for the approach that is most in line with the core accounts methods, and which shares its monetary metric, though without at all excluding the alternative approach of the subjective measurement of well-being, a bit like weather reports that combine objective temperature measurements and how it is felt.

The same sort of approach should guide the measurement of sustainability. In this case, the construction of a summary sustainability

index appears to be much less feasible than for measuring current well-being. However, a mini sustainability dashboard could be useful for coherently bringing together the main components of this sustainability: net savings or investment in the meaning of national accounts are part of it, together with indicators of financial sustainability such as the debt levels of different categories of economic agents, and physical indicators of environmental pressure and other indicators to be defined for quantifying the social component of sustainability.

All this may seem to lead to nothing more than the fourth and final way of going beyond GDP that we have merely mentioned in passing, the dashboard approach, but with the important nuance of introducing into it a dual concern for parsimony and conceptual integration, which are often not really present in the existing dashboards or those in development. □

Link to the Online Appendices: https://www.insee.fr/en/statistiques/fichier/4770146/ES-517-518-519_Blanchet-Fleurbaey_Online_Appendices.pdf

BIBLIOGRAPHY

Accardo, J. (2020). Supplementing GDP: Some Recent Contributions from Official Social Statistics. *Economie et Statistique / Economics and Statistics*, this issue.

Aeberhardt, L., Hatier, F., Leclair, M., Pentinat, B. & Zafar, J.-D. (2020). Does the Digital Economy Distort the Volume-Price Split of GDP? The French Experience. *Economie et Statistique / Economics and Statistics*, this issue.

Alvaredo, F., Chancel, L., Piketty, T., Saez, E. & Zucman, G. (2020). Towards a System of Distributional National Accounts: Methods and Global Inequality Estimates from WID.world. *Economie et Statistique / Economics and Statistics*, this issue.

Arrow, K. J., Dasgupta, P., Goulder, L., Daily, G., Ehrlich, P., Heal, G., Levin, S., Mäler, K. G., Schneider, S., Starrett, D. & Walker, B. (2004). Are We Consuming Too Much? *Journal of Economic Perspectives*, 18(3), 147–172. https://doi.org/10.1257/0895330042162377

Atkinson, A.B. (1970). On the Measurement of Inequality. *Journal of Economic Theory*, 2(3), 244–263. https://doi.org/10.1016/0022-0531(70)90039-6

Blanchet, D., Khder, M.-B., Leclair, M., Lee, R. & Ragache, N. (2018). La croissance est-elle sous-estimée ? Insee Références, *L'économie française - comptes et dossiers, édition 2018*, pp. 59–79. https://www.insee.fr/fr/statistiques/3614240?sommaire=3614262

Boarini, R., Murtin, F. & Schreyer P. (2015). Inclusive growth: the OECD measurement framework. *OECD Statistics working paper* N° 2015/06. https://doi.org/10.1787/18152031

Boarini, R., Murtin, F., Schreyer P. & Fleurbaey M. (2016). Multi-dimensional living standards: a welfare measure based on preferences. *OECD Statistics working paper* N° 2016/05. https://doi.org/10.1787/18152031

Canry, N. (2020). Why and How Should Human Capital be Measured in National Accounts? *Economie et Statistique / Economics and Statistics*, this issue.

Coyle, D. (2014). *GDP: A Brief but Affectionate History*. Princeton: Princeton University Press.

Cling, J.-P., Eghbal-Teherani, S., Orzoni, M. & Plateau, C. (2019). La France et les objectifs du développement durable. Insee Références, *L'économie française - comptes et dossiers, édition 2019*, p. 57–81. https://www.insee.fr/fr/statistiques/4181438?sommaire=4180914

Decancq, K., Fleurbaey, M. & Schokkaert, E. (2015). Happiness, equivalent incomes and respect for individual preferences. *Economica*, 82, 1082–1106. https://doi.org/10.1111/ecca.12152

Decancq, K. & Schokkaert, E. (2016). Beyond GDP: using equivalent incomes to measure well-being in Europe. *Social indicators research*, 126(1), 21–55. https://doi.org/10.1007/s11205-015-0885-x

Fleurbaey, M. & Gaulier, G. (2009). International comparisons by living standards. *Scandinavian Journal of Economics*, 111(3), 597–624. https://doi.org/10.1111/j.1467-9442.2009.01578.x

Fleurbaey, M. & Blanchet, D. (2013). *Beyond GDP. Measuring well-being and assessing sustainability*. Oxford: Oxford University Press.

Gadrey, J. & Jany-Catrice, F. (2016). *Les nouveaux indicateurs de richesse*, 4eme édition. Paris: La Découverte.

Germain, J.-M. & Lellouch, T. (2020). The Social Cost of Global Warming and Sustainability Indicators: Lessons from an Application to France. *Economie et Statistique / Economics and Statistics*, this issue.

Hamilton, K. & Clemens, M. (1999). Genuine Saving in Developing Countries. *World Bank Economic Review*, 13, 33–56. https://elibrary.worldbank.org/doi/abs/10.1093/wber/13.2.333

Hicks, J. (1940). The Valuation of Social Income. *Economica*, 7(26), 105–124. https://doi.org/10.2307/2548691

Jones, C. I. & Klenow, P. J. (2016). Beyond GDP? Welfare Across Countries and Time. *American Economic Review*, 106(9), 2426–2457. https://doi.org/10.1257/aer.20110236

Jorgenson, D. W. (2018). Production and Welfare: Progress in Economic Measurement. *Journal of Economic Literature*, 56(3), 867–919. https://doi.org/10.1257/jel.20171358

Lange, G. M., Wodon, Q. & Carey, K. (eds.) (2018). *The Changing Wealth of Nations 2018: Building a Sustainable Future*. Washington DC: World Bank.

Laurent, E. & Le Cacheux, J. (2016). *Un nouveau monde économique : mesurer le bien-être et la soutenabilité au XXIe siècle*. Paris : Odile Jacob.

Murtin, F., Boarini, R., Cordoba, J. & Ripoll, M. (2015). Beyond GDP: is there a law of one shadow price? *OECD Statistics working paper* N° 2015/05. https://doi.org/10.1787/18152031

Nordhaus, W. & Tobin, J. (1974). Is Growth Obsolete? *In: The Measurement of Economic and Social Performance, Studies in Income and Wealth*, vol. 38. National Bureau of Economic Research. https://www.nber.org/chapters/c3621.pdf

OCDE (2014). *Rapport sur le cadre de l'OCDE pour une croissance inclusive*. Paris: OCDE.

Ravallion, M. (2013). Troubling tradeoffs in the Human Development Index. World Bank, *Policy Research Working Paper* N° 5484. http://hdl.handle.net/10986/3967

Roman, P. & Thiry, G. (2016). The inclusive wealth index. A critical appraisal. *Ecological Economics*, 124, 185–192. https://doi.org/10.1016/j.ecolecon.2015.12.008

Samuelson, P. A. & Swamy, S. (1974). Invariant Economic Index Numbers and Canonical Duality: Survey and Synthesis. *American Economic Review*, 64(4), 566–93. https://www.jstor.org/stable/1813311

Schreyer, P. (2016). GDP. *In:* Fleurbaey, M.& Adler, M. (eds.), *The Oxford Handbook of Well-Being and Public Policy*. Oxford: Oxford University Press.

Sen, A. (1979). The Welfare Basis of Real Income Comparisons: A Survey. *Journal of Economic Litterature*, 17(1), 1–45. https://www.jstor.org/stable/2723639

Stiglitz, J. E., Sen, A. & Fitoussi, J.-P. (2009). *Richesse des nations et bien-être des individus*. Paris: Odile Jacob.

Stiglitz, J. E., Fitoussi, J.-P. & Durand, M. (2018). *Beyond GDP: Measuring what counts for economic and social performance*. Paris: OCDE. https://doi.org/10.1787/9789264307292-en

Triplett, J. E. (2001). Should the Cost-Of-Living Index Provide the Conceptual Framework for a Consumer Price Index? *The Economic Journal*, 111(472), 311–334. https://doi.org/10.1111/1468-0297.00633

United Nations Environment Program (2018). *Inclusive wealth report 2018*. http://hdl.handle.net/20.500.11822/27597

Vanoli, A. (2002). *Une histoire de la comptabilité nationale*. Paris: La découverte, collection Repères.

Weitzman, M. L. (1976). On the Welfare Significance of National Product in a Dynamic Economy. *Quarterly Journal of Economics*, 90, 56–162. https://doi.org/10.2307/1886092

World Commission on Environment and Development - Bruntland Commission (1987). *Our Common Future*. New York: Oxford University Press. https://sustainabledevelopment.un.org/content/documents/5987our-common-future.pdf

World Bank (2006). *Where is the Wealth of Nations? Measuring capital in the 21st century*. Washington DC: The World Bank. http://documents.worldbank.org/curated/en/287171468323724180/pdf/348550REVISED0101Official0use0ONLY1.pdf

Supplementing GDP: Some Recent Contributions from Official Social Statistics

Jérôme Accardo*

Abstract – In its efforts to take the criticisms levelled against GDP into consideration, national accounting benefits from the contribution made by social statistics. In the last decade, it has developed major innovations, such as accounts by household category, and it has provided other useful empirical and methodological tools to address the problem of evaluating domestic production and of taking into account non-monetary dimensions. It has also been very active in the new field of "the economics of happiness" (with the measurement of subjective well-being). This paper offers a critical discussion of the work carried out in these areas by official statistics.

JEL Classification: E01, I31
Keywords: accounts by household category, social indicators, subjective well-being

* *Insee (jerome.accardo@insee.fr)*

This article benefited from comments provided by several readers and two anonymous reviewers. I would particularly like to thank Didier Blanchet, Fabrice Lenglart and André Vanoli for their criticisms and suggestions on an initial version of the text. I am solely responsible for any errors that may remain.

Received 28 June 2018, accepted after revisions on 15 April 2019.
Translated from: "Compléter le PIB : quelques contributions récentes de la statistique sociale"
Citation: Accardo, J. (2020). Supplementing GDP: Some Recent Contributions from Official Social Statistics. *Economie et Statistique / Economics and Statistics*, 517-518-519, 25–39. https://doi.org/10.24187/ecostat.2020.517t.2016

As an indicator of the level of socio-economic development of a country, GDP has many limitations, criticism of which dates back almost as far as the indicator itself (Vanoli, 2002, Chap. 7). The main criticisms essentially focus on three issues: *i)* aggregation: as an aggregate indicator, GDP is not able to reflect phenomena associated with the distribution of flows or stocks among economic units; *ii)* scope: it only takes into account certain transactions, while excluding others, even though their economic nature and significance seem indisputable; *iii)* relevance: it raises very complex questions when the figure is interpreted in terms of social well-being; in turn, these questions raise the further question of its ability to inform public policy.

The report by Stiglitz, Sen and Fitoussi (2009) provided both a summary of these analyses and new impetus for research aimed at going "beyond GDP".

The aim of this paper is to present, among the recent developments made in this area by official statistics, the contribution of social statistics, understood as the collection of data describing the living conditions of individuals at the microeconomic level. The article is not intended as a survey of these questions or to cover all social statistics. It attempts only to provide a detailed presentation of some of the current approaches adopted by national statistical institutes to address the three categories of criticism by explaining the objectives identified, the methods and sources used and the obstacles encountered. Most of the practical illustrations are drawn from work carried out by Insee, with which the author is quite familiar.

Thus, this article will examine, in turn, the work devoted by social statistics over the last decade to taking into account the distribution of household account transactions, the domestic production of services, to incorporate non-monetary dimensions (health status, quality of the natural environment, security, social capital, etc.); and to the direct measurement of well-being ("happiness", "life satisfaction", etc.).

The first part seeks to address the criticism of the aggregate nature of GDP. The second applies to the criticism relating to scope. The last two address issues of interpretability.

1. Beyond the Analysis of Aggregates Alone: Distribution and Decomposition of Accounts by Household Category

In what precise way can the national accounts follow the recommendation of the Stiglitz-Sen-Fitoussi report to "take distributions into account" as far as the household account is concerned? As is well known, national accounts use only aggregates to describe the various transactions in the household account. However, thanks to the ever-growing number of microeconomic sources, the distribution of the corresponding variables is often known. Is it not possible to use these distributions to produce "distributional accounts"? That is, accounts which – in addition to establishing a set of monetary aggregates made consistent, according to internationally standardised methods – could also show their distribution across the population, thereby enabling the accounts to be interpreted in terms of inequality. The idea is particularly appealing: including the analysis of inequalities (inequalities in income, consumption and savings, a redistributive assessment of the tax and benefit systems, effect of public policies, etc.) within the theoretical framework of national accounts guarantees the consistency of the analysis and its exhaustive nature. It also appears very early on in academic thinking concerning National accounts (see Online Appendix C1[1]). We start below by explaining why the complete fulfilment of this objective remains unattainable, then we present the solutions that have nonetheless been explored by national accountants with a view to integrating differences between households into the accounting framework.

1.1. The Practical Impossibility of "Distributional Accounts"

To facilitate the exposition of the problem, a minimal version can be considered, with a household account limited to three (aggregated) transactions: gross (adjusted) income, actual final consumption and savings. Is it possible to create a distributional version of such an account – in other words, to produce this account for each household in a representative sample? Currently, the answer is no.

The obstacle lies in the fact that knowledge of the respective distributions of two variables in the population (in this case, income and consumption) does not make it possible to

1. *Link to the Online appendices at the end of the article.*

calculate the distribution of their sum or of their difference. It is necessary to know their joint distribution. In other words, it is necessary to have the following information available, for at least each household of a representative sample:

- the income of the household members (income from employment, including social taxes, replacement income, social benefits, capital income, etc.);

- transfers in kind from public health bodies that can be individualised (hospital care, reimbursements for healthcare, etc.);

- the educational situation of the household members (level of education, course of study, etc.) to allow the calculation of the public education expenditure from which they directly benefited during the year (based on micro-economic data from the National Education department on costs per pupil according to the type of education);

- a sufficiently precise description of that residence for the households in the sample that own their main residence, so as to be able to apply a satisfactory imputed rent model;

- the total consumption of market services and goods of each household.

All of this represents a great deal of information and collecting it all directly in a single survey would constitute an excessive burden on respondents. The information does exist, however, scattered across household surveys on one side – the Labour Force Survey (*Enquête Emploi en Continu*, or EEC), the European Union Statistics on Income and Living Conditions (EU-SILC), the National Housing Survey (*Enquête Nationale Logement*, or ENL), the Household Income and Expenditure Survey (*Enquête Budget de Famille*, or BDF), the Health Survey (*Enquête Santé*), etc. – and administrative files on the other side – income tax and housing tax files, files on claimants from social security organisations, Annual Declaration of Social Data (*Déclarations Annuelles de Données Sociales*, or DADS), health insurance files, etc.

If all these sources were matched together, the objective would be achieved: we would then have the adjusted gross disposable income (AGDI) and the actual final consumption, and therefore the saving rate[2], of each household in the sample. Such matching is currently only partially feasible, its full implementation being currently hampered by legal obstacles, which themselves reflect the political and philosophical problems raised by this type of "panoptic" project.

Work in progress in France, in a rapidly changing legal context (digital act and act on health data), certainly allows us to hope for significant progress in the coming years. Nevertheless, it should be stressed that these obstacles exist in most countries and, from this point of view, the situation concerning French official statistics is quite favourable compared to that of many countries in which the possibilities of matching survey data with administrative sources are much more limited. However, were distributional analyses possible in only a small number of countries, this "improved" accounting would lack international comparability, which is one of the major strengths of the system of national accounts.

1.2. The Principle of Decomposition of the Household Account

In the absence of broad matching allowing the creation of a complete account for each household in a representative sample, the solution lies in statistical imputation: this method (which is often referred to as "bottom-up") consists of selecting a survey in which the information collected, at household level, on the account's transactions is as extensive as possible. This information is then supplemented by imputing a value for each missing transaction for each household. Imputation is performed using models estimated in the other available sources.

One possible way to complete this process, in practice, is as follows. The primary source is the BDF survey: it provides an estimation of the annual consumption (at a fine level of the product nomenclature) and annual income of each household in the sample, obtained by matching with socio-fiscal administrative sources. This income represents only part of the income taken into consideration by the national accounts. Therefore, social security contributions, resources in kind, imputed rents, etc. must be added to it. These additions are obtained by applying the value predicted by models estimated based on the Tax and Social Incomes Survey (*Enquête revenus fiscaux et sociaux*, or ERFS) and the EU-SILC for social security contributions and income from self-employment

2. This would, of course, be a saving rate that is not rigorously dated, as several sources are only available on a multi-year basis and their availability is not synchronised.

and on the ENL survey for the imputed rents to each household. The same is carried out using models relating to public health and education expenditure to obtain an adjusted income at the household level. The crucial point here is that the explanatory variables in the different models used are also collected in the BDF survey. This condition is clearly essential for applying these models to each household in the BDF survey to estimate the most likely value (given the characteristics taken into account) of the missing components.

The end result is a representative sample in which, formally, resources and expenditure (and thus savings) are fully known for each household in the sample.

However, this file does not allow distributional analysis of the accounts. Indeed, if the sample is large enough, the distribution in the file of a given account transaction, whether expenditure or resources, adequately represents its actual distribution. In contrast, the imputation procedure does not make it possible to ascertain the true joint distribution of the various account transactions, only the joint distribution that is conditional on the explanatory variables used in the models. This makes it impossible to determine the distribution of sums and balances, starting with the distribution of savings.[3] The measures of inequality in adjusted gross disposable income or in actual consumption carried out using this microeconomic file will be biased.[4]

An example may shed some light on the difficulty: household out-of-pocket health expenditure (i.e. what they spend beyond what is covered by the health care system) is difficult to quantify in a household survey. Respondents often find it very difficult to estimate what they have spent on medical care, and they find it even more difficult to determine the portion that was not reimbursed. Here, the Health Surveys constitute the source of reference: as health expenditure is one of their main variables of interest, they devote questioning time to it and, if necessary, they carry out matching between their sample and the health insurance data. This is not the case with the BDF surveys which, consequently, provide an unreliable estimate of this expenditure. The solution is then to impute it for the households in the BDF survey sample based on a model estimated in the Health Survey.

The health expenditure of a household depends on its socio-demographic characteristics (age, income, social category and level of educational attainment of its members) and health-specific variables: health status, medical history and health cover of its members. Of all these factors, those specific to health are of course by far the most explanatory. Information on those factors is collected in the Health Survey, not the BDF survey. The imputation model will therefore have to make do with the usual socio-demographic variables; it will then only be able to explain a fairly small part of the dispersion of health expenditure. In fact, between two households with the same usual socio-demographic characteristics (age, income, qualification, social category, etc.), expenditure can differ greatly if the factors most directly related to health are different. Imputation then amounts to assigning to each of the two households a value selected at random in the Health Survey from among the expenditures of households with the same socio-demographic characteristics.

On average, this procedure is unbiased: it provides, for any given group of households of fixed age, income, etc., that household's true average level of health expenditure. In contrast, it is incorrect in terms of distribution since it assumes that, once these characteristics have been fixed, health expenditure is randomly distributed among households, regardless of the rest of their consumption, in particular. However, with other given characteristics, a very sick person will have a lower final consumption than a person in good health, but a higher health expenditure. Random imputation lacks this correlation and will tend to assign too low a health expenditure to that person and, therefore, an underestimated total consumption.[5] This limitation is inherent in the very principle of imputation. Only the actual collection of the variables for each household makes it possible to obtain their joint distribution.

Although it is not, strictly speaking, possible to determine the true joint distribution (income and consumption) in the population based on the distribution of income on the one hand, and the distribution of consumption on the other, but only an approximation, obtained under the at best rather crude assumption that they are

---

3. To provide a very simplified example: knowing the distribution of consumption C *on the one hand, and that of income* R*, on the other, does not make it possible to determine that of savings* R - C *while it is not known whether the two covariate in the same way (the wealthier a person is, the higher their consumption) or whether, on the contrary, they tend to compensate each other (the wealthier a person is, the more they save).*

4. The direction of the bias has not, a priori*, been determined.*

5. The imputation of health consumption, as with that of health expenditure by public health bodies that can be individualised, raises the exact same difficulty.

independent of each other[6], imputation is nevertheless a method in line with good statistical practice, even if it remains a little cumbersome to implement.

However, it is not the method that has been used in the various studies aimed at decomposing the household account. The practical problem is that the imputations depend on the explanatory variables used in the model. The international comparison of distributions, which is clearly a major objective (an enrichment of the national accounts that would be doomed to lose international comparability would be of limited interest), is reliable only between countries that have rigorously followed the same imputation methodology, i.e. the same models, with the same variables. However, it is virtually impossible to display a core of variables common to all sources used in the different countries that is sufficiently large to be useful.

Consequently, international work on the decomposition of the accounts has fallen back on a much simpler pseudo-matching method (often referred to as "top-down"): it consists of dividing households in each source into groups according to a particular criterion that is present in all sources. Each aggregate of the account is then distributed (using the relevant source) between these different groups.

For example, households can be classified by age group and, for each group, the average value of the missing components in the BDF survey can be calculated using the appropriate sources. A complete account is thus calculated for each age group.[7] The method entails matching average (or "representative") households in a group between the different sources. This is what is referred to as pseudo-matching of sources. It may also be seen as an elementary case of the imputation method, in which the imputation model is reduced to a single explanatory variable, namely the criterion used (in this case, age); incidentally, this confirms that it shares the same limitations as the imputation method. This method has been followed since the beginning by the international working group, coordinated by the OECD, dedicated to the development of accounts by household category. Online Appendix C1 briefly traces the history of the attempts and efforts to decompose national accounts aggregates by category of households.

These considerations call for a number of comments:

- The use of such pseudo-matching to introduce a decomposition of the household account is, in the case of age groups, an old and proven procedure: it is the method used by the generational accounting developed in the 1980s and 1990s by Auerbach & Kotlikoff, an objective that has been taken up again since the 2000s by the promoters of the *National Transfer Accounts Project* (see Online Appendix C2). However, the method can decompose the accounts according to any household classification criterion (gender of the reference person, household size, level of educational attainment of the reference person, etc.), provided that, for each transaction and for all countries, a microeconomic source is available that identifies the households according to that criterion, in a homogeneous manner across sources and countries;

- Breaking down the household account requires taking into account monetary transfers between households (support, donations, etc.), as well as exchanges of market goods and services between them (sales of second-hand vehicles, rentals, etc.);

- The decomposition exercise for the household account aims to enrich the economic description provided by the national accounts. However, one of its associated results is improving the quality of household surveys. Rigorous collation[8] with the accounting aggregates makes it possible to accurately assess the lack of coverage of these surveys in order to try to remedy it or, at least, to take it into account in the analyses;

-The availability of microeconomic sources, which are often only available on a multi-year basis, means that, in principle, it is not possible to perform decomposition of the account each year. Nevertheless, work is currently carried out to address this shortcoming, at least in part (see hereinafter).

1.3. Methodological Issues and Avenues for Progress

Even with the agreed simplifications in relation to the unattainable goal of a complete account at the level of each household, in practice, the decomposition of the household account

6. Strictly speaking, this is their independence conditional on the household description variables used in the imputation models. This is much more plausible than unconditional independence.

7. The total of a transaction across the different age groups should give the aggregate of the account. Otherwise, it is sufficient to recalibrate the source used using the aggregate. The sole function of the source is to provide the profile, not the level.

8. i.e. by ensuring that we are working on the same scope and with the same concepts.

raises several technical difficulties, exposed in Bellamy *et al*. (2009). In this article, the choice has been made to focus on one of them; it gives an idea of the work that still needs to be done to overcome the obstacles to the implementation of a system of accounts by household category, which has the same properties of reliability and international comparability as those under the central framework. It is also an interesting illustration of the differences, in both the objectives and the approach, between the decomposition of the accounts and two important recent approaches that also aim to link the distribution of resources and consumption with the corresponding accounting aggregates: the *National Transfer Accounts*, at the initiative of R. Lee and A. Mason and the *World Income Database* (WID.World), developed by researchers led by T. Piketty. Online Appendix C2 provides a presentation of these two approaches and details their similarities and their discrepancies in comparison with the breakdown of the household account.

The accounts by quintile of living standard show substantial dissaving by the poorest households in every country, except France. At the root of this discrepancy is the excess, in the consumption surveys, in the level of consumption over the level of income for a significant number of households.[9] The BDF survey is no exception in this respect. In this instance, however, the decomposition of the French account by quintile of living standard was based on a specific BDF variable that makes it possible to identify and adjust the responses of households showing aberrant consumption-income discrepancies. The effect of this treatment is considerable. Without it, the lowest quintile would have a dissaving rate of around 20% in France. However, this variable that allows for adjustment is not present in household income and expenditure surveys in most other countries. In addition, this adjustment is merely a practical method, which has the sole merit of simplicity and plausibility. The assumptions on which it is based are open to discussion and, therefore, the results published also include a version with a different, less selective, adjustment method. With this method, the dissaving rate stands around 13% in the first quintile.

That is to say that the use of microeconomic information is not always an immediate operation. This information must be analysed, discussed, arbitrated and without any guarantee to find a satisfactory solution for the problems encountered. Moreover, the solutions possible in a particular information system cannot necessarily be generalised. Undoubtedly, the most satisfactory solution requires a significant improvement in the accuracy of the microeconomic measurement of consumption. But this is an objective that will be difficult to achieve, even in the long term.

The difficulty outlined above is just one example of the problems to be solved. One could also mention the multi-year frequency (at least in the majority of countries) of certain microeconomic sources, such as the consumption survey; can an annual publication of accounts by category nevertheless be envisaged? In what manner?[10]

Another important issue is the accuracy of the accounts. Traditionally, the statistical accuracy of the aggregates of the central account is not considered. As there is no alternative, these aggregates are assumed to be "accurate". In contrast, it is known that microeconomic data from surveys are marked (at least) by a sampling risk, which can be estimated. Is it possible to take this risk into account in order to assess confidence intervals for the differences established between household categories? The work of the OECD Expert Group (see Online Appendix C1) on these methodological issues and others is continuing, with the challenge to provide answers that are not only conceptually and practically satisfactory but also common, so as to arrive at a process for the production of accounts by household category that is as stabilised and standardised as that for the aggregate account.

2. Expansion of the Scope of GDP

2.1. Time Spent on Domestic Work

Of all the expansions of the scope of GDP, the inclusion of the domestic production of services is probably the one that is most in line with the logic governing the indicator: first, the domestic production of goods is already taken into account (self-consumption); second, GDP includes the value of the housing service that owners-occupiers render to themselves (imputed

9. This is a classic finding, both at household level and at the level of groups of households. Consumption econometricians readily explain this by poor measurement of income, assuming that households tend to under-report their resources to the survey (hence the traditional practice in econometric study models of instrumenting income). The problem is, in reality, more profound and more complex, as the excess of consumption over income also appears implausibly widespread when the data on income is of administrative origin, as in the case of the BDF 2010 survey.

10. One possibility that has been explored by Insee recently (Accardo et al.*, 2017) is to fix the disparities between households, as observed in the surveys, but to change the aggregates annually as indicated in the national accounts.*

rents). This is clearly domestic production of services. Furthermore, it generally constitutes a major item in the household account (in France, for example, it accounts for around 13% of final consumption expenditure). Another is that ignoring the domestic production of services can bias international comparisons (this is also one of the justifications for including imputed rents). As the Stiglitz-Sen-Fitoussi report points out, a country in which the level of household production for self-consumption is significant may have a lower GDP than another, in which more goods and services go through the market, while households have the same level of consumption, if their own production is taken into account. For example, Alesina & Ichino (2009) calculate that when all domestic production is taken into account, Italy's GDP per capita rises from 56% to 79% of US GDP; finally, ignoring this (non market) production can lead to an overestimation of GDP growth, as households turn to the market for activities they used to do themselves.[11]

In practice, however, there are many unresolved difficulties in measuring the value of these activities, despite the efforts that have been devoted to it for several decades now:

- the precise scope of the activities to be taken into consideration remains a subject of debate. In principle, the criterion of "delegability" (or third party) is agreed upon. However, its application is often problematic[12] (Gershuny, 2011; Roy, 2012);

- various valuation options are available: at opportunity cost or at the observed market wage for an equivalent task. The latter option is the one most often used, as the former raises quite a number of objections; but it is not necessarily more realistic;[13]

- in the absence of accurate information on the characteristics of the task and the resulting product, their valuation is probably fairly biased.[14]

The estimated value of domestic work not only varies considerably depending on the scope and the valuation option chosen (in a ratio of 1 to more than 3), but in all cases it also represents a substantial sum (up to 50% of GDP according to Roy, 2012). This makes it difficult to include it in the central framework (and suggests instead that it be processed in a satellite account).

The main source of information on domestic activities is the Time Use Survey (*Enquête Emploi du Temps*). The results of the valuation are closely dependent on the information gathered by these surveys and the methods used to collect it. The standard method consists of having a sample of respondents complete a daily diary as they go about their activities.[15] The retrospective survey method[16], which is less costly and cruder, can give results that differ significantly in their level and distribution (Kan, 2008), with a tendency to overestimate the time spent on domestic activities.[17] More elaborate and costly methods (such as the experience-sampling method or continuous observation) also exist, which could be developed in the future thanks to technological developments (online surveys, sensors installed on respondents' mobile phones, etc.). These methods still result in different estimates. Furthermore, the level of accuracy of the information collected is crucial in characterising domestic activities. Knowing all the secondary activities carried out at the same time as the main activity represents a much heavier survey burden for the respondent but is the only way to gain an understanding of all the domestic activities.[18]

---

*11. However, this bias is undoubtedly more limited than the previous one. For example, in France, the average time spent producing domestic services (cooking, cleaning, childcare, etc.) per person (aged 18 or over) per day fell by 28 minutes between 1974 and 2010 (Brousse, 2015, p. 84). Valued using the super-gross minimum wage (*SMIC super brut*) used by Roy (2012), this reduction in domestic production represents, under the maximum hypothesis in which it is entirely externalised in the monetary sphere, a contribution to GDP of €91 billion in 2010. The annual GDP growth, 2% for the period 1974-2010, would then be overestimated by a maximum of about 0.13 percentage points.*

12. Do playing with one's children, DIY and shopping qualify as domestic production, or are they done for the personal pleasure derived from them? Depending on the answer given, time spent on domestic production varies by 50% (Roy, 2012). Similarly, excluding care of one's own body, as done by Roy (2012), rather than including it, as done by Alesina & Ichino (2009), has a very significant impact (a reduction of one hour in the time spent on domestic production per person per day).

13. In particular, its reference to the observed market price is questionable, as there is generally no precisely defined market price for these activities. For example, there is no market, for reasons that are easy to understand, where it is possible to buy 15 minutes of cooking time to prepare the children's ham and mashed potatoes in the evening, the two and a half minutes of time to wash the dishes and the 18 minutes of time to read them a bedtime story. In addition, the fact that the parent(s) who performed these tasks that evening produced a value of exactly 35.5 minutes x the super-gross hourly minimum wage is far from obvious.

14. In practice, in fact, the studies resolve to value these tasks uniformly at or around the minimum wage. However, there is no guarantee that an activity carried out by the household is of a quality comparable with that of activities performed professionally. It should also be noted that imputed rents, the only domestic service currently included in the national accounts, are not set equal to the average actual rent but are determined by taking into account the characteristics of the stock of principal residences occupied by their owners.

15. The time-slot of the diary is variable: in the French survey, it was five minutes until 1998, when it increased to ten minutes. It is 15 minutes in many surveys. Some, such as the Australian survey, leave it blank.

16. It is also known by the name "Stylised time-use items". It is the one adopted by the Labour Force Surveys in most countries.

17. However, this point, which has been observed on British data, is debated; Schulz & Grunow (2011), in contrast, find fairly good consistency between the two methods on German data.

18. For example, the statement "I watch TV" in a daily diary results in that time not being coded as a domestic production activity. However, if the diary also gathers information on secondary activities and if one of them indicates the presence of children under the respondent's care, then the time will be counted, at least in part, as a domestic activity ("Childcare").

Obtaining comparable estimates of the domestic production of services across countries or over time therefore requires a high degree of harmonisation between the surveys used for measurement. This harmonisation is still only partial. Many countries carry out Time Use Surveys with variable, but generally quite widely spaced, frequencies (in France the survey is carried out about every ten years), as those surveys are considered most expensive. At present, Eurostat has managed to coordinate the European countries by getting them to use a common methodology: collection of information using a daily diary, taking secondary activities into account and using a nomenclature of activities. Japan has chosen to collect information via a diary completed gradually. However, the United States, Canada, Australia and New Zealand have chosen (at least in the most recent surveys) to stick to the retrospective method.[19] Generally speaking, despite the many international efforts to standardise nomenclatures[20], they do not always coincide, with risk of different classifications for the same activity.

Clearly, a major effort to harmonise the scope, valuation and measurement methods is still needed before domestic work can be integrated into the preparation of the accounts, with the figure produced having a status comparable to that of the aggregates in the standard account.

2.2. Non-Monetary Dimensions: Health, Safety, Social Capital, Human Capital, etc.

Criticising GDP (and national accounting more generally) for ignoring many dimensions of life that have a value for individuals raises three questions: *i)* Is a quantitative measurement of these values necessary? *ii)* Can such measures be designed and determined? *iii)* How can this information be linked with that provided by GDP?

Statisticians, economists and national accountants are undoubtedly (is it a professional bias?) inclined to answer the first in the affirmative, but this view is not as straightforward as we may think. Just consider Robert Kennedy's famous speech during the 1968 American presidential election campaign[21] in which, clearly, most of the values mentioned did not, in his mind, require quantification. It should also be recalled that economic theory itself stresses "the rather loose nature of the link between overall income and social well-being" (Fleurbaey & Blanchet, 2013, p. 115[22]), which may put into question the utility of quantifying the unquantified.

If, however, the decision is taken to measure these values, questions *(ii)* and *(iii)* can be addressed in two ways: by juxtaposing complementary indicator tables, possibly summarised into composite indicators, with national accounting aggregates or by calculating a monetary equivalent for the non-monetary dimensions that can be directly measured against GDP and other accounting variables.

2.2.1. Synthetic Indicators and Dashboards

The first approach is limited to identifying indicators (in principle, non-monetary indicators) capable of describing the situation of individuals in the dimension (health, safety, democracy, social cohesion, etc.) under consideration. They provide information that complements that provided by the major accounting aggregates (GDP, gross disposable income, consumption, savings, etc.). This approach has been developed since the 1970s. An intuitive approach and GDP limits that have become commonplace explain the ever-increasing demand from decision-makers or the public for these indicators. An ever-increasing abundance of economic and social information that is easier to process explains why supply has been able to keep up. The result has been a flurry of initiatives that (based on pre-existing statistics) constitute sets of indicators meant to compensate for the shortcomings of traditional macroeconomic variables.[23] Among the most recent and significant examples are the European Sustainable Development Indicators (2005), the Sustainable Development Indicators for France (2010), the European Union's 2020 Strategy Indicators (2010), the OECD's Better Life Indicator (2011), the French government's new wealth indicators (2015) and the indicators

19. A surveyor questions a household member by telephone about their activities the previous day.

20. ICATUS nomenclature (UN), HETUS nomenclature (Eurostat), UNECE guidelines in 2013 as well as the work of Gershuny's team (the MTUS project of the CTUR in Oxford).

21. "Yet the gross national product does not allow for the health of our children, the quality of their education or the joy of their play. It does not include the beauty of our poetry or the strength of our marriages, the intelligence of our public debate or the integrity of our public officials. It measures neither our wit nor our courage, neither our wisdom nor our learning, neither our compassion nor our devotion to our country, it measures everything in short, except that which makes life worthwhile".

22. Chapter 4 of their work provides an in-depth analysis of how to give a monetary expression to preferences and how to use it in a normative analysis of well-being.

23. It should be noted that this article does not propose a general epistemological and historical analysis of the indicators (see Noll, 2002, for example), but only presents the main characteristics of the indicator compilations that aim to go beyond the description through only the aggregates of the national economic and social development accounts, and to quantitatively establish a notion of quality of life or quality of growth (for example, sustainability, inclusiveness, etc.).

for the Sustainable Development Goals adopted by the UN in 2017.

The indicators are practically always selected from the vast set of indicators published (or at least publishable) by the various public and private producers of economic and social information, generally in a very pragmatic way, outside of any theoretical framework, after more or less long and complex negotiations between political, administrative, scientific, expert and community representatives. Various authors or organisations have indeed tried to identify general principles for selecting an indicator, but these principles are primarily pragmatic (and sometimes *ad hoc*)[24] and do not provide a theoretical justification for the indicators selected.[25] This explains, at least in part, why the sets of indicators produced are often very disparate.[26] Once the indicators have been identified and collected, the question of how they relate to the usual accounting aggregates is likely to be resolved in two ways. The most simple solution is to make the information available as it is, in the form of a dashboard. It is left up to the user to consider the various messages in front of their eyes and to draw the conclusions they can.

During the 1980s and especially in the 1990s, the advocates of an indicator-based approach readily nurtured the ambition of creating an indicator capable of replacing GDP. In the belief that the place of GDP in the public debate is largely explains by the fact that it is a single figure (therefore easy to remember, easy to quote and allowing countries to be classified), they have tried to summarise sets of indicators into a single so-called "synthetic" index[27]: this has led to the design of the Index of Social Health (Miringoff, 1987; Miringoff & Miringoff, 1998), the Human Development Index (Haq, 1990), the Advanced Quality of Life Index (Diener, 1996), the Weighted Index of Social Progress (Estes, 1997), the CSLS's Index of Economic Well-Being (Osberg & Sharpe, 1998), the Index of Living Standards (Sarlo, 1998), the BIP40 (Inequality Observatory, 2004), etc. The OECD's Better Life Index (2011) is among the most recent.

To allow aggregation of variables describing very heterogeneous phenomena[28] into a single figure, these indices project them linearly (the minimum observed for the variable on 0 and its maximum observed on 100) and then create a simple (like the UN HDI) or weighted average for them. The weights are then either chosen at the discretion of the designer of the index, determined by factor analysis techniques (ISP), left to the user's initiative (the CSLS provides an Excel macro to vary the weight of its IEW and the OECD provides an online application for the Better Life Index[29]).

The procedure is convenient, but *ad hoc* and, technically, it is not without flaws (Gadrey & Jany-Catrice, 2012, p. 41; Accardo & Chevalier, 2005). Above all, the summary produced is problematic to interpret and it is generally recommended not to stick to the index (which is tantamount to acknowledging that it is basically just a convenient artefact) but to take into consideration the information provided by its components.[30]

As the issue of the weighting of the components of a synthetic index is without solution other than conventional, the designers of alternative indicators to national accounts aggregates currently tend to abandon the objective of a single index competing with GDP. The recent initiatives mentioned above are all either of the "dashboard" type or allow the user to choose their weighting preference.

2.2.2. Monetisation

Monetising the non-monetary dimensions is the alternative to tables of indicators and synthetic indicators. This time, the stage of aggregation to GDP is immediate (or almost immediate[31]) and it is obviously the first stage, which consists in assigning a price to things that are readily

---

24. The Indicators Sub-Group of the European Social Protection Committee has thus proposed the following principles: "An indicator should 1) capture the essence of the problem and have a clear and accepted normative interpretation, 2) be robust and statistically validated, 3) provide a sufficient level of cross countries comparability, 4) be built on available underlying data, and be timely and susceptible to revision, 5) be responsive to policy interventions but not subject to manipulation". See also Atkinson et al. *(2002).*

25. To clarify the criticism made here: this lack of theory can be contrasted with the construction of notions such as poverty in living conditions (Townsend, 1979), in which indicators are selected within an explicit conceptual framework. This, at least, makes it possible to question the ability of the indicators to measure what needs to be measured.

26. For example, the ten key European Sustainable Development Indicators (of 130 indicators) produced by Eurostat since 2007 include monetary aggregates (e.g. GDP per capita), counts of people (e.g. number of poor people) or years (e.g. life expectancy) or animal species (e.g. common birds) and tonnes of CO_2.

27. Sen (2003) describes this reasoning in detail, in the case of the HDI.

28. In effect, this involves aggregating rates of infant mortality, unemployment, youth suicide, housing access indicators (ISH) or even the number of doctors per 1,000 inhabitants, the saving rate, an income inequality index and the number of environmental treaties ratified (AQoLI), etc.

29. http://www.oecdbetterlifeindex.org/fr/#/11131111111

30. If only to (try to) understand why the different indices do not match up. For example: "Hence, the probable reason for Canada's fall from first (HDI) to 31st (WISP) in international ranking is the greater breadth of coverage of the WISP – but the complexity of the WISP calculation prevents a clear comparison" (Osberg & Sharpe, 2001).

31. It is not that immediate, insofar as the monetisation is based on a stock rather than an annual flow.

said to be priceless, which is the tricky part of this solution.[32]

Two methods are used to assign a value to a non-monetary asset.

(i) Stated preferences: this method is based on directly questioning a sample of individuals (in principle representative of the population). The questioning can be more or less sophisticated. Individuals may simply be asked: "What do you think this asset is worth?". They may also undergo complex questioning protocols, involving detailed experimental simulations and questions on binary choices or rankings of assets or scenarios, designed to allow formal choice models to be estimated.

(ii) Revealed preferences: here, the method is not based on statements but on observed behaviour. There are two main techniques. The first is implicit costs: the most commonly cited example is transport expenditure incurred by visitors to a nature reserve. This transport cost itself is a lower bound of the value that the public places on this environmental good. In random utility models, it can be used to estimate the value itself (although, it must be said, at the expense of quite a number of other additional assumptions). The second is hedonic prices: this uses as inputs the observed variations in the market price of a good depending on its characteristics. Thus, by observing the different prices of cars, depending on the models and ranges, it is possible to identify, econometrically, the value of a particular vehicle characteristic (speed, driving comfort, fuel economy, etc.) despite there being no specific market for that characteristic. The different methods have been in common use for many years in the field of cost-benefit analysis for the selection of public investments.

The revealed preference method is regularly used in the national accounts: for example, in the valuation of services provided by government bodies at their production cost, which is the implicit cost technique; hedonic methods are commonly used for the valuation of the housing service produced to themselves by owner-occupiers or for the determination of price indices (vehicles, computers, household appliances, etc.) at constant quality. By contrast, national accounts do not use the stated preference method for which, in fact, there is no reference procedure. This method is based on hypothetical choices, raising the crucial question of under what conditions these reported estimates constitute truly relevant information on individuals' preferences.

Thus, there is no consensus on the assessment of non-market dimensions, such as environmental capital, on the statistical value of life or social capital (in the meaning of Putnam, see OECD 2001) in a country. Correlatively, the available studies are rarely comparable and international comparisons are impossible most of the time.

3. Another Paradigm: Subjective Well-Being, Satisfaction and Happiness

The current wave of interest in a direct measure of well-being as perceived by individuals can probably be traced back to the mid-1990s. Easterlin's article in *Journal of Economic Behaviour and Organization* (Easterlin, 1995) is thought to be the starting point. It was essentially a continuation of an article from twenty years earlier (Easterlin, 1973) little noticed at the time. In 1995, however, economists, who are traditionally more inclined to rely on what individuals do rather than on what they say, had become influenced by work on behavioural economics, developed in particular by Thaler, Kahneman and Tversky, who were much more familiar with approaches questioning the canonical model of a *Homo Economicus* and more willing to consider the perception that individuals report about their economic situation.

3.1. The Good Fortune of the Happiness Paradox

Easterlin's idea is to use Happiness Surveys: since 1946, at least, surveys have asked respondents directly whether they consider themselves "happy" (phrased this way or similarly). Looking at average satisfaction calculated based on the responses collected, it seems to have remained stationary over the post-war decades even though, over the same period, GDP per capita has increased by a factor of two, three or even more, depending on the country. This result is the "Easterlin paradox".

As he points out in his 1995 article (p. 37), this paradox was fairly well documented as early as the late 1970s. However, it was not until the 1990s that the interpretation of it as an index to

32. Problems related to the link between monetary value and well-being are left aside here (see Fleurbaey & Blanchet, 2013, chap. 4). They also concern the usual monetary dimensions, those for which the national accounts can use existing market prices.

use against GDP became evident. Previously, the same observation seemed to reinforce economists in their bias against the use of subjective data, a material that the majority of them deemed to be acceptable, at a push, for "soft" disciplines such as sociology and psychology, but not for positive economic science for which only actions can reveal preferences.

This is in stark contrast to the current situation: for more than twenty years now, more and more supporters of the measurement of subjective well-being have been stressing that "the ways in which people value their lives […] should be an integral part of the concept of human well-being"[33], an idea that is also supported by the Stiglitz-Sen-Fitoussi report.

At present, regular survey data on perceived well-being are available for many countries, feeding into the abundant work on an "economics of happiness" that gained momentum in the academic field during the 2000s and in which decision-makers, media and the general public take a keen interest. The OECD, which includes a measure of reported satisfaction in its Better Life Index, has also published guidelines to that end, aimed in particular at national statistical institutes (OECD, 2013). In 2013, a secondary module of the EU-SILC panel, a survey under European regulation, collected the answers of respondents to questions on their well-being and Eurostat published the results.

Several NSIs have taken an interest in this type of indicators; Insee has produced an annual measurement of life satisfaction since 2011 and the British ONS since 2015. The French indicator is also included in the table that the law on new wealth indicators (or "*Loi Sas*"), adopted in April 2015, obliges the government to publish annually.

3.2. Measuring Happiness

The economic theory of happiness actually distinguishes three notions of subjective well-being (OECD, 2013):

- respondents' satisfaction with the life they lead: here, the respondents must produce a global judgement on their life overall or over a more limited period (most often the current period). The issue then boils down to a question. The respondents are generally asked to rate their life on a scale (known as the Cantrill scale) that goes from 0 (very poor rating) to 10 (very good rating);

- "affects": the term is used in psychology to refer to the emotional states of an individual at a given time. The reference measurement method is then the Experience Sampling Method (ESM), in which the participants must keep an accurate account of the following states ("happy", "peaceful", "irritated", "angry", etc.);

- the "eudaimonic" approach: the aim is to take a step back from the hedonism underlying the two previous methods and to have the respondents assess the degree to which they feel fulfilled in their life and the degree to which they think they have effectively developed their physical, intellectual and moral potential.

In practice, the first two approaches are dominant. The measurement of satisfaction is certainly the most widespread, due to the simplicity of its implementation. Nevertheless, recent technological developments offer researchers new options for measuring affects, from applications installed on the respondents' smartphones reminding them of the times of day when they should send information on their emotional state, to medical-type devices that continuously measure their blood pressure, stress, etc., information from which the researcher is supposed to be able to derive an assessment of the respondent's affects throughout the day. It should be noted here that, with this method, the respondent's subjectivity is set aside: this is a physical type of measurement, through use of a measurement tool. Naturally, the question then arises as to whether the usual subjective assessment and this much more objective measurement relate to the same thing.

Judging by its public success and its integration among the instruments guiding public action, "life satisfaction" appears to be particularly well placed to compete with GDP as a measurement of well-being. However, it raises major difficulties, the resolution of which seems all the more difficult since they are probably not taken into account sufficiently by the various users of this indicator. First of all, there is a fundamental theoretical question: what conclusion should be drawn from the level of satisfaction reported? Is it to be understood that maximising satisfaction should be the goal of public policy? On this point, one may object, firstly, that it is by no means obvious that this is the objective of the individuals themselves and, secondly, even

33. Taken from the Recommendations for Measuring Sustainable Development of the joint UNECE/Eurostat/OECD Task Force established by the Conference of European Statisticians in 2014.

if that were the case, such a choice is likely to be criticised from an ethical point of view (Fleurbaey & Blanchet, 2013, p. 169 *et seq*.).

3.3. The Meaning of the Figures

In addition, many problems arise in the practical construction of this measure and in its use, most of which do not seem to be resolved, nor are they in the process of being resolved. They are not due to the subjective nature of the respondents' response, which in itself does not exclude the possibility of developing indicators of proven usefulness, but to the lack of clarity over what the response covers.

Firstly, the wording of the question in statistical surveys is necessarily brief and does not make it possible to specify with what a respondent is satisfied with sufficient precision: his/her life at present, their life since birth, or their life over a shorter period of time? Their own life, their own life and the lives of those they care about or their life as a member of a wider community? And how can one be sure which dimensions of life the respondent takes into account in their assessment?

Similarly, how is it possible to control the respondent's reference point: in other words, to what does the "0" (or the "10") on the scale refer? The worst life ever lived in the history of the world? Or just in the respondent's country at present? Or the worst life that the respondent has a reasonable chance of experiencing personally? Or the worst life that the respondent has actually experienced? Etc.

Finally, what metric is the respondent using? In other words, is the respondent harsh or indulgent in rating their life? What is a "5", a "7" or a "10" actually worth to them? This is what Fleurbaey & Blanchet (2013) refer to as the "calibration problem".[34]

In the absence of a minimum degree of clarification of this set of ambiguities affecting the responses collected, it seems doubtful that it would be possible to give a reliable meaning to the aggregation of the satisfaction ratings reported by a sample of individuals. There is indeed good reason to believe that these ambiguities are of very real practical importance.

Let us thus consider the calibration problem: it is possible to seek to assess (and attempt to control) its importance using a "vignette methodology" (Kapteyn *et al.*, 2009; Angelini *et al*., 2014). This involves brief descriptions of individual situations that the respondent is asked to rate. The distribution of the ratings collected for a single vignette makes it possible to measure calibration differences within the population.

The self-questionnaire that has been included in the EU-SILC panel since 2011 includes eight different vignettes of this kind.[35] They are presented to the panel respondents (more specifically, to those in the second re-interview). Whatever the year of the survey and whatever the vignette considered, it is found that the ratings are about as dispersed (sometimes even significantly more so) as the ratings given by respondents concerning their own lives.

This is quite a remarkable phenomenon: if, for the same situation, the respondents give such divergent ratings, it becomes very risky to give any substantial interpretation to the rating they give to their own situation. Seeing that Respondent A gives their life a satisfaction rating of 5, while Respondent B gives a satisfaction rating of 8, what can be concluded about their actual situation and how they perceive it? It should be noted that the average life satisfaction indicator included in the alternative indicators to GDP provided for by the law of 2015 is not corrected for calibration differences, nor is the indicator calculated for each European Union country by Eurostat, based on the EU-SILC module carried out in 2013.

It is important to underline that even a modest correction can have a highly visible impact, due to the smallness of the differences generally observed between the average satisfaction levels in the various countries: according to data from the EU-SILC 2013[36], the average rating for 30 countries are between 6.2 and 8; a correction of +0.5 therefore represents about 10 places gained in the ranking. Under these conditions, and to use an example, France's mediocre rating

---

34. In psychometrics, the problem is known as "Differential Item Functioning". See for example Osterlind & Everson (2009).

35. Here are two examples of these vignettes:

(No 7) Maria is a veterinarian aged 58. She lives with her husband in a house with a garden. She has three children and five grandchildren who visit her regularly. She plays tennis every weekend. How would you rate Maria's situation in respect of the life she is currently living?

(No 8) Anne is 40 years old. She works as a nursery assistant. She lives with her husband and their three children in a small apartment they rent on the outskirts of the city. The neighbours are quite noisy. Her husband has been unemployed for two years, it is not always easy to make ends meet and this creates tensions in their relationship. She suffers with back pain and has trouble sleeping because, this year, she is working in a difficult class. How would you rate Anne's situation in respect of the life she is currently living?

36. The European Union Statistics on Income and Living Conditions (EU-SILC) is a panel survey carried out annually in Europe, within the framework of a European regulation. The French part of the EU-SILC is referred to as SRCV in French.

in terms of life satisfaction (with an average rating of 7.0, measured in the 2013 EU-SILC module, the 16th place out of 32 countries), a phenomenon that is regularly observed, is easily interpreted either as an indication that life in France is less pleasant than in its neighbouring countries, or as the mark of a national mood that is more gloomy than elsewhere.[37] In the absence of rigorous calibration, these interpretations appear fragile, to say the least.

Of course, the interpretation of this dispersion as an issue of calibration can be questioned. The available data do not make it possible to rule out that, in reality, individuals use the scale in exactly the same way and that it is simply their conceptions of what is pleasant, bearable, painful or intolerable in life that differ.[38] However this objection does not answer questions about the relevance of measurements based on self-assessment of subjective well-being.[39] The heterogeneity of preferences can even be viewed as further calling into question their usefulness, as substantially different preferences from one individual to the next make it very difficult to interpret the collected ratings, for which it is not clear not only to which exact situations they relate but also what, in these situations, is judged positively and what is judged negatively by each respondent.

For the promoters of the economics of happiness, measuring well-being is measuring what people think of their own happiness.[40] This approach has become quite widely accepted in recent years, resulting in high demand for data. Social statistics has acted with remarkable responsiveness, but it cannot confine itself to producing figures. It must also enable users to understand their nature, their scope and their limitations. From this point of view, the measurement of subjective well-being still requires significant clarification efforts. While going beyond GDP is clearly a necessity, it is still necessary to know exactly where we stand once we have gone beyond.

* *
*

Looking "beyond GDP" may simply mean not limiting oneself to that indicator but taking other socio-economic indicators into consideration in the economic analysis. However, one could also set a more ambitious goal of "going beyond GDP", i.e. developing a conceptual and operational framework that integrates other information in a coherent way without losing that provided by the existing framework. Each of the three avenues of research described in the article illustrates not only the benefit of such an approach, but also its difficulty.

Disaggregation of GDP is fully in line with the accounts approach. Admittedly, conceptual questions arise, particularly with regard to the scope concerned: should all the components of GDP be disaggregated? Or only the household account (but including consumption)? There is also the issue of the statistical unit: should the distribution be measured at the level of the individual or at the level of the household? The answers depend on the objectives set. However, the more complex issue remains a practical one. It is the issue of the sources of information on income and consumption distributions: how reliable are they, what is their availability, how comparable are they over time and between countries. The use of sources of social statistics (surveys and data from government departments) is on the rise, particularly over the last ten years or so, and undeniable progress has already been made. This progress is expected to speed up. It can be expected that within ten years, various distributions (income, consumption, wealth and savings) that are fully consistent with the framework of the accounts will be available in many countries.

Taking into account socio-economic dimensions not included in GDP but deemed necessary to make judgements on well-being is more difficult to combine with the traditional accounts approach and the conception and construction of GDP. Monetisation is certainly the approach most directly consistent with them, as it makes it possible to produce an expanded GDP, in the continuity of the traditional GDP and subject to the same analyses. However, this raises conceptual and technical issues (for example, the dependence of the valuation obtained on the method chosen and the unavoidable and numerous additional assumptions), the solution to which does not appear to be forthcoming.

The subjective well-being approach is surely the one that poses the greatest difficulties: how can the national accounts integrate the information

37. Algan et al. *(2018) thus speak of the "exception of the French malaise".*
38. It is not possible, however, with the existing data, to estimate the respective weights of calibrations and preferences in the heterogeneity of the vignette evaluations.
39. For more general information on these methodological issues, see the OECD manual (OECD, 2013, op. cit.*) which provides detail on them.*
40. "Self-reported happiness has turned out to be the best indicator of happiness" (Frey & Stutzer, 2002).

it produces? For, whatever the merits of the "economics of happiness", the advantage of the research it generates and the insights it can shed on socio-economic behaviour and the fundamental problems of the economics of well-being[41], it is clear that its purpose is not of the same nature as those developed in the national accounts. Indeed, the collection of subjective preferences is an interesting tool for the monetisation of non-monetary dimensions and can thus contribute to their inclusion in the expanded GDP. However, notions of happiness, subjective well-being, life satisfaction, etc. still raise, in terms of the clarity of concepts as well as comparability and traceability of measurements, too many issues that are difficult to resolve and which, whether they can be resolved or not, will always clearly fall outside the scope of national accounts. □

41. See, for example, Layard (2005), one of the main representatives of the field, and Clark (2018) for a presentation of research developments.

Link to the Online Appendices: https://www.insee.fr/en/statistiques/fichier/4770148/ES-517-518-519_Accardo_Online_Appendices.pdf

BIBLIOGRAPHY

Accardo, J. & Chevalier, P. (2005). Les indicateurs synthétiques. *In*: Verger, D., Bas revenus, consommation restreinte ou faible bien être : les approches statistiques de la pauvreté à l'épreuve des comparaisons internationales. *Économie et Statistique*, 383-384-385, 33–35. https://doi.org/10.3406/estat.2005.7192

Accardo, J., Billot, S. & Buron, M. L. (2017). Les revenus, la consommation et l'épargne des ménages par grande catégorie entre 2011 et 2015. *Insee Références – L'économie française - Comptes et dossier - édition 2017*. https://www.insee.fr/fr/statistiques/2894028?sommaire=2894036

Alesina, A. & Ichino, A. (2009). *L'Italia fatta in casa. Indagine sulla vera ricchezza degli italiani.* Milano: Mondadori.

Algan, Y., Beasley, E. & Senik, C. (2018). *Les Français, le bonheur et l'argent*. Paris: Éditions Rue d'Ulm.

Angelini, V., Cavapozzi, D., Corrazzin, L. & Paccagnela, O. (2014). Do Danes and Italians rate life satisfaction in the same way? Using vignettes to correct for individual-specific scale biases. *Oxford Bulletin Economics and Statistics*, 76(5), 643–666. https://doi.org/10.1111/obes.12039

Atkinson, T., Cantillon, B., Marlier, E. & Nolan, B. (2002). *Social Indicators: The EU and Social Inclusion*. Oxford: Oxford University Press.

Bellamy, V., Consalès G., Fesseau, M., Le Laidier, S. & Raynaud, E. (2009). Une décomposition du compte des ménages de la comptabilité nationale par catégorie de ménage en 2003. Insee, *Document de travail* N° G2009/11. https://www.insee.fr/fr/statistiques/1380884

Brousse, C. (2015). La vie quotidienne en France depuis 1974. Les enseignements de l'enquête Emploi du temps. *Économie et Statistique*, 478-479-480, 79–117. https://doi.org/10.3406/estat.2015.10559

Clark, A. E. (2018). Four Decades of the Economics of Happiness: Where Next? *Review of Income and Wealth*, 64(2), 245–269. https://doi.org/10.1111/roiw.12369

Conference of European Statisticians (2014). *Recommendations on Measuring Sustainable Development*. UNECE. https://www.unece.org/publications/ces_sust_development.html

Easterlin, R. A. (1973). Does money buy hapiness? *The Public Interest*, 30, 3–10.

Easterlin, R. A. (1995). Will raising the incomes of all increase the happiness of all? *Journal of Economic Behavior and Organization*, 27(1), 35–47. https://doi.org/10.1016/0167-2681(95)00003-B

Fleurbaey, M. & Blanchet, D. (2013). *Beyond GDP: Measuring Welfare and Assessing Sustainability.* Oxford: Oxford University Press.

Frey, B. & Stutzer, A. (2002). *Happiness and Economics: How the Economy and Institutions Affect Human Well-Being*, Princeton: Princeton University Press.

Gadrey, J. & Jany-Catrice, F. (2012). *Les nouveaux indicateurs de richesse.* Paris: La Découverte.

Gershuny, J. (2011). Time-Use Surveys and the Measurement of National Well-Being, mimeo, Centre for Time-use Research, Department of Sociology, University of Oxford, september.

Kan, M. Y. (2008). Measuring housework participation: The gap between « stylised » questionnaire estimates and diary-based estimates. *Social Indicators Research*, 86, 381–400. https://doi.org/10.1007/s11205-007-9184-5

Kapteyn, A., Smith, J. P. & Van Soest, A. (2009). Comparing life satisfaction. Rand Corporation Publications Department, *Working Paper* N° 623-1. https://www.rand.org/pubs/working_papers/WR623-1.html

Layard, R. (2005). *Happiness: Lessons from a New Science*. London: Penguin Books.

Noll, H. H. (2002). Social Indicators and quality of life research: background, achievements and current trends. *In*: Genov, N. (Ed.), *Advances in Sociological Knowledge over Half a Century.* Paris: International Social Science Council.

OECD (2001). *The Well-Being of Nations. The Role of Human and Social Capital.* Paris: OECD Publishing. https://doi.org/10.1787/9789264189515-en

OECD (2013). *OECD Guidelines on Measuring SubjectiveWell-being.* Paris: OECD Publishing.

Osberg, L. & Sharpe, A. (1998). An Index of Economic Well-Being for Canada. CSLS Conference on the State of Living Standards and Quality of Life in Canada, October 30-31, Ottawa, Ontario.

Osberg, L. & Sharpe, A. (2001). Trends in Economic Well-Being in Canada in the 1990s. *The Review of Economic Performance and Social Progress*, 1, 233–247.
https://econpapers.repec.org/RePEc:sls:repsls:v:1:y:2001:aslo

Osterlind, S. J. & Everson, H. T. (2009). *Differential item functioning*. Thousand Oaks, CA: Sage Publishing.

Roy, D. (2012). Le travail domestique : 60 milliards d'heures en 2010. *Insee Première* N° 1423.
https://www.insee.fr/fr/statistiques/2123967

Sen, A. (2003). L'indice de développement humain. *Revue du MAUSS*, 21(1), 259–260.
https://doi.org/10.3917/rdm.021.0259

Schulz, F. & Grunow, D. (2011). Comparing Diary and Survey Estimates on Time Use. *European Sociological Review*, 28(5), 622–632. https://doi.org/10.1093/esr/jcr030

Stiglitz, J.E., Sen, A., Fitoussi, J.-P. (2009). Rapport de la Commission sur la mesure des performances économiques et du progrès social.
https://www.vie-publique.fr/rapport/30513-mesure-performances-economiques-et-progres-social

Townsend, P. (1979). *Poverty in the United Kingdom.* London: Allen Lane and Penguin Books.

Vanoli, A. (2002). *Une histoire de la comptabilité nationale.* Paris: La Découverte.

Towards a System of Distributional National Accounts: Methods and Global Inequality Estimates from WID.world

Facundo Alvaredo*, Lucas Chancel, Thomas Piketty***, Emmanuel Saez**** and Gabriel Zucman******

Abstract – This paper briefly presents the methodology of Distributional National Accounts (DINA), which distributes total national income and total wealth among all individual residents. With DINA, we can estimate inequality statistics and growth by income and wealth groups that are consistent with aggregate growth from National Accounts. This methodology has been recently applied to a number of countries, and the data produced are available from WID.world. The paper summarizes the initial empirical findings. We observe rising top income and wealth shares in nearly all countries in recent decades, but the magnitude of the increase varies substantially, thereby suggesting that different country-specific institutions and policies matter. We combine countries' statistics to estimate global inequality since 1980. Global inequality has increased since 1980 in spite of the catching up of large emerging countries like China and India. This has been driven by the income growth of top world earners.

JEL Classification: D31, D33
Keywords: inequality, distribution, income, wealth, national accounts

PSE, IIEP-UBA-Conicet and INET at Oxford (alvaredo@pse.ens.fr); **PSE and Iddri (lucas.chancel@psemail.eu); ***PSE (piketty@psemail.eu); ** UC Berkeley and NBER (saez@berkeley.edu; zucman@berkeley.edu)*

The authors acknowledge funding from the Berkeley Center for Equitable Growth, the European Research Council (Grants 340831 and 856455), the Ford Foundation, the Sandler Foundation, the Sloan Foundation, and the Institute for New Economic Thinking. This article draws from recent work in Alvaredo et al. *(2016, 2017, 2018, 2018b). We are grateful for comments from two anonymous reviewers.*

Received on 26 June 2018, accepted after revisions on 16 July 2019.
Citation: Alvaredo, F., Chancel, L., Piketty, T., Saez, E. & Zucman, G. (2020). Towards a System of Distributional National Accounts: Methods and Global Inequality Estimates from WID.world. *Economie et Statistique / Economics and Statistics*, 517-518-519, 41–59. https://doi.org/10.24187/ecostat.2020.517t.2018

Rising inequality has attracted considerable interest among academics, policy-makers and the general public in recent years. Yet we still face important limitations in our ability to measure the changing distribution of income and wealth, both within and between countries, and at the world level. In this paper, we discuss novel methods to develop a System of Distributional National Accounts, DINA (Alvaredo *et al*., 2016) and present new findings about global inequality dynamics that follow this general framework.

The development of economic statistics is a historical lengthy process that involves economic theory, the limits of available data, the construction of a body of conventions, and the agreement of the community of scholars. Macroeconomic aggregates (GDP, national income) from the System of National Accounts (SNA) are the most widely used measures of economic activity. In the beginning, national accountants were also experts in distributional issues, as the inter-linkages between the estimation of national income and its distribution were clearly recognised. However, the focus of the SNA has so far always been on the main sectors in the economy, only distinguishing results for the household sector as a whole, and not providing insights into disparities within the household sector. Partly as a result of these developments, the discrepancies between levels and growth rates displayed in national accounts and the ones displayed in micro statistics and underlying distributional data have been growing in all dimensions: income, consumption, wealth. Scholars have been aware of the discrepancies, and have provided a list of general reasons behind them, but systematic and coordinated action to put them in a consistent framework has just started.[1]

One reason why this work has only begun recently is clear: it is not a simple task. A renovated approach to the measurement of economic inequality should rebuild the bridges between distributional data available from micro sources and national accounts. This is the main goal of the World Inequality Database project (WID.world) pursued through DINA: to provide annual estimates of the distribution of income and wealth using concepts that are consistent with the macroeconomic national accounts. In this way, the analysis of growth and inequality can be carried over in a coherent framework.

The article is structured as follows. In section 1, we start by discussing the current limitations when measuring and understanding inequality, and by describing the reasons for the development of a System of Distributional National Accounts. In section 2, we summarize the concepts and methods used (and proposed) for the estimation of DINA series. In sections 3 through 5, we present selected findings on income inequality, private *vs.* public wealth to income ratios, and wealth inequality. In section 6, we discuss new estimates of global inequality (also presented in Alvaredo *et al*., 2018). To conclude, we identify pathways for further progress.

1. Towards a System of Distributional National Accounts

Renewed interest in the long-run evolution of the distribution of income and wealth has given rise to a flourishing literature over the past 20 years. By combining historical tax and national accounts data, a series of studies has constructed time-series of the top income share for a large number of countries (see Piketty 2001, 2003 for France, Piketty & Saez, 2003 for the United States, and the two multi-country volumes on top incomes edited by Atkinson & Piketty 2007, 2010; see also Atkinson *et al.*, 2011 and Alvaredo *et al.*, 2013 for surveys of this literature). To a large extent, this literature has followed the pioneering works and methods of Kuznets (1953) and Atkinson & Harrison (1978), extending it to more countries and years. As these projects generated a large volume of data, intended as a research resource for further analysis as well as a source to inform the public debate on inequality, the data have subsequently been made public through the World Top Incomes Database, WTID (Alvaredo *et al.*, 2011-2015), and now the World Inequality Database (WID.world) (see Box for a brief history of the WID.world project).

The progress made in the last two decades meant an enormous step forward in the field of applied inequality studies. However, despite the latest developments and endeavours, we still face important limitations when measuring, analysing, and understanding economic inequality. Addressing the following concerns is at the core of the DINA project. First and most important, there is a large gap between national accounts (NA) – which focus on macro totals and growth – and inequality studies – which focus on distributions using survey and tax data.

1. *Social accounts matrices are a related precedent.*

Box – **History of the WID.world project**

By combining historical tax and national accounts data, a series of studies has constructed time-series of the top income share for a large number of countries (see Piketty 2001, 2003 for France, Piketty & Saez, 2003 for the United States, and the two multi-country volumes on top incomes edited by Atkinson & Piketty, 2007, 2010. See also Atkinson, Piketty & Saez, 2011 and Alvaredo *et al.*, 2013 for surveys of this literature). These projects generated a large volume of data, intended as a research resource for further analysis, as well as a source to inform the public debate on income inequality. To a large extent, this literature has followed the pioneering work and methodologies of Kuznets (1953) and Atkinson & Harrison (1978) on the long-run distribution of income and wealth, extending it to many more countries and years.

The World Top Incomes Database-WTID (Alvaredo *et al.*, 2011-2015) was created in January 2011 to provide convenient and free access to all the existing time series generated by this stream of work. Thanks to the contributions of over a hundred researchers in a clear synergetic framework, the WTID expanded to include time-series on income concentration for more than 40 countries, spanning most of the 20th, the early 21st centuries and, in some cases, going back to the 19th century. The key innovation of this research was to exploit tax and national accounts data in a systematic manner. This permitted the estimation of longer and more reliable time-series on the top income shares than previous inequality databases (which generally rely on self-reported survey data, with usually large under-coverage and under-reporting problems at the top, and limited time span).

These new series had a large impact on the discussion of global inequality. In particular, by making it possible to compare the shares captured by top income groups (e.g. the top 1%) over long periods of time and across countries, they contributed to reveal new facts, and refocus the discussion on rising inequality. Although the top income share series have contributed to improve our understanding of inequality trends, they suffer from important limitations (Atkinson *et al.*, 2011). In particular, they cover only the top part of the distribution; they are based only on fiscal income, which can diverge from national income because of tax exempt income, tax avoidance and evasion; finally, they focus on pre-tax income inequality and are therefore silent on redistributive effects of public policies between and across countries.

In December 2015, the WTID was subsumed into the WID.world, the World Wealth and Income Database (relabelled the World Inequality Database in March 2017). In addition to the WTID top income shares series, the first version of WID.world included an updated historical database on the long-run evolution of aggregate wealth-income ratios and on the changing structure of national wealth and national income first developed in Piketty & Zucman, 2014 (see also Piketty, 2014 for a historical interpretation on the basis of this material, and of the top income shares time-series). The name of the database changed from WTID to WID.world in order to reflect the extension in scope of the database, and the new emphasis on both wealth and income. In January 2017, a new website was launched (www.wid.world), with better data visualisation tools and more extensive data coverage. The World Inequality Lab was also created then, with the mission of maintaining and expanding WID.world, coordinating the statistical operations of the network (now with over 120 researchers around the world, in universities, research centres, official statistics offices, and tax offices) and publishing the World Inequality Report-WIR every two years (the first volume WIR2018 (Alvaredo *et al.*, 2018) was released in December 2017).

The discrepancies can be seen both in the level of income, wealth, and consumption, as well as in the observed growth rates of the economic aggregates (see, for example: Bourguignon, 2015; Deaton, 2005; Nolan *et al.*, 2018; Ravallion, 2003); they can attain particularly high levels in developing countries. National income is larger and has been growing faster than the other income concepts traditionally used to study inequality. Such gaps make it hard to assess how macroeconomic growth is distributed across income groups, and to address questions such as: what fraction of economic growth accrues to the bottom 10%, the bottom 50%, the middle 40%, and the top 10% of the distribution? How much of the rise in income inequality owes to changes in the share of labour and capital in national income, and how much to changes in the dispersion of labour earnings, capital ownership, and returns to capital?

Second, a substantial fraction of national income (e.g. about a third in the USA and half in several European countries) is redistributed through taxes, transfers, and public spending on services such as education, police, and defense. Yet we do not have a comprehensive measure of how the distribution of pre-tax income differs from the distribution of post-tax income, making it hard to assess how government redistribution affects inequality.

Third, existing inequality statistics use the tax unit (when they mostly rely on tax data) or the household (when they are based on surveys) as the unit of observation. As a result, we do not have a clear view of how long-run trends in income concentration are shaped by the major changes in women's labour force participation – and, in general, gender inequality – that have occurred over the past century.

Fourth, it is not an easy task to predict whether the observed trend of rising concentration of wealth will continue. In the long run, steady-state wealth inequality depends on the inequality of saving rates across income and wealth groups, the inequality of labor incomes and rates of returns to wealth, and the progressivity of income and wealth taxes. How have these factors affected the process of wealth accumulation in the past, and what can they tell about potential future dynamics? Numerical simulations show that the response of steady-state wealth inequality to relatively small changes in these structural parameters can be rather large (Saez & Zucman, 2016; Garbinti *et al.*, 2016). In our view, this instability reinforces the need for increased data quality to allow the dynamics of income and wealth to be properly studied and understood.

Fifth, the move from national states considerations to the study of inequality at the regional and global level requires an acceptable level of homogeneity of statistics across countries. Distributional information published by national statistical offices cannot be aggregated in a simple way. These limitations also apply to provinces within a country.[2]

A renovated approach to the measurement of economic inequality consistent with NA should overcome the limits of the existing series, and re-build the bridges between distributional data available from micro sources and national accounts aggregates more systematically than done in the past. This is our main and overall objective: to produce a System of Distributional National Accounts – which includes the theoretical principles as well as the statistics for all countries in the world –, and to use the newly created series to make progress in the understanding of the inequality phenomena. We propose to combine national accounts, tax, and survey data to build DINA, that is, series on the distribution of total national income and national wealth for the longest possible period and, ideally, for all the countries in the world. The series should be homogeneous across countries and along time as in the internationally agreed SNA. In this way, the analysis of growth and inequality can be carried over in a coherent framework.

The DINA project involves extending the past developments into three main directions. First, the project aims to cover developing countries and not only developed countries (which were the majority in WTID); in recent years, tax information has been released in a number of emerging economies, including China, Brazil, India, Mexico, and South Africa. Second, WID.world intends to provide more and updated series on wealth-income ratios and the distribution of wealth, and not only on income. Third, we aim to cover the entire distribution of income and wealth, and not only of top groups (as was the case in the WTID). The overall long-run objective is to produce a set of Distributional National Accounts, which are the main focus of this paper.

A main methodological contribution is the production of synthetic micro-files: individual level data that are not necessarily the result of direct observation but rather estimations that reproduce the observed distribution of the underlying data. They include – whenever possible – the joint distribution of age, gender, marital status, numbers of dependent children, and provide information on income and wealth. This synthetic micro-files of pre-tax and post-tax income (and wealth) consistent with macro aggregates, ideally contain all the variables of the national accounts as well as synthetic adult individual observations that are obtained by statistically matching tax and survey data, and by making explicit the assumptions about the distribution of income (and wealth) categories for which there is no directly available source of information, and which are being imputed.[3] By construction, the totals in these micro-files add up to the national accounts totals, while the distributions are consistent with those in the underlying distributional information (tax data, surveys, etc.). The synthetic micro-files can be used to compute a wide array of distributional statistics (labour and capital income earned, taxes paid, transfers received, wealth owned, etc.). The long-run aim is to release income and wealth synthetic DINA micro-files for all countries on an annual basis. Such data could play a critical role in the public debate, and be used as a resource for further analysis by various actors in civil society and in the academic, business and political communities.

It is worth stressing that the WID.world and DINA have both a macro and a micro dimension. Homogenous time-series should cover both the macro-level structure of national income and wealth, as well as the micro-level distribution.

---

2. Even in Europe, comparing national inequality trends and analyzing the dynamics of regional inequalities is far from straightforward; see Blanchet et al. *(2019), who discuss some of the difficulties arising in the production of DINA for thirty-eight European countries.*

3. Naturally, the assumptions will be, in many cases, specific to the countries and years under study, and dependent on the institutional arrangements as well as on the data available. See Piketty et al. *(2018) and Garbinti* et al. *(2018) for synthetic files for the USA and France respectively.*

By doing so, we hope to contribute to the reconciliation of inequality measurement and national accounting, i.e. the micro-level measurement of economic and social welfare and the macro-level measurement. In some cases, this may require revising central aspects of key national accounts concepts. By combining the macro and micro dimensions of economic measurement, we are following a very long tradition. In particular, it is worth recalling that Simon Kuznets was both one of the founders of US national accounts (and author of the first national income series), and also one of the first scholars to combine national income series and income tax data to estimate the evolution of the share of total income going to top fractiles in the USA over 1913-1948 (Kuznets, 1953).[4] This line of research continued with Atkinson & Harrison (1978), who combined historical inheritance tax data with capital income data to study the long-run evolution of the distribution of wealth in Britain over 1922-1972. We are simply pushing this effort further by trying to cover more countries and years.

Such an ambitious long-term objective – annual distributional national accounts for both income and wealth and for all countries in the world – will require a broad international and institutional partnership. The initial set of methodological principles and recommendations are being set by on-going work in the first version of the DINA Guidelines (Alvaredo *et al*., 2016). There are still many methodological decisions to be made and agreed upon. It took four decades from the 1910s to the 1950s before scholars (Kuznets, Kendrick, Dugé, Stone, Meade, Frankel) could hand over the estimation of national income to official statistics bodies. It also took a long time (from the 1950s to the 2000s) before official national accounts were able to include standardised wealth accounts. In fact, the first consistent guidelines for balance sheets – covering stocks of assets and liabilities – appear in the SNA manuals of 1995 and 2008 (in some key countries, such as Germany, the first official stock accounts were released only in 2010). Along the same lines, the development of a system of DINA could to take some time before consensus among scholars and the statistical community is reached.

We should stress at the outset that our methods and time-series are imperfect, fragile and subject to revision. The WID.world DINA project attempts to combine the different data sources that are available (in particular tax data, survey data, and national accounts) in a systematic way. We also try to provide a very detailed and explicit description of our methodology and sources, so that other users can contribute to improving them. But our time-series and methods should be viewed in the perspective of a long, cumulative, collective process of data construction and diffusion, rather than as a finished product.

2. Distributional National Accounts: Concepts and Methods

The concepts and methods used in the WTID series were initially presented in the two collective volumes edited by Atkinson & Piketty (2007, 2010), and in the corresponding country chapters and research articles. Despite our best efforts, the units of observation, the income concepts and the Pareto interpolation techniques were never made homogenous over time and across countries. Moreover, for the most part attention was restricted to the top income decile, rather than the entire distribution of income and wealth. In contrast, the DINA time-series aim to be homogenous across these dimensions (or at least to make much more explicit the remaining heterogeneity) and, most importantly, to provide more detailed and comprehensive measures of inequality. In the DINA series, inequality is always measured using homogenous observation units, and taxable income reported on fiscal returns is systematically corrected and upgraded in order to match national accounts totals separately for each income category (wages, business income, etc.) using various sources, imputation methods and techniques to align the micro and macro data. WID.world aims to provide series on wealth (and not only on income) and on the entire distribution (and not only on top shares).

The two main data sources used in the DINA continue to be income tax data and national accounts (just like in the WTID series), but we use these two core data sources in a more systematic and consistent manner, with harmonized definitions and methods, and together with other sources such as household income and wealth surveys, inheritance data, estate and wealth tax data, as well as the wealth rankings in "rich lists" compiled by the press. In most cases, the general trends in inequality depicted in the WTID series will not be very different in DINA series.[5]

4. Kuznets (1953) was preceded by ten years by Frankel & Herzfeld (1943), who made estimates of the European income distribution in South Africa based on the income tax returns, making use of control totals from the census of population and from the national accounts.

5. Results of these comparisons are available for France (Garbinti et al.*, 2018) and the United States (Piketty* et al.*, 2018).*

The following elements are key in the construction of DINA:

- The unit of observation (adult individual with equal split of income among married couple, adult individual with own individual income).

- The income concepts (pre-tax national income, pre-tax factor income, post-tax disposable income, post-tax national income, and fiscal income) and the wealth concepts (personal wealth, private wealth, public wealth, and national wealth).

- The methods employed to reconcile income tax returns and household survey micro files with NA, as well as with wealth inequality sources.

- The methods employed to produce synthetic micro files.

- The methods that can be used in the case of countries and time periods with more limited data sources.

In this section, we briefly refer to the units of observation, the income and wealth concepts, and the case of countries and years with limited data.[6]

2.1. The Unit of Observation

One of the limitations of the WTID series was the lack of homogeneity in the micro-level observation unit. WTID series were constructed by using the 'tax unit' (as defined by the tax law of the country at any given point in time) as the observation unit. In joint-taxation countries like France or the United States, the tax unit has always been defined as the married couple or the single adult. This is problematic, since variations in the share of single people in the population, or in the extent of assortative mating in couples could potentially bias the evolution of income inequality in various and contradictory ways. In other countries, the tax system switched to individual taxation in the last decades (e.g., in 1990 in the United Kingdom), which creates other discontinuities in the WTID series (see Atkinson, 2005, 2007).

In order to correct for these biases, the DINA series try to use homogenous observation units. Generally speaking, the benchmark unit is the adult individual. Whenever possible, we also aim to estimate distributions that can be decomposed by age, gender and number of dependent children. One key question is how to split income and wealth between adults who belong to a couple (married or not) and/or to the same household. To the extent possible, we want to produce two sets of inequality series: equal-split-adults series and individualistic-adults series. In the equal-split series, we split income and wealth equally between adults who belong to the same couple. In the individualistic series, we attribute income and wealth to each individual income recipient and wealth owner (to the extent possible). Both series are equally valuable. They offer two complementary views on different dimensions of inequality. The equal-split perspective assumes that couples redistribute income and wealth equally between their members. This is arguably a very optimistic: bargaining power can be typically very unequal within couples. But the opposite perspective (zero sharing of resources) is not realistic either, and tends to underestimate the resources available to non-working spouses (and therefore to overestimate inequality in societies with low female participation in the labor market).

Regarding the equal-split series, an important question is whether we should split income and wealth within the couple (narrow equal-split) or within the household (broad equal-split). In countries with significant multi-generational cohabitation (e.g. grandparents living with their adult children), this can make a significant difference. In countries where nuclear families are prevalent, this makes relatively little difference.

Finally, when we look at the inequality of post-tax disposable income, we also introduce dependent children into the analysis, in order to be able to compute the child related cash and in-kind transfers to the parents.

The issues are more complicated for capital income flows. In joint-taxation countries, capital income is usually not reported separately for both spouses, and we generally do not have enough information about the marriage contract or property arrangements to split capital income and assets. So we simply assume in our benchmark series that each spouse owns 50% of the wealth of a married couple and receives 50% of the corresponding capital income flow. If and when adequate data sources become available, we might be able to offer a more sophisticated treatment of this important issue.

6. We invite the interested reader to consult the DINA Guidelines for the complete documentation, and a thorough (though on-going) investigation of details, problems, limitations and challenges.

2.2. The Income and Wealth Concepts

2.2.1. Income Concepts

Other major limitation of the WTID time-series was the lack of homogeneity of the income concept and its dependence on the tax laws of each country. In contrast, the concepts used in DINA series are defined in the same manner in all countries and time periods, and aim to be independent from the tax legislation. We use four basic pre- and post- tax income concepts to measure inequality: *i)* pre-tax national income; *ii)* pre-tax factor income; *iii)* post-tax disposable income; and *iv)* post-tax national income.[7] All of them are anchored on the notion of national income: GDP minus capital depreciation plus net income received from abroad, defined by using the same concepts as those proposed in the latest international guidelines on national accounts, as set forth by the 2008 UN SNA. However, in attributing income to the household sector we apply a broader definition, as we also distribute the income of the other sectors in the economy (i.e. corporations, general government and non-profit institutions), rather than focusing on the household sector as defined in SNA.

Despite the usual focus on GDP, national income is a more meaningful concept for two reasons. First, capital depreciation is not economic income: it does not allow one to consume or accumulate wealth. Allocating depreciation to individuals would artificially inflate the economic income of capital owners. Second, including foreign income is important, because foreign dividends and interest are sizable for top earners.

Importantly, we include corporate retained earnings – the fraction of after-tax corporate profits which is not distributed to shareholders – in our measures of income. They can be sizable and vary significantly over time or across countries, so their omission can lead to deficient estimates of the level and trend in income concentration. The key reason for adding undistributed profits (or at least a fraction of them) to personal income is because undistributed profits should be considered as income for the owners of corporations. Undistributed profits are an income flow in the Hicksian sense: they make the owners of corporations wealthier. Depending on the tax system, shareholders may prefer to accumulate profits in corporations rather than to receive dividends (e.g. because this may allow them to realize capital gains by selling shares at a later stage, and by doing so they might pay less taxes than what they would have paid on the corresponding dividends). We only include the fraction of corporate retained earnings that accrue to resident households, i.e., we subtract the retained earnings in domestic firms that are foreign-owned (and, symmetrically, add the retained earnings in foreign firms owned by domestic residents). This adjustment is particularly important for low-tax countries, which tend to have high profits (and in particular high retained earnings) in foreign-owned firms (Tørsløv *et al.*, 2018).

By construction, pre-tax national income and pre-tax factor income are both equal to national income at the aggregate level, but they are not the same at the individual level and in terms of distribution. The central difference is the treatment of pensions, which are counted on a contribution basis in pre-tax factor income, and on a distribution basis in pre-tax national income. We tend to favor the pre-tax national income concept for our benchmark series for pre-tax inequality, but pre-tax factor income inequality also provide complementary information. Both series should be produced. The key reason why we prefer the pre-tax national income inequality series is that it is less affected by the age structure of the population. We aim to define pre-tax national income so as to satisfy the following neutrality condition: in a hypothetical economy with 100% replacement rates for pensioners, the cross-sectional inequality of pre-tax national income should be the same whether it is measured within the entire population (including pensioners) or within the working-age population.

Post-tax disposable income is defined as pre-tax national income, minus all taxes on production, income and wealth, plus social assistance benefits in cash. In order to compute post-tax national income, we add social transfers in kind.

2.2.2. Wealth Concepts

In the same way as for the income concepts, our wealth concepts refer to the NA guidelines, based on which we define personal wealth, private wealth, public wealth, corporate wealth, and national wealth.[8]

---

7. We also keep the fiscal income definition associated with the top income share series in Atkinson & Piketty (2007, 2010) and Alvaredo et al. *(2011-2015).*

8. Readers are referred to the DINA Guidelines Appendix, where we provide the formulas linking the DINA income and wealth definitions to the SNA 2008 classification codes.

We should make clear at the outset that our choice of using NA income and wealth concepts for distributional analysis certainly does not mean that we believe that these concepts are perfectly satisfactory or appropriate. Quite the contrary: our view is that NA statistics are insufficient and need to be improved. In particular, one of the central limitations of official GDP accounting is that it does not provide any information about the extent to which the different social groups benefit from GDP growth. The other reason for using NA concepts is simply that they represent the only existing systematic attempt to define notions such as income and wealth in a common way, which (at least in principle) can be applied to all countries independently from specific legislation.

2.2.3. Countries and Years with Limited Income and Wealth Data: Simplified DINA

The construction of DINA series is very demanding in terms of data and other information. Countries do not usually have all the sources required, the limitations being very pronounced in many countries/years. This problem was also at the center of the development of NA: designing the SNA meant accepting that the standards could not be set at the level of the best; their implementation had to be feasible in less well-advanced countries. Methods (labeled here as "Simplified DINA") need to be developed in the case of countries and time periods with more limited sources, typically on the basis of income tax tabulations rather than income tax micro-files, and/or with income tax data covering only a subset of the population, and/or inadequacy of income tax data (e.g. due to exemptions on capital incomes).

Some of the methods that can be applied in such circumstances can be found in recent work on DINA for China (Piketty *et al.*, 2017) and France (a country with detailed tax data but where only income tax tables are available prior to 1970; see Garbinti *et al.*, 2018).[9] Piketty *et al.* (2019) further develop, for the USA, a simplified methodology that starts from the fiscal income top income share series and makes very basic assumptions on how each income component from national income that is not included in fiscal income is distributed.

3. Income Inequality Dynamics: Countries and Regions

The methods proposed in the DINA project have already been applied to several countries: the United States in North America; France in Europe; China, India and Malaysia in Asia; Brazil in South America; Russia; and the Middle East. The new series combine national accounts, survey, and fiscal data in a systematic manner in order to estimate the distribution of pre-tax national income (including tax exempt capital income and undistributed profits).[10]

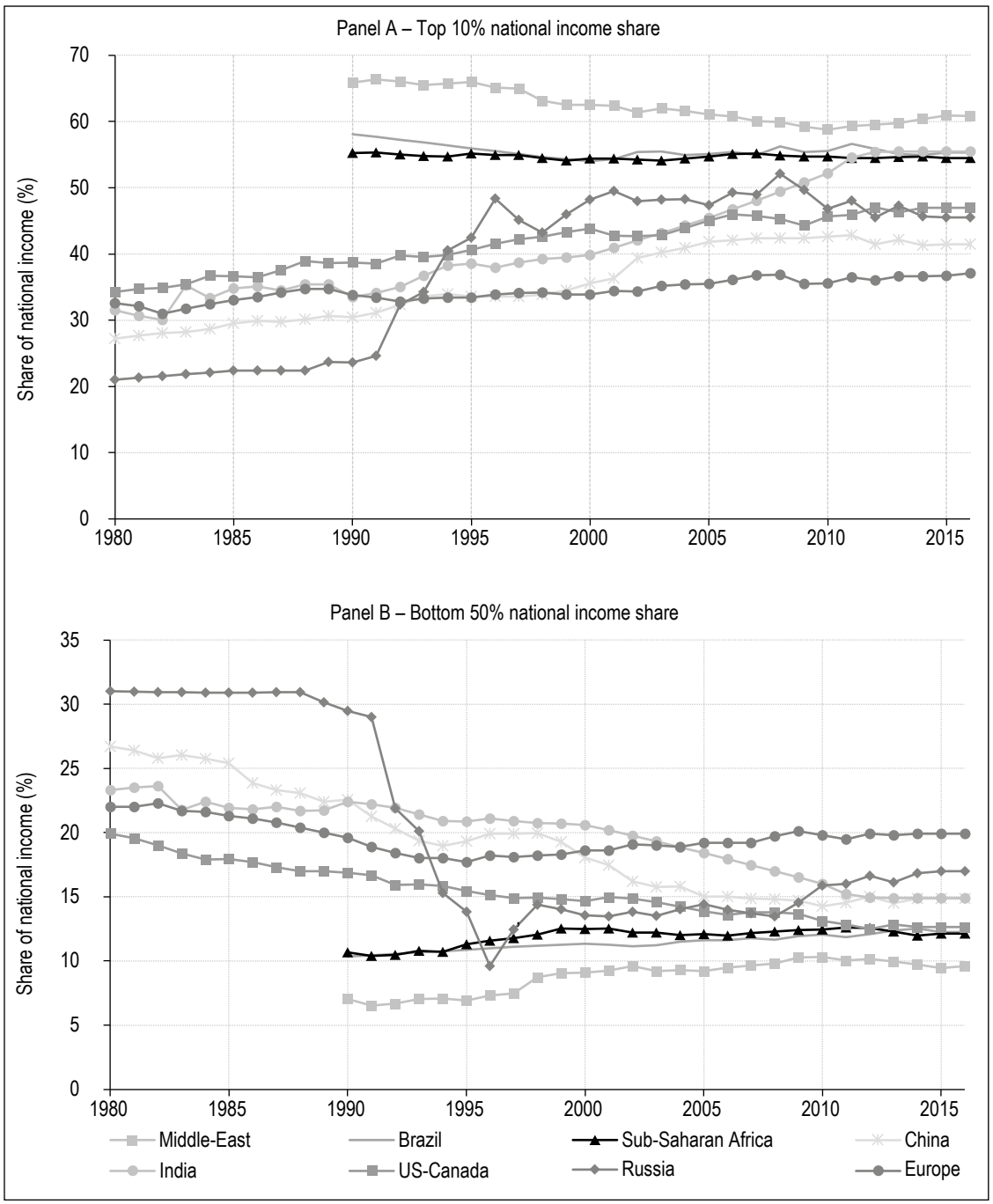
Figure I displays the evolution of inequality in various countries and regions based on the new estimates. As shown in panel A, the top 10% income share has increased almost everywhere since 1980, but with large variations in magnitude. In Europe, the rise was moderate. It was much more marked in North America, India, China, and Russia. By 2016, the top 10% income share stands at about 41% in China, 46% in Russia, 47% in North-America, and 56% in India. The rise in inequality correlates with policy changes in each country: the Reagan revolution in the United States, the transition away from communism in China and Russia, the shift to a deregulated economy in India. Policies and institutions matter: rising inequality cannot be viewed as a mechanical, deterministic consequence of globalization or technological change, as most economic models assume.

There are exceptions to the general pattern of increasing inequality. In the Middle East, Brazil, and sub-Saharan Africa, income inequality has remained relatively stable at extremely high levels since 1990, the first year for which we can construct estimates for these regions. In effect, for various historical reasons and in contrast to the other countries shown in Figure I, these regions, despite local developments, never went through the post-war egalitarian regime and have always been at the world's high-inequality frontier.

As shown in the panel B of Figure I, the share of income accruing to the bottom 50% looks like the mirror image of the top 10% income share. The bottom 50% income share is lowest in places where the top 10% share is highest (Middle East, Brazil, Sub-Sahara Africa) and vice-versa (Europe). The bottom 50% share has also fallen most in countries where the top 10% has increased the most (Russia, China, India,

9. See Blanchet et al. *(2017) and http://WID.world/gpinter for technical details on Pareto curves and the corresponding interpolation techniques.*
10. We refer the reader to the country-specific articles; they can be found in the Library section of WID.world: for the Middle East, see Alvaredo, Assouad & Piketty (2019); for Brazil, Morgan (2017); for India, Chancel & Piketty (2017); for Russia, Novokmet et al. *(2018). For details on the methods to go from country inequality to regional inequality, see Alvaredo* et al., *2018.*

Figure I – **Distribution of income**



Notes: Share of total national income earned by the top 10% and bottom 50% of adults in various countries and regions from 1980 to 2016. Income is before taxes and transfers but after the operation of public and private retirement and unemployment insurance systems. For married couples, income is split equally across spouses.
Sources: WID.world.

and the United States). It has remained stable in places where the top 10% income has also been stable.

The combination of tax and survey data leads to markedly revise upwards the official inequality estimates of China. We find a corrected top 1% income share of around 13% of total income in 2015, *vs.* 6.5% in survey data. We stress that our estimates should likely be viewed as lower bounds, due to tax evasion and other limitations of tax data and national accounts in China. But they are already more realistic and plausible than survey-based estimates, and illustrate the need for more systematic use of administrative records, even in countries where the tax

administration is far less than perfect. Figure I shows that China had very low inequality levels in the late 1970s, but it is now approaching North America. In particular, we observe a collapse of the bottom 50% income share in the US-Canada between 1980 and 2016, from 20% to 12% of total income, while the top 1% income share rose from 11% to 20%. In contrast, and in spite of a similar qualitative trend, the bottom 50% share remains higher than the top 1% share in 2015 in China, and even more so in France.[11]

In light of the massive fall of the bottom 50% pre-tax incomes in US-Canada, our findings also suggest that policy discussions about rising global inequality should focus on how to equalize the distribution of primary assets, including human capital, financial capital, and bargaining power, rather than merely discussing the ex-post redistribution through taxes and transfers. Policies that could raise the bottom 50% pre-tax incomes include improved education and access to skills, which may require major changes in the system of education finance and admission; reforms of labor market institutions, including minimum wage, corporate governance, and workers' bargaining power through unions and representation in the board of directors; and steeply progressive taxation, which can affect pay determination and pre-tax distribution, particularly at the top end (see, for example: Piketty *et al.*, 2014; Piketty, 2014).

The comparison given above illustrates how the *DINA* series can be used to analyze the distribution of growth across income classes. The Table below decomposes income growth within China, Europe, India, Russia, and North America, by income group. Real average national income per adult grew at very different rates in the five regions from 1980 to 2016: an impressive 831% in China and 223% in India, a moderate 40% in Europe, 34% in Russia, and 63% in US-Canada. In all these countries, income growth is systematically higher for upper income groups. In China, the bottom 50% grew 417% while the top 0.001% grew more than 3,750%. The gap between the bottom 50% and the top 0.001% is even more important in India. In Russia, the top of the distribution had extreme growth rates too while bottom 50% incomes fell; this reflects the shift from a regime in which top incomes were constrained by the communist system towards a market economy with few regulations limiting top incomes. In line with Figure I, Europe stands as the region with the lowest growth gap between the bottom 50%, the full population, and the top 0.001%. In China, top groups have enjoyed very high growth, but aggregate growth was also so large that even the bottom 50% average income grew markedly. This is likely to make rising inequality much more acceptable. In contrast, in the US-Canada, there was very little growth left for the bottom 50% (+5%).

4. Private *vs.* Public Wealth-Income Ratios

Next, we present findings on the evolution of aggregate wealth on Figure II. We observe a general rise of the ratio between net private wealth and national income in nearly all countries in recent decades. It is striking to see that this long-run finding was largely unaffected by the 2008 financial crisis. It is also worth stressing the unusually large rise of the ratio

11. These series refer to pre-tax, pre-transfer inequality. Post-tax, post-transfer series (not discussed here) reinforce these conclusions, at least regarding the USA-France comparison; see Bozio et al.*, 2018.*

Table – **Real income growth and inequality, 1980-2015, in %**

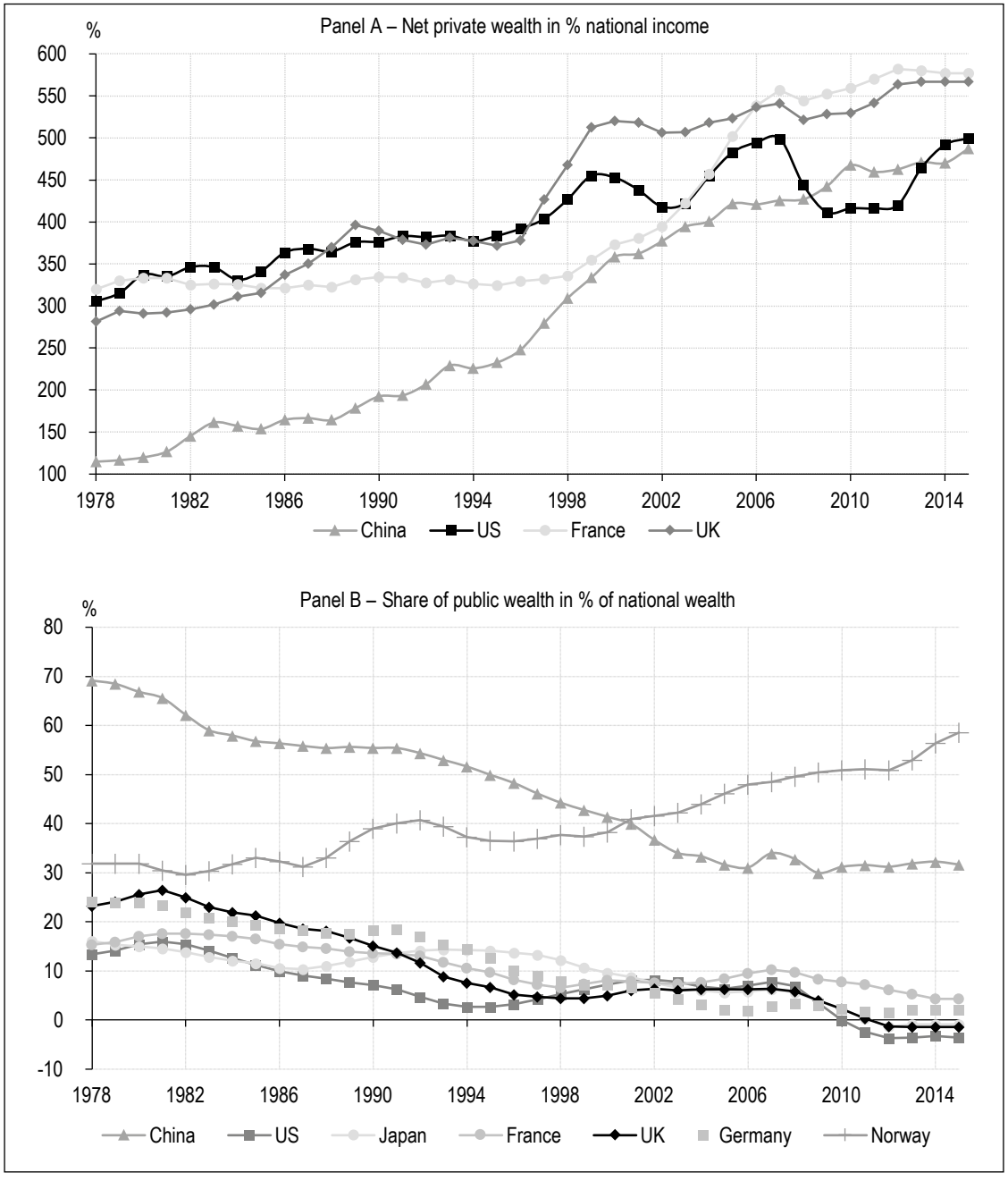
Income group (distribution of per-adult pretax national income)	China	Europe	India	Russia	USA-Canada	World
Full population	831	40	223	34	63	60
Bottom 50%	417	26	107	-26	5	94
Middle 40%	785	34	112	5	44	43
Top 10%	1,316	58	469	190	123	70
including Top 1%	1,920	72	857	686	206	101
including Top 0.1%	2,421	76	1,295	2,562	320	133
including Top 0.01%	3,112	87	2,078	8,239	452	185
including Top 0.001%	3,752	120	3,083	25,269	629	235

Notes: Distribution of pre-tax national income (before taxes and transfers, but including pensions and unemployment insurance) among adults. Corrected estimates combining survey, fiscal, wealth and national accounts data. Equal-split-adult series (income of married couples divided by two).
Sources: WID.world.

for China (panel A). According to our estimates, net private wealth was a little above 100% of national income in 1978, while it is above 450% in 2015. The private wealth-income ratio in China is now approaching the levels observed in the USA (500%) and in the UK and France (550-600%).

The structural rise of private wealth-income ratios in recent decades is due to a combination of factors, which can be decomposed into volume factors (high saving rates, which can themselves be due to ageing and/or rising inequality, with differing relative importance across countries, combined with growth slowdown), and relative asset prices and institutional factors, including the increase of real estate prices (which can be due to housing portfolio bias, the gradual lift of rent controls, and the lower technical progress in construction and transportation technologies as compared to other sectors) and stock prices (which can reflect higher power of shareholders

Figure II – **Private *vs.* public wealth-income ratios**



Notes: Net private wealth is personal plus non-profit wealth. Net public wealth is public assets minus public debt.
Sources: WID.world.

leading to the observed rising Tobin's Q ratios between market and book value of corporations).

Another key institutional factor to understand the rise of private wealth-income ratios is the gradual transfer from public wealth to private wealth. This is particularly spectacular in the case of China, where the share of public wealth in national wealth dropped from about 70% in 1978 to 35% by 2015 (panel B). The corresponding rise of private property has important consequences for the levels and dynamics of inequality of income and wealth. In rich countries, net public wealth (public assets minus public debts) has become negative in the USA, Japan and the UK, and is only slightly positive in Germany and France. This arguably limits government ability to redistribute income and mitigate rising inequality. The only exceptions to the general decline in public property are oil-rich countries with large public sovereign funds, such as Norway.

5. Wealth Inequality Dynamics

In this section we present findings on wealth inequality on Figure III. We stress that currently available statistical information on the distribution of wealth and cross-border assets are highly imperfect in today's global economy. More transparency and better access to administrative and banking data sources are sorely needed if we want to gain knowledge of the underlying evolutions. In WID.world, we combine different sources and methods in a very transparent way in order to reach robust conclusions: the income capitalization method (using income tax returns), the estate multiplier method (using inheritance and estate tax returns), wealth surveys, national accounts, rich lists and generalized Pareto curves. Nevertheless, our series should still be viewed as imperfect, provisional, and subject to revision. We provide access to our data files and computer codes so that everybody can use them and contribute to improve the data collection.[12]

We observe a large rise of top wealth shares in the USA and China in recent decades, and a more moderate rise in France and the UK. A combination of factors explains these different dynamics. First, higher income inequality and severe bottom income stagnation can naturally explain higher wealth inequality in the USA. Next, the very unequal process of privatization and access by Chinese households to quoted and unquoted equity probably played an important role in the very fast rise of wealth concentration in China, particularly at the very top end. The potentially large mitigating impact of high real estate prices should also be taken into account. This middle class effect is likely to have been particularly strong in France and the UK, where housing prices have increased significantly relative to stock prices.

Given all these factors, it is not an easy task to predict whether the observed trend of rising concentration of wealth will continue. In the long run, steady-state wealth inequality depends on the inequality of saving rates across income and wealth groups, the inequality of labor incomes and rates of returns to wealth, and the progressivity of income and wealth taxes. Numerical simulations show that the response of steady-state wealth inequality to relatively small changes in these structural parameters can be rather large (see Saez & Zucman, 2016; Garbinti *et al.*, 2016). In our view, this instability reinforces the need of increasing transparency about the dynamics of income and wealth.

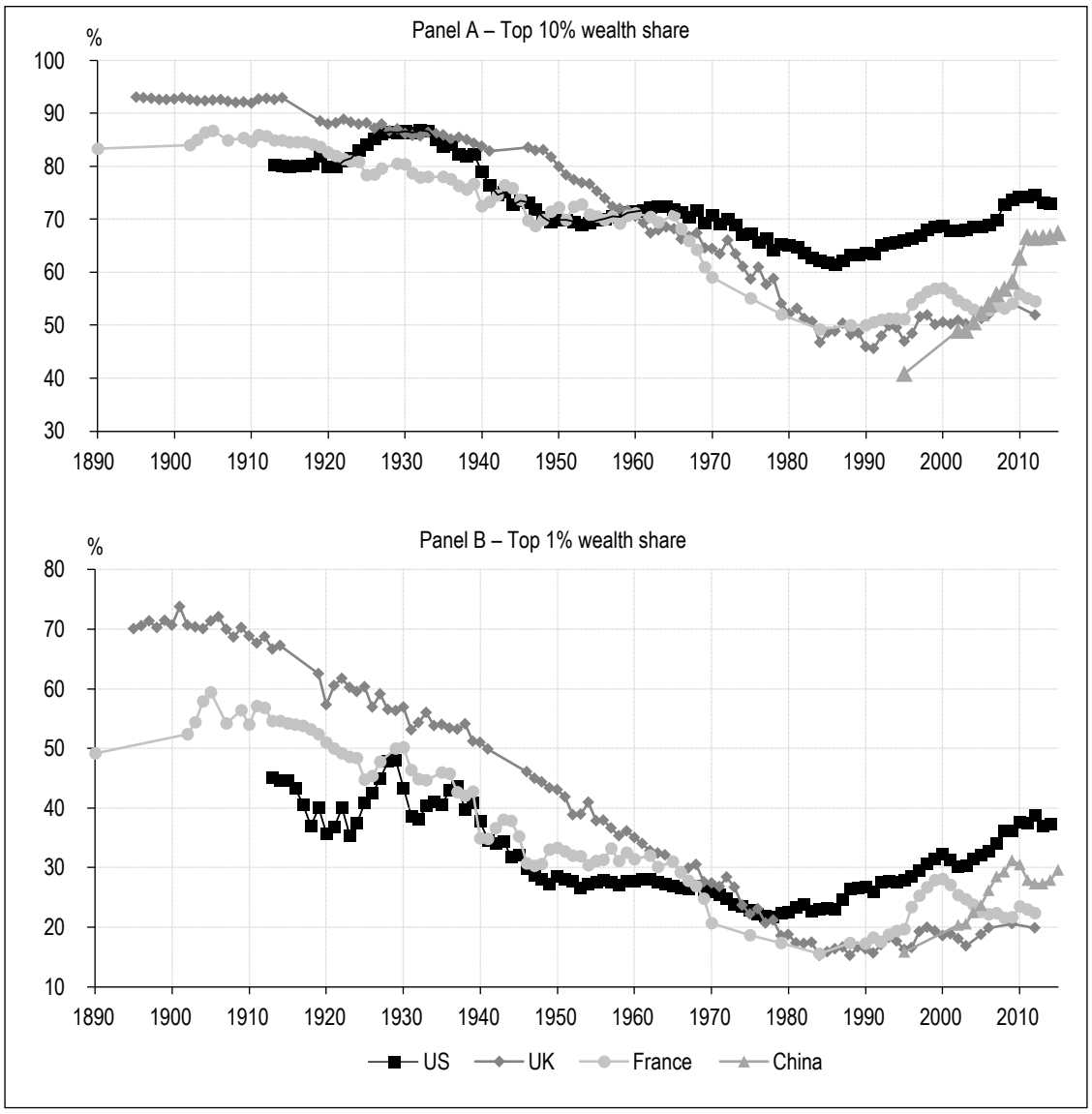
6. Global Income Inequality Dynamics

The dynamics of global inequality has also attracted growing attention in recent years. This, in part, should not be surprising, as it reflects the recognition that the distribution of income and wealth are not only determined at the national state level, but also (and necessarily) at the world level. As we have discussed in previous sections, inequality has been increasing in many countries, but large emerging countries (India, China) are catching up, with the effect of driving global inequality down. Recent studies, based on adjusted household survey data, provide valuable estimates (Lakner & Milanovic, 2015; Anand & Segal, 2008, 2017; Liberati, 2015; Ortiz & Cummins, 2011). Surveys, however, are not uniform across countries; they do not capture high incomes well, and are not consistent with macroeconomic totals. Such limitations remind again of the need of developing DINA series.

Using simple assumptions, we estimate the evolution of incomes in the rest of the world (that is, in the countries and regions not covered yet by the DINA estimates discussed in section 4) so as to distribute 100% of global income. We start with aggregate national income and adult population in all countries and assume that countries with missing inequality information have the same level of inequality as other

12. We refer to the country-specific papers for detailed discussions; see Saez & Zucman, 2016; Alvaredo, Atkinson & Morelli, 2016, 2018; Garbinti et al.*, 2016; Piketty* et al.*, 2017.*

Figure III – **Top 10% and Top 1% wealth share in China, USA, France and UK, 1890-2015**



Notes: Distribution of net personal wealth among adults. Corrected estimates (combining survey, fiscal, wealth and national accounts data). For China, USA and France, equal-split-adult series (wealth of married couples divided by two); for UK, adult series.
Sources: USA: Saez & Zucman (2016); UK: Alvaredo, Atkinson & Morelli (2017, 2018); France: Garbinti *et al.* (2016); China: Piketty *et al.* (2017).

countries in their region. This is obviously an over simplification and our estimates will be refined as better data become available for more countries. Robustness tests and novel results using more detailed distributional information for missing countries suggest that our findings appear to be robust to these simplifications.[13] We stress that this exercise on income aggregation at the world level is possible mainly thanks to the fact that the DINA income concept is homogeneous across countries.

A powerful way to visualize the evolution of global income inequality dynamics is to plot the rate of growth at each percentile following Lakner & Milanovic (2015). We do this in Figure IV. The top percentile of the global income distribution earns over 20% of total global income today, and has captured 27% of total income growth from 1980 to 2016 (these growth rates are obtained once all the individuals of the different regions are pooled together using purchasing power parity exchange rates). To reflect its outsized importance, we further split it into 28 smaller groups: P99-99.1,…, P99.8-99.9, P99.9-99.91,..., P99.98-99.99,

13. The methodological details and robustness checks are presented in Chancel & Gethin (2017); all data and programs are available from WID.world. Estimates for Europe are discussed in Blanchet et al. *(2019).*

P99.99-99.991,…, P99.999-100. Growth rates are low at the very bottom due to low growth in the poorest countries (mostly in sub-Saharan Africa). Growth rates are quite high in percentiles 20 to 60 due to the high growth in large emerging countries (China and India). They are low in percentile 70 to 90 due to the modest growth of the incomes of the poor and middle classes in advanced economies. Finally, they are extremely high among top earners due to the explosion of top incomes in many countries. Therefore, this curve has the shape of an elephant (Lakner & Milanovic, 2015), with a long trunk.

Table shown earlier presents in an alternative way the growth rates of different groups for the world as a whole (as in Figure IV, we use purchasing power parity exchange rates to pool incomes together). Average global growth is relatively low (60%) compared to emerging countries' growth rates. At the world level (and contrary to what is observed in most countries), growth rates do not rise monotonically with income. Instead, we observe high growth for the bottom 50% (94%), low growth in the middle 40% (43%), and high growth for the global top 1% (101%), and especially the top 0.001% (235%).

Figure V shows the evolution of the global top 1% and bottom 50% income shares between 1980 and 2016. The global top 1% income share rose from about 16% in 1980 to more than 22% in 2007. It was then slightly reduced to 20.4% in 2016. The bottom 50% income share oscillated around 9% with a very slight increase between 1985 and 2016. Throughout the period, the top 1% earns in total about twice as much income as the bottom 50%, a group by definition 50 times more numerous. Hence, incomes of the global top 1% income are on average 100 times those of the global bottom 50%. Another notable finding is that neither high growth in emerging countries since 2000 nor the global financial crisis of 2008 stopped the rise in global income inequality.

Whether future growth in emerging countries will be enough to revert this trend is a key question that we now discuss. The right side of Figure V displays different possible global income inequality scenarios until 2050. The number of variables that we consider in our analysis is limited. This makes our projections straightforward and simple to understand, but it obviously limits their predictive power. Our projections are based on combining the demographic projections of the United Nations (UNDESA, 2017) with the OECD growth forecasts (OECD, 2017) and simple assumptions on how growth will be distributed within each country.[14] We consider three scenarios on growth distribution within countries. All three scenarios have the same between-country inequality evolutions (i.e., a given country has the same average income growth rate in all three scenarios).

Our first scenario represents an evolution based on "business as usual", that is, we assume that economic growth in each country will be distributed across percentiles in the same way as it has been distributed since 1980. For instance, the bottom 50% income earners in China captured 13% of total growth over the 1980-2016 period. We thus assume that the bottom 50% earners in China will capture 13% of growth up to 2050. The second scenario illustrates a high within-country inequality setting; it assumes that all countries will follow the same inequality trajectory as the United States did over the 1980-2016 period. The third scenario considers a low inequality trend; it assumes that all countries will follow the same inequality trajectory as the European Union did over the 1980-2016 period.

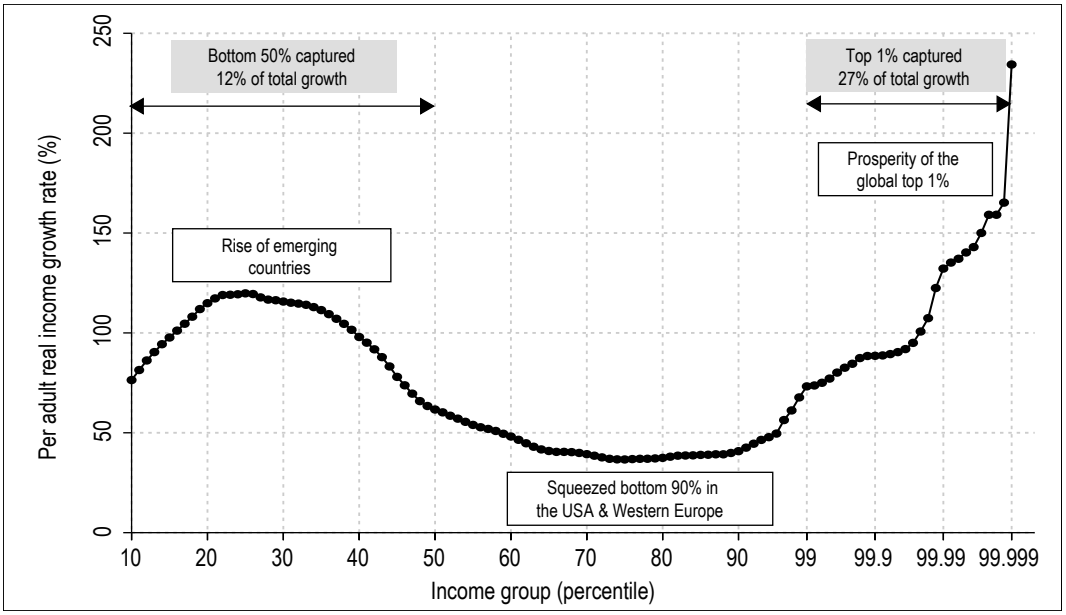
Under the business-as-usual scenario, the income share of the bottom 50% of the world population slightly decreases from approximately 10% today to less than 9% in 2050. The top 1% share rises from less than 21% today to more than 24% of world income. Global inequality thus rises steeply in this scenario, despite strong growth in emerging countries. The progressive catching-up of low-income countries would not be sufficient to counterbalance the worsening of within-country inequality at the current rates.

In the US-style inequality scenario, the global top 1% would earn 28% of global income by 2050, while the bottom 50% would earn 6%, less than in 1980 (before large emerging countries started to catch up with the industrialized world). In this scenario, the increase in the top 1% income share is largely, but not entirely, made at the expense of the bottom 50%.

The last scenario shows that global inequality can be reduced if all countries align on the

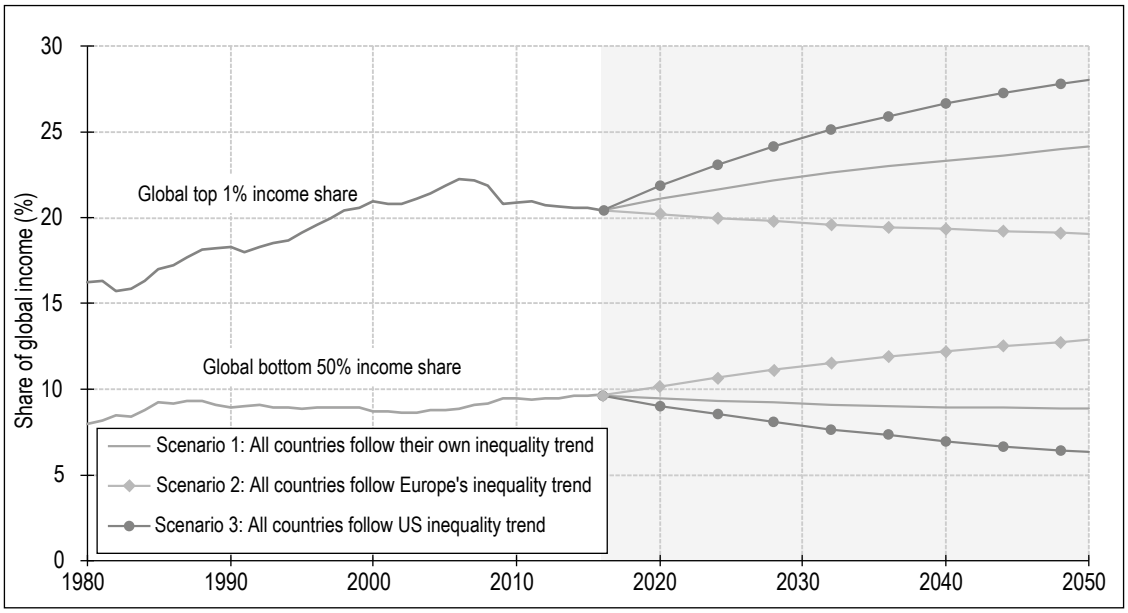
14. The growth rates we use are more optimistic than the rates assumed by the OECD to compute their total global income in 2050 for Africa, Latin America, and Asia. Assuming higher growth rates increases the force of convergence between countries, and hence tends to reduce global inequality. Therefore, we take a conservative approach to the rise of global inequality in the coming decades (for details see Alvaredo et al.*, 2018, and Chancel & Gethin, 2017).*

Figure IV – **Total income growth by percentile across all world regions, 1980-2016**



Notes: The vertical axis shows the total real income growth between 1980 and 2016 for each percentile of the global distribution of income per adult. The bottom 10 percentiles are excluded as their income levels are close to zero. The top 1% is divided into smaller groups (up to the top 0.001%) so as to better account for its share in total global growth captured.
Sources: WID.world.

Figure V – **Top 1% *vs.* bottom 50% shares of global income, 1980–2050**



Notes: This figure displays the global top 1% and bottom 50% income shares with data from 1980 to 2016, and projections from 2016 to 2050 under three scenarios for inequality: 1. Business as usual, 2. European scenario, 3. USA scenario; e.g. if all countries follow the inequality trajectory of the USA between 1980 and 2016 from 2017 to 2050, the income share of the global top 1% will reach 28% by 2050.
Sources: WID.world.

European inequality trajectory – or more equitable ones. The bottom 50% income share would rise from 10% to 13% in 2050, whereas the top 1% would decrease from 21% to 19% of total income. Even more equitable growth trajectories would be needed for the global bottom 50% share to catch up with the top 1% income share by mid 21st century.

We should stress again that there is much to be improved in the data underlying such projections. As DINA become available for more

countries and more years, we will be able to refine our understanding of global income inequality dynamics. What these scenarios suggest, however, is that global inequalities are likely to remain substantial in the coming decades.

* *
*

We stress that global inequality dynamics involve strong and contradictory forces. We observe rising top income and wealth shares in nearly all countries in recent decades. But the magnitude of the increase varies substantially, thereby suggesting that different country-specific policies and institutions matter. High growth rates in emerging countries reduce between-country inequality, but this in itself does not guarantee acceptable within-country inequality levels, and does not ensure the social sustainability of globalization. Access to more and better data is critical to monitor global inequality dynamics, as this is a key building block both to properly understand the present as well as the forces that will dominate in the future, and to design potential policy responses.

There are a number of limitations in the data sources we are using to create DINA statistics that we would like to explicitly mention.

First, the scope of individual fiscal income (i.e. income as reported through tax-based sources) has deteriorated over time as many countries have chosen to exclude large components of capital income from the individual income tax. Countries such as Sweden and Germany have moved to a dual income tax system where capital income is taxed separately at a flat rate. Other countries have carved out large exemptions, such as tax-preferred life insurance accounts in France. As a result, the quality of the (necessary) imputation of capital income deteriorates. However, in most cases, the government still receives – or could collect at very low cost – information on exempted capital income on an individual basis. Countries such as Denmark for example do tax dividends and capital gains separately from other income, but it is still possible to merge both data sources at the individual level. Additionally, administrative wealth data are much sparser than income tax data because progressive wealth taxation is much less prevalent than progressive income taxation. Yet it would be possible to gather and collect wealth data at very low cost. Such data would be invaluable to measure wealth inequality but would also help with the administration of the progressive income tax. Once again, Denmark provides a good illustrative example: even if the country abolished its wealth tax in 1997, data on balances of individual financial accounts are still collected for the administration of the individual income tax on capital income.

Second, survey data could be greatly improved if they were systematically linked to administrative data.[15] Linkage with administrative data is useful both for sampling and for data quality. The US Survey of Consumer Finances is one of the most successful examples of the value of using administrative tax data to oversample the top of the wealth distribution and to capitalize investment incomes for the improvement of both the sampling framework and the accuracy of the estimates.[16] In this sense, survey data and administrative data should become complements instead of being viewed as competitors.[17]

Third, administrative data can be very defective in situations where large parts of the economy are informal. This is the still the situation in many emerging countries today. In such cases, surveys remain necessary to cover the full population.

As we stressed at the beginning of the paper, the production of distributional national accounts can only be sustained over time with the collaboration between national accountants, tax departments, statisticians and academic researchers within and across countries. □

---

15. Blanchet et al. *(2018) provide a method to reweight surveys using tax data where both sources are not yet linked.*

16. The Enquête Patrimoine *in France also applies oversampling strategies based on administrative data; this could be further improved by taking additional external information from the capitalization of investment incomes, and by using administrative data on assets.*

17. Meyer et al. *(2015) document a noticeable and worrying rise in unit non-response, item non-response, and measurement error in a number of USA household surveys. Those threats to survey quality seem to be a widespread phenomenon across countries.*

BIBLIOGRAPHY

Alvaredo, F., Assouad, L. & Piketty, T. (2019). Measuring lnequality in the Middle East 1990-2016: The World's Most Unequal Region? *Review of Income and Wealth,* 65(4), 695–711. https://doi.org/10.1111/roiw.12385

Alvaredo, F., Atkinson, A. B., Piketty, T. & Saez, E. (2013). The Top 1% in International and Historical Perspective. *Journal of Economic Perspectives*, 27(3), 3–20. https://www.doi.org/10.1257/jep.27.3.3

Alvaredo, F., Atkinson, A. B., Piketty, T. & Saez, E. (2011-2015). The World Top Incomes Database. https://40ddec72-a-62cb3a1a-s-sites.googlegroups.com/site/alvaredo/Home/WTID_archived_2015_11.pdf

Alvaredo, F., Atkinson, A. B. & Morelli, S. (2016). The Challenge of Measuring UK Wealth Inequality in the 2000s. *Fiscal Studies*, 37(1), 13–33. https://doi.org/10.1111/j.1475-5890.2016.12084

Alvaredo, F., Atkinson, A. B., Chancel, L., Piketty, T., Saez, E. & Zucman, G. (2016). Distributional National Accounts (DINA) Guidelines: Concepts and Methods used in the World Wealth and Income Database. WID.world *Working Paper* 2016/2, version 9 June 2017. https://wid.world/document/dinaguidelines-v1/

Alvaredo, F., Atkinson, A. B. & Morelli, S. (2017). Top Wealth Shares in the UK over more than a century. CEPR *Discussion Paper* 11759. https://cepr.org/active/publications/discussion_papers/dp.php?dpno=11759

Alvaredo, F., Atkinson, A. B. & Morelli, S. (2018). Top Wealth Shares in the UK over more than a century. *Journal of Public Economics*, 162, 26–47. https://doi.org/10.1016/j.jpubeco.2018.02.008

Alvaredo, F., Chancel, L., Piketty, T., Saez, E. & Zucman, G. (2017). Global Inequality Dynamics: New Findings from WID.world. *American Economic Review: Papers & Proceedings*, 107(5), 404–409. https://doi.org/10.1257/aer.p20171095

Alvaredo, F., Chancel, L., Piketty, T., Saez, E. & Zucman, G. (2018). *The World Inequality Report 2018*. Cambridge, USA: Harvard University Press. http://wir2018.wid.world/.

Alvaredo, F., Chancel, L., Piketty, T., Saez, E. & Zucman, G. (2018b). The Elephant Curve of Global Inequality and Growth. *American Economic Review: Papers and Proceedings*, 108, 103–108. https://doi.org/10.1257/pandp.20181073

Anand, S. & Segal, P. (2008). What Do We Know about Global Income Inequality? *Journal of Economic Literature*, 46(11), 57–94. https://doi.org/10.1257/jel.46.1.57

Anand, S. & Segal, P. (2017). Who are the Global Top 1%? *World Development*, 95, 111–126. https://doi.org/10.1016/j.worlddev.2017.02.001

Atkinson, A. B. (2005). Top Incomes in the U.K. over the 20th Century. *Journal of the Royal Statistical Society*, 168(2), 325–343. ttps://doi.org/10.1111/j.1467-985X.2005.00351.x

Atkinson, A. B. (2007). Measuring Top Incomes: Methodological Issues. *In:* Atkinson, A. B. & Piketty, T. (eds.), op. cit., chap.2, pp.18–43.

Atkinson, A. B. & Harrison, A. (1978). *Distribution of Personal Wealth in Britain*. Cambridge, UK: Cambridge University Press.

Atkinson, A. B. & Piketty, T. (2007). *Top Incomes over the Twentieth Century. A Contrast Between European and English Speaking Countries*. Oxford: Oxford University Press.

Atkinson, A. B. & Piketty, T. (2010). *Top Incomes. A Global Perspective*. Oxford: Oxford University Press.

Atkinson, A. B., Piketty, T. & Saez, E. (2011). Top Incomes in the Long-Run of History. *Journal of Economic Literature*, 49(1), 3–71. https://doi.org/10.1257/jel.49.1.3

Blanchet, T. & Chancel, L. (2016). National Accounts Series Methodology. WID.world *Working Paper*. https://wid.world/document/dinaguidelines-v1/

Blanchet, T., Chancel, L. & Gethin, A. (2019). Income inequality in Europe, 1980-2017. WID.world *Working Paper* 2019/06. https://wid.world/document/bcg2019-full-paper/

Blanchet, T., Fournier, J. & Piketty, T. (2017). Generalized Pareto Curves: Theory and Applications. WID. world *Working Paper* 2017/3. https://wid.world/document/blanchet-t-fournier-j-piketty-t-generalized-pareto-curves-theory-applications-2017/

Blanchet, T., Morgan, M. & Flores, I. (2018). The weight of the rich: improving surveys using tax data. WID. world *Working Paper* 2018/12. https://wid.world/document/the-weight-of-the-rich-improving-surveys-using-tax-data-wid-world-working-paper-2018-12/

Bourguignon F. (2015). Appraising income inequality databases in Latin America. *Journal of Economic Inequality*, 13(4), 557–578. https://doi.org/10.1007/s10888-015-9304-4

Bozio, A., Garbinti, B., Goupille-Lebret, J., Guillot, M. & Piketty, T. (2018). Inequality and redistribution in France 1990-2018. Evidence from post-tax Distributional National Accounts. WID.world *Working Paper* 2018/10. https://wid.world/document/inequality-and-redistribution-in-france-1990-2018-evidence-from-post-tax-distributional-national-accounts-dina-wid-world-working-paper-2018-10/

Chancel, L. & Gethin, A. (2017). Building a global income distribution brick by brick WID.world Technical Note 2017/5.
https://wid.world/document/building-global-income-distribution-brick-brick-wid-world-technical-note-2017-5/

Chancel, L. & Piketty, T. (2017). Indian income inequality 1922-2015. From British Raj to Billionaire Raj. WID.world *Working Paper* 2017/11. https://wid.world/document/chancelpiketty2017widworld/

Deaton, A. (2005). Measuring Poverty in a Growing World (or Measuring Growth in a Poor World). *The Review of Economics and Statistics*, 87(1), 1–19. https://doi.org/10.1162/0034653053327612

Frankel, S. H. & Herzfeld, H. (1943). European income distribution in the Union of South Africa and the effect thereon of income taxation. *South African Journal of Economics*, 11(2), 121–136.

Garbinti, B., Goupille-Lebret, J. & Piketty, T. (2016). Accounting for Wealth Inequality Dynamics. Methods, Estimates and Simulations for France, 1800-2014. WID.world *Working Paper* 2016/5.
https://wid.world/document/b-garbinti-j-goupille-and-t-piketty-wealth-concentration-in-france-1800-2014-methods-estimates-and-simulations-2016/

Garbinti, B., Goupille-Lebret, J. & Piketty, T. (2018). Inequality Dynamics in France, 1900-2014: Evidence from Distributional National Accounts (DINA). *Journal of Public Economics*, 162, 63–77.
https://doi.org/10.1016/j.jpubeco.2018.01.012

Kuznets, S. (1953). *Shares of Upper Income Groups in Income and Savings*. NBER.

Lakner, C. & Milanovic, B. (2015). Global Income Distribution: From the Fall of the Berlin Wall to the Great Recession. *World Bank Economic Review,* 30(2), 203–232. https://doi.org/10.1093/wber/lhv039

Liberati, P. (2015). The World Distribution of Income and Its Inequality, 1970–2009. *Review of Income and Wealth*, 61, 248–273. https://doi.org/10.1111/roiw.12088

Meyer, B. D., Mok, W. & Sullivan, J. (2015). Household surveys in crisis. *Journal of Economic Perspectives*, 29(4), 199–226. https://doi.org/10.1257/jep.29.4.199

Morgan, M. (2017). Falling Inequality beneath Extreme and Persistent Concentration: New Evidence for Brazil Combining National Accounts, Surveys and Fiscal Data, 2001-2015. WID.world *Working Paper* 2017/12. https://wid.world/document/extreme-persistent-inequality-new-evidence-brazil-combining-national-accounts-surveys-fiscal-data-2001-2015-wid-world-working-paper-201712/

Nolan, B., Roser, M., & Thewissen, S. (2018). GDP per capita versus median household income: What gives rise to divergence over time and how does this vary across OECD Countries? *Review of Income and Wealth.*
https://doi.org/10.1111/roiw.12362

Novokmet, F., Piketty, T. & Zucman, G. (2018). From Soviets to oligarchs: inequality and property in Russia 1905-2016. *Journal of Economic Inequality*, 16(2), 189–223. https://doi.org/10.1007/s10888-018-9383-0

OECD (2017). GDP long-term Forecast. http://dx.doi.org/10.1787/d927bc18-en.

Ortiz, I. & Cummins, M. (2011). Global Inequality: Beyond the Bottom Billion – A Rapid Review of Income Distribution in 141 Countries. *UNICEF Social and Economic Working Paper*.
https://www.unicef.org/socialpolicy/files/Global_Inequality_REVISED_-_5_July.pdf

Piketty, T. (2001). *Les hauts revenus en France au XXème siècle: Inégalités et redistributions 1901–1998*. Paris: Grasset.

Piketty, T. (2003). Income Inequality in France, 1901–1998. *Journal of Political Economy*, 111(5), 1004–1042. https://doi.org/10.1086/376955

Piketty, T. (2014). *Capital in the 21st Century.* Cambridge: Harvard University Press.

Piketty, T. & Saez, E. (2003). Income Inequality in the United States, 1913-1998. *The Quarterly Journal of Economics,* 118(1), 1–39. https://doi.org/10.1162/00335530360535135

Piketty, T., Saez, E. & Stantcheva, S. (2014). Optimal Taxation of Top Labor Incomes. *American Economic Journal: Economic Policy,* 6(1), 230–271. https://doi.org/10.1257/pol.6.1.230

Piketty, T., Saez, E. & Zucman, G. (2018). Distributional National Accounts: Methods and Estimates for the United States. *The Quarterly Journal of Economics*, 133(2), 553–609. https://doi.org/10.1093/qje/qjx043

Piketty, T., Saez, E. & Zucman, G. (2019). Simplified Distributional National Accounts. WID.world *Working Paper* 2019/1.
https://wid.world/document/simplified-distributional-national-accounts-wid-world-working-paper-2019-01/

Piketty, T. & Zucman, G. (2014). Capital is Back: Wealth-Income Ratios in Rich Countries, 1700-2010. *Quarterly Journal of Economics,* 129(3), 1255–1310. https://doi.org/10.1093/qje/qju018

Piketty, T., Yang, L. & Zucman, G. (2017). Capital Accumulation, Private Property and Rising Inequality in China 1978-2015. WID.world *Working Paper* 2017/6. Published (2019): *American Economic Review*, 109(7), 2469–2496. https://doi.org/10.1257/aer.20170973

Ravallion, M. (2003). Measuring aggregate welfare in developing countries: how well do national accounts and surveys agree? *The Review of Economics and Statistics*, 87, 645–652.
https://doi.org/10.1162/003465303322369786

Saez, E. & Zucman. G. (2016). Wealth Inequality in the United States: Evidence from Capitalized Income Tax Data. *The Quarterly Journal of Economics,* 131(2), 519–578. https://doi.org/10.1093/qje/qjw004

Tørsløv, T., Wier, L. & Zucman, G. (2018). The Missing Profits of Nations. NBER *Working Paper* No. 24701.
https://www.nber.org/papers/w24701

UNDESA (2017). UN Population Prospects. https://esa.un.org/unpd/wpp/

Zucman, G. (2014). Taxing Across Borders: Tracking Personal Wealth and Corporate Profits. *Journal of Economic Perspectives*, 28(4), 121-148. https://www.aeaweb.org/articles?id=10.1257/jep.28.4.121

Why and How Should Human Capital be Measured in National Accounts?

Nicolas Canry*

Abstract – There is currently a significant divergence in the way in which education expenditure is perceived in economic theory and in national accounting: the former treats it as investment, the latter as consumption. In fact, the accounting framework is still structured around two major production factors (labour and physical capital), whereas human capital appears to be essential if certain current phenomena are to be perceived accurately, notably the resurgence of inequality in certain countries. This paper presents the work undertaken to incorporate human capital into national accounts and explores the two main methods used: that based on costs (inputs) and that based on income (output). We go on to use the inputs method to estimate the savings rate of USA, French and British households when education and health expenditure is transferred to investment. Only the inclusion of health expenditure would enable the USA savings rate to be redressed significantly.

JEL Classification: E01, E21, E24
Keywords: human capital, national accounting, savings, inequality

* *CES (Centre d'Économie de la Sorbonne) – Université Paris 1 Panthéon-Sorbonne (nicolas.canry@univ-paris1.fr)*

I would like to thank two anonymous reviewers for their comments and their suggestions.

Received on 30 June 2018, accepted after revision on 16 June 2019.
Translated from: "Pourquoi et comment mesurer le capital humain dans la comptabilité nationale ?"
Citation: Canry, N. (2020). Why and How Should Human Capital be Measured in National Accounts? *Economie et Statistique / Economics and Statistics*, 517-518-519, 61–79. https://doi.org/10.24187/ecostat.2020.517t.2023

Although "human capital" was first evoked by Adam Smith (1776), the articles of Schultz (1961, 1962) and Becker (1962) have contributed decisively to the inclusion of this concept in modern economic theory. In the view of these authors, agents' education expenditure is an investment with a view to the accumulation of a stock of knowledge, namely human capital. Conversely, in national accounting, the education expenditure of the various institutional sectors (households, public administrations) is a consumption expenditure: agents consume (and therefore destroy) an education service, meaning that this operation does not give rise to the accumulation of any assets: human capital does not appear in the accounts of the agents' assets or, *a fortiori*, in the national accounts.

In fact, this conceptual divergence between the national accounting framework and the economics theoretical framework seems to be particularly salient in certain ongoing economic debates. Two important points should be mentioned in this regard:

- The dualism of the labour market, particularly the consequences it can generate in terms of income inequality. Many studies carried out in recent years seem to suggest that capital income is of secondary importance in the resurgence of income inequality in recent decades and that human capital now plays a much more central role in the genesis of such inequality. Other analyses indicate, conversely, that human capital at best only explains some of the inequalities currently being observed. In fact, the (international) accounting framework, which is still founded today on a productive model based on two factors, labour and physical capital, fails to support the economic theory with empirical data, which is nonetheless crucial to these issues.

- Currently, the accounting framework still adopts a fairly narrow definition of household investment, which is limited to the household's acquisition of immovable property. Extending the scope of household investment expenditure to education or health would in fact have a direct impact on the estimation of household savings. However, household saving behaviour (as well as these factors) is, again, at the centre of numerous macroeconomic debates of recent decades, whether considering the regular fall in savings rates of USA households since 1980 or, conversely, the notably high level of savings rates of Chinese households (Chamon & Prasad, 2010), which is probably largely responsible for the "saving glut" identified in 2005 by Bernanke (2005).

This paper focuses on all of these issues, exploring in particular how national accounting could incorporate human capital into its accounts. The first section sets out an overview of the research studies relating to the resurgence of income inequality over the past thirty years or so in numerous developed countries, placing human capital at the heart of this key economic debate. After having briefly looked at the way in which human capital is perceived by economic theory (section 2), we set out the empirical studies undertaken over several decades, with a view to estimating human capital accounting series, exploring in particular the two principal methods – costs (or input) method and income (or output) method – used in these studies. These two approaches generally result in substantially different estimates. As the output method is very cumbersome to implement, we set out certain results of recent studies which produced human capital series using this method, notably for the USA. The final section focuses on the input method to construct alternative indicators to the savings rate of USA, French and British households, once their education and then health expenditure is deducted from consumption. While such an approach seems to have an impact on savings rate levels, the effect on changes to those levels appears modest, education expenditure having remained relatively stable (as a percentage of gross domestic product – GDP) in the countries under consideration.

1. The Role of Human Capital in the Resurgence of Income Inequality

Following the seminal paper of Solow (1956) on economic growth, human capital rapidly came to be considered as an essential contributing factor to growth. Denison (1962) establishes a positive correlation between the Solow residual and education, thereby paving the way for an extension of economic growth factors and for the first attempts to estimate human capital and its returns. These studies concluded with the analysis of Mankiw *et al.* (1992), who offer an extended version (incorporating human capital) of Solow's model, which they then estimate for an international cross-section, valuing human capital using secondary school enrolment rates in the countries under consideration.

Human capital also plays a key role in the analysis of inequality. The resurgence of significant income inequality in certain developed countries since 1980 has therefore resulted in various academic studies in recent years. Indeed, the resurgence of inequality in

numerous English-speaking countries over the past three decades has led to challenges to the idea, illustrated in the "Kuznets curve", that the relationship between development and inequality is bell-shaped.[1]

Historically, and rather simplistically, it could be said that the national accounting framework lent itself perfectly to the analysis of inequality, given that such inequality was based principally on the distinction between a minority of the population, drawing its wealth from capital income (capital income characteristically being highly concentrated at that time) and the rest of the population, receiving income from labour. In these circumstances, income inequality remained closely linked to primary income distribution, and therefore entirely in line with the national accounting framework. The fall in inequality observed in most developed countries during the first half of the twentieth century is due to the spectacular fall in the income held by this small minority at the top end of the income scale, which Piketty (2001), following Keynes, classes as euthanasia of the rentiers (there are multiple causes: war, the 1929 crisis, increasing use of progressive taxation). While capital income has represented a relatively stable share of income throughout the twentieth century, it is now distributed across a much wider spectrum of the population: it is substantially "diluted" across a relatively sizeable middle class.

As of 1970, inequality resurfaces but does not in any way seem to be connected (at least until recently) to a revival of rentiers. The factors behind this resurgence are now fairly clear, with two principal explanations generally given. Firstly, globalisation: following the Hekscher-Ohlin-Samuelson (HOS) model, international specialisation is based on the factors available to the different economies; accordingly, rich countries, having substantial physical and human capital available, will specialise in goods intensive in these factors (high technology sectors, etc.) whereas developing countries specialise in sectors intensive in unskilled labour. This rise in inequality in developed countries (to which this section is devoted) can nevertheless be combined with a reduction in inequality at a global level, resulting from a fall in inequality between developed countries and developing countries (Bourguignon, 2015). Secondly, technical progress: new information and communication technology is produced by qualified workers (IT professionals, engineers, etc.) and replaces unskilled labour – as well as, increasingly, the routine tasks of intermediary professions (see Autor *et al*., 2008) in production chains. Technical progress is therefore biased in favour of skilled labour (Acemoglu, 2002).

Until recently, most academic studies concluded that biased technical progress was, by far, the principal factor explaining the increase in inequality (Berman *et al.*, 1994). However, recent studies are more nuanced and show the growing influence of globalisation on income inequality in countries, notably in the USA (Acemoglu *et al*., 2016). In fact, the two factors put forward are based on the same market mechanisms: the increase in inequality in developed countries results from the fall in demand for unskilled labour and the corresponding increase in demand for skilled labour (the curves move in a similar way in both explanations, but the causes of the "shocks" differ). Therefore, it is clear that the inequality trends are now occurring even within the sphere of labour earnings alone and that, whatever explanation is adopted, they are caused by differing trends in the demand for unskilled labour (simple labour) and skilled labour (human capital). However the national accounting framework in force is founded implicitly on a production function based on two major factors: labour and physical capital. It is therefore less suited to the analysis of the interaction at play within the domain of labour earnings itself, between skilled and unskilled labour. Furthermore, this framework is all the more outdated given that the boundary between pay for labour and pay from capital now appears to be increasingly blurred, shareholders seeking, within a principal-agent relationship, to bring managers' interests in line with their own: performance-related bonuses, stock-options, etc.

According to Goldin & Katz (2010), the resurgence of inequality can only be accurately perceived by focusing exclusively on company demand for human capital: the supply of human capital, which partially depends on the level of investment in education by the public authorities, must also be taken into consideration. While the skill biased technical change theory focuses on the specific features of ICT, affecting the demand for skilled and unskilled labour in different ways for the past thirty years or so, Goldin and Katz argue that, conversely, the

1. According to this curve, economic expansion is initially associated with an increase in inequality (between those instigating the expansion and benefiting from it fully and the rest of the population). Subsequently, this inequality narrows, the entire population ultimately benefiting from the economic development in terms of both productivity and pay, through diffusion and generalisation.

increase in demand for human capital is not recent: the difference between the 1950-1980 and 1980-2010 periods is due primarily to changes in the supply of human capital: regular growth until 1980 (in the case of the USA) therefore occurred alongside the increase in supply, but was followed by stabilisation. It is therefore the "race between education and technology" which explains the increase in inequality, demand for skilled labour growing more rapidly than the stock of human capital since 1980. According to Verdugo (2014), this analysis also explains the trajectory of income inequality in France since 1950: inequality effectively widened until 1965 and then narrowed, remaining relatively stable after 1980. Unlike in the USA, however, the investment effort in education was relatively late in France and took place primarily in the 1950s and 1960s; there was a delayed impact on the supply of human capital, which explains the growth in income inequality during the so-called "thirty glorious years". Conversely, and contrary to the USA, the investment effort in education continued in the 1980s and 1990s, which may explain why France was spared the return of inequality in recent decades.

The study undertaken by Autor (2014) confirms very clearly that in the United States, the speed of the rise in real incomes is closely linked to the level of study, notably since the start of the 1980s, which confirms the role of human capital in the widening of inequality. Piketty (2013), on the other hand, points out that, in the USA, the upper percentile of employees, that of super-managers or managers of large groups, has monopolised a very substantial share of the increases in the national wage bill for the past thirty years. Other interpretations of inequality have therefore been put forward: according to Gabaix & Landier (2008) in particular, the increase in Chief executive officers' (CEOs) remunerations can be explained by large groups competing to recruit the most talented individuals, as only they are able to respond to a constantly changing, increasingly unstable environment. However, not everyone agrees with this analysis: Bertrand & Mullainathan (2001) demonstrate that managers' pay is governed more by luck than by their performance (companies benefiting from noticeable positive shocks which are entirely distinct from managers' strategy pay those managers better than companies which do not have the benefit of such shocks). For these authors, the asymmetric information relationship between shareholders and managers enables managers to determine their own pay in many situations. More generally, authors such as Piketty (2013) and Krugman (2007) consider the institutional, or "sociological", dimension, which encouraged both the surge in income for super-managers and the decline in income at the lower end of the distribution of earnings (particularly in the USA): fall in the real minimum wage (Lee, 1999), erosion of the power of trade unions (Lemieux, 2008), capacity of the current economic elites to modify, to their advantage, social norms (on this point, see Akerlof, 1980) which were put in place long ago, notably during the second world war, the time of the great compression (see Goldin & Margo, 1992) and in the immediate post-war era, to limit wage dispersion.

Lastly, certain authors have placed human capital at the heart of inequality, whilst others downplay its explanatory power in the current era, when other factors seem to play an equally essential part: talent, luck, social norms. Human capital is therefore undoubtedly at the heart of the inequality debate, but measuring it often remains problematical. This is also one of the principal criticisms made by Weil (2015a) of Piketty's book "Capital in the Twenty-First Century": the analysis in the book is based on an empirical study which is most impressive, but never seeks to develop human capital series.

Before moving on to the question of the valuation of human capital within the framework of national accounting, the following section briefly explains how it is perceived in economic theory.

2. Human Capital in Economic Theory

The introduction of human capital into the marginalist framework of economic theory goes back to the work of Schultz (1961, 1962) and Becker (1962, 1964): the individual adopts maximising behaviour to determine the optimal level of education (schooling) he or she should attain. The marginal return to human capital is assumed to be decreasing or (which means the same thing) its marginal cost is increasing[2]: the stock of knowledge which can be acquired through education is limited (at least at a given point in time); the closer the individual gets to the "frontier' of knowledge, the harder it is to acquire marginal knowledge and the greater the (intellectual) effort required. Assuming (which

2. It should be noted that several important contributions to the theory of endogenous growth (for example Lucas, 1988 or Romer, 1990) advance a theoretical framework combining reducing human capital returns at a private (or microeconomic) level and constant, or increasing, returns at a social (or macroeconomic) level, as a result of the existence of a positive externality relating to the stock of human capital.

is highly theoretical) that human capital is a discrete "variable" – that is, it can be divided into distinct units which can be accumulated – the acquisition of each additional unit requires more time to be spent on training than the previous unit, which results in an increasing marginal cost (of human capital). Added to this principle is the fact that the marginal financing of education is also generally increasing: often free or subsidised in the initial years, higher education has to be paid for in many countries and can also require the agent to incur debt, etc. It is important to include in the costs of education the additional costs of transport or accommodation associated with education (notably higher education) and not to overlook its opportunity cost, notably foregone earnings as a result of the decision to continue to study, therefore delaying entry into the labour market; consideration can also be given to the time devoted by parents to their children's educational success.

Moreover, and although investment in human capital is generally reduced to education expenditure, Becker notes that for it to be effective, human capital needs to be "carried" by individuals in good health: a broad vision of investment expenditure should therefore incorporate health expenditure and even agents' expenditure on food (health expenditure notably enables agents' life expectancy and therefore, presumably, their intertemporal utility would be increased). In a first step, the analysis is limited to education expenditure.

In human capital theory, earnings reflect both unskilled labour (which would be achieved without any qualification) and the human capital acquired by the agent, namely the premium associated with qualification, or skill-premium. Here again, it may be assumed, on a highly theoretical basis, that the market will set a "skill-premium rate" representing pay for a unit of human capital. In other words (although this is one of several possible models), the earnings w received by an employee can be broken down as follows:

$$w = w_L + h.w_H \quad (1)$$

where w_L is the earnings rate for unskilled labour, w_H is the "skill-premium rate" and h is the number of units of human capital accumulated by the agent.

The agent will therefore seek to determine the optimal number h of units of human capital which he/she must accumulate considering as a given the skill-premium rate w_H and assuming an increasing marginal cost of this human capital. Using a marginalist calculation, the agent can therefore compare the cost and income associated with any additional unit of human capital which he/she may obtain. To determine this income, the fact that the associated gain $\partial h.w_H$ will be received by the agent throughout his/her working life must of course also be taken into consideration: it is therefore necessary to compare cost and the discounted amount of additional income generated by this additional cost.

The agent pursues his/her studies for as long as the (discounted) marginal income exceeds the marginal cost. At the point of equilibrium, marginal cost and marginal income are equal but average income is clearly quite likely to exceed average cost. In theory, however, the gain associated with the acquisition of human capital therefore encourages new (young) agents to accumulate human capital: this additional supply (of human capital) on the qualified labour market ultimately causes a fall in w_H, meaning that in terms of dynamics, the gain associated with human capital will reduce, or even disappear altogether: in the long term equilibrium, average cost and average revenue (and therefore ultimately total cost and revenue) are equal.

In theory, therefore, the value of the human capital accumulated by the agent can be estimated by valuing either the costs of education he/she pays, or (since the result should be the same) the discounted income flows generated by his/her level of education. The first approach represents a costs-based valuation of human capital (input method), the second an income-based valuation (output method).

Of course, in reality, all agents do not have the same capacity to access human capital, for reasons which are potentially very varied: different cultural baggage inherited from parents, different personal predispositions, or the existence of an imperfect financial market, making funding impossible for certain people. Moreover, the risk associated with investment in human capital (notably failure at school) may dissuade certain risk-averse agents from undertaking study, unless this risk is offset by a high premium (Abraham, 2010). All of these factors reduce the aggregated investment volume and explain the continued discrepancy between marginal return and marginal cost at the equilibrium. In these circumstances, the discounted average income from the human capital exceeds

its average cost at the equilibrium, meaning that these two approaches give different results, the output method therefore giving an estimate of human capital which is higher than for the inputs method.

3. Attempts to Value the Investment in Human Capital and Its Stock

The inputs method consists of valuing the stock of capital acquired by agents using the overall cost of the studies pursued by agents. This overall cost represents the sum of the production cost of non-market education services supplied by the public sector and the value of the market production of education sold by private entities. To estimate the stock of capital in the economy, it is then necessary to construct an investment time series and then aggregate this time series data, determining a depreciation rate for human capital.

A first difficulty associated with this method relates to the fact that the cost of study must also include the opportunity cost associated with pursued training, namely the total discounted earnings foregone by agents in order to pursue their studies. The time spent by parents helping their children with their school work must also be valued. One of the first studies carried out using this method was conducted by Kendrick (1976), who estimates that opportunity cost represents at least half of the total education costs. Another difficulty associated with this method is the distinction between price effect and volume effect (as is often the case in services): what share of the increase in production costs over time is attributable to an improvement in the "quality" of the education system? Notwithstanding these difficulties, this method has the benefit of being relatively easy to implement. The estimation of the stock of human capital is based, in this method, on a prior valuation of the investment for successive periods. However, this method requires a rate of depreciation of human capital to be determined.

The income approach (discounted lifetime income approach) is far more technical. It was first proposed and applied to the USA economy by Jorgenson & Fraumeni (1989), then it was refined in numerous subsequent studies, and now features in this context in the System of National Accounts 2008 (European Commission *et al.*, 2009) or in the very comprehensive United Nations Guide on Measuring Human Capital (UNECE, 2016). This method is based on determining the value at which an individual could resell, at any time, the human capital he/she has accumulated if it was not "embodied" in the person. The method of valuing human capital is therefore identical to that applied to a financial asset. In these circumstances, this method starts by valuing the stock of human capital in the economy for successive periods. Gross investment in human capital corresponds to the additional discounted future income received by all agents undertaking a further year of studies during the period under consideration.

Assuming that agents can work for a maximum of N periods (it is assumed here that agents reach age 1 when they are at working age) but that they can decide to initially dedicate between 0 and n (from N) periods ($n < N$) to training. n is an integer and the level of human capital attained is given by the number of years of study undertaken: $h = 0,1,2\ldots n$.

In theory – and making the simplified assumption that continuous training during working life is impossible – computing at a date t the value of the stock of capital of an agent who has already entered working life (having finished studying), aged a and having a level of training h involves estimating future income earned throughout his/her remaining working life:

$$KH_{A,a,h,t} = \sum_{i=0}^{N-a} \frac{h.\left(w_h\right)_{t+i,a+i}}{\left(1+r\right)^i} \qquad (2)$$

where $KH_{A,a,h,t}$ is the discounted value (in t) of the stock of human capital of a working individual (A), aged a and having a level of training $h \leq n$, $\left(w_h\right)_{t,a}$ is the annual skill-premium in period t of an agent aged a having accumulated h units of human capital during his/her training; r is the discount rate and N is the end of the agent's working life (retirement). It is immediately clear that the older an individual, the more his/her future income flow is reduced, thus decreasing the value of his/her human capital which depreciates over time, falling to nil when the individual reaches retirement age. Accordingly, net investment in human capital in t is deduced by the difference between the stocks of human capital estimated in (t+1) and the stocks estimated in t (Christian, 2010; McGrattan, 2010).

In practice, the current income in t of the older cohorts will be used to value all future income of agents (Figure I): to calculate, for example, $\left(w_h\right)_{t+1,a+1}$, the data available in t will therefore be used, that is $\left(w_h\right)_{t,a+1}$. It is simply assumed that, for a given level of training, the earnings

(or skill-premium) of an individual aged a increase at the constant rate g over time, so that $(w_h)_{t+1,a} = (1+g)(w_h)_{t,a} \ \forall t$

In these circumstances, equation (2) above can be rewritten as follows:

$$KH_{A,a,h,t} = h.(w_h)_{t,a} + \sum_{i=1}^{N-a} \frac{h.(w_h)_{t+i,a+i}}{(1+r)^i} = h.(w_h)_{t,a} + \sum_{i=0}^{N-(a+1)} \frac{h.(w_h)_{t+1+i,a+1+i}}{(1+r)^{i+1}} \quad \forall a < N$$

$$KH_{A,a,h,t} = h.(w_h)_{t,a} + \sum_{i=0}^{N-(a+1)} \frac{h.(w_h)_{t+i,a+1+i}(1+g)}{(1+r)^{i+1}} \quad \forall a < N$$

Finally, we have:

$$KH_{A,a,h,t} = h.(w_h)_{t,a} + \frac{1+g}{1+r} \sum_{i=0}^{N-(a+1)} \frac{h.(w_h)_{t+i,a+1+i}}{(1+r)^i} = h.(w_h)_{t,a} + \frac{1+g}{1+r} KH_{A,a+1,h,t} \ \forall a < N \quad (3)$$

It should be noted that, in this theoretical model, an agent who does not pursue any training ($h = 0$) has a stock of human capital of nil which will not prevent him/her from receiving earnings income w_L as payment for the unskilled labour undertaken throughout his/her life (according to (1)).

Calculating the value of the stock of capital in t of an agent undertaking his/her studies in theory involves determining the maximum level of training that he/she wishes to or will attain and then estimating, as for assets, his/her future income once he/she starts working life, which will be received throughout his/her professional career:

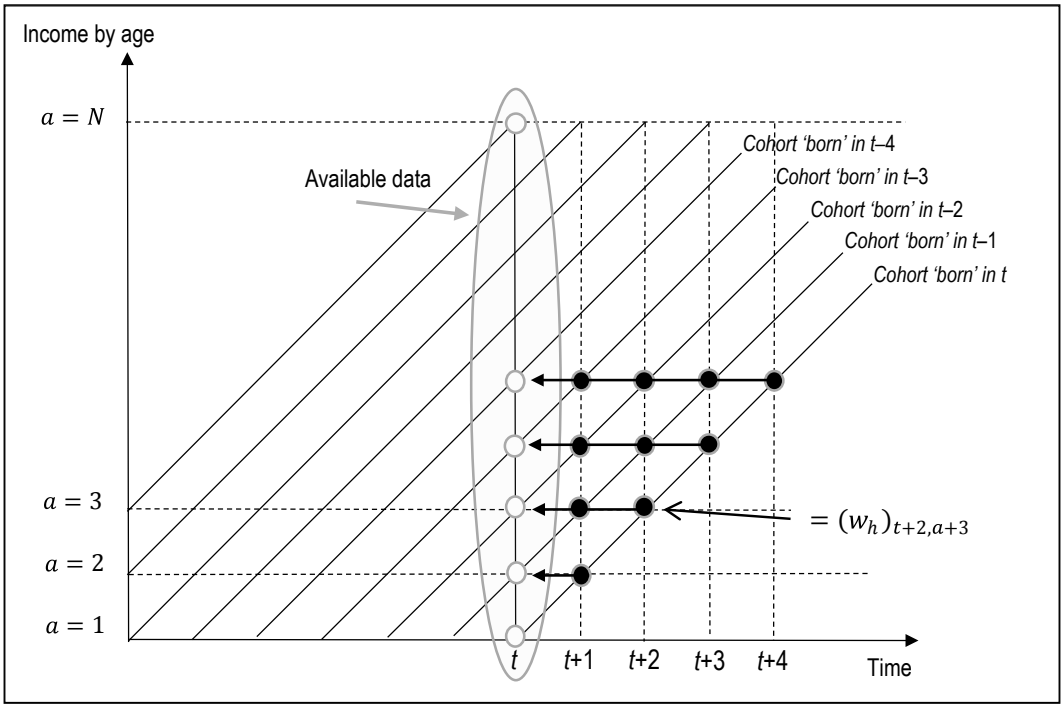
$$KH_{E,a,h,t} = \sum_{i=0}^{N-a} \frac{h.(w_h)_{t+h-a+1+i,h+1+i}}{(1+r)^{h-a+1+i}} \quad (4)$$

where $KH_{E,a,h,t}$ is the discounted value of the stock of human capital of an individual undertaking study (E), aged a (in t) and having an expected final level of training of $h \leq n$.

Furthermore, contrary to what has been stated above, agents will not necessarily live for N working periods and reach retirement age because they may die during their working life. It is therefore also necessary to calculate the one-year survival rates of agents of different ages. The one-year survival rate of an agent aged a is therefore recorded as $s_{a,a+1}$ ($a < N$). Agents may also become unemployed or decide to withdraw from the labour market. It is therefore necessary to keep account of the rates of employment $e_{a,h}$ for each group under consideration (by age and level of training).

Another empirical difficulty relates to the determination of the optimal level of training of an agent undertaking study in the period t under consideration. Here again, the statistics for previous generations will be used to estimate the

Figure I – **The methodology of the income-based method: Reconstitution of a fictional cohort**



likelihood enr_{y+1} that students already having y years of study ($y < n$) will continue their studies for one further year.

Lastly, it is assumed that the survival rates at each age as well as the likelihood of study being continued at each level of training are constant over time: therefore, the available data on past cohorts can be used to estimate these future values.

Empirically, equation (3) is therefore written as:

$$KH_{A,a,h,t} = e_{a,h}h.\left(w_h\right)_{t,a} + s_{a,a+1}.\frac{1+g}{1+r}.KH_{A,a+1,h,t} \quad (5)$$

Similarly, it can easily be shown that (4) is written:

$$KH_{E,a,h,t} = s_{a,a+1}\left[enr_{h+1}.\frac{1+g}{1+r}.KH_{E,a+1,h+1,t} + \left(1-enr_{h+1}\right).\frac{1+g}{1+r}.KH_{A,a+1,h,t}\right]$$

Lastly, the aggregate human capital equates to:

$$KH_t = \sum_{a=1}^{N}\sum_{h=0}^{n}\left(KH_{E,a,h,t} + KH_{A,a,h,t}\right)$$

Note also that separate stocks of human capital are generally constructed for each gender.

In their 1989 paper, using data constructed using this method, Jorgenson & Fraumeni assert that investment in human capital represents four times the gross fixed capital formation (GFCF) appearing in the national USA accounts. The value of human capital is also likely to correspond to at least seven times the value of the stock of "traditional", non-human capital, again estimated in the national accounts. The study undertaken by Liu (2014) in fifteen OECD countries shows that, in most countries, the ratio of the value of human capital (estimated using this method) to nominal GDP varies between nine and eleven; the value of human capital represents between four and seven times that of non-human capital. Liu (2014) also shows that this method enables an index for the volume of capital to be determined, based in particular on structural effects and their progression in each population group considered: changes in the percentage of the population attaining each level of study, structure by age, employment rate, and structure by gender within each group.

Additionally, this method allows for a comparison of stocks of capital for each level of qualification in each country: the growing divergences observed in certain countries are explained by an increase in the earnings differential but, in certain cases, the divergence also results from the growing numbers of people accessing higher levels of study.

A number of difficulties associated with this method of estimation can nevertheless be identified. A first important criticism addressed by Weil (2015b) is that Jorgenson & Fraumeni (1989) ultimately reduce to two (or rather, maintain at two) the number of factors in the economy's production function: physical capital and human capital. "Unskilled", unqualified labour has disappeared altogether, to the extent that the authors use the entire earnings received by agents in calculating the discounted income flows. With the notation used in this paper, Jorgenson and Fraumeni estimate $w = w_L + h.w_H$, without restricting payment for human capital to the sole component $h.w_H$. This approach can be justified in "our" developed economies where schooling is compulsory in childhood and no individual is now entirely devoid of human capital. This observation, however, sits uncomfortably with a reality in which the productivity of young people leaving the education system early stands at a low level. It may prove relevant to retain the distinction between unskilled labour and human capital, notably to analyse income inequality, although measuring pay for unskilled labour can be problematic: how can the threshold between unskilled labour and skilled labour be determined? Weil (2015b) is of the view that pay for unskilled labour currently represents around half of global pay for labour (unskilled and skilled).

Abraham (2010) offers a detailed analysis of the difficulties associated with the technique used by Jorgenson & Fraumeni. Four essential aspects must be mentioned:

i) Use of a reconstituted fictional cohort: in this method, the income at the age of 60 of a young person aged 20 today is estimated using the current income of persons currently aged 60 (with the same level of education) to which a growth trend is applied (associated with gains in productivity), being constant over 40 years. This hypothesis assumes that returns on education are constant (or increase in a constant manner) over time, which is far from certain: the quality of teaching provided may change (or may have changed) over time; moreover, this hypothesis does not take into consideration the potential dynamic effects: for example, high returns to human capital today could increase the desire of young generations to seek further training, which will reduce the returns on education in future.

ii) No stock of capital is determined for children aged under 15. Any agent capable of working at the time of the valuation is effectively considered to carry human capital. However, such an assumption is open to debate, the discounted future income being capable of valuation (as expected income) as of the individual's date of birth (Christian, 2017). Similarly, any person who leaves the labour market, even temporarily, reduces the stock of human capital in the economy, which is far from satisfactory.

iii) The results are affected (primarily in terms of level rather than in terms of progression) by estimates of the discount rate r and growth rate g for wages (Liu, 2014). In fact, there is no reason for g to be constant over time or, most importantly, to systematically have the same effect across the entire wage (or qualifications) structure: technical progress may, at certain times, further increase the productivity of skilled or unskilled workers.

iv) Even more fundamentally, the income-based approach treats any increase in wages as growth in the value of human capital. Any wage differential based on the levels of training attained is explained entirely by the human capital differential. The relevance of this hypothesis is certainly open to question and one might wonder in particular what wage would be earned by qualified workers if they had not pursued any studies: when the wages of individuals having different levels of training are compared, selection bias can exist in the constitution of samples of both skilled and unskilled individuals. Is it not the case that certain individuals have individual attributes (personal talent, "cultural" or "social" capital inherited from parents) explaining why it is easier for them to pursue studies but also why they receive a higher level of pay than the rest of the population if they decide not to pursue their studies? If these individuals are more skilled, part of the wage differential could certainly be explained by their individual attributes. The studies of Gabaix & Landier (2008) may be referred to again here: in a relatively stable economic environment, it is not always necessary to discriminate between "talented" skilled persons and "untalented" skilled persons; in a fast-changing world where a company's success depends on its constant capacity to be innovative or flexible, qualification is not enough and companies will look for "talent" at least as much as "skills": wages will rise, but it is the payment for individual attributes which increases, not payment for human capital; in some aspects, this reasoning may bring to mind the signalling theory of Spence (1973), according to which investment in human capital is simply used to signify the intrinsic attributes of agents, without substantially increasing their productivity. Similar reasoning can undoubtedly be applied to the changes to wages norms for the super-managers referred to by Piketty in his book (2013): the spectacular increase in very high earnings is due more to the ability of a small minority (sometimes also highly qualified) to monopolise a very substantial share of the income from innovation than to a true increase in the intrinsic productivity of their human capital. From this perspective, the most emblematic empirical case is the finding that, at given level of human capital, the value of the capital stock of men is higher than that of women (Liu, 2014)! While part of this differential seems to be explained by women's lower participation rates, a significant part remains attributable to wage differentiation, which is difficult to explain by purely economic mechanisms.

Lastly, and to mitigate this final comment, everything depends on the definition attributed to human capital. The OECD (2011), for example, advances a fairly wide definition, incorporating both skills acquired as well as individual attributes, whether innate or inherited (UNECE, 2016). The method of Jorgensen & Fraumeni (1989) may therefore offer the opportunity to reveal the differentials in returns on human capital between sub-groups of the population (which remains, moreover, to be explained). Conversely, with a narrower definition of human capital, the fundamental question is to establish whether earnings reveal the productivity associated with human capital. The earnings differential between two individuals having the same level of education should not therefore be attributed to human capital. These two representations can, however, be reconciled if we consider that the alternative factors explaining increases in wages (talent, capacity to monopolise a profit, etc.) are often complementary to human capital, on which they rely in order to operate fully.

Accordingly, each of the two methods (input and output) offers advantages and disadvantages: the output method certainly enables focus to be placed on national trends, and potential divergences in returns between different sub-groups of the population to be analysed; the inputs method is undoubtedly easier to implement, requiring a smaller amount of data (primarily national accounting data), which facilitates international, or regional, comparisons. In

fact, estimates using the discounted income method lead to far higher valuations of the stock of human capital. This divergence may be explained, at least in part, by factors already referred to in the previous section (imperfections in the financial market, agents' risk aversion, etc.). But it certainly reveals the valuation difficulties which continue to affect both these methods: likely under-estimation of the inputs in the former (notably due to the existence of opportunity costs which are difficult to measure), likely over-valuation of the output associated with human capital in the latter (high sensitivity of the result to the discount rate, over-valuation of the skill-premium, etc., see Abraham, 2010; Fraumeni, 2011; UNECE, 2016).

4. How Should Education and Health Expenditure Be Allocated to the Gross Fixed (Human) Capital Formation for the Purposes of National Accounting?

The final part of this paper is dedicated to an analysis of the effect on the savings rate of the reallocation of certain items of consumption expenditure to investment expenditure. In fact, there are many alternative ways of measuring the savings rate, depending on whether or not durable goods (motor vehicles, large domestic appliances, etc.) are included in consumption, whether or not capital gains taxes are deducted from disposable income or non-redistributed company profits are paid back to individuals (Reinsdorf, 2007). Moreover, it has been established (Galiana *et al.*, 2017) that the household savings rate is affected by institutional factors such as the retirement regime (distribution *vs.* capitalisation) or taxation (direct *vs.* indirect). We do not attempt here to set out exhaustively the various empirical definitions and measurements of savings, but rather it aims to focus on the specific impact of human capital expenditure.

The introduction of human capital into a national accounting framework *via* a satellite account is explained by the United Nations Economic Commission for Europe (UNECE, 2016). In this satellite account, two alternative methods are proposed: it is assumed either that the institutional sectors meet the costs of education produced by human capital, or that households produce this capital themselves. To do this, they accordingly invest "intermediate goods for the production of human capital", which are primarily produced by other institutional sectors and used by households as intermediate consumption in their production activity. In the first case, the agents (public administrations, companies, etc.) no longer produce an education service (effectively consumed by households) but instead produce human capital directly, which is subsequently purchased in the form of investment (GFCF) by households – whereas, within the accounting framework defined by the System of National Accounts 2008 (European Commission, 2009), education expenditure is systematically recorded under consumption expenditure of the institutional sectors. In this case, the portion of this (household) investment expense which is imputed is funded by a resource of the same amount which is itself imputed (primarily from the public sector), recorded as a transfer of capital. In the second method, the production of human capital is imputed in the household account, estimated at its production cost, such cost including the "intermediate inputs" (for the production of human capital) produced by the other institutional sectors, and the time devoted by students to pursuing their studies (opportunity cost, recorded in uses in the household account under the form of mixed income). The intermediate consumption imputed (appearing under uses in the household production account) is the subject of a social transfer in kind (for an identical amount) from the sector which produced these "inputs" (principally public administrations). Households' market spending on education is transferred from their final consumption expenditure to intermediate consumption, as it is now associated with their human capital production activity. The time devoted to study increases households' disposable income and savings. The entire production of human capital (for its own account) is ultimately a GFCF household expense. In the first method, the resource imputed is a transfer of capital, which does not therefore affect either the household's disposable income, or savings (this transfer effectively takes place "downstream", in the capital account). In the second method, the resource imputed is a social transfer in kind (representing the amount of the intermediate inputs "purchased" from other institutional sectors), which increases both households' disposable income and savings. In aggregate terms, both methods therefore result in the same volume of national savings, but the first method increases the administrations' savings, unlike the second, which increases households' savings.

The next part focuses on the impact of education (and then health) expenditure on the savings rates of households alone, relying on the second

method of the satellite account for human capital presented above.[3] We also make a simplified estimation of the production value of human capital, which does not take account of the opportunity costs of education (similar methodology on this point to that of Kokkinen, 2008). This assumes that production of human capital and the associated intermediate consumption are of the same value, meaning that households' disposable income and savings are only increased by the amount of social transfers in kind (associated with education) from the public sector (and in no circumstances by an increase in mixed household income). It therefore suffices to transfer the (actual) consumption expenditure on education to households' GFCF, with a significant impact on the volume of their savings.

This type of analysis is especially interesting for countries where the household savings rate has varied significantly in recent decades. This is the case in particular for the USA, where the savings rate has fallen markedly since the early 1980s. We therefore concentrate firstly on the USA situation. Subsequently, estimates for two major European countries, Great Britain and France, are presented.

4.1. The USA Case

The fall in the savings rate in the USA is generally explained by wealth effects (Bostic *et al*., 2009), households' ease of access to credit, or imitation phenomena causing a large proportion of the American middle classes to increase its expenditure to attain a lifestyle akin to that of the most well-off, whose income has increased far more quickly than average (Barba & Pivetti, 2009), although most of these points seem to fail to entirely resolve the "puzzle" of USA savings (Guidolin & La Jeunesse, 2007).

In the next section, we return to Becker's "wide" definition of investment expenditure in human capital: therefore, firstly (market) education expenditure (narrow definition of human capital) is added to household savings followed, secondly, by health expenditure (wider definition).[4]

In the household account of the USA National Income and Product Accounts (NIPA), available household income includes the social benefits in kind represented by the Medicare and Medicaid public health schemes; moreover, this income is calculated prior to payment of contributions (including employer contributions) to private health insurance. These contributions are recorded, net of payments received, as consumption expenditure, under "health insurance". Consumption expenditure on health appearing in the household account therefore incorporates all "actual" household expenditure, except for that received on a non-market basis by the public health authorities.

Analysis of the household account alone shows that household education expenditure, albeit low, has grown significantly, from 0.6% of GDP in 1960 to 1.5% in 2017, having therefore increased 2.4 times. However (Figure II), this progression is not sufficiently marked to significantly modify their savings rate and, notably, to alter its course, the divergence between 1960 and 2018 nevertheless falling from 3.4 points to 2.3 percentage points (the divergence therefore narrowed by a third between these two dates).

It may be interesting at this stage to compare these results with the series for human capital and net investment in human capital recently constructed by Christian (2016)[5], using the output method on USA data. The author distinguishes "market" human capital from "non-market" capital (valuation of the opportunity cost of education). He also values the capital stock of persons aged over fifteen (active human capital) and that of children (nascent human capital). Irrespective of the coverage of the study, the data constructed by the author shows that net investment in human capital declined slightly (as a percentage of GDP) between 1975 and 2013 (Figure III). The national USA accounts conversely show that total (gross) expenditure associated with education compared with GDP has increased by around 0.5 percentage points during the same period (further, the portion of this expenditure met directly by households has increased slightly). However, Christian's estimates principally confirm that the non-market portion of this investment is on average over twice as high as the market portion, although it is true to say that this component is defined by the author very widely as household domestic production.

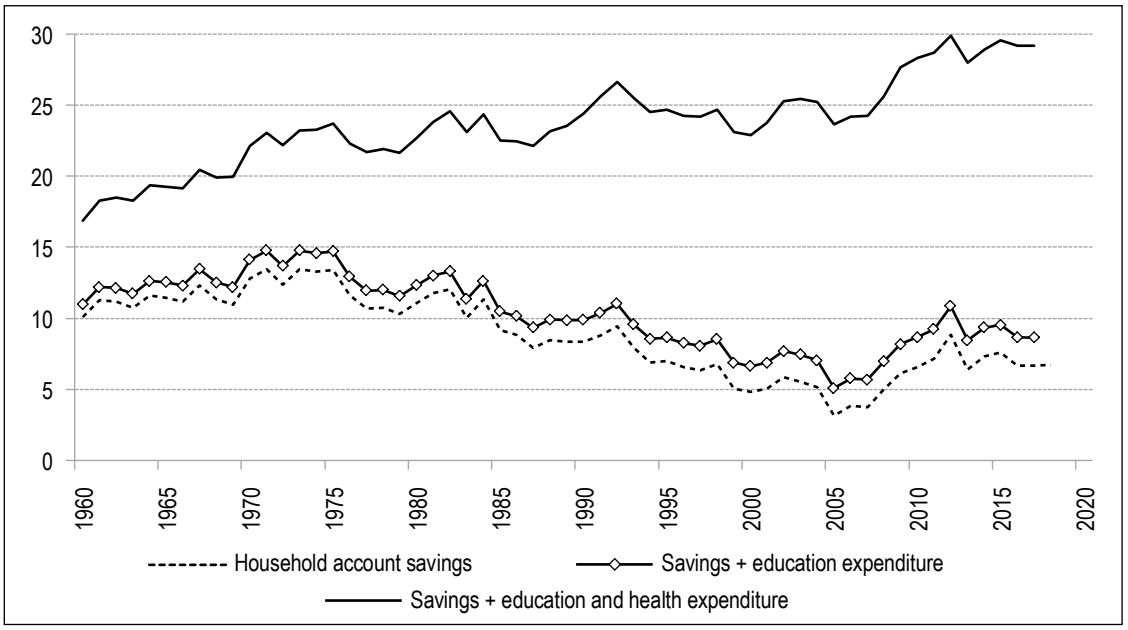
---

3. In this paper we only deal with education expenditure, ignoring the treatment of vocational training expenses incurred by businesses.

4. Up to this point, we have of course adopted a narrow definition (limited to education expenses), or a very narrow definition (excluding from human capital cultural capital which is difficult to acquire during studies, and attributes specific to agents). Moving to a wide definition of human capital at this stage, although this is suggested by Becker, therefore represents one of the limitations associated with this "exercise".

5. The values estimated by Christian (2016) are net rather than gross, unlike for Jorgenson & Fraumeni (1989); according to the author, the calculation of net values explains the significant differences in level obtained in the series produced, by comparison with those of Jorgensen & Fraumeni (1989).

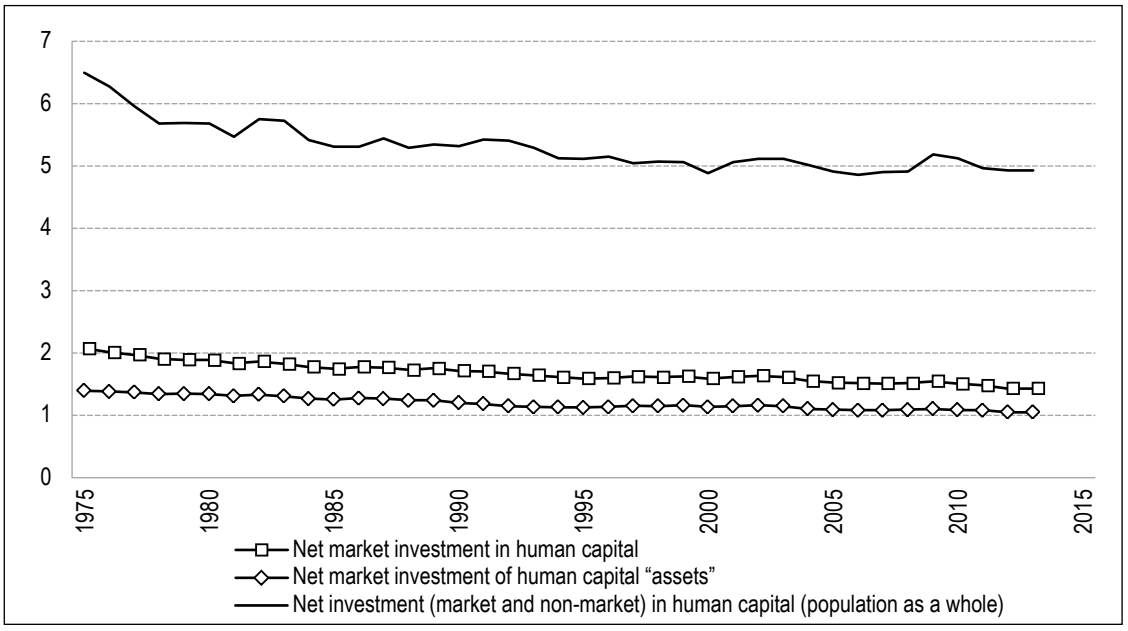
Figure II – **Calculations of savings rate (as a percentage of gross disposable income) for different definitions of savings, using the USA household account only, 1960-2018**



Sources: Bureau of Economic Analysis. NIPA. Author's calculations.

It should be noted that, in figure II, education expenditure does not include the opportunity costs associated with the pursuit of studies. If, like Kendrick (1976), we assume that these costs are proportionate to the "actual" expenditure, the impact of this omission on the progression of the savings rate remains modest. In the estimations made by Christian (2016) using the inputs method, the burden of this imputed expenditure (in total household education expenditure) has reduced; but in his estimates using the output method, this burden increases

Figure III – **Different indicators of net investment in human capital (Christian, 2016) as a percentage of GDP (NIPA), 1975-2013**



Sources: Christian (2016) and Bureau of Economic Analysis, NIPA. Author's calculations.

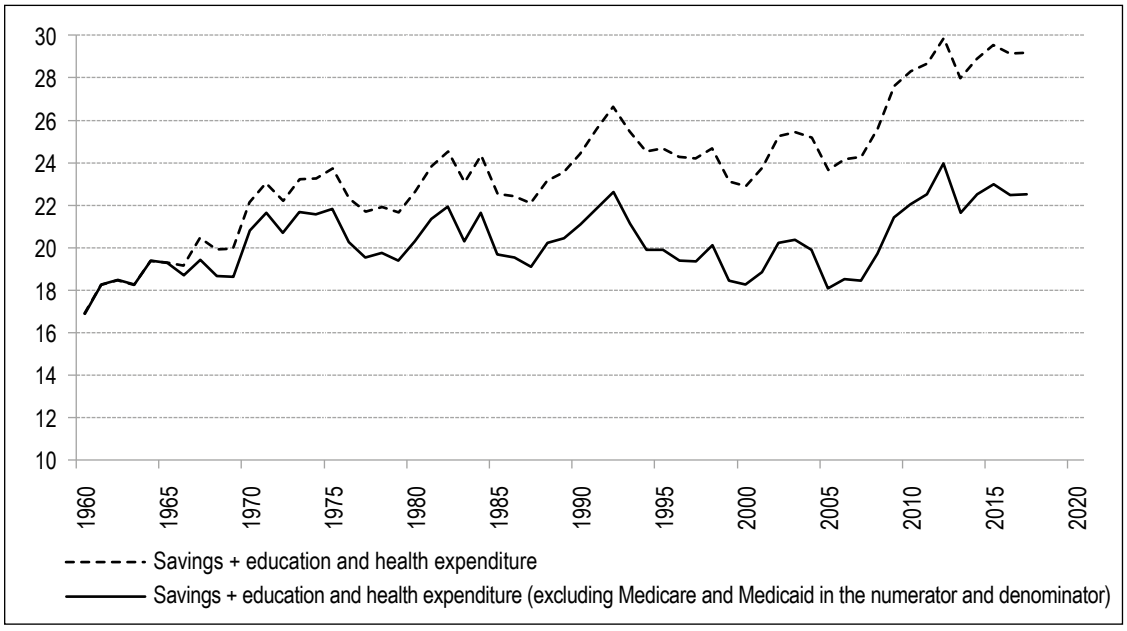
slightly. Moreover, the data provided by UNECE (2016) for Canada demonstrates that the portion of this expenditure imputed (in total education expenditure of Canadian households) has seen a spectacular increase, rising to as much as 11% of Canadian GDP in 2010, compared with only 2.2% in 1981 (on the situation in Canada, see in particular Gu & Wong, 2015). Accordingly, it is not entirely impossible that a refined estimation of the opportunity costs associated with education would, if taken into account, further adjust the savings rate of USA households (by the input method). The perception that, during recent decades, parents have attached growing importance to their children's educational success and have therefore "invested" more (primarily in non-monetary terms) in their "education" increases the likelihood of this happening.

The addition of health expenditure to savings, which undeniably reflects a far wider definition of human capital than that adopted so far in this paper, adjusts and even marginally reverses the trend in savings rates. The savings rate has also been computed excluding the Medicare and Medicaid schemes set up in 1965 from both the numerator and the denominator (Figure IV); this is therefore closer to the definition of gross disposable income (GDI) according to the European system of accounts, given that these schemes are social transfers in kind, which are included in adjusted gross disposable income (AGDI) but not in disposable income. The result is that this indicator is much more stable (around 20%) over a long period.

Conversely, households' "actual final consumption" of education and health can be calculated by adding the consumption expenditure of the public sector for these two items to theirs. In the accounts of the public administrations, final consumption expenditure (FCE) and investment expenditure (GFCF) per function are effectively available. It is therefore easy to access the education and health FCE of these public administrations.

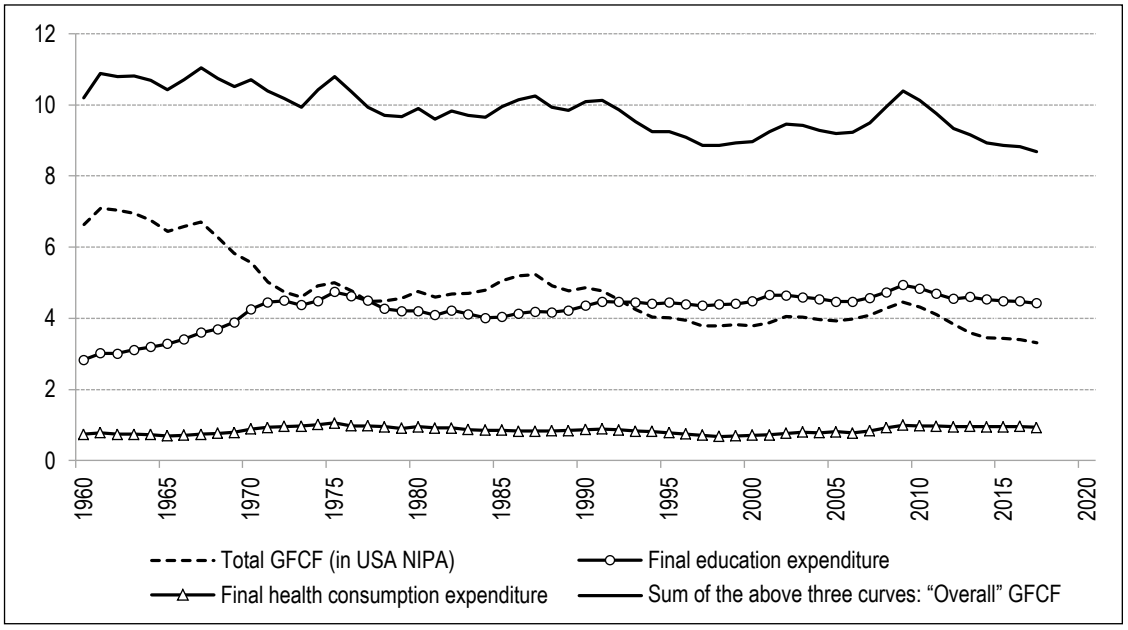
On the basis of the General government's accounts, it is possible, in the first instance, to estimate a "global" investment indicator for General government (related to GDP), by including their education and health FCE in their GFCF (Figure V). The ratio of (GFCF General government) / GDP provided by the NIPA shows a declining trend: since 1960, it has fallen by 3.3 percentage points. At the same time, consumption expenditure on education of General government (compared with GDP) has risen by 1.6 points, primarily between 1960 and 1970. Their health FCE has increased marginally (+0.2 points of GDP). Lastly, with this new measurement of the investment of General government, the declining trend is clearly less pronounced: only -1.5 points.

Figure IV – **Medicare and Medicaid included/not included in household health expenditure, 1960-2017**



Sources: Bureau of Economic Analysis, NIPA. Author's calculations.

Figure V – **GFCF expenditure of USA General government, as a percentage of GDP, 1960-2017**

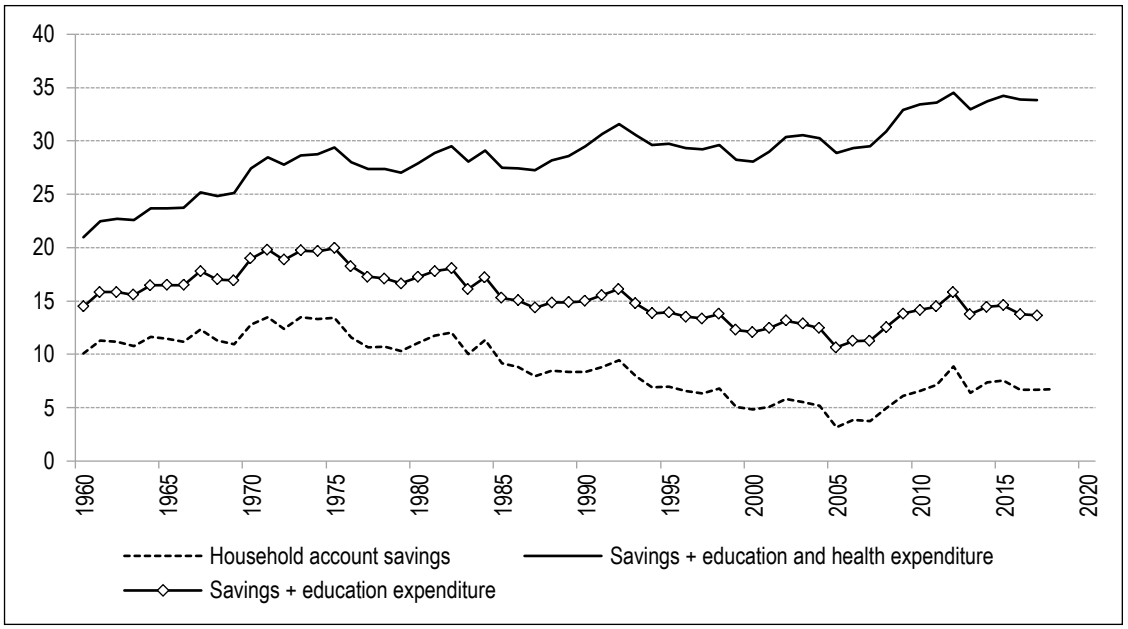


Sources: Bureau of Economic Analysis, NIPA. Author's calculations.

At this stage, household savings rates can therefore be calculated on the basis of their AGDI, by moving (by means of social transfers in kind added to their disposable income) all education and then health FCE of the public sector to the household account (Figure VI). The results obtained are quite similar, in terms of trends, to those computed using gross disposable income, although it is worth noting that the savings rate including household actual final consumption (AFC) on education practically returned to its 1960 level (14.5%) in 2017 (13.6%).

To conclude this section, the savings rate of USA households is only modestly redressed by the

Figure VI – **Different calculations of household savings rate (as a percentage of the "adjusted" GDI) using the accounts of USA households and General government, 1960-2018**



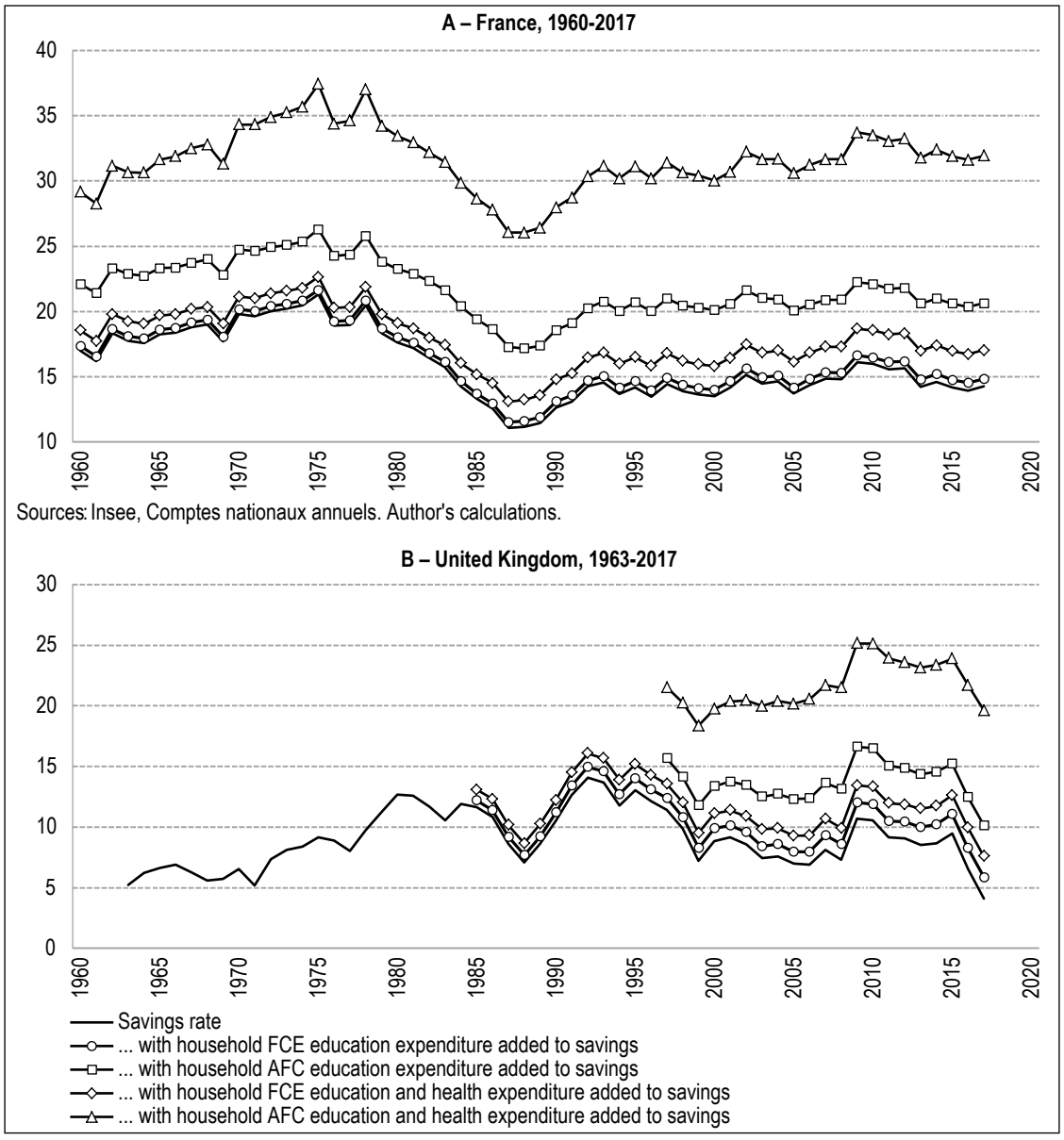
Sources: Bureau of Economic Analysis, NIPA. Author's calculations.

transfer solely of the education FCE to investment expenditure; the adjustment is improved for the indicator transferring education AFC rather than the FCE, but, in any event, the fall in the savings rate between 1980 and 2008 remains significant. This "restrictive" definition of investment in human capital does not in any circumstances explain the fall in the savings rate of USA households over almost 40 years. If, however, health expenditure is also added to investment, this savings rate is redressed significantly, therefore remaining relatively stable between 1960 and 2016.

4.2. The Case of Two European Countries: Great Britain and France

This final section sets out, still using the input method, alternative indicators for savings rates, once education (and then health) expenditure has been incorporated into the savings of French and British households. In both these countries, the household savings rate has seen no significant fall in recent decades, although the French rate experienced some fairly dramatic changes between 1975 and 1990; the rate for Britain is certainly fairly cyclical, but has remained stable

Figure VII – **Different calculations of savings rate (as a percentage of GDI and of AGDI) using the accounts of households and General government, 1960-2017**



Sources: Office for National Statistics, UK National Accounts. Author's calculations.

overall since 1963. The savings rate is computed, firstly, in the same way as for the USA, using the household account only, "transferring" their education and then health FCE to their savings; secondly, we proceed in a similar way but firstly construct a household AGDI, restricting social transfers in kind (from the public sector) solely to education and then health expenditure.

In the case of France, as national education expenditure has been very stable (around 4.5% of GDP) for several decades (and the portion of this expenditure paid by households has itself been both low and relatively stable), the alternative indicators have progressed in very similar ways to the "standard" savings rate (the curve incorporating education FCE into savings is in fact almost identical to the savings rate in the national accounts). With regard to health expenditure, household AFC has increased from 2.4% of GDP in 1960 to 6.8% in 2017, household FCE having increased far less rapidly (increasing from 0.9% to 1.3% during the same period): only the savings rate compared with the adjusted income (including actual health consumption) is therefore marginally redressed (Figure VII-A).

For Great Britain, data are only available on household education and health FCE since 1985 and on household AFC since 1997. The results obtained, however, appear similar to those obtained for France, the impact of alternative measures relating essentially to the indicator levels (Figure VII-B).

* *
*

While economic theory considers education expenditure as an investment, it is treated as consumption expenditure in the national accounting framework. This paper emphasises the point that for a long time economic work have been undertaken with a view to bringing the human capital factor into national accounts, some by adopting the input method (estimation of a stock based on investment expenditure incurred), others the output method (discounted income flows generated by the constitution of a stock). It is important to unerline in this conclusion how difficult this evaluation is, whichever method is adopted, which undoubtedly explains in part why, notwithstanding the attempts presented in this paper, national accountants have until now chosen not to take this step. Nevertheless, it seems that such an approach would enable national accounting frameworks to become more aligned with certain key debates between economists and would, undoubtedly, contribute to these debates and perhaps enable some areas of controversy to be resolved.

The construction of data on human capital clearly demonstrates that this is a production factor at least as important today as physical capital, and that this factor must be taken into account if we are to properly understand the productive dynamics of developed economies. We have used what is known as the input method to estimate a savings rate for USA, French and British households where education and health expenditure are considered as investment expenditure. The savings rate of USA households, which fell between 1980 and 2008, is only modestly redressed when only education FCE is transferred to investment expenditure (reduction of one third of the decline observed between 1960 and 2018). This may be explained, at least partially, by the fact that imputed education expenditure is not taken into consideration in the method adopted in this paper. Lastly, it emerges that health expenditure must be incorporated into investment in human capital if this savings rate is to be significantly redressed (and stabilised). For the two European countries considered, Great Britain and France, the alternative indicators have an impact on the savings levels, but not on their evolutions. □

BIBLIOGRAPHY

Abraham, K. G. (2010). Accounting for Investments in Formal Education. *Survey of Current Business*, 90(6), 42–53. https://fraser.stlouisfed.org/files/docs/publications/SCB/2010-19/scb_201006.pdf.

Akerlof, G. A. (1980). A Theory of Social Custom, of Which Unemployment May be One Consequence. *The Quarterly Journal of Economics*, 94(4), 749–775. https://doi.org/10.2307/1885667

Acemoglu, D. (2002). Directed Technical Change. *The Review of Economic Studies*, 69(4), 781–809. https://doi.org/10.1111/1467-937x.00226

Acemoglu, D., Autor, D., Dorn, D., Hanson, G. H. & Price, B. (2016). Import Competition and the Great US Employment Sag of the 2000s. *Journal of Labor Economics*, 34(1), 141–198. https://doi.org/10.1086/682384

Autor, D. H., Katz, L. F. & Kearney, M. L. (2008). Trends in U.S. Wage Inequality: Revising the Revisionists. *Review of Economics and Statistics*, 90(2), 300–323. https://doi.org/10.1162/rest.90.2.300

Autor, D. H. (2014). Skills, Education, and the Rise of Earnings Inequality Among the 'Other 99 Percent. *Science*, 344(6186), 843–851. https://doi.org/10.1126/science.1251868

Barba, A. & Pivetti, M. (2009). Rising household debt: Its causes and macroeconomic implications—a long-period analysis. *Cambridge Journal of Economics*, 33(1), 113–137. https://doi.org/10.1093/cje/ben030

Becker, G. S. (1962). Investment in Human Capital: A Theoretical Analysis. *Journal of Political Economy*, 70(5), 9–49. https://doi.org/10.1086/258724

Becker, G. S. (1964). *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. Chicago: University of Chicago Press.

Berman, E., Bound, J. & Griliches, Z. (1994). Changes in the Demand for Skilled Labor within U.S. Manufacturing: Evidence from the Annual Survey of Manufacturers. *The Quarterly Journal of Economics*, 109(2), 367–397. https://doi.org/10.3386/w4255

Bernanke, B. (2005). The Global Saving Glut and the U.S. Current Account Deficit, *The Federal Reserve Board,* Remarks by Governor Ben S. Bernanke at the Sandridge Lecture, Virginia Association of Economists Richmond, Virginie. https://www.federalreserve.gov/boarddocs/speeches/2005/200503102/

Bertrand, M. & Mullainathan, S. (2001). Are CEOs Rewarded for Luck? The Ones Without Principals Are. *The Quarterly Journal of Economics*, 116(3), 901–932. https://doi.org/10.1162/00335530152466269

Bostic, R., Gabriel, S. & Painter, G. (2009). Housing wealth, financial wealth, and consumption: New evidence from micro data. *Regional Science and Urban Economics*, 39(1), 79–89. https://doi.org/10.1016/j.regsciurbeco.2008.06.002

Bourguignon, F. (2015). *The Globalization of Inequality*. Princeton: Princeton University Press.

Chamon, M. & Prasad, E. S. (2010). Why Are Saving Rates of Urban Households in China Rising? *American Economic Journal: Macroeconomics*, 2(1), 93–130. https://doi.org/10.1257/mac.2.1.93

Christian, M. S. (2010). Human Capital Accounting in the United States, 1994–2006. *Survey of Current Business*, 90(6), 31–36. https://fraser.stlouisfed.org/files/docs/publications/SCB/2010-19/scb_201006.pdf, p. 31.

Christian, M. S. (2016). Net Investment and Stocks of Human Capital in the United States, 1975-2013. *BEA Working Paper*.
https://www.bea.gov/research/papers/2016/net-investment-and-stocks-human-capital-united-states-1975-2013

Christian, M. S. (2017). Net Investment and Stocks of Human Capital in the United States, 1975-2013. CSLS, *International Productivity Monitor,* 33, 128–149. http://www.csls.ca/ipm/33/Christian.pdf

Denison, E. F. (1962). *The Sources of Economic Growth in the United States and the Alternatives before Us*. Committee for Economic Development. New York.

European Commission, IMF, OECD, United Nations & World Bank (2009). *System of National Accounts 2008*. New York. https://unstats.un.org/unsd/nationalaccount/docs/sna2008.pdf

Fraumeni, B. M. (2011). Human Capital Accounts: Choice of Rates and Construction of Volume Indices. NBER, *Working Paper* N° 16895. https://doi.org/10.3386/w16895

Gabaix, X. & Landier, A. (2008). Why has CEO Pay Increased So Much? *The Quarterly Journal of Economics*, 123(1), 49–100. https://doi.org/10.1162/qjec.2008.123.1.49

Galiana, L., Lafféter, Q. & Simon, O. (2017). Les taux d'épargne dans la zone euro : comment les comparer, comment les expliquer ? *In:* Insee (Eds), *Insee références, L'économie française - Comptes et dossiers édition 2017*, pp. 97–114. https://www.insee.fr/fr/statistiques/2894030?sommaire=2894036

Goldin, C. & Katz, L. F. (2010). *The Race between Education and Technology*. Cambridge: Harvard University Press.

Goldin, C. & Margo, R. A. (1992). The Great Compression: The U.S. Wage Structure at Mid-Century. *The Quarterly Journal of Economics*, 107(1), 1–34. https://doi.org/10.2307/2118322

Guidolin, M. & La Jeunesse, E. A. (2007). The Decline in the U.S. Personal Saving Rate: Is It Real and Is It a Puzzle? *Federal Reserve Bank of St. Louis Review*, 89(6), 491–514. https://doi.org/10.20955/r.89.491-514

Gu, W. & Wong, A. (2015). Productivity and Economic Output of the Education Sector. *Journal of Productivity Analysis* 43(2), 165–182. https://doi.org/10.1007/s11123-014-0414-y

Jorgenson, D. W. & Fraumeni, B. M. (1989). The Accumulation of Human and Nonhuman Capital, 1948-84. In: R. E. Lipsey and H. Stone Tice (Eds), *The Measurement of Saving, Investment and Wealth*, pp. 227–285. Chicago: University of Chicago Press. https://www.nber.org/chapters/c8121.pdf

Kendrick, J. W. (1976). *The Formation and Stocks of Total Capital*. New York: Columbia University for NBER. https://www.nber.org/books/kend76-1

Kokkinen, A. (2008). Human Capital and Finland's Economic Growth in 1910–2000. *30th General Conference of The International Association for Research in Income and Wealth* (IARIW). http://www.iariw.org/papers/2008/kokkinen.pdf

Krugman, P. (2007). *The Conscience of a Liberal*. New York and London: W. W. Norton & Company.

Lucas, R. E., Jr. (1988). On the Mechanics of Economic Development. *Journal of Monetary Economics*, 22(1), 3–42. https://doi.org/10.1016/0304-3932(88)90168-7

Lee, D. (1999). Wage Inequality in the United States During the 1980s: Rising Dispersion or Falling Minimum Wage? *The Quarterly Journal of Economics*, 114(3), 977–1023. https://doi.org/10.1162/003355399556197

Lemieux, T. (2008). The changing nature of wage inequality. *Journal of Population Economics*, 21(1), 21–48. https://doi.org/10.1007/s00148-007-0169-0

Liu, G. (2014). Measuring the Stock of Human Capital for International and Intertemporal Comparisons. *In:* D. W. Jorgenson, J. Steven Landefeld, & P. Schreyer (Eds), *Measuring Economic Sustainability and Progress*, pp. 493–544. Chicago: University of Chicago Press.

Mankiw, G., Romer, D. & Weil, D. N. (1992). A contribution to the empirics of economic growth. *The Quarterly Journal of Economics*, 107(2), 407–437. https://doi.org/10.2307/2118477

McGrattan, E. R. (2010). Comment on Michael Christian's "Human Capital Accounting in the United States, 1994–2006". *Survey of Current Business*, 90(6), 37–41. https://fraser.stlouisfed.org/files/docs/publications/SCB/2010-19/scb_201006.pdf, p. 37.

Mira, M. & Liu, G. (2010). The OECD Human Capital Project: Progress Report. *31st General Conference of The International Association for Research in Income and Wealth* (IARIW). http://www.iariw.org/papers/2010/5Mira.pdf

OECD (2010). *The OECD Human Capital Project: Progress Report*. Paris.

Piketty, T. (2001). *Les Hauts revenus en France au 20e siècle : inégalités et redistribution, 1901-1998*. Paris: Grasset.

Piketty, T. (2013). *Le capital au XXIe siècle*. Paris : Seuil. English edition, 2014: *Capital in the Twenty-First Century*. Harvard University Press.

Reinsdorf, M. B. (2007). Alternative Measures of Personal Saving. *Survey of Current Business*, 87(2), 7–13. https://apps.bea.gov/scb/pdf/2007/02%20February/0207_saving.pdf

Romer, P. M. (1990). Endogenous Technological Change. *Journal of Political Economy*, 98(5), 71–102. https://doi.org/10.1086/261725

Schultz, T. W. (1961). Investment in Human Capital. *American Economic Review,* 51(1), 1–17. http://la.utexas.edu/users/hcleaver/330T/350kPEESchultzInvestmentHumanCapital.pdf

Schultz, T. W. (1962). Reflections on Investment in Man. *Journal of Political Economy*, 70(5), 1–8. https://doi.org/10.1086/258723

Smith, A. (1776). Recherches sur la nature et les causes de la richesse des nations. Paris. Flammarion.

Solow, R. M. (1956). A Contribution to the Theory of Economic Growth. *The Quarterly Journal of Economics*, 70(1), 65-94. https://doi.org/10.2307/1884513

Spence, M. (1973). Job Market Signaling. *The Quarterly Journal of Economics*, 87(3), 355–374. https://doi.org/10.2307/1882010

UNECE - United Nations Economic Commission for Europe (2016). *Guide on Measuring Human Capital*. Task Force on Measuring Human Capital. New York, Geneva: United Nations. https://www.unece.org/fileadmin/DAM/stats/publications/2016/ECECESSTAT20166_E.pdf

Verdugo, G. (2014). The Great Compression of the French Wage Structure, 1969-2008. *Labour Economics*, 28, 131–144. https://doi.org/10.1016/j.labeco.2014.04.009

Weil, D. N. (2015a). Capital and Wealth in the 21st Century. *The American Economic Review*, 105(5), 34–37. https://doi.org/10.1257/aer.p20151057

Weil, D. N. (2015b). Capital and Wealth in the 21st Century. NBER, *Working Paper* N° 20 919. https://doi.org/10.3386/w20919

The Social Cost of Global Warming and Sustainability Indicators: Lessons from an Application to France

Jean-Marc Germain and Thomas Lellouch*

Abstract – In order to meet the Paris agreements, significant financial resources must be incurred, which are evaluated here using a macroeconomic model combining a criterion of intergenerational distribution of the climate effort and assumptions on decarbonisation technologies. The results show that, for France, the current greenhouse gas emissions trajectory is unsustainable, in the sense that in order to reach the carbon neutrality commitment in 2050, the annual level of climate spending would have to increase very substantially, to 4.5% of GDP from the current 1.9%. These evaluations make it possible to deduce a social price of carbon or a value for climate action, which has been increased significantly compared to previous evaluations such as those of the Stiglitz-Stern commission, in line with the results of the Quinet Commission in 2019. Such evaluations of the emissions trajectory and the social price of carbon could be the entry point for environmental economic accounting that includes the degradation of natural assets caused by economic activities.

JEL Classification: Q01, Q54, Q56, E01, E21, O13
Keywords: sustainability, climate change, carbon price, adjusted net savings

* *Insee (jean-marc.germain@insee.fr; thomas.lellouch@insee.fr)*

This work has benefited from useful discussions and enriching comments from Didier Blanchet, two anonymous reviewers and the participants in the 5 may 2020 seminar of the Insee Economic Studies Department (DEE) and, more specifically, Xavier Timbeau, who discussed the presentation of an earlier version.

Received on 4 October 2019, accepted after revisions on 4 June 2020.
Translated from: "Coût social du réchauffement climatique et indicateurs de soutenabilité : les enseignements d'une application à la France"
Citation: Germain, J.-M. & Lellouch, T. (2020). The Social Cost of Global Warming and Sustainability Indicators: Lessons from an Application to France. *Economie et Statistique / Economics and Statistics*, 517-518-519, 81–102. https://doi.org/10.24187/ecostat.2020.517t.2024

While the global temperature has experienced a very clear increase since the 1980s, the scientific consensus is now established and recognises that human activities have an impact on global warming through greenhouse gas (GHG) emissions. In exchange, global warming will cause damage to human societies and natural environments, and the risks of abrupt and irreversible damage increase with the degree of warming.

In this context, the international framework for combating climate change has been considerably strengthened in recent years, particularly with the Paris Agreements in 2015 (COP 21) which define a shared goal of limiting the rise in the average temperature of the planet to "well below 2°C above pre-industrial levels". This goal is based in particular on the work of the IPCC, which shows that the risks of damage become very high in scenarios involving a rise in temperature above 2°C (IPCC, 2015). Various nations are also beginning to make individual commitments by setting targets for reducing GHG emissions within a certain time. In the case of France, the goal of achieving carbon neutrality in 2050 was set by law in 2019 and the climate goals are reflected in the National Low Carbon Strategies (*stratégies nationales bas carbone* - SNBC), which consist of a GHG emissions reduction trajectory and measures to be implemented to achieve that objective. These strategies give rise to implementing decrees that set three-yearly carbon budgets (annual quantities of emissions that must not be exceeded). The scale of the efforts needed to achieve these goals, their distribution over time and the consequences for standards of living and their sustainability remain points of debate.

Thus, the question of the climatic sustainability of growth arises and the aim of environmental economic accounting is precisely to provide the data that allows this key issue to be analysed. Unlike traditional areas of national accounting, in which values, prices and volumes are measured, environmental matters are characterised by the absence of prices or by the fact that the latter do not reflect the value of assets (natural resources, biodiversity, the climate, etc.) or liabilities (pollution and global warming). Environmental economic accounting involves replacing market prices with a social value. In this respect, the Paris Agreement constitutes a turning point in the sense that the objective of human societies, in terms of climate, can now be considered fixed: to limit global warming to 2°C and to achieve carbon neutrality by 2050. In the language of environmental economic accounting, this Agreement is the benchmark for placing a value on carbon.

Translating our shared climate goal into action requires being able to predict the different possible economic and climatic trajectories in accordance with individual efforts. Using a macroeconomic model, created on the basis of realistic assumptions on decarbonisation technologies and the distribution of efforts across generations, we evaluate the optimal emission reduction trajectories for France and the world, as well as a measurement of the annual climate change mitigation effort. This model also makes it possible to determine a carbon value in France, revisiting the results of the Quinet Commission (Quinet, 2019). By significantly raising the carbon price in comparison with previous evaluations, the report of the Quinet Commission was an important moment in the debate on the social valuation of climate action. Our results go even further in this direction and lead us to consider the Quinet prices as minimums, in view of the goal of achieving carbon neutrality by 2050.

Modelling GHG emission reduction trajectories allows us to evaluate climate sustainability. However, it is more complex to measure sustainability in a general sense. The Commission on the Measurement of Economic Performance and Social Progress had, moreover, abandoned this ambition and its report recommended separating the two dimensions of economic sustainability and environmental sustainability (Stiglitz *et al*., 2009), thus rejecting approaches such as those based on inclusive wealth or adjusted net savings, which seek to evaluate overall sustainability by massing all the economic and natural "capital" that is transferred from one generation to the next.

However, progress concerning the carbon price and the estimation of decarbonisation technologies invites a review of the subject, by re-evaluating overall sustainability in France and worldwide, when the degradation of natural capital is valued using the new carbon price estimates.

After a description of the simplified climate economics model (section 1), we will focus on evaluating climate sustainability by comparing the actual trajectory of GHG emissions to that which would be required to meet the goals set by the Paris Agreements and by measuring the scale of the effort required (section 2). We will

then look at the resulting estimates of the social value of climate action (section 3), followed by addressing the issue of sustainability in a broad sense, through an evaluation, at both national and global level, of adjusted net savings and inclusive wealth (section 4), before concluding.

1. Evaluating the Effort to Combat Climate Change and Distributing the Effort across Generations: Analytical Framework

Evaluating the sustainability of the economic development trajectory essentially entails making future projections, both medium and long term, and therefore requires the use of modelling. There are many models, both national and international, that integrate environmental concerns, whether centrally or peripherally. They can comprise several hundred equations and are all useful for simulating, in the short or medium term, the impact of targeted or sector-specific measures. Their sophistication also has a cost, which is to make it more difficult to identify the assumptions that fundamentally determine their results. As Robert Solow noted in the introduction to his "A Contribution to the Theory of Economic Growth" (Solow, 1956), the strength of an economic model sometimes lies less in its complexity than in its ability to formulate central assumptions that offer the right compromise between simplicity and realism. It is on these central assumptions, which form the core of the integrated economy-climate model reactor, that we focus here (Figure I).

1.1. Greenhouse Gas Production and Emissions

In this spirit of focusing on the critical factors, here we consider a stylised model of an economy with capital K_t and labour L_t as factors of production and a Cobb-Douglas production function: $Y_t = K_t^{\alpha}(A_t L_t)^{1-\alpha}$. The work and technical progress A_t are exogenous and grow according to an exponential law.[1] In each period, households save a proportion s_t of the national income, feeding into the stock of capital, which depreciates each year at the rate of δ. The law for the development of physical capital is therefore: $K_{t+1} = K_t + s_t Y_t - \delta K_t$

The interconnection between the climate and the economy is essentially reflected in two elements:

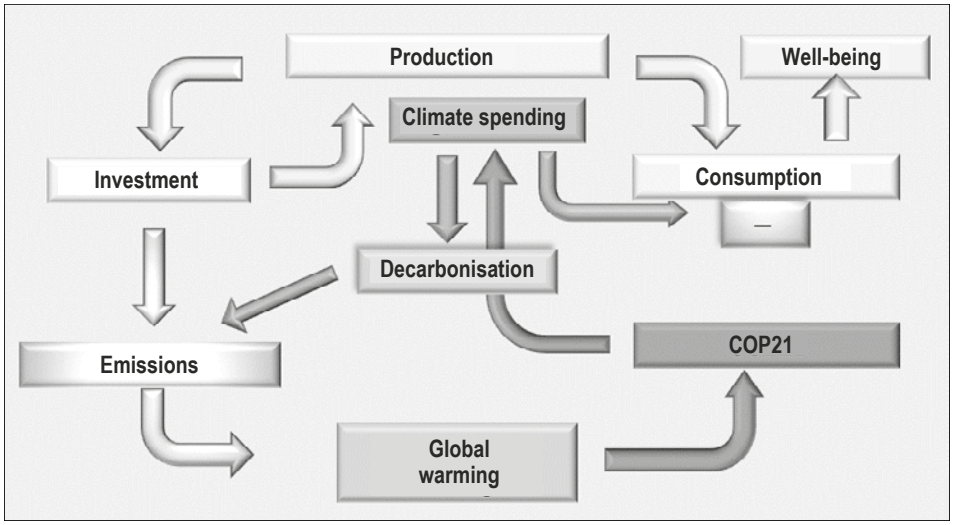
(i) economic activities are responsible for greenhouse gas emissions $E_t = \sigma_t Y_t$, where σ_t represents the carbon intensity of the economy;

(ii) climate spending D_t, in favour of decarbonisation technologies, can reduce carbon intensity and thus limit the growth of emissions. This spending reduces consumption by the same amount $C_t = Y_t - s_t Y_t - D_t$.

In each period, the public authorities have the opportunity to act on the two levers that are $\Lambda_t = D_t / Y_t$, the proportion of climate spending in GDP and s_t, the savings rate. They do so by seeking out the economic trajectory, compatible with the climate

1. *For the global model, the annual population growth rate decreases gradually over time, to reach a global population of around 10 billion inhabitants in 2050.*

Figure I – **Main determining factors of the model**



objective, which maximises a previously defined inter-temporal utility function.

1.2. Damage Function and Climate Target

One of the central issues is to evaluate the optimal GHG emissions target. The pioneering work of Nordhaus (1977), who built a Dynamic Integrated model of Climate and Economy (DICE), provides some initial elements of an answer. By precisely expressing the damage function as a function of global temperature, this type of model makes it possible to calculate an optimal trajectory, both economic and climatic. The greenhouse gas emissions goal appears endogenous to the overall model: this is the cost-benefit approach (Figure II-A).

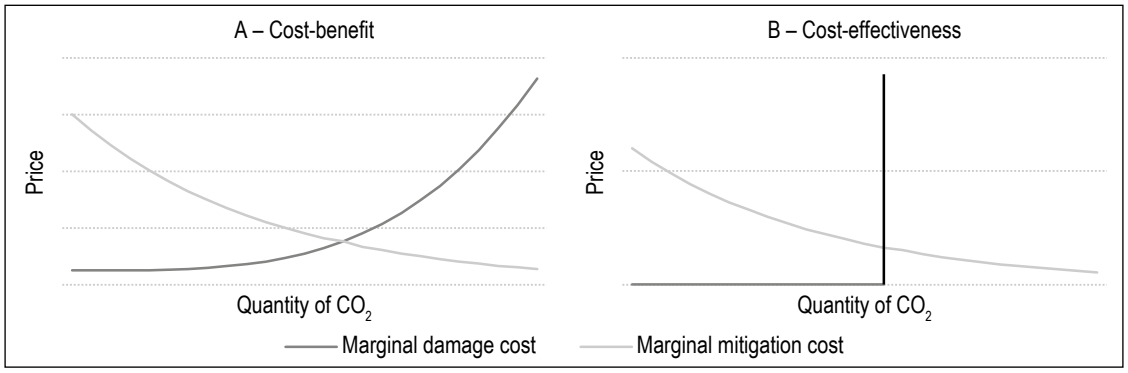
While this approach is natural from a theoretical point of view, it is particularly difficult to implement in practice due to the extreme difficulty of determining a monetary value for climate damage. There are commercial costs (such as the erosion of productivity and destruction of productive capital), but there are also non-commercial costs (such as the loss of biodiversity and destruction of ecosystems) that are much more difficult to value properly. In addition to marginal damage, there is the issue of the risks of serious and irreversible damage, or even collapse, which are generally not taken into account. The result is an underestimation of the damage and, consequently, economic policy recommendations that accommodate an unreasonable level of global warming. This is the case for the damage function of the DICE model, which is certainly quadratic as a function of temperature, but with such a low coefficient that the climatic optimum is achieved for a temperature of around +4°C compared to pre-industrial levels, which seems particularly optimistic, especially in view of the latest work of the IPCC.

In this respect, there is a before and an after 2015. The work of the IPCC has made it possible to form a scientific consensus on the consequences of global warming and the need to limit warming since the pre-industrial era to 2°C, which implies a cap on emissions over a certain time period. Other models therefore treat as given the goals of limiting the rise in temperature set by the international community (IPCC, Paris Agreements, etc.) and, accordingly, of reducing GHG emissions. This is particularly the case, out of necessity, with national models, since climate balances only make sense at global level. This second category of model is used to evaluate national and/or global trajectories. The principle is to set an exogenous protective goal of reducing emissions, then to quantify the spending trajectory necessary to achieve that goal. The damage function is therefore implicitly defined by the climate goal: before reaching the goal, the damage is zero or only slightly increasing; it becomes infinite if the goal is passed. We then speak of a cost-effectiveness approach (Figure II-B).

For France, the climate goal is currently defined by the 2019 Energy-Climate Law. The goal is to achieve net zero emissions (NZE), i.e. carbon neutrality, by 2050, by combining a division of emissions by a factor of around F=7 compared to 1990 levels and a doubling of the capacity of the carbon sink,[2] increasing it from 40 to

2. Reservoir that stores atmospheric carbon using a natural or artificial mechanism. Carbon sinks are essentially the oceans and forests, as well as CO_2 capture and sequestration projects.

Figure II – **Cost-benefit and cost-effectiveness approaches**



Reading note: Graph A shows the shape of the curves for the marginal damage cost (increasing with the quantity of CO_2 emitted) and mitigation (decreasing with the amount of CO_2 emitted). Graph B shows a new shape for the damage cost curve, which becomes infinite from a certain emission threshold, corresponding to the exhaustion of the carbon budget.

80 million tonnes per year. This goal follows on from an initial goal of dividing emissions by a factor of *F*=4 by 2050 in comparison with 1990 emissions levels, defined by the 2015 Energy Transition Law.

1.3. Technologies for Mitigation and Decarbonisation of the Economy

As the difficult issue of damage valuation is discarded by the *ex-ante* definition of an emission reduction goal, it is indeed the development of decarbonisation technologies that becomes a central assumption of the model. What is the cost of the GHG emission reduction technologies, known as the "mitigation cost", that will need to be used? In other words, what is the law for the development, between now and 2050, of the carbon intensity of the economy as a function of climate spending?

There is a broad consensus based around the idea that, the lower the carbon intensity, the more costly it is to reduce emissions, simply because the cheapest decarbonisation techniques are implemented first. This invites us to use a general law for the development of carbon intensity in accordance with climate spending that takes the following form: $\sigma_{t+1} = \sigma_t \left(1 - \varepsilon\left(\sigma_t\right) \Lambda_t\right)$, where $\varepsilon\left(\sigma_t\right)$ is a growing function of σ_t. At a given level of GDP, the lower the emissions, the more expensive it is to "mitigate" a given amount of CO_2. Here we use a simple functional form: $\varepsilon\left(\sigma_t\right) = \varepsilon\sigma_t^{\theta-1}$ where ε and θ are the parameters to be defined.

Two approaches are theoretically possible for assessing these parameters. The first approach is macroeconomic and econometric. It would consist of inter-temporal and inter-country regressions. Unfortunately, to date, the lack of sufficient data on climate spending does not allow this. This underlines how useful it would be if progress could be made very quickly in establishing environmental economic accounting. There is already a framework, the System of Environmental Economic Accounting (SEEA), which is a set of standards defined by the UN Statistical Commission and modelled in its architecture on the SNA (System of National Accounts) which governs the public accounts of nations.

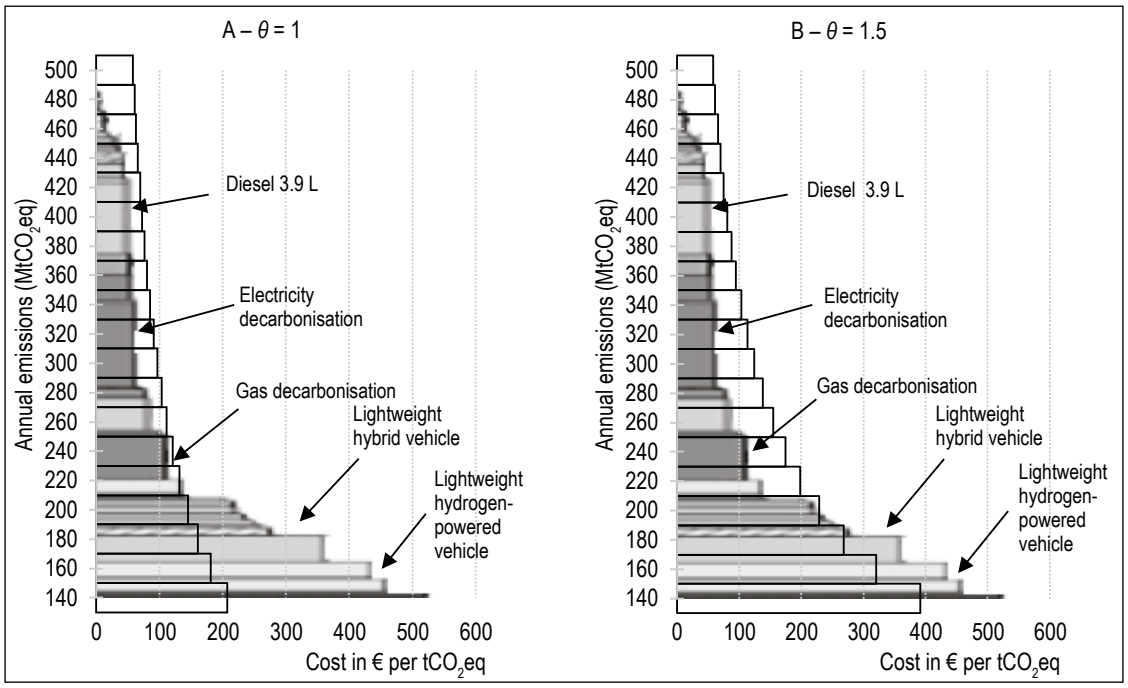
The other approach is microeconomic and parametric, based on the average mitigation cost curves for economy decarbonisation technologies. As its name suggests, this method consists of calculating the cost/effectiveness ratio of the different technologies (housing insulation, wind power, hydrogen-powered cars, etc.), which is the ratio between the total costs of implementation and the total emissions avoided. This method is implemented in France by the Ministry for the Environment using the TITAN model (formerly D-CAM), which ranks technologies in ascending order of cost and derives a curve comparing unit cost and total mitigation potential.

Figure III compares the average mitigation costs obtained from the technico-economic studies and those obtained with our carbon intensity development assumption for both cases $\theta = 1$ or 1.5, with a value of $\varepsilon\sigma_0^{\theta-1} = 1.5$.[3] This approximation tends to validate both the nature of the mitigation equation and the value of the parameter ε. For example, the technologies planned by the SNBC in the area of annual emissions of around 150 $MtCO_2eq$ (e.g. lightweight hydrogen-powered vehicles) have an average mitigation cost of €450, which is quite close to the average macroeconomic cost for $\theta = 1$ (€370). In general, our development assumption is consistent with the available microeconomic evaluations.

Some studies also presuppose the discovery of a so-called "backstop" technology that can be deployed on a large scale to absorb greenhouse gases and that is partly an alternative to reducing emissions. Such technologies, particularly bioenergy with carbon capture and storage (BECCS), are currently being tested. They aim to generate so-called negative CO_2 emissions by intercepting the release of CO_2 into the atmosphere and redirecting it to geological storage sites. Nevertheless, the path to widespread use of such technology remains very long, making this possibility rather uncertain in the medium term. Furthermore, there is no consensus regarding the cost of such a technology, with estimates in the literature ranging from one hundred to several thousand euros per tonne of CO_2, or on the possibility of large-scale deployment. In view of our study period of up to 2050, which is relatively short given the time required to industrialise production of such technology, we assume that its use will remain marginal.

3. To determine a value for this parameter, we proceed from the observation that emissions per € of GDP have fallen by an average of 2.5% per year over the last ten years, which is a slight acceleration of the reduction compared to the previous two decades (2% per year). Climate spending, in turn, is evaluated at €41.4 billion for France in 2018, or 1.8% of GDP, which is a slight increase compared to the beginning of the decade (€34.4 billion, or 1.6% of GDP). It is on this basis that we can estimate a value $\varepsilon = 1.5$*, which is equal to the ratio between the average reduction in the carbon intensity of GDP over 2013-2018 (2.5%) and the average climate spending between 2011 and 2017 as a % of GDP (1.7%).*

Figure III – **Comparison of average technico-economic (D-CAM) and macroeconomic mitigation curves for cases $\theta = 1$ and $\theta = 1.5$**



Reading note: The solid bars represent the average mitigation costs of the different D-CAM technologies calculated by the CGDD; the hollow bars trace the average mitigation cost corresponding to the macro-technological equation $\dot{\sigma}_t/\sigma_t = -\varepsilon\,\sigma_t^{\theta-1}\Lambda_t$.
Sources: Baptiste-Perrissin & Foussard (2016) for D-CAM, authors' calculations.

1.4. Optimality and Intergenerational Equity

Once the climate goal has been defined, the path to take towards that goal must be determined, taking into account the intergenerational equity of the climate spending trajectory that allows for emissions reduction. Which generations should pay for the climate? Is it preferable to make the entire adjustment now, being prepared to lower per capita consumption today and then returning to an upward trajectory in the future, or would it be better to spread the adjustment over the first decade, for example, if there is a greater preference for the present?

It is customary in the models to formalise this issue by using the framework set out by Hotelling (1931) on the economic analysis of exhaustible resources. The "Hotelling rule" stipulates that the income drawn from an exhaustible resource must develop exponentially, at a rate equal to the interest rate, until the resource is exhausted.[4]

This approach leads to two pitfalls. First of all, while the carbon budgets allocated to each country under the Paris Agreement are akin to an exhaustible resource, the fact that decarbonisation technologies exist means that governments have the option to somewhat "extend" the resource. Thus the Hotelling rule does not apply directly, but this pitfall is easily overcome by integrating the additional control variable of climate action into the optimal programme. The second pitfall is a type of contradiction between the method and the goal. Since the Brundtland Report (1987), the goal has been to promote sustainable development, defined as a form of development that meets the needs of current generations without compromising the ability of future generations to meet their own needs. It is paradoxical in this context to define the corresponding economic programme as the maximisation of the inter-temporal satisfaction of current generations.

4. The interest rate r is in turn determined by Euler's canonical equation, which is $r = \rho + n + \tau\gamma$ *where* ρ *is the rate of preference for the present,* n *is the population growth rate and* γ *is the rate of technical progress, while* τ *is the inverse of the elasticity of the utility function. The Euler equation derives from a Ramsey optimisation programme for the present and future utility flow for consumption* $\sum_{t=0}^{T}\beta^t u(c_t)$*, where* c_t *is the per capita consumption,* u *is a concave function and* β *is a discount factor that reflects a preference for the present. For clarification, with a preference rate of 2%, a technical progress rate of 1%, a population growth rate of 1% and an elasticity of utility of consumption of 0.5, Euler's canonical equation results in an* r *rate of 5%.*

The Brundtland doctrine is more in line with the idea, formalised by Arrow *et al.* (2012), that a sustainable trajectory is one in which well-being should not decline. If monetary satisfaction of generation t is equated with $V_t = \left[C_t / L_t\right] / \left(1+\rho\right)^t$, where ρ is a parameter taking into account the effects on perceived monetary well-being of the passage of time alone (Easterlin, 1974), then, as in the steady state C_t / L_t growing at the rate of γ, $V_t = \left[C_0 / L_0\right] \left[(1+\gamma) / (1+\rho)\right]^t$ is increasing when ρ is less than γ and decreasing if it is not. The more ρ is high and the lower the technical progress, the more disadvantaged future generations will be. If the public authorities aim to achieve Brundtland-style sustainable development, they can express it by setting a parameter $\rho = \gamma$ in the collective utility.

A relatively simple way to express this idea is to define the programme of the public authorities as the determination of the level of the control variables (climate spending Λ_t and savings rate s_t) making it possible to maximise monetary well-being, equated to the discounted per capita consumption of the worst-off generation.

In analytical terms, the aim is to maximise the inter-temporal utility defined by:

$$\max_{\Lambda_t, s_t} \left\{ \min_t \left[C_t / L_t\right] / \left(1+\rho\right)^t \right\}$$

When this parameter is equal to the growth of technical progress and the savings rate is constant over the period, this optimisation programme also leads to a ratio of Λ_t for climate spending over GDP constant during the period. In this specific case, the optimal path to the goal follows an intuitive notion of generational equity, according to which the effort required at each date follows a uniform distribution over time. It would therefore be a matter of making the adjustment from the initial period, or at least as quickly as possible, and then ensuring that all generations have a constant level of climate spending as a percentage of GDP.

With the assumptions described above (exogenous emission goal, law for the development of carbon intensity and the intergenerational equity criterion), we are equipped to examine the issue of sustainability in its various aspects, both climatic and economic. In particular, we will define the concept of the climate sustainability of the economy according to an equity/effectiveness approach, starting from the concept of a sustainable trajectory corresponding to a trajectory that satisfies the following two conditions: *(i)* compliance, by 2050, with a ceiling goal for annual greenhouse gas emissions; *(ii)* a distribution of climate efforts over time that protects future generations.

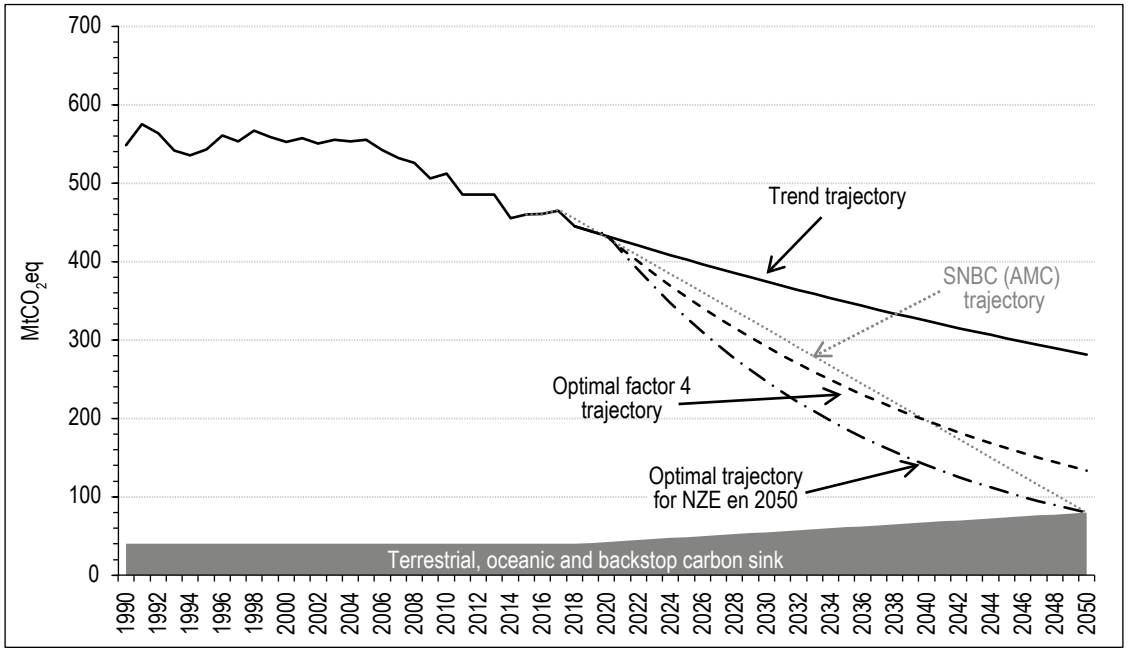
2. Carbon Emissions Trends that are Currently Incompatible with our Climate Commitments

2.1. CO_2 Emission Reduction Trajectories

To begin, we describe the results for France corresponding to the assumption $\theta = 1$ concerning the decarbonisation technologies, i.e., let us recall, in which climate spending mitigates the carbon intensity of production according to the relationship $E_{t+1} / Y_{t+1} = E_t / Y_t \ (1 - \varepsilon\Lambda_t)$. This optimal trajectory is plotted (Figure IV) for reductions by factors 4 and 7 of SNBC-1 and 2. The graph also plots (a) the trend trajectory assuming 1.5% annual growth and a decline in carbon intensity consistent with the maintaining of the current climate effort and (b) the AMC trajectory as notified to the European Commission (with AMC standing for "*avec mesures complémentaires*" in French, referring to a scenario with complementary measures that have not yet been approved). The accumulated levels of emissions per sub-period are also provided in Table 1. It is specified that the emissions in question are in all cases emissions within the national territory, also called the "national inventory". These emissions are those that are the subject of international commitments and it is for this reason that they are used in this article; however, they should be distinguished from the notion of a "carbon footprint", which measures the emissions related to our lifestyle, including greenhouse gas emissions associated with our imports.

Past achievements and the trend trajectory appear to be well above the two optimal trajectories by a factor of 4 and 7 and the SNBC budgets. The first budget for SNBC-1 was slightly exceeded (458 $MtCO_2eq$ compared with the planned 440 $MtCO_2eq$) and, above all, the trend scenario would then clearly diverge from the planned trajectories: 2030 would be at 68% of the 1990 level instead of 57% of the SNBC-2 and 2050 would be 3.5 times higher than the carbon neutrality goal (281 $MtCO_2eq$ instead of the planned 80 $MtCO_2eq$). In addition, the AMC scenario notified to the European Commission would meet the carbon neutrality goal, but at a quasi-linear pace, therefore making it different from the optimal scenario of intergenerational equity defined above.

Figure IV – **Greenhouse gas reduction trajectories in France**



Reading Note: Assuming the reduction in emissions by 2050 by a factor of 4, emissions in 2030 are expected to be 53% of 1990 levels. As at the same date, this percentage should be 45% for a target of net zero emissions in 2050, while the trend trajectory suggests a ratio of 68%.
Sources: CITEPA, authors' calculations.

Table 1 – **Planned, trend and optimal carbon budgets by sub-period**

Période	Planned emissions (low carbon strategy)	Actual and trend emissions	Optimal emissions (Factor 4 in 2050)	Optimal emissions (Neutrality in 2050)
2015-2018 (1st budget*)	422	458	458	458
2019-2023 (2nd budget*)	399	427	417	408
2024-2028 (3rd budget*)	359	397	346	311
2029-2033 (4th budget**)	300	369	285	234
2034-2038 (AMC***)	244	343	235	177
2039-2043 (AMC***)	185	320	194	133
2044-2048 (AMC***)	127	298	160	101
2050 (AMC***)	80	281	137	80

Sources: *SNBC2015, **SNBC2020, *** 2019 Government projection with complementary measures.

In the steady state, it is possible to formulate a simple rule that makes it possible to gauge whether the carbon trajectory is meeting its goal by dispensing with the solution of a model. In this case, economic activity grows at a constant rate g, and the optimal trajectory of carbon emissions, as we have just defined it, obeys a simple law of decreasing at a constant rate that we call Γ. Indeed, if climate spending represents a constant proportion Λ of GDP, the carbon intensity σ decreases at a constant rate of $\varepsilon\Lambda\%$, since $d\sigma/\sigma = -\varepsilon\Lambda$. As a result, GHG emissions decrease at a constant ratew of $\Gamma = \varepsilon\Lambda - g$. The value to be assigned to Γ is then deduced directly from the GHG reduction factor F in relation to the starting year, and from the number of years T before the set deadline, with condition $(1+\Gamma)^T = F$ leading to $\Gamma = F^{1/T} - 1$. Thus, for France, where the aim is to reduce emissions from 439 to 80 $MtCO_2eq$ between 2019 and 2050, F=5.48 of a T duration of 31 years, $\Gamma = 5.48^{1/31} - 1 = 5.6\ \%$. This means that once emissions are decreasing by less than 5.6% per year, climatic sustainability is not ensured, in the sense that either carbon neutrality will not be achieved on time or the effort is too spread out over time.

This rule is not fully an accounting rule: it is indeed a case of moving from point A to point B in time *T*, but with a rate of progression resulting from the equity rule defined above, constant as a percentage and therefore in level, moving faster at the beginning and slower at the end than the straight line. Nevertheless, it is very useful for providing clarification and determining orders of magnitude, because it tells us how much the GHG emissions should be decreased immediately and sustainably to restore a sustainable trajectory (in the same way as sustainability indicators, such as the tax gap).

Using the variant $\theta = 1.5$ for the decarbonisation technologies would imply a slightly modified distribution of effort (Figure V). In this case, the rule just stated does not apply, the rate of reduction is not constant and simulations must be used. Unsurprisingly, however, the bearish profile of the new trajectory is more pronounced at the beginning of the period.

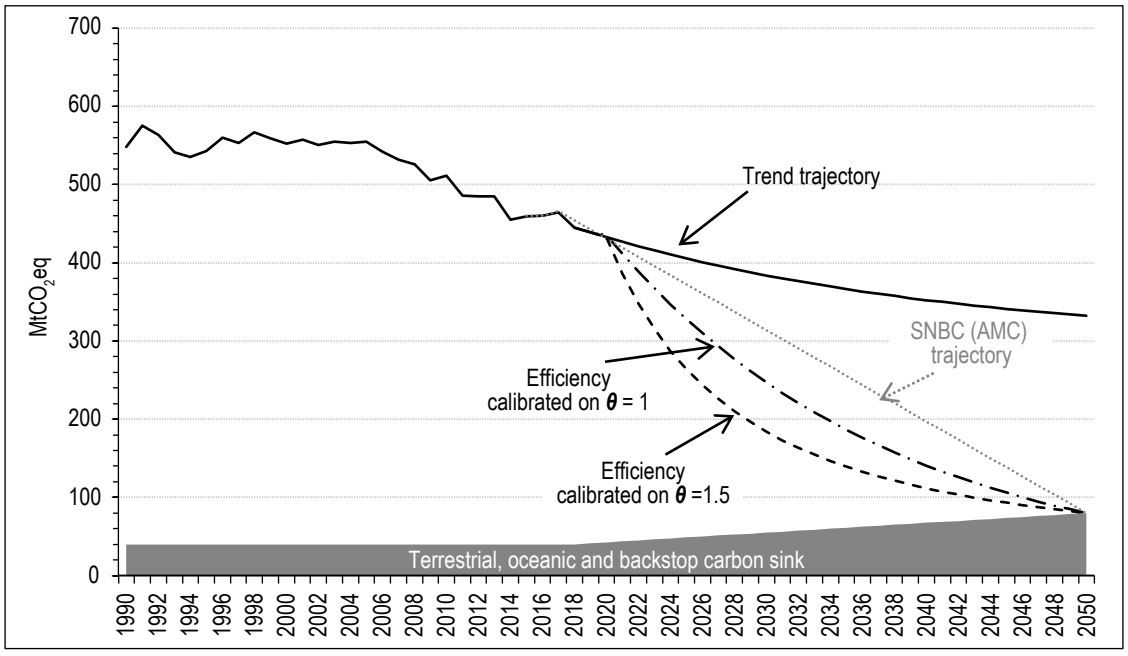
We can come back here to the recommendations of the Stiglitz Commission on Carbon Prices (Stiglitz *et al*., 2009) for the measurement of sustainability. It recommended that the "environmental aspects of sustainability deserve a separate follow-up based on a well-chosen set of physical indicators. In particular, there is a need for a clear indicator of our proximity to dangerous levels of environmental damage". The monitoring of GHG emissions is perfectly in line with this goal as far as the climate challenge is concerned and it conveys a message that appears here without appeal: in terms of climate, our trajectory is not sustainable. France, although not the worst placed among the richest countries, emits ten times more GHG (439 $MtCO_2eq$) than it absorbs (40 $MtCO_2eq$). The projections show a likely downward trend in the coming years, but one clearly insufficient for a return to equilibrium in the time necessary. At global level, the situation appears even more critical: the trend is upward, whereas emissions need to be divided by a factor of 4 by 2050 to contain warming at 1.5°C.

Finally, it should be recalled that, despite a drop in the carbon inventory, France's footprint has continued to grow, which means that emissions produced within the national territory have been gradually replaced by imported emissions. Figure VI shows the different possible projections depending on whether France (France NZE + World BAU), the rest of the world (France BAU + World NZE) or both (France NZE + World NZE) respect the climate goals of achieving carbon neutrality by 2050 (see Online Appendix C1, link to the Online Appendices at the end of the article).

2.2. Climate Spending

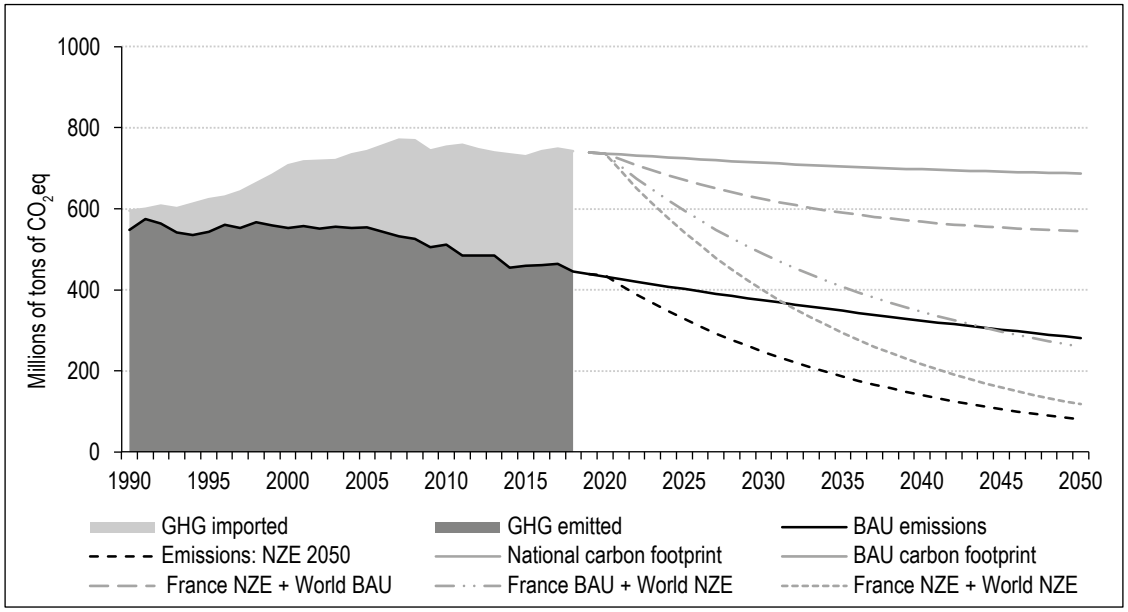
With the trajectories for returning to carbon neutrality having been established, our model

Figure V – **Optimal trajectories towards carbon neutrality by efficieny of decarbonisation technologies**



Sources: CITEPA, author's calculations.

Figure VI – **France's carbon inventory and footprint**



Sources: INSEE, CITEPA, authors' calculations

makes it possible to directly quantify the costs of adhering to them. For France, annual spending associated with the optimal trajectory would amount to 4.5% of GDP, corresponding to around €100 billion,[5] which represents an increase of more than a factor of 2 compared to the current spending evaluated, for the state, businesses and households, with just over €45 billion spent in 2018 (1.9% of GDP) by the Institute for Climate Economics (I4CE, 2019). This represents a significant, but not impossible, effort: in relation to the population, the amount is around €1,500 per capita instead of the current €600.

Again, we can reveal a simple rule for economies in a steady state, between the optimal national carbon effort and economic growth. It should be remembered that in this case and where $\theta = 1$, the constant rate Γ of reduction of GHG emissions is equal to $\varepsilon\Lambda - g$. As a result, the effort that ensures compliance with the goal is $\Lambda^* = [g + \Gamma] / \varepsilon$. This relationship teaches us, for example, that the current effort of 1.9% of GDP, if not increased in the coming years, would not be compatible with achieving carbon neutrality by 2050 unless GDP falls at a rate of 2.7% per year.[6]

These results are sensitive to the assumptions used, particularly concerning decarbonisation technologies, the rule for intergenerational effort sharing and also the economic growth rate, which is considered exogenous in this model. Table 2 illustrates the sensitivity of the level of annual effort required under the growth and energy efficiency scenarios.[7] Thus, the annual climate spending may increase from around €65 billion in 2018 (zero growth and optimistic on efficiency) to €165 billion (growth of 1.5% and prudent on efficiency).

At global level, although the emission reduction factor required to achieve carbon neutrality is slightly lower than that required for France, the projected growth is higher and, in the end, the global financial effort would be of the same order and even slightly higher than that to be made nationally, as percentage points of GDP: our model results in a climate effort rate of 5.1% of global GDP, compared to 4.5% at national level for France. In contrast, the change in scale is much larger, with global climate spending likely to be less than 1% of global GDP at present.[8]

---

5. Very precisely, €105 billion in 2019, which would then develop in value like the GDP.

6. Indeed, $g = \varepsilon\Lambda - \Gamma = 1.5 * 1.9\% - 5.6\ \% = -2.75\%$

7. The results are tested for a value of ε ranging from 1 (prudent scenario) to 2 (optimistic scenario), with the so-called "central" scenario corresponding to $\varepsilon = 1.5$.

8. $681 billion in 2016, according to the 2018 report of the Standing Committee on Finance (SCF) of the Conference of the Parties (COP) of the UNFCCC (SCF, 2018), for a global GDP of $76,000 billion, which equates to 0.9%. It should be noted that this figure is consistent with a value of ε_M=1.5, as it implies a reduction in carbon intensity of 0.9X1.5=1.25 per year, which is more or less the trend observed (-1.2% per year over the 2008-2018 period).

Table 2 – **Sensitivity of climate spending (as a % of GDP and in billions of € in 2018) to the growth and energy efficiency assumptions**

Energy efficiency scenario \ Growth scenario	1.5%		1.0%		0.0%	
Prudent	6.9 %	(€157 bn)	6.4 %	(€147 bn)	5.6 %	(€129 bn)
Central	4.5 %	(€104 bn)	4.3 %	(€97 bn)	3.7 %	(€85 bn)
Optimistic	3.4 %	(€77 bn)	3.2 %	(€72 bn)	2.8 %	(€63 bn)

2.3. Saving Strategies

To conclude this section, here we examine four variants that depart from the assumption of a constant savings rate and vary the rules for the development of consumption and climate effort (Figure VII).

The first column (scenario S1) corresponds to the trajectories that we have described so far: as the exogenous savings rate is constant, both GDP and capital remain on their regular growth path, hence a constant *K/AL* ratio, where *AL* represents labour plus the factor of technical progress. By construction, the consumption per unit of efficient labour remains constant after the initial adjustment, which implies a constant discounted standard of living $V_t = \beta^t C_t / L_t$ with $\beta = 1/(1 + \rho)$.

Scenarios S2 and S3 maintain the assumption of a constant climate effort rate, but with an endogenous savings rate, which varies across time. More precisely, the savings rate is the consequence of the choice of consumption, which derives from an intertemporal optimisation programme. The two scenarios are different in the choice of the utility function that will be maximised (see Online Appendix C2). In the second scenario (S2), this is a max/min type of optimisation, which implies a constant level of consumption by unit of efficient labour, once the initial adjustment is realised. The savings rate is gradually reduced to bring capital to its new steady state,[9] which corresponds to a slight decrease of the average standard of living compared to the reference trajectory. In the third scenario (S3), the consumers seek to maximise $\sum_{t=0}^{T} \beta^t \frac{c_t^{1-\tau}}{1-\tau}$, with a finite parameter τ[10] involving a substitution between current and future consumption (in contrast to the Brundtland approach of the public authorities which corresponds to an infinite τ). They chose to reduce more strongly their savings in the initial period to smooth the downfall of consumption caused by a constant climate effort over the period.

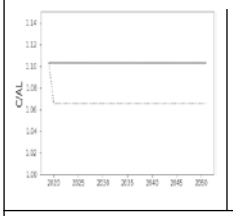
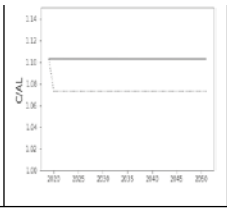
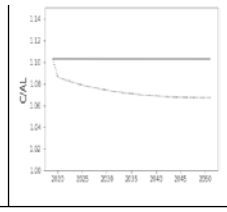
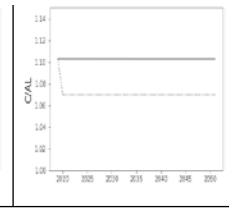
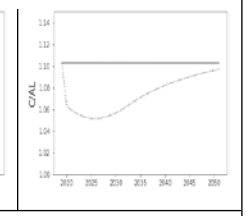
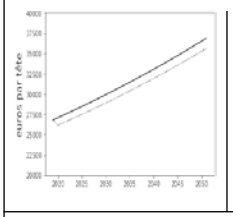
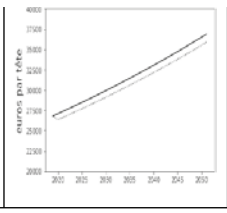
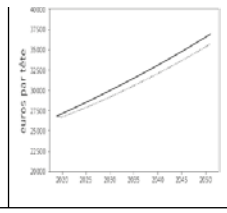
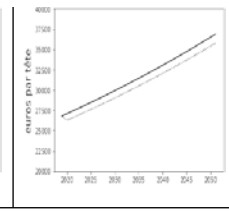
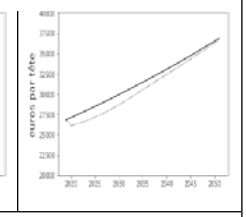
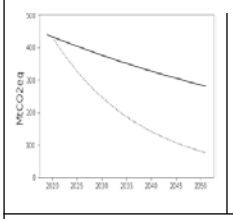
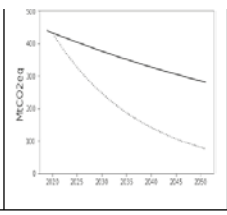
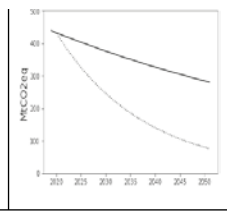
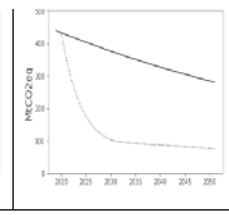
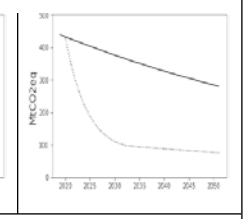
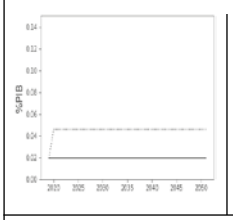
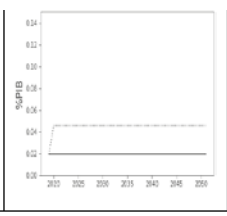
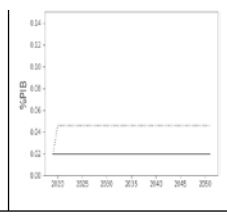
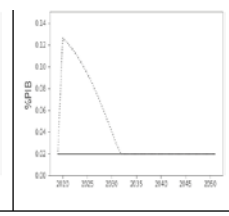
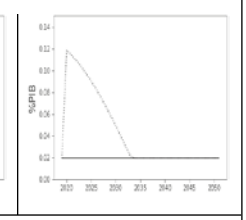
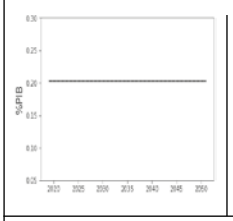
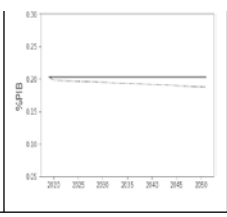
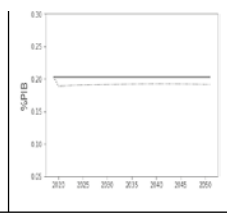
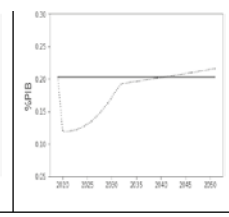
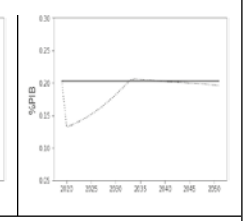
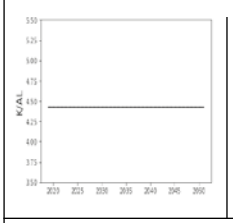
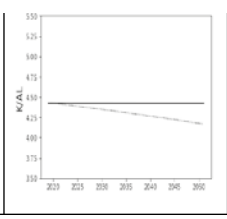
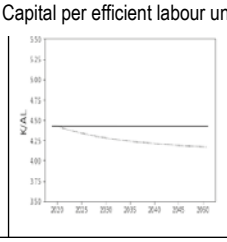
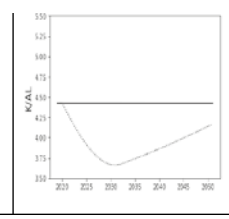
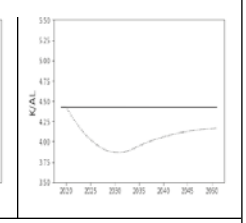
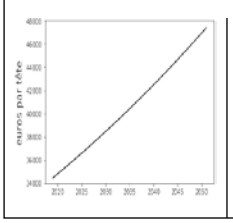
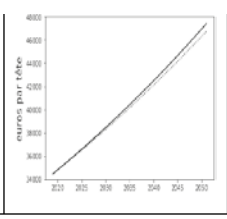
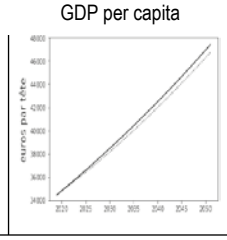
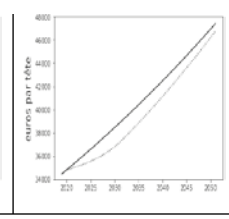
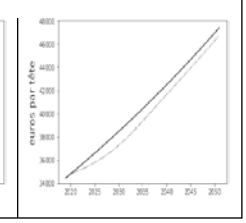
The fourth and fifth scenarios (S4 and S5) make the climate effort rate endogenous. The difference between scenarios S4 and S5 lies in the rule for the development of consumption per efficient work unit resulting either from a max/min type of optimisation programme (S4), or that of more impatient consumers (S5). If the end point is the same for both emissions and the capital goal, trade-offs can be made over time between investment, climate effort and consumption. The optimal trajectory corresponds to a much faster decarbonisation, with carbon neutrality being achieved by 2030; this assumes a higher climate effort until that time in the reference scenario, and lower thereafter; this effort is cushioned, symmetrically, by an immediate reduction in the savings rate, before it returns to its initial trajectory. The growth of both GDP per capita and capital per capita is slowed down before, once decarbonisation is complete, resuming its course towards the new steady state. This latter trajectory, due to the scale of the adjustments it implies, is undoubtedly not the most likely, but it has the merit of showing the possibility of a faster reduction in CO_2 emissions – thus further limiting global warming – without harming standards of living, by taking action on the savings rate.

All the scenarios presented here display a reduction in consumption per capita the first year during the initial adjustment, due to a significant rise of climate effort. This initial effort in consumption is largely offset later on

9. As the savings rate is endogenous, in order to solve the optimal public authority programme, it is necessary to define the economic output goal. Our simulations here are based on the goal that the economy, in 2050, will be in its new regular state, integrating a permanent decarbonisation effort equal to the optimal effort of the period 2020-2050. The need to decarbonise the economy increasingly constantly means reducing total factor productivity and thus reducing the optimal K/AL ratio.

10. τ set at 2, the standard value in the literature.

Figure VII – **Economic and climatic trajectories under different savings scenarios**

S1 - Constant saving rate and climate effort	S2 – Euler consumption (maximin) and constant climate effort	S3 - Euler consumption (τ=2) and constant climate effort	S4 - Euler consumption (maximin) and endogeneous climate effort	S5 - Euler consumption (τ=2) and endogeneous climate effort
Consumption per efficient labour unit				
				
Consumption per capitae				
				
Emissions				
				
Climate spending				
				
Savings rate				
				
Capital per efficient labour unit				
				
GDP per capita				
				

by an increase of consumption by capita, which grows like technical progress. Nevertheless, in order to prevent the risk that growth is lower than expected, or decarbonisation more expensive, there is a clear interest to bring forward the efforts at the beginning of the period. If our utility function invites to do all the ajustment efforts as quickly as possible, the adjustment can also be smoothed on several years to avoid the negative initial shock in consumption per capita.

3. A New, Higher Carbon Price, in Line with the Goal of Achieving Carbon Neutrality by 2050

3.1. The Social Value of Climate Action

Based on evaluations of the overall cost of the decarbonisation strategies, it is then possible to move on to the determination of a carbon price. It is known that market mechanisms are of little use in placing a value on the cost of CO_2 emissions. The fundamental reason for this is that CO_2 has no extraction cost, unlike, for example, the gas and mining industries: because it is neither sold nor purchased, CO_2 has no price. Since 2005, there has been a European market for CO_2 quotas, the EU Emissions Trading System (EU ETS). However, firstly, it concerns only around 5,000 companies, representing 45% of emissions, and, secondly, the allowances allocated to them are insufficiently binding for the price on this market to reflect a social value. Thus, between 2013 and 2018, the CO_2 allowances, known as EUAs (European Union Allowances) traded at around five euros per tonne of CO_2.

At what level, then, should the carbon price be set? It is necessary to go back to the basics of the climate economy. CO_2 emissions have a cost because they are responsible for global warming and, therefore, cause damage to the economy. Climate action has value because investing in decarbonisation technologies will prevent future generations from suffering the now well-documented negative consequences of rising temperatures. This is why the Quinet Commission wished to refer to the notion of the "value of climate action" (Quinet, 2019).

This general principle being set, the term of "social" price of carbon can correspond to a number of notions, which need to be considered with caution in comparisons, as well as in the use that can be made from the estimated valued of the models. Talking of a social price requires before all to clarify what is meant by "social". In other words, what is the objective fixed by the society with regard to climate change, that the fixation of such carbon price can contribute to. There are essentially two approaches: an "accounting approach" and a "cost approach". The first one, based on the volume-price split of the optimal climate spending, consists in dividing such spending by the current GHG emissions, allowing thus to measure at which price to charge, implicitly or explicitly, carbon emission in order to reach the target of carbon neutrality in a equitable repartition of efforts among generations. The second approach consists in dividing the optimal climate spending by the cumulated flow of current and future emissions avoided. It is thus a logic of incentives targeting the evolution of behaviours towards decarbonisation: it is the viewpoint of the Quinet Commission, aimed at integrating the climatic dimension in the measurement of socio-economic cost of investments.

The two notions are of course linked, and can be made consistent with each other. We will nevertheless put forward the first approach, which seems to be the most effective and robust, given the uncertainty in the measure of the cumulated flow of avoided emissions, and notably the actualisation rate.

In practice, the social value of carbon covers a very wide range of climate policies, ranging from carbon taxes and emission allowances to the imposition of thermal standards for buildings, the cost of which is covered partly by the owners and partly by public support such as tax cuts, and the financing of public transport by local authorities and their transport authorities. To confuse the social value of carbon with a carbon tax is to confuse policies to combat global warming with their funding. Furthermore, both in France and everywhere else, carbon taxes so far represent only a minority share of the climate effort.

3.2. Estimates of the Social Cost of Carbon

With the meaning of the notion of the social value of climate action – or the social cost of carbon – having been clarified, its calculation follows directly from its definition, as this value – according to the "accounting approach" – must verify equality at every point of the optimal trajectory: $P_t^{co2} E_t^* = \Lambda_t^* Y_t^*$ where E_t^*, Y_t^* and Λ_t^* refer to emissions, GDP and climate effort along this trajectory, respectively. Stated in this way, the social value of carbon would amount for France to around €250 in 2020, €500 in 2030, €1,010 in 2040 and €2,050 in 2050 for the objective of carbon neutrality (Table 3).

Table 3 – **Social value of climate action for the objective of carbon neutrality in 2050**

	2020	2030	2040	2050
National values (€/ton of CO_2eq)				
Model results				
Accounting approach	247	501	1,014	2,052
For the record, with the objective of Factor 4 in 2050	*188*	*320*	*547*	*937*
Cost approach (actualisation rate of 5%)	127	258	522	1,057
Values retained by the Quinet Commission 2019				
Cost approach	88	250	500	775
Models used by the Quinet Commission				
ThreeME Model		143	1,128	2,389
NEMESIS Model		185	784	(*)1,934
POLES Model		351	845	3,515
TIMES Model		228	465	2,451
Global values (€/ton of CO_2eq)				
Model results				
Accounting approach	72	161	359	801
IPCC Estimates				
IPCC 1.5°C		284	497	872
IPCC 2°C		139		440

(*) Value for the year 2045.
Sources: Quinet Commission (2019), authors' calculations

The models used nationally as well as those used by the IPCC, also tend to produce even higher evaluations. Our estimates for the objective of carbon neutrality correspond to a quasi-doubling of the social value of carbon compared to the factor 4 goal that prevailed until 2018. This can be understood easily if we come back to the formation of this value: since $P_t^{co2}E_t^* = \Lambda_t^* Y_t^*$, the price ratio P_t^{F7} / P_t^{F4} can be decomposed as a product $\left[\Lambda_*^{F7} / \Lambda_*^{F4}\right] \times \left[\mathrm{E}_t^{*F4} / \mathrm{E}_t^{*F7}\right] \times \left[\mathrm{Y}_t^{*F7} / \mathrm{Y}_t^{*F4}\right]$. In a scenario where the two GDP trajectories would be the same, we would have $P_{2050}^{F7} / P_{2050}^{F4} = 7/4 \times \left[\Lambda_*^{F7} / \Lambda_*^{F4}\right]$. Given that effort Λ_*^{F7} is obviously higher[11] than that, Λ_*^{F4}, corresponding to a factor of 4, P_{2050}^{F7} is approximately equal to $2 \times P_{2050}^{F4}$. Let us stress that this doubling of the price does not necessarily mean a doubling of the optimal climate spending because, at the same time, the GHG reduction is also faster.[12]

If we now measure the social value of carbon according to the "cost approach" and with an actualisation rate of 5% on the measure of future avoided emissions, we obtain the amounts of €127 in 2020, €258 in 2030, €522 in 2040 and €1,057 in 2050. The orders of magnitude are comparable to those proposed by the Quinet report, specifically €250 in 2030, €500 in 2040 and €775 in 2050.[13] Our simulations tend to confirm the very strong revaluation made by the Quinet report (Quinet, 2019), as opposed to the estimates commonly accepted previously, such as the one proposed in 2017 by the Stiglitz-Stern Commission (Stigltiz *et al*., 2017),[14] which was €70 to €100 in 2030, not to mention the values still used by the World Bank (World Bank, 2011) or the UNDPD (UNU-IHDP, 2012) : $30 or €25.5, to calculate net savings and adjusted net savings, to which we will return later and which seem out of scale.

11. For France, Λ_*^{F7}=4.5% and Λ_*^{F4}=3.5%.

12. For France, $\Lambda_*^{F7}/\Lambda_*^{F4}$= 4.5 / 3.5 = 28%;

$P_{2050}^{F7} / P_{2050}^{F4} = 7/4 \times \left[4.5 / 3.5\right] = 2.25$.

13. While the Quinet Commission re-evaluates the social price of carbon in light of the new neutrality goal, it considered that the results of the technico-economic and macro-sectoral models used become less sound from 2040, or even 2030, and therefore decided to cap the price afterwards, in view of the technological uncertainties in the medium term.

14. The authors have nevertheless clarified that their estimated cover only one part of the social value of carbon: "This commission concludes that the explicit carbon-price level consistent with achieving the Paris temperature target is at least US$40–80/$tCO_2$ by 2020 and US$50–100/$tCO_2$ by 2030, provided a supportive environment policy is in place" *(Stiglitz* et al. *2017, p. 3).*

Finally, we estimate a global carbon price, which is a priori not the same as the national price. In fact, if we start from the definition of the social value of carbon, the ratio between the global and national levels can be written as follows:[15] $P_t^{MD} / P_t^{FR} = \left[\Lambda_*^{MD} / \Lambda_*^{FR}\right]\left[\sigma_t^{*FR} / \sigma_t^{*MD}\right]$. However, as we have seen, global and national climate efforts represent a comparable proportion of GDP (5.1% and 4.5%, respectively), the same cannot be said for carbon intensity (CO_2/GDP ratio), which is 720 g per € of GDP[16] at global level, compared with 189 g per € in France, which is a ratio of 1 to 3.8. The global value comes out of our simulations at €161 per tonne of CO_2 in 2030, €359 in 2040 and €801in 2050, which is broadly in line with the simulations of the IPCC. Indeed, the average IPCC values for the goal of limiting the temperature increase to 1.5°C (i.e. a scenario with a 33% probability of exceeding 1.5°C), a goal that would require achieving carbon neutrality in 2050, come out at €284 in 2030, €497 in 2040 and €872 in 2050.

3.3. Carbon Price Accounting

Continuing with our endeavour to express simplified rules under the assumption of stable growth at rate g, we can establish two new rules concerning the social value of climate action, according to the "accounting approach". To obtain a reduction Γ in emissions at a given date, if GDP grows at the rate g, a reduction in the carbon intensity of production is needed at the rate of $d\sigma/\sigma=\Gamma+g=r$. Since $d\sigma/\sigma=-\Lambda\varepsilon$, this requires constant climate spending when expressed as a percentage of GDP of $\Lambda=r/\varepsilon$. We deduce therefore an initial price, which is equal to the initial spending per tonne of GHG emitted, is $P_0^{co2} = (r/\varepsilon)\times(Y_0 / E_0)$. On any given date, this same price will be $P_t^{co2} = (r/\varepsilon)\times(Y_t / E_t) = \left(rY_0 e^{gt}\right) / \left(\varepsilon E_t e^{-\Gamma t}\right)$ or even $P_t^{co2} = P_0^{co2} e^{rt}$ and a price that thus increases at the rate r, with this growth reflecting the increasing difficulty of continuing to reduce emissions as carbon intensity declines.

Two new rules can therefore be established. The first is that, as at the initial date, the social value of carbon is at least equal to $P_0^{co2} = (r/\varepsilon)\times(Y_0 / E_0)$, where Y_0 and E_0 are the initial GDP and level of CO_2 emissions, respectively. The second is that, along this trajectory, the social value of carbon follows a law of exponential growth at the rate $r = g + \Gamma$, where g is the GDP growth rate and Γ is the annual percentage reduction in emissions goal.[17]

The latter rule is similar to a Hotelling rule, which stipulates that the price of a scarce resource must develop exponentially, to compensate for scarcity. It specifies the rate of development. In the case of France, this rate is 7.4% for the factor 7 goal and 5.5% for the factor 4 goal. By way of comparison, the Quinet Commission uses a rate of 7.2% between 2030 and 2040 and the averages of the simulations used by the IPCC corresponding to r=5.5% between 2030 and 2050 (see Table 3); in contrast, the underlying rates of TIMES, POLES, NEMESIS and ThreeME are significantly higher (between 12% and 13% for the first three and 16% for ThreeME), reflecting either a rule of equity less favourable to future generations, or a more optimistic view regarding the progressiveness of decarbonisation costs, or a combination of the two.

3.4. Towards a Concept of Climate Debt?

Once the social cost of carbon has been defined, it becomes possible to consider several monetary indicators to describe the climate situation, beginning with two "climate debt" indicators.

It is possible to start by examining, following a forward-looking approach, the costs to be paid in the future to return to the goal trajectory, i.e. the discounted cumulative sum of future climate spending needed to achieve the goal. In other words, this is the amount of financial resources that would need to be held in reserve to achieve the goal without having to drain future incomes. This is an important concept because it reflects the idea that every euro not spent on climate investment today will be passed on to future generations. Here we will use the term implicit climate debt to refer to this indicator.[18] This is, in effect, a forward-looking concept, similar to the concept of implicit liabilities used for other types of public spending, such as pensions, the discounted equivalent of the stream of future spending necessary to honour a commitment made. In the scenario of achieving carbon neutrality by 2050, with a rate

15. *Indeed,* $P_t^{MD,FR} = \Lambda_*^{MD,FR} Y_t^{*MD,FR} / E_t^{*MD,FR} = \Lambda_*^{MD,FR} / \sigma_t^{*MD,FR}$.
16. *The ratio is 613 g per $ which, assuming the dollar converts to €0.85, equates to 720 g per €.*
17. $\Gamma = F^{1/T} - 1$, *see above.*
18. *Implicit debt is a cumulative sum not to be confused with the annual mitigation effort. By way of analogy, if we differentiate between the annual mitigation effort as a percentage of GDP and the effort actually achieved today, we come closer to the notion of a tax gap or discounted funding gap, which reflects the amount as a percentage of GDP for the improvement of the structural balance that would have to be made in a sustainable manner to bring the public debt back to a sustainable trajectory.*

of preference for the present equal to the rate of growth of the economy, the implicit climate debt amounts to around 150% of 2018 GDP, and can be measured at first order simply by the number of years to achieve neutrality multiplied by the annual climate spending as a percentage of GDP.

Another approach for "climate debt" would follow a backward-looking approach, by measuring the costs not paid in the past, as these terms are often used to express the notion that the burden of past inaction is passed down to future generations. As this debt has neither a creditor nor a debtor, its definition is normative. Nevertheless, once a value of climate action is defined for the future, it is a natural candidate to be used to value insufficient past efforts. Consequently, we propose defining climate debt as the sum of past net emissions, valued at the current social price of carbon according to the "accounting approach".[19] This corresponds to the simple idea that, regardless of when the CO_2 was emitted, it contributes to climate disruption in the same way[20] and it must be valued at the same level. This concept can also be linked to the idea of debt that developed countries, "historical" polluters, would have accumulated towards to less developed countries, which remains a fundamental question when it comes to the repartition of decarbonisation efforts at global level. Climate debt since 1990 has been estimated at €3,475 billion, which is also close to 150% of GDP and represents around €50,000 per capita. This debt can then be projected and compared to a climate debt ceiling corresponding, for example, to the level making it possible to achieve the goal of +2°C (Figure VIII).

The development paths of these two indicators are linked. Indeed, each year the unpaid costs (resulting in positive gross emissions) will be added to the costs of returning to carbon neutrality. They may also be monitored year after year, based on an official carbon price set by the public authorities and updated each time the database of the national accounts changes. Indeed, the climate debt indicator can constitute a steering tool for public authorities from an equity/efficiency perspective: it makes it possible to measure both the deviation from the carbon neutrality goal and the fair distribution of the effort between generations, with an insufficient effort in one year having to be compensated for in the following year(s).

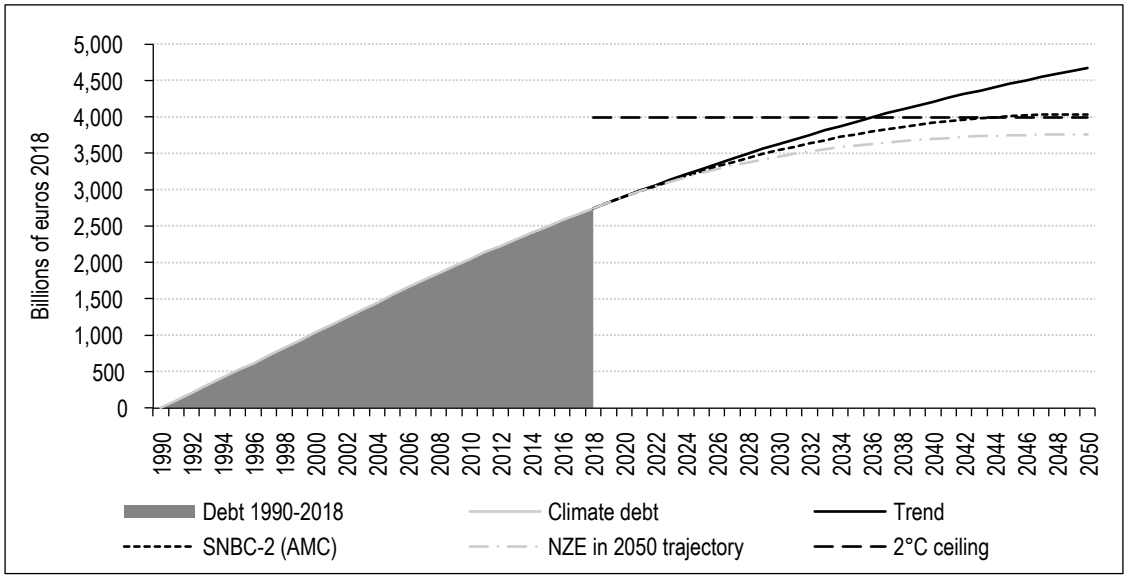
---

19. Formally, as at date t_0, its accumulated variation $\Delta D_{t0}^{\tilde{T}}$ is defined in relation to an initial date $\tilde{T}$ by $\Delta D_{t0}^{\tilde{T}} = \int_{\tilde{T}}^{t0} P_{t0}^{co2} \left(E_s - \bar{E}\right) ds$,

where E_s are the GHG emissions as at date s and $\bar{E}$ is the terrestrial and oceanic carbon sink. Financial climate debt can therefore be calculated simply as the accumulative sum of net emissions (physical "debt"), multiplied by the social value of carbon.

20. This amounts to disregarding the time taken for GHGs to disappear spontaneously, which is legitimate, as this is effectively a very long time in view of the time periods considered here.

Figure VIII – **Climate debt since 1990 (backward looking approach)**



Sources: CITEPA, authors' calculations.

4. A Net Savings Rate Adjusted for the Social Climate Cost That is Now Negative, a Sign of a World That Would Gradually Become Poorer

4.1. Beyond Climate: Broader Approaches to Sustainability

The answer to the question of whether the current emissions regime is compatible with meeting national commitments is, therefore, clearly no. We are far from the goal trajectory for greenhouse gas emissions. If we consider that failure to respect these commitments exposes us to major environmental risks, it can be said that we are consuming more natural resources than nature is capable of bearing. With the exception of a few climate sceptics, this assessment is widely shared: the notion of strong sustainability of economic development, which requires that each generation leave natural, physical and human capital at least equal, in each of these dimensions, to that which it has inherited, is not being fulfilled in respect of the environmental dimension.

Is it useful to supplement this message with indicators measuring what the literature describes as "weak" sustainability? This is what the indicators do when aggregating the developments of these different types of assets, leaving room for the idea that a decline in one type of asset could be offset by an increase in another. In reality, this is not the case for the climate, if the damage is irreversible, because then a marked deterioration of the environment cannot be offset by an accumulation of physical capital. It is this observation that calls for the selection of several sustainability indicators, isolating in particular those having a vital impact for mankind (pollution and global warming), as proposed in the Stiglitz report. However, this should not stop us from looking also at global sustainability indicators.

Enriching national accounts with such indicators is nevertheless a long-standing issue and there has been no shortage of proposals to do so. Conceptually, Hicks (1946) introduced the notion of real income, which he defined as the maximum consumption allowed without deteriorating the capital stock, which can therefore be interpreted as a concept of sustainable consumption. It was the Brundtland Commission (1987) that definitively placed it on the international agenda, defining it as the need to "satisfying the needs of the current generation without compromising the capacity to satisfy the needs of future generations". It was during this period that Cobb & Daly (1989) introduced an indicator of sustainable well-being, the ISEW (Indicator of Sustainable Economic Welfare) – also called green GDP – which includes the cost of environmental deterioration as well the issues of leisure and human capital. However, green GDP does not resolve the issue of global sustainability. To do so, "what we need", as the Stiglitz report pointed out, "is an evaluation of the distance between our current situation and the sustainable goals [...] in other words, we need indicators of over-consumption or under-investment" (Stigltiz *et al.*, 2009, p. 73), with both notions being understood in a broad sense.

The analytical framework linking inclusive wealth and adjusted net savings is best suited to solve this issue (see Online Appendix C3). Measuring sustainability in this way is precisely the objective that the World Bank has been pursuing since the 2000s by calculating an "adjusted" net savings indicator for most countries (World Bank, 2006, 2011, 2018). That work is based on the Hicksian idea that a sustainable trajectory, defined as a trajectory in which monetary well-being – comprehensive wealth – never decreases, is one in which the adjusted net savings are always positive. In concrete terms, the adjusted net savings (ANS) calculated by the World Bank can be written in the form GS - FCC + EDU - ENV, where GS is gross national savings, FCC is fixed capital consumption, EDU is education spending,[21] and ENV is the cost of environmental damage. Five factors are taken into account for the latter: the depletion of forest, oil and mining resources, global warming and air pollution.

The World Bank estimates global adjusted net savings at 10.7% in 2016, for gross savings of 25.9%. Environmental deterioration only accounts for -2.6% of GDP. Despite the apparent breadth of the spectrum of damage taken into account, the adjustments made by the World Bank to measure environmental deterioration are very small at global level. They are virtually imperceptible in the case of France.[22] In particular, the financial valuation of global warming is greatly underestimated, based on a social price

21. ANS, in contrast, does not take into account the depreciation of educational capital (which leads to an overestimation of educational savings in developed countries) or the quality of education.

22. For France, the ANS figure is 7.1%: gross savings is 20.3%, from which 17.7% is deducted for fixed capital consumption, equating to 2.2% of the net savings; education spending has a positive impact accounting for 4.9% of GDP, while the environment contributes negatively, with -0.4%.

of carbon of only €25.5 per tonne of CO_2.[23] This issue should therefore be re-examined here in light of the new evaluation of the social price of carbon that we have just reviewed.

4.2. Net Savings Adjusted for Climate Repair Costs

Here, we focus on climate issues. The data are taken from the World Bank's database for net savings EN_t and greenhouse gas emissions E_t; the carbon sink $\overline{E}$ is based on the SNBC-2 at national level,[24] and remains constant at 10 $MtCO_2eq$ at global level. With a view to simplify, the carbon price used – for France and at global level – is the average IPCC value for the objective of +1.5°C, "backcasted" for 2019, i.e. $180 per $MtCO_2eq$ or €153 per $MtCO_2eq$ (Figure IX).

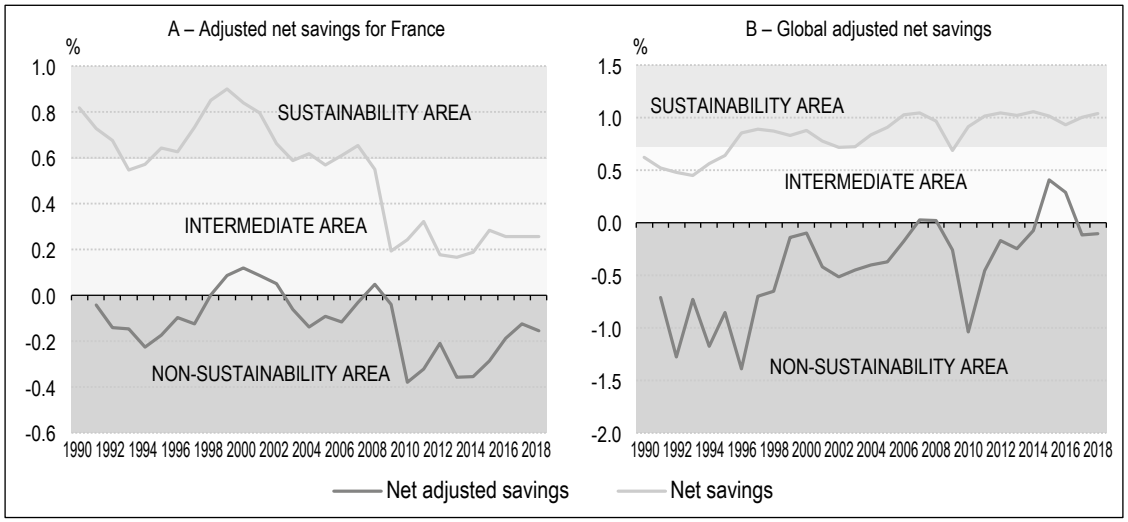
The adjusted net savings rate then appears negative at global level. Even if, after reaching a low point close to -13% in 1996, it has since recovered, mainly due to the rise of the Chinese economy, which has a high savings rate, it remains significantly negative on average over the last two decades. In France, the adjusted net savings rate has also been negative since the beginning of the 1990s. Contrary to the conclusions of the World Bank,[25] growth thus appears to be unsustainable, even in the so-called "weak" sense of the term, i.e. considering the substitutions between physical and natural capital. Not only are we using more resources than nature is capable of regenerating, but the wealth we leave behind does not compensate for the costs of repairing climate damage.

4.3. Inclusive Wealth, Integrating Climate Debt

The stock concept of inclusive wealth can be associated with this concept of flow. Inclusive wealth is defined as the sum of the different forms of capital weighted by the implicit price of each of them. Here we consider physical capital and climate "capital". The capital stock is created using a permanent inventory from 1975, i.e. by assuming a capital/output ratio of 2.8 in 1975. This calculation is performed using net savings data from the World Bank.

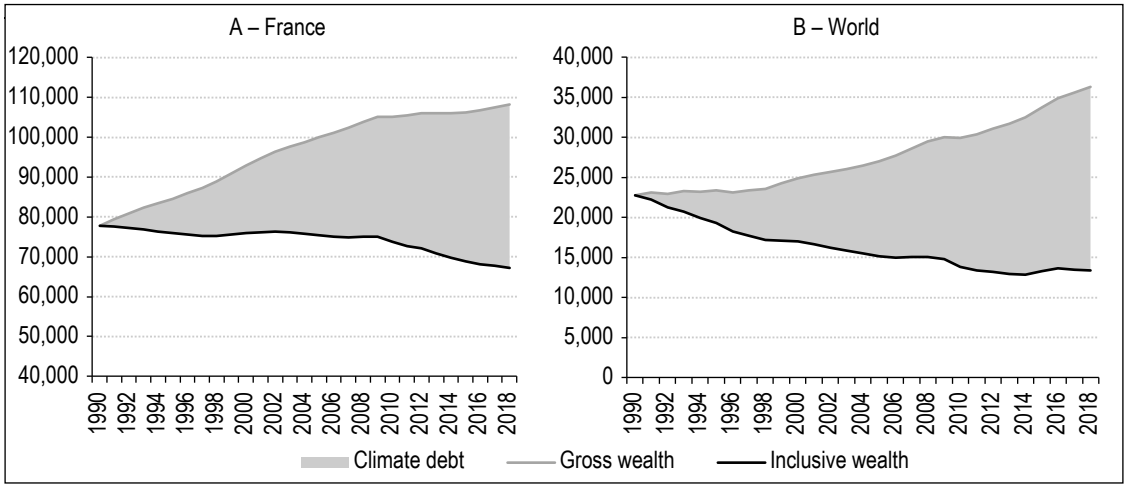
23. In the rest of the article, we will express all carbon "prices" in euros per tonne of CO_2. Carbon prices sometimes also refer to a price per tonne of carbon and not CO_2. The shift from the first to the second is done by multiplying by 0.275: as the atomic mass of carbon is 12 and that of oxygen is 16, there is 12/44 of a tonne of carbon in a tonne of CO_2. A price of €20 per tonne of carbon is therefore equivalent to a price of €5.5 per tonne of CO_2.
24. 40 $MtCO_2eq$ in 2020, rising slightly to 80 $MtCO_2eq$ in 2050.
25. The main reason for the discrepancy with the World Bank estimates is a difference in the assessment of the carbon value. The World Bank also takes into account, contrary to this article, the accumulation of human capital, which is valued at the level of public education spending. This choice may seem optimistic in the sense that it seems more appropriate to use permanent inventory methods, considerably reducing the impact on adjusted savings, particularly when the school-leaving age stops rising, as has been the case in France for the past two decades. Furthermore, it is more than likely that the positive valuation of the accumulation of human capital would be more than offset by the negative consideration of biodiversity loss - the other major environmental concern - without altering the message of the unsustainability of the current economic trajectory.

Figure IX – **National and global adjusted net savings (in % of GDP)**



Reading Note: The different areas apply to the adjusted net savings curve. The sustainability area corresponds, on average, to an area of simultaneous growth of natural and physical capital (strong and weak sustainability). In the non-sustainability area, on the one hand, natural capital is declining and, on the other, the repair costs are higher than the increase in income (strong and weak non-sustainability). In the intermediate area, adjusted net savings are positive but below the average net savings value – this means that natural capital is declining overall (weak but not strong sustainability).
Sources: World Bank Data, authors' calculations.

Figure X – **National and global wealth integrating climate debt (in euros per inhabitant)**



Sources: World Bank Data, authors' calculations.

The climate is taken into account from 1990 onwards. This is a normative choice consistent with the one we have made for the evaluation of climate debt and, of course, with the choice of the COPs since Kyoto to make it the reference for all the processes associated with them. It can also be considered that from that date, the fight against global warming became a social goal, and that continuing to emit more GHGs than the planet is able to absorb has become a debt for future generations.[26]

In 2019, readjusted inclusive wealth per capita amounts to approximately €57,500 in France and €13,175 at global level, corresponding to wealth (capital) of €109,000 for France and €31,450 at global level, respectively (Figure X). The difference represents the financial value of carbon debt since 1990. In both cases, wealth, extended to include natural resources (in this case, the climate), is in decline, which has been more pronounced in France since the 2008 crisis due to the drop in gross savings, with that drop being more attenuated at global level due to the emergence of China.

Inclusive wealth is constructed as the cumulative sum of adjusted net savings over time. Adjusted net savings measure instantaneous sustainability, which is interesting in itself. However, one year of negative net savings can be compensated the following year by a positive year; inclusive wealth takes into account developments in both the medium and long term. In our construction there is a very simple relationship between inclusive wealth, capital and financial climate debt (according to the backward looking approach), with the first being simply the difference between the other two. These latest results shed new light on the work that concluded that most rich countries were sustainable, based on these adjusted net savings and inclusive wealth indicators. A correctly calibrated carbon price indeed leads to the opposite conclusion.

* *
*

In this article, we have endeavoured to re-evaluate the issue of the climate sustainability of economic development, at national level for France and globally. On a theoretical level, considerable progress has been made since the late 1990s, with an important milestone around the work of the Stiglitz Commission (Stiglitz *et al*, 2009).

In the language of the theories of well-being and sustainability, COP 21 and its continental and national iterations have placed a social value on climate action. Societies now consider CO_2 emissions beyond terrestrial and oceanic absorption capacities as a cost for future generations. And this gives mitigation efforts a value, the value of climate action. The other essential factor for assessing climate sustainability is of a technico-financial nature. Translating the CO_2

26. Another option would be to go further back, to the moment when GHGs exceeded the capacity of the global carbon sink, i.e. in the 1950s to 1960s.

emission reduction goals into financial terms requires knowing the cost of the techniques and technologies in relation to their potential to decarbonise the economy. In this respect, too, the magnitudes are now beginning to be better established, contributing to the reliability of both macro-sectoral and technico-economic models.

Therefore, we have proposed a macroeconomic framework that makes it possible to evaluate the optimal GHG emission reduction trajectories with constraints in terms of intergenerational equity and the development of decarbonisation technologies. This dual movement to clarify the climate goal and technico-financial knowledge seems sufficient for us to be able to assign a reasonably reliable price to carbon. We have demonstrated that this value for France should be positioned, for the objective of carbone neutrality, at around de €120 to €250 today; €250 to €500 in 2030; €500 to €1,000 in 2040; €1,000 to €2,000 in 2050. These estimates are globally consistent, in terms of their order of magnitude, and when we analyse comparable concepts of social price of carbon, with the IPCC estimates or those established by the Quinet Commission (2019) and the models on which they are based: they constitute the high end of the range. To meet France's climate commitments, i.e. to achieve carbon neutrality by 2050, the climate spending effort should be increased to 4.5% of GDP each year. The global effort should be on a comparable scale (5.1% of global GDP).

Finally, these increased social values of carbon shed new light on the evaluation of the sustainability that we are accustomed to describing as weak, i.e. the – otherwise rightfullly controversial – issue of determining whether, despite everything, and thus despite the environmental damage, the balance for future generations would be "positive", given the continuous improvement of the average standard of living. The global net savings rate, adjusted to take account of climate damage, is negative over the entire period under review. Over the past three decades, the world is thought to have become poorer not richer, with the cost of human activities on the climate being thought to have outweighed the accumulation of both private and public capital. Inclusive wealth, aggregating natural and physical capital, is in decline. Even in the weak sense of the term, we are on a trajectory of unsustainability and, in reality, we have been on it for several decades, and only a change of scale in the economy's decarbonisation policies is likely to correct it.

We can only stress, for a definitive conclusion, the implications of recent advances and clarifications, both theoretical and empirical. Determining a value for the carbon price is a major issue for steering public policies, and the implementation of environmental economic accounting, would be likely to shed light on the public debate. However, such public accounting would also be useful for guiding individual choices on consumption, production and travel. One possibility would be to set a social value for climate action or a social cost of carbon by law,[27] which would also specify how the value is to be used, which could range from systematic labelling or inclusion in the business accounting standards to more binding measures, such as inclusion in public procurement contracts or setting a minimum price for CO_2 emissions allowance trading. □

27. We have distinguished the social value or carbon from the carbon tax. As a matter of fact, the carbon tax is only one of the possible levers of climate policies and raises significant questions of fiscal justice without establishing the behavioural effects, which would assume price elasticities to be established.

Link to the Online Appendices : https://www.insee.fr/en/statistiques/fichier/4770154/ES-517-518-519_Germain-Lellouch_Online_Appendices.pdf

BIBLIOGRAPHY

Arrow, K.J., Dasgupta, P., Goulder, L., Mumford, K. & Oleson, K. (2012). Sustainability and the Measurement of Wealth. *Environment and Development Economics*, 17, 317–353. https://doi.org/10.1017/S1355770X12000137

Baptiste-Perrissin, F. & Foussard, A. (2016). Trajectoires de transition bas carbone au moindre coût. Ministère de Environnement de l'Énergie et de la Mer. *Théma du SEEIDD*. https://www.ecologique-solidaire.gouv.fr/sites/default/files/Th%C3%A9ma%20-%20Trajectoires%20de%20transition%20bas%20carbone%20au%20moindre%20co%C3%BBt.pdf

Bruntland (1987). *Our common future*. World Commission on Environment and Development. New-York: Oxford University Press. https://sustainabledevelopment.un.org/content/documents/5987our-common-future.pdf

CGDD (2020). Chiffres clés du Climat : France, Europe et Monde. *Datalab.* https://www.statistiques.developpement-durable.gouv.fr/sites/default/files/2019-11/datalab-62-chiffres-cles-du-climat-france-europe-monde-edition2020-novembre2019_0.pdf

Cobb, J. & Daly, H. (1989). *For the Common Good: Redirecting the Economy toward Community, the Environment, and a Sustainable Future* (2nd updated edition). Boston: Beacon Press.

Easterlin, R. (1974). Does Economic Growth Improve the Human Lot? Some empirical evidence**.** *In:* David & Reder (Eds.), *Nations and Households in Economic Growth, Essays in Honor of Moses Abramovitz*, pp. 89–125. New York: Academic Press. https://doi.org/10.1016/B978-0-12-205050-3.50008-7

GIEC (2015). Changements climatiques 2014 – Rapport de synthèse. 5ème rapport du groupe d'experts intergouvernemental sur l'évolution du climat. https://www.ipcc.ch/site/assets/uploads/2018/02/SYR_AR5_FINAL_full_fr.pdf

Hicks, J. R. (1946). *Value and Capital: An Inquiry into Some Fundamental Principles of Economic Theory* (2nd Edition). Oxford: Clarendon Press.

Hotelling, H. (1931). The Economics of Exhaustible Resources. *Journal of Political Economy*, 39(2), 137–175. https://doi.org/10.1086/254195

I4CE (2019). Panorama des financements climat, édition 2018. Institute for Climate Economics. https://www.i4ce.org/wp-core/wp-content/uploads/2019/09/I4CE-Panorama2019.pdf

Nordhaus, W. D. (1977). Economic Growth and Climate: The Carbone Dioxide problem. *American Economic Review*, 67, 341–346. http://www.econ.yale.edu/~nordhaus/homepage/homepage/carbondixoideproblem.pdf

Quinet, A. (2019). La valeur de l'action pour le climat. Rapport de la commission présidée par Alain Quinet. Paris: France Stratégie. https://www.strategie.gouv.fr/sites/strategie.gouv.fr/files/atoms/files/fs-2019-rapport-la-valeur-de-laction-pour-le-climat_0.pdf

SCF (2018). Rapport du Standing Committee on Finance. Convention-cadre des Nations unies sur les changements climatiques. https://unfccc.int/fr/process/bodies/constituted-bodies/standing-committee-on-finance-scf/documents

Solow, R. (1956). A Contribution to the Theory of Economic Growth. *The Quarterly Journal of Economics*, 70(1). 65–94. https://doi.org/10.2307/1884513

Stiglitz, J. E., Sen, A. K. & Fitoussi, J.-P. (2009). Rapport de la Commission sur la mesure des performances économiques et du progrès social. https://www.vie-publique.fr/sites/default/files/rapport/pdf/094000427.pdf

Stiglitz, J. E. & Stern, N. (Dir.) (2017). Report of the High-Level Commission on Carbon Prices. https://static1.squarespace.com/static/54ff9c5ce4b0a53decccfb4c/t/59244eed17bffc0ac256cf16/1495551740633/CarbonPricing_Final_May29.pdf

UNU-IHDP. (2012). *Inclusive wealth report 2012: measuring progress toward sustainability*. Cambridge University Press.

World Bank (2006). *Where is the Wealth of Nations? Measuring Capital for the 21th Century*. Washington DC: World Bank. http://documents.worldbank.org/curated/en/287171468323724180/pdf/348550REVISED0101Official0use0ONLY1.pdf

World Bank (2011). *The Changing Wealth of Nations? Measuring Sustainable Development in the New Millennium*. Washington DC: World Bank.
http://documents.worldbank.org/curated/en/630181468339656734/pdf/588470PUB0Weal101public10BOX353816B.pdf

World Bank (2018). *The Changing Wealth of Nations: Building a Sustainable Future*. Washington DC: World Bank. https://openknowledge.worldbank.org/handle/10986/29001

A Comparison of Deflators for Telecommunications Services Output

Mo Abdirahman*, Diane Coyle, Richard Heys*** and Will Stewart******

Abstract – Data usage in the UK expanded by nearly 2,300% between 2010 and 2017, yet real Gross Value Added for the telecommunications services industry fell by 8% between 2010 and 2016, while the industry experienced one of the slowest rates of recorded productivity growth. The disconnect between rapid technological improvements and the measured economic performance of the industry can largely be explained by the deflators applied to nominal output. We contrast two methodologically distinct options: the first consists in strengthening the existing Services Producer Price Index for Telecommunication Services, the second in measuring price changes through the average price per unit of data for various telecommunication services. The key distinction between these options can be considered as contrasting a revenue weighted index with one that can be seen as a volume weighted index. Using these methods, we conclude that telecommunications services prices fell by between 37% and 96% from 2010 to 2017, considerably more than the current deflator. The real output of the sector will therefore have been considerably higher than indicated by current statistics.

JEL Classification: E01, L16, L96
Keywords: technological progress, telecommunications, deflators

* *Intellectual Property Office (UK) (mohamed.abdirahman@ipo.gov.uk);* ** *University of Cambridge and Economic Statistics Centre for Excellence (ESCOE) (dc700@cam.ac.uk);* *** *Office for National Statistics (UK) (richard.heys@ons.gov.uk);* **** *The Institution for Engineering and Technology (IET) (w.stewart@ieee.org)*

We thank the following IET working group members for their invaluable contributions: Ahmed Kotb, Hannah Conway, Stephanie Baxter and Jeremy Watson. Thanks also to Anna Ardanaz-Badia, Ash Loakes, Chris Payne, Emma Howley, Gaganan Awano, George Agbugba, Hannah Evans, James Scruton, John Jeremy, Helen Sands, Kat Pegler, Marilyn Thomas, Mark Stephens, Monique Sidhu, Sanjiv Mahajan, Robert Kent-Smith and Thomas Lewis – all with the ONS; to David Mark Harrison, Gary Clemo, Jonathan Porter and Max Fernando (Ofcom); Tim Miller and Tony Lavender (Plum Consulting); Jonathan Haskel and Peter Goodridge (Imperial College); Frances Cairncross (Oxford University), Marko Melolinna (Bank of England), Miles Elsden (Massive Dynamics), Nicolas Oulton (LSE/Centre for Macroeconomics), Daniele Viappiani (HMT), Nigel Brown (Cabinet Office), Peter Ladkin (Causalis); and finally to two anonymous reviewers.

Received on 27 June 2018, accepted after revisions on 9 April 2019.
Citation: Abdirahman, M., Coyle, D., Heys, R. & Stewart, W. (2020). A Comparison of Deflators for Telecommunications Services Output. *Economie et Statistique / Economics and Statistics*, 517-518-519, 103–122. https://doi.org/10.24187/ecostat.2020.517t.2017

Users of National Accounts data usually want to analyse data in real terms for purposes such as comparison through time. This requires the deflation of nominal values. Statistical offices calculate National Accounts deflators in compliance with international guidance, but there are well-known challenges in constructing deflators, in particular how to treat new goods entering the consumption basket, quality change which may change the price as well as the nature of the product, and products reaching 'corner solutions', such as where prices fall to zero, or where consumption at a given price is without limit. These particularly affect high-tech and digital products, as engineering progress has been rapid over the last twenty years, and big increases in usage have been accompanied by large declines in unit price. This paper explores these issues with respect to telecommunications services[1], as the industry clearly manifests these challenges.[2]

This paper considers the deflation of telecommunications services as currently used in the construction of the output measure of Gross Domestic Product (GDP).[3] Telecommunication services have experienced extremely rapid technological change in recent years, and the issues debated in the literature are particularly acute here. Both Office for National Statistics (ONS) and EU-KLEMS data suggests the telecommunications sector has seen one of the slowest rates of productivity growth in recent years, and yet to a telecommunications engineer this is at odds with the extremely rapid technological progress it has experienced. The sector has also experienced rapid demand growth observed in terms of the volume of data usage, but not total industry revenues. As demonstrated in ONS (2018), the telecommunications sector has gone from being one of the two fastest growing industries in the United Kingdom in terms of productivity in the pre-Great Recession period, to being one of the two seeing the biggest decline, recording negative productivity growth in 2012-2017. This has led some (including official reports such as Bean, 2016) to suggest the official deflators understate 'true' declines in the price of such products, and therefore that real economic growth may be understated.

Our contribution is to show that both a modest improvement in the current method for constructing the output deflator for the product and a more radical alternative method deliver an estimated price decline of between 37% and 96% over an eight-year period, compared with the 3% price increase shown by the current deflator. Our alternative improvements to the current price index for telecommunications services, taking account of broadband data services suggest that the real output of telecommunications services in the UK (and likely other countries too) will have been significantly understated in recent years.

Similar issues are likely to apply to several sectors where digital technologies have driven improvements in services, but they are dramatic in the case of telecommunications services. Recent years have demonstrated exponential growth in the quantity of data transmitted via telecommunications networks. Intuitively, this huge gain in data transmission performance at constant or declining cost should represent a significant gain in real output. This paper does not venture into the complexities of new digital goods, or boundary issues concerning where they are produced (see for example Coyle, 2017), but focuses on a simpler question: the measurement of telecommunication services output in real terms and what difference alternative approaches for calculating deflators would make.

We consider both an improvement of the current methodology and an alternative data usage driven approach. These provide wide estimated bounds, so we also consider the degree to which market structure and technological change in the sector may lead to convergence between the two methods over time. At present, the price per unit of data differs significantly between services; for instance, it is more expensive to convey the same amount of data via SMS message than an 'over-the-top' service such as *WhatsApp*. This may be a transitional phenomenon. Convergence over time in the price per unit of data charged currently for different communications services can be expected, primarily through competition between differently priced close substitutes: where customers are currently charged a different

1. Telecommunication services comprises four sub-categories in the International System of Industrial Classification of All Economic Activities (ISIC) 2008 system: Wired telecommunications activities (6110), Wireless telecommunications activities (6120), Satellite telecommunications activities (6130) and Other telecommunications activities (6190). Note, however, that the deflators we are comparing in this paper are product level deflators. They are therefore informative about price changes for the product telecommunications services, rather than price changes for the industry as a whole. Whilst most telecommunications services are produced by the same industry, some of that activity also takes place in other industries.

2. In 2016 the Office for National Statistics (ONS) joined with leading economists and engineers in the Institution of Engineering and Technology (IET) to review this issue. A previous ONS article (Heys & Awano, 2016) outlined some of the key conceptual issues in scope.

3. GDP can be calculated by the Output, Expenditure or Income approaches. To ensure that the three approaches yield the same estimate in practice, National Accountants use a balancing process.

price per unit of data this should ultimately lead to the lower cost substitute becoming dominant and winning market share, as long as there is enough competition in the market. Convergence would make a data usage based unit value index a more meaningful proxy deflator. We present evidence that such convergence is under way, although for now it would be too early to recommend a switch from the current (but improved) methodology to the aggregate unit value index we calculate.

The two options exemplify a key difference between the engineering and economic approaches: economists observe a variety of products with different prices and weights in a basket of goods, delivered via the means of data transmission; engineers observe the telecommunications service sector delivering a single product – data transmitted, which has a variety of uses in delivering different services – which has experienced a clear and substantial fall in cost per bit of data through time. Our first option presents a relatively cautious updating of the current deflator in line with current international norms and standards, notably adding important components to the basket of goods in scope. The second option starts from the engineering perspective that there is a single service – data – and thus considers a data usage driven approach by translating all services into a single measure of the volume of data and using the revenue per unit of data as the deflator.

The results are striking in either case. Both approaches suggest substantially faster price decline than the present deflator. We find that prices of telecommunications services have fallen by between 37% and 96%. This is significantly lower than the current deflator suggests and implies that the real growth of telecommunications services in the national accounts has been understated. We also present some potential amendments to our two approaches that may help narrow this range.

This paper is structured as follows. First, we set out the context. Secondly, we discuss the engineering issues in terms of the differences between the various telecommunication services and how to represent the output of all services in terms of bits transported. Then we present the methodology for calculating the current deflator, and the two alternative options; and we discuss their strengths and weaknesses. Finally, we discuss the results and some potential future improvements.

1. Context

The UK fixed line telecommunications market is concentrated[4], with BT and Virgin Media having a market share of around 53% in 2017. A number of smaller providers account for the remaining 47% but these usually use the BT (Openreach) network.

Fixed line telecommunications service contracts are often bundled contracts where customers usually purchase broadband with a phone line at the minimum. However, unlike mobile phone contracts, these fixed line contracts do not always have an inclusive allowance of voice calls. As a result, we find that the revenue weights of phone calls decline significantly, as data enabled applications have emerged as substitutes. The monthly contract fee also includes line rentals but these are no longer invoiced separately and just subsumed in the bundled price. Some bundles have also evolved to include non-telecoms services, such as TV packages. However, the revenues used in this analysis exclude all non-telecommunications services revenues. This ensures that our resulting telecommunications services deflator is not biased by the inclusion of non-related revenue.

The mobile telecommunications market is equally concentrated in the UK[5] with the largest two providers controlling around 56% of the market at the end of 2015 and the largest four operators controlling around 85%. The remaining 15% of the market was served by a number of smaller virtual network operators who use the networks of the larger operators. Mobile services contracts are provided on either a pre-pay or post-pay basis. Post-pay contracts are predominately provided on a bundled tariff basis, which contain a pre-determined allowance of calls, texts and data usage. Whilst pre-pay contracts are usually based on a usage basis, these increasingly give the option of purchasing monthly bundles of calls, texts and data.

The bundling of different telecoms services into the monthly price makes it difficult to observe true revenue weights for the different mobile services. This is because the mobile operators do not break down the bundled revenue into the different components. We therefore have to apply the strong assumption that the unit

4. https://www.ofcom.org.uk/__data/assets/pdf_file/0012/110154/Q3-2017-Telecoms-Data.pdf [Data in Table 2 on page 4; retrieved: 04 December 2018].
5. https://www.ofcom.org.uk/__data/assets/pdf_file/0026/26648/uk_telecoms.pdf [Figure 4.21 on page 154; retrieved: 04 December 2018].

prices for the different services in the bundle equal the out of bundle unit prices for these services. However, voice and text services are often offered on an unlimited basis and newer bundles therefore focus on increasing the data allowance. This in turn limits the share of mobile data in the out of bundle revenue, and distorts the calculations of our revenue weights. The resulting differentials in unit prices between the different mobile services do therefore not necessarily indicate substantial differences in consumer values between the different services, particularly the conventional voice and text services compared to the newer data services.

1.1. What are Telecommunications Services?

Users primarily perceive that they are buying digital products and services of many kinds, from movies to banking services, rather than buying their transportation per se. However, in engineering terms communications, whether traditional telephony, TV/video, banking or social/text networking, is essentially a bit-transport service. An analogy would be that the domestic user may use water to wash, clean, cook and a variety of other purposes, but the water supplier sees only the quantity of water being piped to each home, with charges being driven by the volume of water consumed and the fixed costs of the network. For ordinary physical products they would expect that any transportation necessary to cost an amount relating to specific characteristics such as the product's size and weight, rather than the intrinsic value of the product itself (with some exceptions). Data services in the UK are provided by data bits transmitted to consumers via either fibre or wireless connections. These are weightless and essentially non-physical, but otherwise the analogy remains.

The cost of a fibre network is typically dominated by the fixed costs of installation[6], which has not changed much in recent years. However, the data rate achieved on a single installed fibre has risen by some 10^{10} times (from 0.1 MBit/s to 1 Petabit/s) for 'champion' results[7] between 1960 and 2015. Similarly, the data rate for widely installed systems rose 10^6 times between 1980 and 2015 (from about 1Mbit/s to about 1Terabit/s).[8] These improvements each broadly equate to a fairly steady log growth gradient of 150% per annum or 5,000-6,000% per decade.[9] Although there has been some levelling off in the champion rates in recent years, these are considerably higher than the installed rates. This means that large further gains in the installed rates remain possible.

1.2. Measuring Price Change

The key question in this market is therefore how to conceptualise and measure the fundamental communications product, 'data', encapsulating broadband (fixed and mobile) data and all other telecommunications services (phone calls, text messages, etc.). The question concerns the appropriate volume units of measurement, taking into account quality change and hence the appropriate price deflator to apply to nominal output to permit volume estimation to occur.

This question sits within a family of similar recent questions concerning measurement of the digital economy. These however only re-open, in a particularly acute manner, older debates. Innovation is the defining characteristic of the digital economy, either in the form of new products and services, improved quality and variety, or new business models (such as digital platforms), and can be clearly observed in the changes described above in recent years in the telecommunications sector. Innovation in general has long posed a challenge to the construction of price indices, as elegantly summarised by Diewert (1998): "The basic problem is that traditional index number theory assumes that the set of commodities is fixed and unchanging from period to period, so that like can be compared to like."

Considerable attention has therefore been paid to how innovation should be treated in price indices, and the extent to which this diverges from normal practice in statistical offices.

The naïve approach is to use a unit value index, calculated using total revenue and total volume for a particular service. Unit value indices are both dependent on the choice of units deployed, and need the goods to be broadly homogenous as otherwise the price series might be biased. This is because the unit price captures both price and quantity changes. Only

---

6. Meaning civil engineering (construction) for the most part.
7. Champion results are those achieved in best case experimental systems. See Ellis et al. *(2016).*
8. These gains in volume for similar or falling cost should deliver equivalent gains in productivity. Indeed if we were producing bags of sugar instead of digital bits it would. Today's annual sugar consumption in the UK would, if spread evenly across the national surface area amount to barely more than the thickness of an oil film on water (4 microns or about 1/30 the diameter of a human hair or optical fibre). However, if the gains since 1980 in installed fibre systems were applied to sugar the UK would now be covered by an extra depth of four metres of sugar each year.
9. Interestingly, similar to Moore's Law.

if the products are completely homogeneous, and a shift in consumption therefore occurs for some reason other than substitution for product characteristics, is there no bias.[10] Statistical offices sometimes use unit value indices for pragmatic reasons but economic theory favours other methods to generate required indices. The traditional Laspeyres index is one such, and answers the question: How much would a given consumer with given preferences need today to make her as well off as she was yesterday still consuming yesterday's basket of goods? It therefore forms an upper bound because it does not take into account consumer substitution when the relative prices of goods change.[11]

From the perspective of economic theory, the price index should preferably answer a subtly different question: How would a hypothetical consumer evaluate the two different sets of prices and goods? What is the compensating variation that keeps the consumer on the same indifference curve, given price changes and substitutions? For instance, suppose a laptop cost £1,000 in both 2012 and 2017 but the 2017 laptop had much better performance characteristics such as speed and memory. It is possible that a given consumer would be equally satisfied in 2012 and 2017, given what is available on the market and her (socially influenced) expectations (and hence the intuitive appeal of unit value comparisons). However, to reflect the real growth through innovation, the price ought to record a decline; there has been an increase in value received as consumer surplus. Hence economists prefer a superlative index such as the Fisher index, which approximates the theoretical cost of living index that keeps consumers' utility constant. However, such superlative indices require expenditure data for the current period that is usually unavailable when price indices are being calculated. The Laspeyres (or Lowe[12]) index is therefore typically used in practice (either with fixed weights or annually updated weights).

Given standard practice, there are several ways of reducing the potential bias from new goods and quality change, employed to differing degrees by statistical offices, particularly after the Boskin Commission Report (1996). One is to update the index weights frequently. Another is to introduce new goods into price indices more swiftly than had previously been the practice, to better capture the rapid price declines that often occur in the early years of the product lifecycle. A third, often seen as the gold-standard solution to the problem of adjusting for rapid quality change, is hedonic adjustment based on regressions on definable characteristics, in order to link prices per unit "to a yardstick more nearly relevant to its intrinsic utility".[13] For instance, hedonic regressions for computer prices might include processor speed, RAM, hard drive capacity, screen resolution, presence of a built-in camera and so on. In effect, products become bundles of more fundamental characteristics, allowing comparison of the price of comparable bundles of these characteristics. However, hedonic adjustment is typically applied to a few goods experiencing rapid change in their quality or characteristics, accounting for a small proportion of the consumption basket (0.39% in the UK[14]), in part because of the significant data requirements. To be a solution to the bias, hedonic adjustment also requires the assumption that the price contribution of different components equals their marginal contribution to consumers' valuation of the product.

There is an extensive literature on both the new goods problem and the hedonic approach. On the topic of new goods, the introduction of broadband as a product has attracted noticeable interest. The common approach in these studies is to evaluate quality-adjusted prices using hedonic regressions (Griliches, 1961). Williams (2008) considers internet access prices in the United States for the period December 2004 to January 2007. The study uses 135 price quotes from the BLS' CPI database and constructs hedonic functions where the main quality characteristic is bandwidth. Williams finds that quality adjusting the internet access price index makes little difference. Greenstein & McDevitt (2010) use a sample of over 1,500 price quotes for the period 2004 to 2009 obtained from a private consultancy. They use this to construct a hedonic model where the main quality characteristic is the download and upload speed. They find that quality adjusted prices fell by around 3%-10% in the period. This was a steeper decline than the official measure but still much smaller than the quality-adjusted price changes for other products such as computers.

10. Equally, there is not really an index number problem in that case.
11. Conversely, the Paasche will form a lower bound, looking back from today's basket of goods.
12. The Lowe will exceed the Laspeyres in a period when there are long-term trends in relative prices and consumers are substituting to lower priced items.
13. Adelman & Griliches (1961).
14. This figure relates to the Consumer Price Index.

However, hedonic studies have limitations, which is why in this instance we have not followed this approach. Hausman (2003) discusses some limitations of hedonic regressions in general. He argues that prices in imperfectly competitive markets are determined by demand, cost and the degree of competition in the market, and that hedonic regressions often fail to separate out these factors. In addition, even in the case where a hedonic regression might be acceptable, Hausman argues that it is difficult to identify all the product characteristics that are needed. This is especially relevant where the product characteristics are changing rapidly.

More fundamental, in terms of practical application, there is also a question about the completeness of product characteristics used in the hedonic regression. Bandwidth and upload/download speeds, while important, are not individually sufficient to explain price and quality changes of broadband. Other factors such as data caps, speed limitations ('throttling') at peak times, latency (the degree of time delay between the person transmitting and the person receiving), and geographical coverage are important quality considerations of the broadband service itself. In addition, even the bandwidth needs to be treated carefully as there is a difference between advertised and actual bandwidth. Advertised speeds can remain static whilst actual download and upload speeds improve, and vice versa. Furthermore, actual bandwidth cannot be captured in hedonic functions, as the actual speeds cannot be observed on an individual service contract level. These shortcomings of the hedonic approach can be overcome by the unit value approach under certain circumstances which could apply to the telecoms services industry.

It is also difficult to construct representative baskets of broadband service contracts, given the complexity of pricing in the industry and the wide range of available tariffs and options available and their dynamic nature. The use of a basket of goods approach in constructing a price index is therefore questionable in this case.

1.3. Alternative Methods to Deal with New Goods and Quality Change

One of the results of the rapid technological change in the telecoms services industry is that the volume weights for the different services differ significantly from their respective revenue weights. For example, while data services are weighted very highly in volume (as measured by bits for all services), the weight of data services in revenue is much lower. A similar problem is observable in the price of drugs. When generic versions of a drug enter the market, the price index is hardly affected, even though the price of generic drugs is much lower (Griliches, 1994). This is because the price index usually uses revenue weights. The incumbents often maintain a large share in the revenue while generics account for the bulk of volume.[15] Griliches & Cockburn (1993) note that the revenue-weighted official prices are a poor measure of the prices actually paid for goods which consumers regard as near-perfect substitutes, but the index treats as distinct goods even as the consumer substitution occurs over time. In the standard index for price change between periods 1 and 0, the revenue weight used for the old good is:

$$\frac{Q^1_{old}.P^1_{old}}{Q^1_{old}.P^0_{old} + Q^1_{new}.P^0_{new}}$$

This amount by which this will overstate the contribution of the old good is related to the change in quantity purchased of the new good and its average reservation price, as the 'true' weight is:

$$\frac{Q^1_{old}.P^1_{old}}{Q^1_{old}.P^0_{old} + Q^1_{new}.P^0_{new} + \left(Q^1_{new} - Q^0_{new}\right).p^r}$$

where p^r is the average reservation price for the new good. Clearly as the quantity substitution by consumers from old to new occurs, the revenue weight on the old good declines and the problem eventually disappears. However, even in the approach proposed in Griliches & Cockburn (1993) does not disregard revenue weights. Instead, the data usage approach in this paper is more closely aligned with Nordhaus (1994, 2007).

Price indices, even hedonically adjusted, will anyway fail to capture the consumer surplus due to the introduction of a new good into the market. Feldstein (2017) argues that the failure to consider new products and their impact on consumer value is an even greater source of bias than the failure to account for quality changes. It is difficult to time the inclusion of new goods in a price index and estimate the impact on consumer value using conventional methods. In theory, and in practice in a few instances, it is possible to estimate the demand curve and hence the reservation price at which demand is zero,

15. Although a key question is why the incumbent products are able to maintain this price differential; is this because of some unobserved characteristic or because of a poorly functioning market where consumers are not reacting fully to new price signals.

when the good is first introduced (Hicks, 1940; Hausman, 1996, 2003). Hausman also shows this reservation price can be approximated using an estimate of the own-price elasticity of demand. This approach requires current expenditure data, and imposes significant data requirements.

An alternative approach is to measure the cost of the service characteristic directly. This approach has been applied to lighting (lumen hours) and computer processing (computations per second) by Nordhaus (1994, 2007), who constructed long run series of directly observed engineering measures of performance and estimated corresponding supply costs per unit of light or computation. To the extent that mark-ups remain constant, changes in prices charged should be closely linked to cost changes. By measuring the price of the fundamental service characteristic (light or computations), instead of measuring the price of the goods delivering the characteristic, this approach should capture quality changes and the value of new goods, as long as mark-ups do not change much for reasons such as varying degrees of competition, for instance. The analogy in our case would be the engineering costs of transmitting a unit of data. However, it is usually much more difficult to collect the costs of such supply-side characteristics over time rather than market prices of goods, particularly for complex network services such as communications.

The alternatives to the hedonic approach also indicate substantial upward bias in conventional price indices. However, both involve painstaking statistical and econometric work and are not practicable for the regular calculation of official price indices. A key question we consider here is whether a reliable service characteristic – bits of data transported – can be measured in a way which is both conceptually useful and relatively easy to construct. However, there seems to be no completely satisfactory practical solution to the potential upward bias in price indices in the case of goods and services where there is significant innovation.

This issue remains a live one: see for example Bean (2016), and work in the US such as Byrne & Corrado (2017) and Groshen *et al.* (2017). Ahmad *et al.* (2017) attempt to gauge the scale of the problem by applying different countries' deflators to other countries to see if the magnitude of the resultant volume change is large enough to merit further work. They find that the impacts are relatively small. The weakness of this approach is that comparing a variety of upwardly biased deflators may not expose any commonly shared bias from a more correctly specified deflator.

1.4. New Methodological Challenges in Telecommunications

Importantly, in telecommunications services technological change means there is convergence between services both from a network perspective and from the perspective of users. For example, voice calling (once called telephony) is still distinct in terms of how it is handled and charged for by the network (and also, mostly, by regulators), but from a user perspective it is increasingly equivalent to services like *Skype* and *WhatsApp* that provide voice calls on the 'data' network, which is subject to a different pricing regime. The same is true of texting; indeed the word once meant SMS but now covers any of a wide range of text-chat services that in fact use the data network, but have the same (or better) functionality for the user. This means there are significant price differences for similar services, particularly when converted into price per data bit. There can still be major cost differentials between similar bit rates carried on different network services and at different ranges.[16] It is likely that the kind of service people use on their devices, fixed or mobile, will continue to shift rapidly in ways that are generally hard to predict.

This therefore leads to some key questions for our construction below of an index based on units of data:

- How long will different products (telephony, texting, data usage), all of which are essentially end-presentations of the same product (data), continue to be regarded as different services by users?

- How long will price differentials exist for these products?

- As cheaper substitutes become available, how long will providers continue to supply these services in the old mode; in other words,

16. Use of the data network is generally cheaper and normally distance-price-insensitive. There can be other differences that are important to the user such as the use of encryption and the blending with video and picture transmission, but the overall effect is to make all services look like bit transport from a network perspective. The phone network has clear guidelines on the maximum latency allowed, to avoid the sort of difficulty that makes voice 'calls' using geostationary satellites as often seen on TV so unsatisfactory. Data network based voice calling services like Skype *once had similar problems, but overall improvements in networks have largely solved these to the extent that broadcasters sometimes prefer them to traditional telephones.*

how long will telephony providers deliver telephony distinct from data rather than port across to using a IP protocol technology delivering the same user service using less data and at lower cost?

- Is it therefore appropriate or not to consider, for example, *Skype* and telephony as substitutes?

Boiling these down, therefore, presents a new challenge to price indices methods, namely, what happens when, rather than an old good being replaced by a new good, multiple old goods converge into a single new good? For example, if *Skype* and telephony increasingly converge which price does one take as the relevant one for the base period, or indeed does one weight these together? If one does, should this be weighted by revenue or volume? The following section considers both of these in the options presented. Option A is an improved Services Producer Price Index (SPPI), using the same methodology as at present (which employs unit value indices), and uses revenue weights. Option B is a unit value index based on data usage. In assuming perfect substitutability, this latter data usage approach is fundamentally driven by volume weights and would in theory reflect pure cost-based changes. Given the caveats about this assumption, discussed above, it should be interpreted as a downwardly biased estimate of the change in prices that would keep consumer utility constant.

Our two options can be considered respectively as upper and lower bounds to some ideal constant utility index, perhaps a hedonically-adjusted superlative index. Before exploring these new methods, however, we first discuss the current method in the UK for constructing the output deflator for telecommunication services.

1.5. Current Method

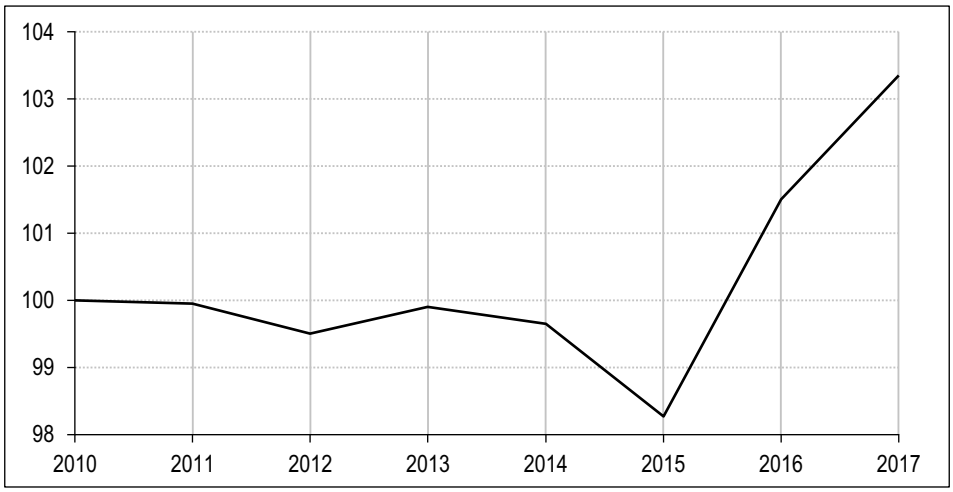
In the UK, the ONS currently deflates telecommunications services output at the domestic aggregate level[17] using an index which comprises two components: the product level index of the Consumer Price Index (CPI) covering Telecommunications Services and Equipment; and the product level index of the Services Producer Price Index (SPPI) covering Telecommunications Services. These are weighted around two-thirds CPI and one-third SPPI in the current deflator.

Between 2010 and 2017, the overall product deflator for telecommunications services in the UK has increased by around 3% (Figure I), despite substantial technological advances in that period (such as the shift from 3G to 4G technology).

An explanation for the trend in the overall deflator can be found by looking at the trend in its two component indices (Figure II). While the SPPI shows a general downward trend, the CPI declines until around 2008 and then rises. Since CPI is more heavily weighted in the output deflator, this has driven the composite deflator

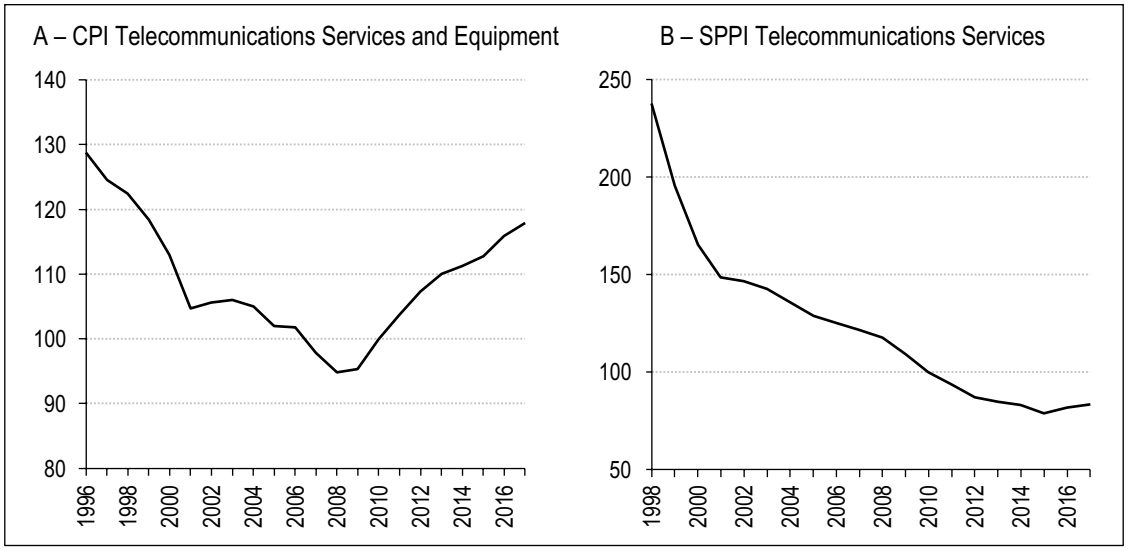
17. *Import and exports are treated separately.*

Figure I – **UK telecommunications deflator**



Note: 2010=100. Current product deflator for telecommunications services (CPA 61) in the UK.
Sources: ONS.

Figure II – **Components of GDP(output) deflator in the UK**



Note: 2010=100. The CPI Telecommunications Services and Equipment weight in deflator is 66%; the SPPI Telecommunications Services weight in deflator is 34%.
Sources: ONS.

(Figure I) to be broadly flat and then began to rise after 2015.

While this approach meets international standards, it is a methodology borne out of pragmatic decisions needed to deliver an appropriate deflator for the sale of telecommunications services to both businesses and consumers in the UK. These are:

1) The shares of the CPI (reflecting business-to-consumer sales) and SPPI (reflecting business-to-business sales) reflect broad usage patterns in the UK economy, but may not be reflective at the product level. For example, the shares of business and consumer usage may differ for different call types, messaging services and data usage.

2) The inclusion of the CPI is necessary because the SPPI conceptually captures only business-to-business transactions and therefore excludes consumer sales. However, output should be deflated in basic prices, and whilst the CPI reflects business-to-consumer transactions, it does so on a purchaser's[18] (rather than basic[19]) prices basis. This does not strictly map to the price of interest, the basic price of telecommunications services output before logistics, retail and margins.

3) The CPI product level index is a Laspeyres type index which captures both telecommunications services and equipment goods, despite the product group to be deflated including only services. The CPI and the product group that is deflated are also classified using different systems that do not easily map to each other.[20] This pragmatic compromise may introduce potential biases.

4) Whilst many of the CPI item level[21] indices are constructed using the traditional 'basket of goods' approach, a notable exception to this is the item level index for mobile phone charges, which includes Pay As You Go and contract charges. Due to the complex pricing structures and range of tariffs in the market, it is difficult to construct a representative basket of tariffs. Instead, this item is constructed using a "basket of consumers" approach recommended by Eurostat.[22] The ONS obtains representative consumer usage profiles from the UK's telecommunications regulator, the Office of Communications (Ofcom). For each consumer profile, the ONS identifies the price for the cheapest available tariff from the main service

18. That is, after non-deductible taxes, subsidies and relevant wholesale and retail margins and separately invoiced insurance and transport charges.
19. Also referred to as 'factory-gate' prices. That is the price before taxes, subsidies, margins and transport costs.
20. The CPI is based on the Classification of Individual Consumption According to Purpose (COICOP) while the national accounts product classification is based on the Classification of Products by Activity (CPA). The SPPI classification is based on CPA.
21. Item level indices are below product levels indices. For example, the item level index for Smartphones would form part of the product level index for Telecommunications Services and Equipment.
22. http://ec.europa.eu/eurostat/documents/272892/7048317/HICP+recommendation+on+telecoms+-+June+2015

providers. These are then weighted together using expenditure shares which are also supplied by Ofcom.[23] This approach has problems, particularly when quality change of more expensive contracts needs to be taken into account. The cheapest tariff is often based on old technology while the price of the new technology declines and the old technology is phased out. In this case, significant price movements in tariffs based on new technologies are missed, even if most people are using the new technology.[24] Likewise, other quality aspects such as coverage would also be omitted since these cannot be determined on an individual tariff basis as they depend on network and geographical region. As a result, actual quality changes might not be reflected in the price index, even when using hedonic methods.

5) With the exception of smartphones, none of the item level indices in the CPI: Telecommunications Equipment and Services are hedonically adjusted to control for quality change within the twelve month life of the basket of goods before new products are selected. In a fast-moving sector where contract design can change significantly and quickly this is a major weakness.

6) There are methodological differences in the way ONS constructs the product level CPI and SPPI, as well as differences in the construction of item level indices within the CPI. While the CPI: Telecommunication Services and Equipment is constructed as a price index, the SPPI: Telecoms Services is a unit value index. The ONS obtains administrative data sets from Ofcom. This includes volume and revenue of calls (by type) and text messages. A unit value (or average price) is then calculated for each item and aggregated up, based on revenue weights. The data for fixed line telecommunications only captures business telephony but the mobile data captures the entire market. Since the SPPI at present only attempts to cover business-to-business transactions, an assumption is made about the proportion of the total mobile phone revenue that is due to business use.

7) The SPPI has not been kept fully up to date with the pace of change in the sector. A notable absence from the SPPI is mobile and broadband data.

2. Alternative Deflators

Irrespective of the two options we present in this paper, the ONS is committed to reviewing and updating the current deflator, not only stimulated by the work described here and the digital economy agenda, but also by mandated changes through the implementation of the European Union's Framework Regulation Integrating Business Statistics (FRIBS). The FRIBS agenda requires expanding the scope of the SPPI to cover business-to-all transactions, not just business-to-business. This suggests that the ONS, alongside the two options presented below has a *de minimis* alternative of moving to exclusively using a business-to-all SPPI and dropping the CPI component from the output deflator. This would resolve issues 1-6, but would still leave issue 7 unresolved, which would be unsatisfactory.

2.1. Option A: An Improved SPPI

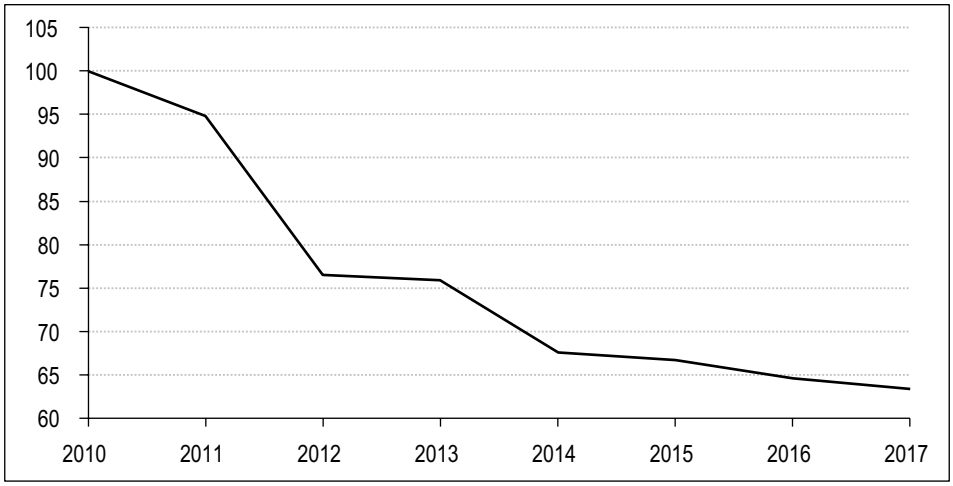
The current SPPI treats voice and text as distinct services, and does not include data services. Adding data into the basket presents one immediate route for improving this deflator and meeting issue 7. Therefore, under this option broadband and mobile data are added to voice and text in the current SPPI. To reflect the potentially large difference in consumer values, we construct granular unit value indices and aggregate them together using revenue weights. This is largely based on the current SPPI but with major differences: the index includes mobile and broadband data, uses a business-to-all transactions basis, and is annually chain linked. Removing the CPI component from the deflator and using the improved SPPI produces an index showing that telecommunications services prices have declined by around 37% between 2010 and 2017 (Figure III).

This method presents key benefits, as it is readily comparable to other deflators and represents a cautious improvement to the existing methodological framework. By constructing granular item level indices and aggregating them up, this method also accommodates the possibility that the different telecoms services remain heterogeneous products rather than perfect substitutes. However, the key weakness of this deflator is that it does not reflect the significant technical and quality improvements in the industry from

23. For details, see the CPI Technical Guide (page 58-60): https://www.ons.gov.uk/ons/guide-method/user-guidance/prices/cpi-and-rpi/cpi-technical-manual/consumer-price-indices-technical-manual--2014.pdf

24. It should be noted that even when a representative basket of tariffs can be constructed, hedonic adjustments would still raise some issues. For example, the headline speed for a tariff (which is often used in the hedonic adjustment) might remain constant while actual achieved speed increases (or indeed decreases, for example due to increased contention).

Figure III – **Improved SPPI deflator**



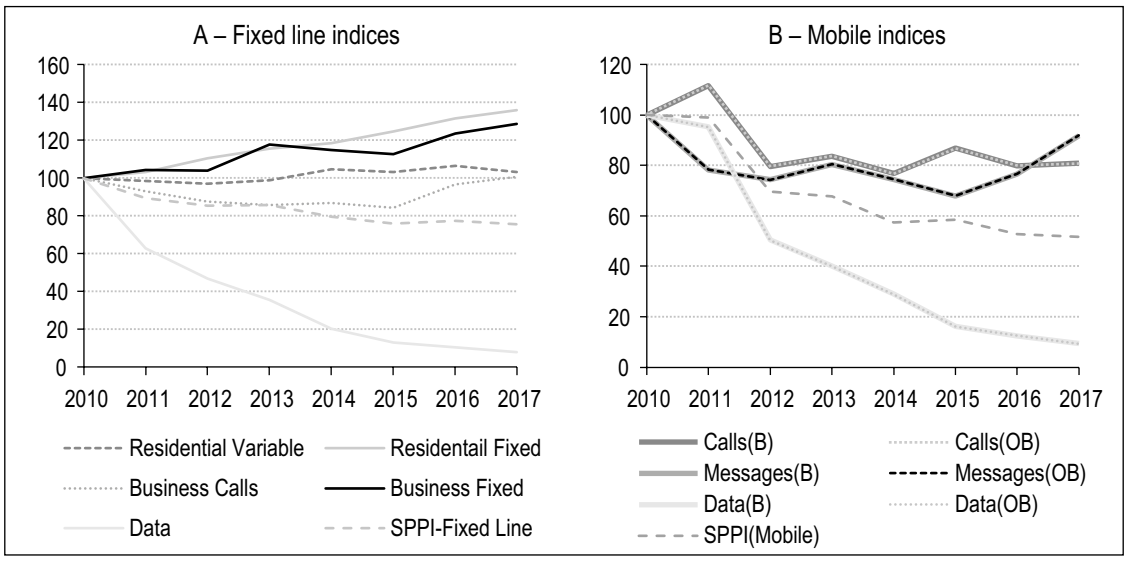
Note: 2010=100.
Sources: Authors' calculations.

an engineering perspective. This is because the deflator uses revenue weights which results in data services having a limited impact on the overall index movement. This is at odds with the engineering perspective, which regards data services as driving technical progress in the industry.

The breakdown of this deflator into the item level indices shows a significant difference in the price movement of the data elements and the voice and texts indices for both fixed line and mobile services (Figure IV). The data items thereby show substantial price decreases but are lower weighted and thus only have a limited impact on this overall SPPI index.

A particular challenge is the treatment of fixed line access charges. While the revenue from voice, texts and data can be divided by the volume of minutes, texts and bits, the denominator to construct unit values for access charges is the number of subscribers as this is the closest to a quantity measure for the access charges. As a result, the item indices for access charges show an increase in prices but different patterns for residential and business subscribers. For residential subscribers, the revenue of line

Figure IV – **Breakdown of improved SPPI deflator**



Note: 2010=100. In the breakdown of the Mobile Index, OB refers to out-of-bundle charges and B refers to bundled charges. The two indices are thereby the same as we assume that bundled and out-of-bundle charges are the same.
Sources: Authors' calculations.

rental has increased much faster than the number of subscribers. For businesses, the number of subscribers declined substantially but the corresponding revenue decline from access charges was less pronounced.

While access charges and the treatment of bundled items are areas that warrant further attention (see Online Appendix C2 for technical details; link to the Online appendices at the end of the article), a general feature of option A is that compared to option B below it places a lower weight on the contributions of broadband and mobile data. This is due to the impact of substantial price differences between the different services through the revenue weights; access, voice and text charges currently contribute a higher share of telecoms revenue. A raw increase in data consumption therefore has a limited impact on the Option A deflator, whereas substitution away from voice and text services toward data-driven alternatives such as *Skype* and *WhatsApp* manifest as a price increase.

2.2. Option B: Data Usage Approach

An alternative approach is to incorporate the engineering perspective on the industry's output, seeing the primary service of the industry as the transfer of data, and as such converting different services into comparable measures of units of data, that is bits or bytes[25], used to deliver the service itself. From a network perspective, there is little difference between a voice call and, say, a *Skype* or *WhatsApp* call, beyond the differences in bit/s that they use. We have accessed sectoral expertise to identify the factors to convert voice and text services into generic data services, using a number of simplifying assumptions[26] (Table):

- for text, we ignore shorter/longer messages and 'emoticons' for simplicity and assume all texts are 140 characters long, although many modern text systems will use more characters;

- a traditional voice call can reduce the data rate to a 'holding' level if both ends happen to be silent, and many systems exploit the relative tendency for both ends not to be speaking together, but we do not adjust for this;

- similar arguments apply to picture and video compression, which will depend upon the characteristics of the particular images involved, and will also likely change over time with technical developments.

Whilst for most services the total number of bits moved within the service period is the dominant consideration, other characteristics also matter. For example, latency (the total end-to-end transmission delay) is important in voice calls and some other services, as is coverage – i.e. whether or not you are in range of a transmission point. However, in most cases, these considerations are modest compared to the basic cost-per-bit-moved. This can be seen, for example, through the frequent use of satellite systems with extremely long latency. We do not therefore consider other characteristics, besides the cost-per-bit-moved, at present. Other traditional cost factors, such as transport range, are much less significant in modern digital communications.[27]

This conversion of voice, texts and data services into a common volume measure (petabytes of data) reveals that broadband and mobile data account for the vast majority of volume. It also shows that output, as measured by data transmitted, has increased 2,300% between 2010 and 2017, primarily driven by the increase in broadband and mobile data volumes. The volume of voice calls and text messages has been decreasing since 2010. This is either due to a drop in demand or, more likely, due to a substitution away from traditional telephony toward data-driven applications.

25. One byte equals eight bits.
26. Differences due to these simplifications are modest compared to the scales involved.
27. Although this was always true to an extent disguised by the relative pricing of, for example, international telephone calls.

Table – **Data conversions**

Medium	Bytes / kBytes rate	Other factors	Aggregate Bytes/ kBytes required
Voice	32 kBit/s each way	×2 for a two-way call /8 to convert kBits to kBytes ×60 to convert seconds to minutes	480 kBytes per minute
Text	1 byte/character	×140 as maximum of 140 characters per text.	140 Bytes per text

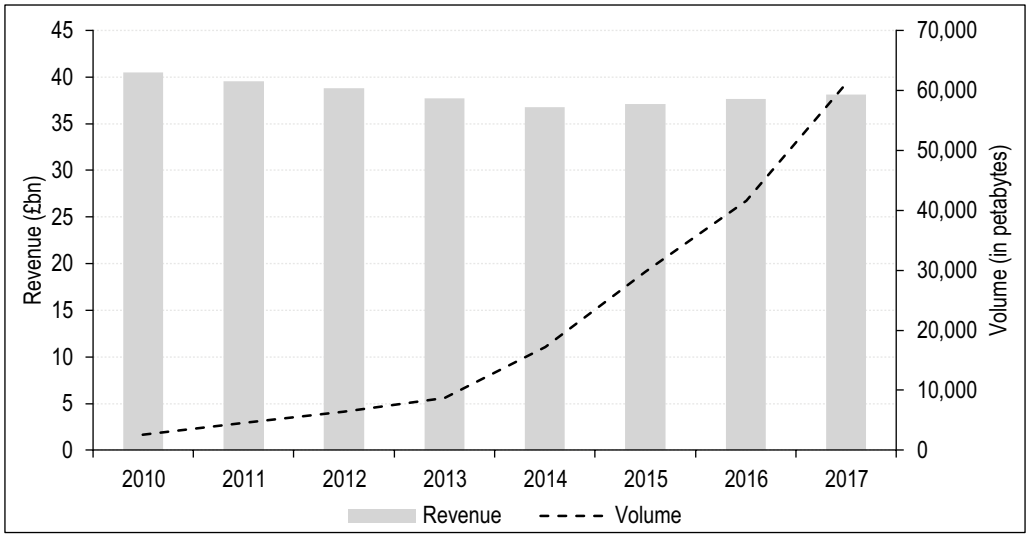
Notes: Authors' assumptions.

By 2017, around 99.8% of total volume was estimated to be broadband and mobile data. This is in stark contrast to the revenue weights, where broadband and mobile data account for around 40% of the total in 2017. In contrast to the exponential increase in volumes, total industry revenue fell by around 6% between 2010 and 2017 (Figure V). This is mainly due to a 47% decrease in wholesale revenues. Retail revenues increased by around 9% in the same period.[28]

Option B uses an aggregate unit value, which divides total revenue[29] in the industry by the total data volume.[30] This unit value index represents the average price per bit transported. Between 2010 and 2017 this measure suggests that telecommunications services prices decreased by around 96% (Figure VI). The increase in data volume, with revenue broadly staying flat, is seen as a volume increase and a price decrease. Likewise, a substitution away from more expensive voice calls and text messages towards cheaper services such as *Skype* and *WhatsApp* is also seen as a volume increase and a price decrease.

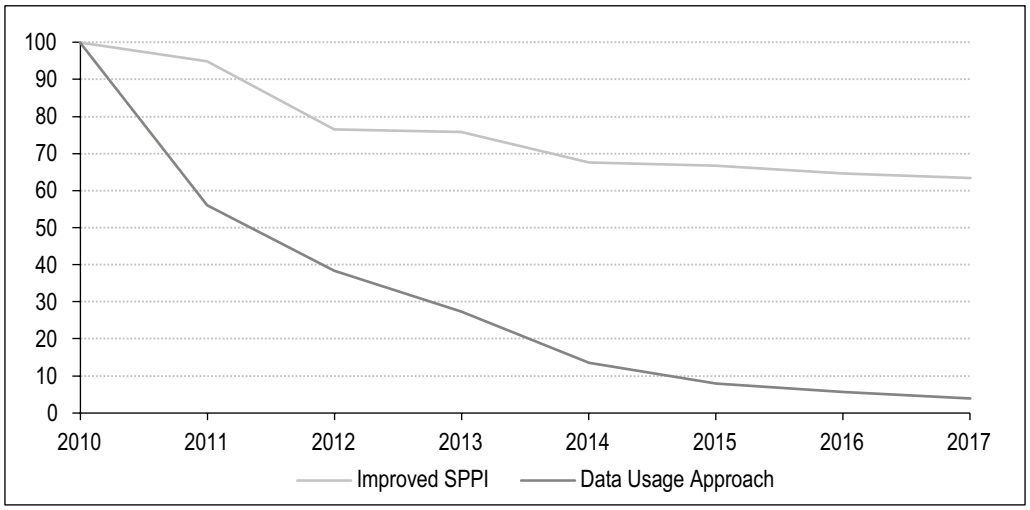
28. See Appendix A1 for details.
29. The total revenue figures exclude non-communications revenue such as TV bundles.
30. See Online Appendix C1 for details. The total volume excludes wholesale and corporate volumes. This does not impact on the main results; see Appendix A2 for details.

Figure V – **Revenue and volume in telecommunications services industry**



Sources: Ofcom and Authors' own calculations

Figure VI – **Comparing improved SPPI (Option A) and data usage (Option B) deflators**



Note: 2010=100.
Sources: Authors' calculations.

The merit of Option B is that it better reflects the significant technical advances and quality improvements observed in telecoms services, and is capable of capturing in a simple measure a variety of quality aspects without further adjustment: increased coverage, for example, allows more people to get access to telecommunications services and thus increases data traffic. Likewise, an increase in speed increases volume as users can consume more data in any given time period. Finally, future changes in technology may be more easily reflected in a data usage based deflator. This is because, as long as the service is defined as the transport of data, any new technology or service will be adding to the volume of data. The impact that the new service will have on prices is then determined by its impact on total revenue relative to its impact on total volume.

The key weakness with this option is that it takes no account of the differential prices currently paid for different communication services. This is vital as consumers do appear to assign different values to the different services, reflected in the differences in prices. However, whether the prices truly reflect consumer utility from different telecoms services can be questioned. Our initial analysis indicates that phone calls cost many multiples per data unit of the equivalent data service, for example by looking at out of bundle charges. While there could be a stronger preference for traditional call and text services, it seems unlikely that the strength of preferences alone could explain the observed magnitude of the difference in prices.

3. Discussion

Our results show a substantial difference between Option A, the improved SPPI, and Option B, the data usage approach, although both reveal a large decline in prices compared to the current methodology. While both deflators are improvements compared to the current method, their incremental impact on real output growth in the sector would differ significantly in terms of magnitude. The key question is whether it is possible to narrow this wide range and so deliver a method that might be applied with confidence in the national accounts.

Two possible extensions to narrow the range are: first, to consider quality adjustment of the SPPI Index, using some of the characteristics of telecommunications not captured presently, such as coverage and latency; secondly, to consider whether the data usage approach can be improved by making an allowance for the fixed infrastructure element in both the delivery and the pricing, which has been increasing in recent years. The index presented here attributes all the costs to the data transmitted. These improvements might help to narrow the gap between these two approaches, but we may need to start with a more basic question, namely why they show such different results in the first place?

The market for communication services is in a period of rapid innovation, resulting in changes both in pricing and consumer behaviour (including significant growth in data usage), thanks to the remarkable engineering advances. The use of an aggregate unit value measure such as the data usage approach, for all that it is not a true price index without the assumption of homogeneity, is probably closer than the Laspeyres to many people's intuition about the effect of advances in communication services on their economic welfare; but to the degree these advances are not reflected in the narrowing of price differentials, we must ask if there may be other reasons for these price differentials which we need to take into account.

3.1. Explaining Price Differentials

In practice, when there are new or improved goods, there will be a period of gradual consumer substitution away from the old goods. The diffusion of digital hardware is typically rapid, with reasonably short replacement cycles, but consumer habits and know-how may take time to catch up. The Boskin Report noted that in a typical product cycle, a new version enters the market at a higher price than old models. When they nevertheless gain market share, "we can conclude that it was superior in quality to the old model by more than the differential in price between the two". This is not the situation across the board in communications, where there is a mix of:

- higher quality and higher price in some services (such as 4G versus 3G for mobile calls and data);

- new, lower prices services substituting for existing ones (such as VOIP versus fixed or mobile telephony, or Rich Internet Applications such as *WhatsApp* versus SMS);

- bundling of different services, and 'convergence' of services, making price and quality comparisons difficult for consumers (and statisticians).

A possible explanation for the price difference therefore lies in product differentiation in a

less than perfectly competitive market. Some specific services may additionally benefit from network effects that would not be captured in market prices. One conceivable unobserved characteristic is the degree to which voice calls and text messaging applications act as platforms, benefiting from significant network effects. While special software or apps might be needed to make a phone call using the data service, the network's own platforms allow the consumer to immediately reach a greater number of people. Once alternative platforms achieve significant market penetration, they become viable alternatives with their own network effects. This, for example, is the case for *WhatsApp*, which reached over a billion users in 2016.[31] However, to get to this stage, consumers need to know about the existence of cheaper and better platforms. We could therefore be experiencing a disequilibrium situation where consumers need time to learn about these alternative platforms.

Furthermore, traditional platforms can be bundled with the equipment. For example, all smartphones come bundled with a telephone and text messaging app which uses the more expensive services of the telecoms provider. Tariffs too come in bundles usually including an allowance of minutes, texts and data. Since consumers cannot opt out of the voice and text elements, they might continue to use these traditional services. Consumers may also have difficulty in comparing prices across differently structured bundles. There are surely large information asymmetries.

The data usage approach clearly presents a lower bound estimate. This is particularly the case if consumers are substituting traditional voice and text services for data driven ones because they feel poorer and so are switching to cheaper and (by assumption) lower quality alternatives. However, the alternative platforms can be superior in that they provide users with additional information and functionality. *WhatsApp* (and other messaging apps) for example indicate if a message has been read and allow users to set up status messages that help their peers know whether someone is available to be contacted. Likewise, if consumers attached lower values to general data usage, such as streaming videos or browsing the internet, then these too should have lower weight in the deflator. However, it is not clear that consumers do indeed attach lower values to data services. For after all, data consumption, along with the usage of data based alternatives to traditional phone calls and text messaging, has been increasing substantially.

3.2. Convergence

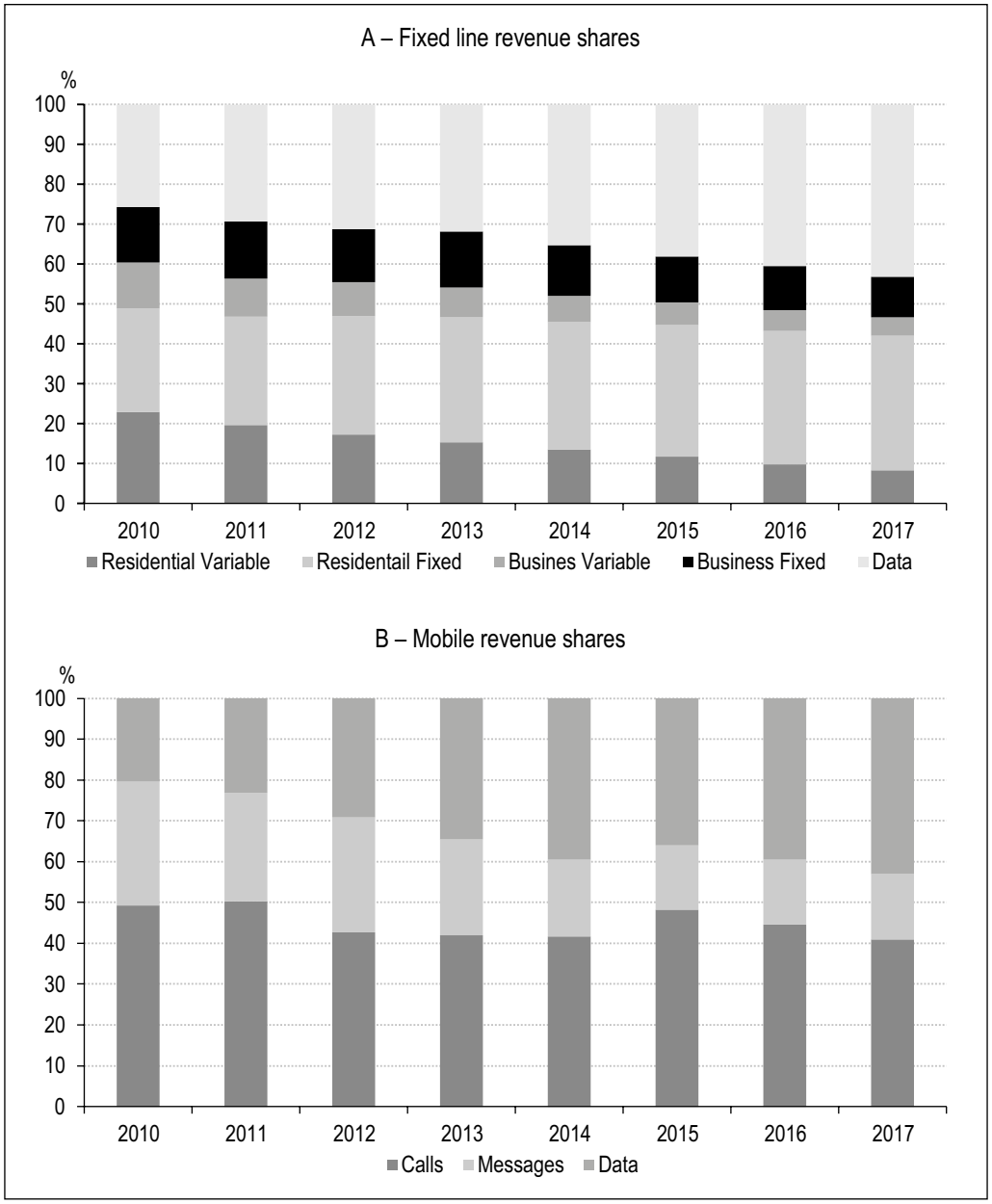
Despite these caveats, it does not seem unreasonable to assume a high and rising degree of substitutability between different forms of telecommunication services as users' behaviour adapts, rather than assuming none – which is the alternative given that we do not have both prices and quantities. In this context, the question is then the degree of homogeneity of voice, text, and data services. Price differentials between these suggest substantial differences from a consumer value perspective. However, from a network perspective, the different services are broadly similar in that they all involve the transportation of data, using the same transmission lines and networks. Having said that, it is clear that this is a transitional phase, both in technology and in consumer behaviour; and in addition that there might be heterogeneous characteristics of voice telephony that some people will continue to buy, such as reliability or coverage.

While the improved SPPI and the data usage approach appear substantially different at present, in future they might converge. The share of total revenue due to data usage increased between 2010 and 2017 for both fixed line and mobile telecommunications (Figure VII). For example, we estimate that broadband data accounted for around 26% of total fixed line revenue in 2010, but by 2017 this had increased to 43%. Similarly, we estimate that mobile data accounted for around 20% of total mobile revenue in 2010, increasing to 43% by 2017, whilst in both mobile and fixed line telecommunications, the share accounted for by voice calls and text messaging decreased. If this trend continues, the revenue and volume weights for the different services could converge. This would mean that the (revenue weighted) improved SPPI and the (volume weighted) data usage approach would converge.

On the face of it, this could favour option A over option B. Since the improved SPPI is chain linked, it could become equivalent to the data usage approach, although this would require further work to establish how to chain link when existing products are converging to become a single, new product. However, until the movements in the two deflators converge, there would continue to be a question of which provides the most unbiased 'true' value of the deflator and hence real volume in the industry.

31. http://www.bbc.co.uk/news/technology-35459812 [Retrieved: 21 July 2017].

Figure VII – **Fixed line and mobile revenue shares (weights for the indices)**



Sources: Authors' calculations.

At the moment, a specific obstacle to convergence is the existence of access charges, now incorporated into bundled prices. While the share of call charges for businesses and residential households decreased from around 35% in 2010 to 13% of total fixed line revenue in 2017, the share of residential and business access charges increased from around 40% to 44% in the same time period (Figure VII). If this trend does not reverse, the two deflators as presently modelled will continue to diverge, as we have no effective way to apportion access charges beyond using the number of subscribers, suggesting the need to incorporate access charges into the data usage model as a cautious way forward.

* *
*

The constant utility approach that informs price theory sits uncomfortably with the practical use of price indices based on specific products to calculate real output and productivity for the national accounts. In the early debate about hedonic prices, Milton Gilbert observed that if quality adjustments fully reflected utility, resulting in lower price indices, a bikini would represent equivalent output to a voluminous Victorian bathing costume, "And should this trend reach its limit of no costumes at all, we would have to say that swimsuit production

had not fallen, even though the industry was out of business." Zvi Griliches replied that the concept of goods made no sense independent of a utility framework, and one would not say the Victorians were better off because they had bulkier swimsuits (quoted in Stapleford, 2009, p. 322). Both perspectives have their appeal, which suggests that the choice of approach and index might depend on whether they are the answer to a question about production or whether in fact the question does not concern output and productivity at all but is instead an aspect of economic welfare.

Our contribution in this paper has been to show that a sensible improvement to the current method for calculating a price index for telecommunications services, taking account of broadband data services, results in an index that has declined substantially more in recent years than the current index. However, this will still be an upward-biased deflator, as it does not sufficiently take account of increasing consumer utility due to new goods. An alternative unit value methodology inspired by the engineering improvements and price declines for data transmission results in an index that declines dramatically more. This understates the 'true' price of the communications services concerned to the extent it does not reflect either consumer attributions of value for service characteristics or attributes such as market structure and price differentiation. It is nevertheless informative about the supply-side efficiency of the services.

Improvements to the current price index for telecommunications services, taking account of broadband data services in both options analysed suggest that the real output of telecommunications services will have been significantly understated in recent years. As these are an intermediate input into other sectors, there will be consequential implications for the sector distribution of output, but potentially also for real GDP. We have focused on telecommunications services but similar considerations may apply to other service sectors experiencing rapid digital innovations. □

Link to Online Appendices: https://www.insee.fr/en/statistiques/fichier/4770156/ES-517-518-519_Abdirahman-etal_Online_Appendices.pdf

BIBLIOGRAPHY

Adelman, I. & Griliches, Z. (1961). On An Index of Quality Change. *Journal of the American Statistical Association*, 56(295), 535–548. https://doi.org/10.1080/01621459.1961.10480643

Ahmad, N., Ribarsky, J. & Marshall, R. (2017). Can potential mismeasurement of the digital economy explain the post-crisis slowdown in GDP and productivity growth? *OECD Statistics Working Papers,* 2017/09. https://doi.org/10.1787/a8e751b7-en

Bean, C. (2016). *Independent Review of UK Economic Statistics*. https://www.gov.uk/government/publications/independent-review-of-uk-economic-statistics-final-report

Boskin Commission (1996). *Final Report of the Commission to Study The Consumer Price Index*. Washington D.C: U.S. Government Printing Office. https://www.finance.senate.gov/imo/media/doc/Prt104-72.pdf

Byrne, D. & Corrado, C. (2017). ICT Prices and ICT Services: What do they tell us about Productivity and Technology? *Finance and Economics Discussion Series* 2017-015. Washington: Board of Governors of the Federal Reserve System. https://doi.org/10.17016/FEDS.2017.015r1

Coyle, D. (2017). Do-it-yourself digital: Searching for clues to solve the productivity puzzle. *Economic Statistics Centre of Excellence Discussion Paper.* Number 2017-1. https://www.escoe.ac.uk/wp-content/uploads/2017/02/ESCoE-DP-2017-01.pdf

Diewert, W. (1998). Index Number Issues in the Consumer Price Index. *Journal of Economic Perspectives,* 12(1), 47–58. https://doi.org/10.1257/jep.12.1.47

Ellis, A., Suibhne, N., Saad, D. & Payne, D. (2016). Communication Networks Beyond the Capacity Crunch. *Philosophical Transactions of the Royal Society* A374: 20150191. http://dx.doi.org/10.1098/rsta.2015.0191

Feldstein, M. (2017). Underestimating the Real Growth of GDP, Personal Income, and Productivity. *Journal of Economic Perspectives,* 31(2), 145–164. https://doi.org/10.1257/jep.31.2.145

Greenstein, S. & McDevitt, R. (2010). Evidence of a Modest Price Decline in US Broadband Services. *NBER Working Papers* 16166, National Bureau of Economic Research. https://doi.org/10.3386/w16166

Griliches, Z. & Cockburn, I. (1993). Generics and New Goods in Pharmaceutical Price Indexes. *NBER Working Papers* 4272, National Bureau of Economic Research. https://doi.org/10.3386/w4272

Griliches, Z. (1961). Hedonic Price Indexes for Automobiles: An Econometric Analysis of Quality Change. *The Price Statistics of the Federal Governement*, 173–196. National Bureau of Economic Research. http://www.nber.org/chapters/c6492.pdf

Griliches, Z. (1994). Productivity, R&D, and the Data Constraint. *American Economic Review,* 84(1), 1–23.

Groshen, E., Moyer, B., Aizcorbe, R. & Friedman, D. (2017). How Government Statistics Adjust for Potential Biases from Quality Change and New Goods in an Age of Digital Technologies: A View from the Trenches. *Journal of Economic Perspectives*, 31(2), 187–210. https://doi.org/10.1257/jep.31.2.187

Heys, R. & Awano, G. (2016). Measuring output in the Information Communication and Telecommunications industries, ONS. https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/labourproductivity/articles/measuringoutputintheinformationcommunicationandtelecommunicationsindustries/2016

Hausman, J. (1996). The Valuation of New Goods under Perfect and Imperfect Competition. *In:* eds Bresnahan, T., F. & Gordon, R., J., *The Economics of New Goods,* pp. 207–248. NBER. http://www.nber.org/chapters/c6068.pdf

Hausman, J. (2003). Sources of Bias and Solutions to Bias in the Consumer Price Index. *Journal of Economic Perspectives*, 17(1), 23–44. https://doi.org/10.1257/089533003321164930

Hicks, J. (1940). The Valuation of Social Income, *Economica*, 7(26), 105–124. https://doi.org/10.2307/2548691

Nordhaus, W. (1994). Do Real Output and Real Wage Measures Capture Reality? The History of Lighting Suggests Not. *Cowles Foundation Discussion Papers* 1078, Cowles Foundation for Research in Economics.

Nordhaus, W. (2007). Two Centuries of Productivity Growth in Computing. *The Journal of Economic History*, 67(1), 128–159. https://doi.org/10.1017/S0022050707000058

Office for National Statistics (2018). Division level labour productivity estimates: February 2018'. ONS, Available at https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/labourproductivity/articles/divisionlevellabourproductivityestimates/january2018#the-effect-of-reallocation-on-output-per-hour-growth

Stapleford, T. (2009). *The Cost of Living in America*. Cambridge University Press. https://doi.org/10.1086/ahr.117.2.565

Williams, B. (2008). A Hedonic Model for Internet Access Service in the Consumer Price Index. *Monthly Labor Review*, 131(7), 33–48. https://www.bls.gov/opub/mlr/2008/article/hedonic-model-for-internet-access-service-in-the-consumer-price-index.htm

APPENDIX 1

BREAKDOWN OF REVENUE AND VOLUME IN THE TELECOMMUNICATIONS INDUSTRY

Table A1-1 – **Revenue breakdown (in £bn)**

	2010	2011	2012	2013	2014	2015	2016	2017
Wholesale services	10.1	8.9	7.8	7.0	6.0	5.9	5.4	5.4
Retail fixed	12.6	12.4	12.4	12.6	13.0	13.5	14.3	14.7
Retail mobile	15.1	15.4	15.8	15.5	15.2	15.2	15.4	15.6
Corporate data services	2.7	2.8	2.7	2.6	2.5	2.5	2.5	2.5
Total	40.5	39.5	38.8	37.7	36.7	37.1	37.6	38.1

Notes: 'Corporate data services' comprises web hosting, Ethernet, IP VPN, digital leased line, corporate VoIP and frame relay/ATM services; wholesale mobile comprises wholesale mobile voice, messaging and data services, mobile voice and SMS termination revenue and wholesale inbound roaming revenue (i.e. revenue from overseas operators when their subscribers use UK networks).
Sources: Ofcom Communications Market Reports 2016, 2017 and 2018.

Table A1-2 – **Volume breakdown (in Petabytes)**

	2010	2011	2012	2013	2014	2015	2016	2017
Total Voice	122	116	113	109	105	104	104	97
Texts	0.018	0.021	0.021	0.018	0.015	0.014	0.013	0.011
Fixed Line Broadband	2,352	4,223	6,017	8,208	16,495	28,751	40,234	59,280
Mobile Data	79	99	239	347	542	880	1,270	1,877
Total	2,553	4,438	6,369	8,664	17,142	29,735	41,607	61,254

Notes: Fixed line Broadband and Mobile Data figures are extrapolated for 2010.
Source: Authors' calculations.

APPENDIX 2

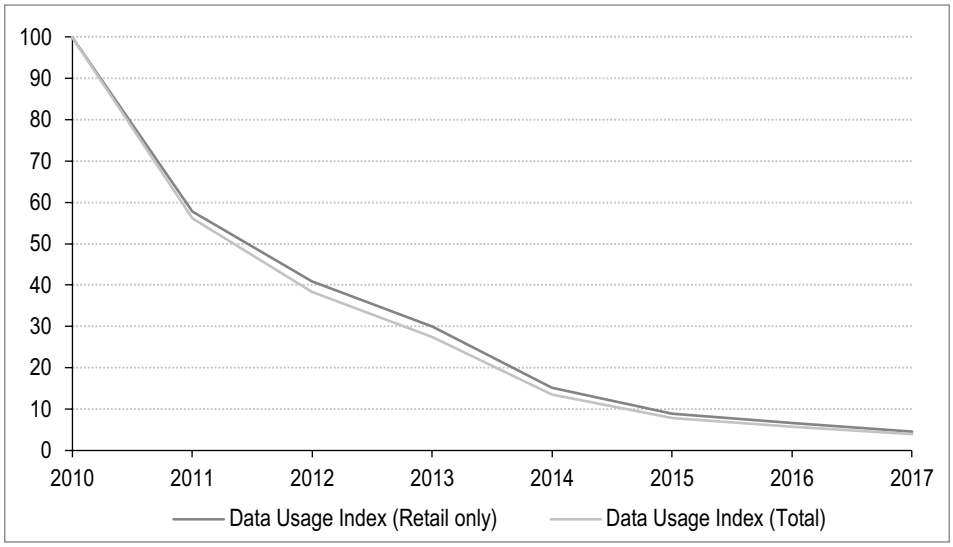
DATA USAGE APPROACH USING RETAIL REVENUES ONLY

Some of the volume data for the data usage approach is limited to retail volumes. Whilst we capture revenue from wholesale and corporate data services, the corresponding volumes are more difficult to identify.

Corporate Data Services for example are often delivered through digital lease lines and the volume of usage is often not measured. Likewise, wholesale volumes, i.e. services that telecoms services providers buy from each other, often have different billing arrangements from the retail market and the volumes are not always readily available.

However, this limitation does not have a substantial impact on our results. When comparing the data usage approach that we use in this paper to an adjusted deflator that only uses retail revenue, we find that there is a minimal difference between the two, with the retail only version of the data usage approach being 1-2.5 index points higher (Figure A2). This is because, while the retail revenues constitute the bulk of telecommunications services revenue, wholesale revenues have been declining at a much faster rate. The inclusion of wholesale and corporate revenues could however bias our results if their corresponding volumes have a significantly different trend from the retail volumes. Further work is required to ascertain these trends and identify suitable datasets for wholesale and corporate data volumes.

Figure A2 – **Data usage approach with different revenue bases**



Note: 2010=100. The Retail index only excludes Wholesale and Corporate Data Services revenue.
Sources: Authors' calculations.

Does Measurement of Digital Activities Explain Productivity Slowdown? The Case for Australia

Derek Burnell and Amani Elnasri*

Abstract – The post 2004 slowdown in productivity growth in developed nations has led to speculation that mismeasurement of digital activities within the national accounts may be responsible. The Australian Bureau of Statistics' (ABS) modelling of potential missing output confirms the findings of Syverson (2017), Ahmad & Schreyer (2016) and Byrne, Fernald & Reinsdorf (2016) that unrecorded digital activities were of insufficient magnitude to explain the productivity slowdown. While there may be room for improvement in data sources and methods more broadly, conceptually digital activities are captured in the National Accounts framework.

JEL Classification: E23, O3, O4
Keywords: productivity slowdown, digital activities, missing output, output mismeasurement

* *Australian Bureau of Statistics (derek.burnell@abs.gov.au)*

The authors would like to thank Paul Roberts and Katrina Richardson for helpful comments and Grace Kim for leading the quantitative study early on and members from the ABS National Accounts Branch who contributed by providing data needed to undertake the simulations.

Received on 29 June 2018, accepted after revisions on 23 May 2019.
Citation: Burnell, D. S. & Elnasri, A. (2020). Does Measurement of Digital Activities Explain Productivity Slowdown? The Case for Australia. *Economie et Statistique / Economics and Statistics*, 517-518-519, 123–137. https://doi.org/10.24187/ecostat.2020.517t.2022

The post 2004 slowdown in productivity growth in developed nations has led to speculation that mismeasurement of digital activities within the System of National Accounts (SNA) may be responsible.

Previous quantitative analysis examined whether the underlying concepts, sources and methods adequately captured emerging digital activities. Syverson (2017) estimated the counterfactual output required to offset labour productivity slowdown in the U.S. economy, and then tested whether estimation of new digital intensive activities could be of sufficient magnitude to account for the counterfactual output. Syverson concluded that the counterfactual argument that "true productivity growth has not slowed (or slowed considerably less than measured) since 2004 was not supported". The author's empirical results showed potential missing digital activity output was substantially lower. Rather, it was more likely that productivity measurement problems may exist for certain digital product classes on a smaller scale than hypothesised.

Other studies by the OECD and the IMF support these findings and evidence from Byrne *et al.* (2016), Nakamura & Soloveichik (2016), and Cardarelli & Lusinyan (2015) supports this view. For example, Byrne *et al.* (2016) note that mismeasurement of Information Technology (IT)-related goods and services was not confined to the post 2004 period. Rather, it was substantial prior to 2004 as well as more recently. Furthermore, rising import penetration for computers and communications equipment implies that U.S. domestic production, which matters for Gross Domestic Product (GDP) growth, has slowed. Using different approaches and data, and pointing to output saving technologies, Nakamura & Soloveichik (2016), and Cardarelli & Lusinyan (2015) argue that the slowdown more likely reflected a true reduction in the rate of technological growth rather than mismeasurement. Nonetheless, as digital activities increase, international collaboration between statistical agencies will become increasingly important. In particular, adopting good methods for price and volume estimation across similar digital technologies to facilitate like-for-like comparisons.

This paper extends Syverson's approach to the Australian economy. However, instead of labour productivity, the paper tests the counterfactual for Multifactor Productivity (MFP), which similarly slowed from 2003-04 as in the U.S. experience. The paper examines the case for potential missing digital output in the Retail industry, and digital peer-to-peer (P2P) activities in the Transport, Postal and Warehousing, Information, Media and Telecommunication and Finance and Insurance Services industries. In addition, the paper examines whether Australian results are sensitive to: *(i)* shorter mean asset lives accompanied by the One-hoss shay age-efficiency assumption for IT capital services; and *(ii)* backcasting quality adjusted internet prices to reflect the volume of data provided.

The rest of this paper proceeds by describing the Australian perspective on concepts, sources and methods. The missing output required by digital activities to explain the slowdown in Australia's MFP growth are estimated for specific industries. Other potential sources of the Australian productivity decline are discussed, followed by some concluding remarks.

1. The Australian Perspective

Conceptually, digital activities are included in the SNA framework whenever they give rise to measurable and recordable transactions. Ahmad & Schreyer (2016) and Byrne *et al.* (2016) clarified that while the SNA framework is robust in concept, digital activities enable some economic activity that was traditionally paid for to now be carried out as unpaid household work, and therefore no production is recorded. Similarly, the SNA was never intended to capture the willingness to pay (consumer surplus) for freely available goods and services. The Australian System of National Accounts (ASNA) framework is consistent with SNA so, by extension, is also robust in concept.

1.1. Sources and Methods

In practice, digital activities are recorded when the sources and methods are adequate. Most Australian businesses report to the Australian Taxation Office (ATO). The ABS uses this data in combination with data directly collected (such as the ABS annual Economic Activity Survey), to ensure that there are no significant undercoverage issues for the data reported for Australian resident businesses. Similarly, to minimise undercoverage in economic transactions between residents and non-residents, the ABS makes extensive use of administrative data (such as from the Australian Customs Office) and cross checks sources with the demand-side, like the Household Expenditure Survey.

Data from a variety of sources are confronted and contrasted within a Supply and Use Tables (SUTs) framework to estimate the current price estimates across the production, expenditure and income measures of GDP. The SUTs framework is a powerful tool to improve the coherence of the economic information system. SUTs reconcile how the supply of products within the economy within an accounting period are used for intermediate consumption, final consumption, capital formation or exports. The SUTs permit an analysis of markets and industries and allow productivity to be studied at this level of disaggregation. The ABS's SUTs are estimated for both current prices and volumes.

The chain volume estimates (and their associated deflators) in the ASNA are also confronted, particularly across the expenditure and production accounts. To facilitate improvement in the estimation of chain volume measures in services industries, over the last five years, the ABS has had an active program to ensure that more representative prices are available as services activities become more influential in the economy.

1.2. An Examination of Digital Data Sources

In this context, Ahmad & Schreyer (2016) noted that digital activities may open doors to new solutions to adequately capture source data. This has tended to be the experience in the ABS. For example, digitally sourced scanner data (transaction data) has been included in the ABS's consumer price index (CPI) dataset for a number of years. The approach has the advantage of increasing sample sizes (thereby lowering the sampling error) and helping to price heterogeneous products and services more accurately. The ABS is also acquiring consumer price data digitally via 'web-scraping' technologies.[1]

The coverage of services prices has also improved over time. For example, ride-sharing, shared accommodation and digital products and services (e.g. streaming services) were implemented recently into the CPI. However, accurately separating the price and volume components for services activities remains a challenging area for statistical agencies. Cross-country comparisons of the spread in the range of price growth for certain activities suggest that more effort is required in accounting for quality change for similar service in a more consistent manner. For example, the OECD's National Consumer Price Indexes for Telecommunications show wide-ranging growth patterns over the 13 years to 2015, falling around 40% for Italy to a rise of nearly 30% for Canada. The divergences between countries suggest that, at least for digital goods and services of similar characteristics irrespective of nation, more effort is required to separate the price and volume components to record quality changes more consistently.

Moreover, this paper recognises the importance to review more model-intensive areas of the national accounts, such as capital stock and productivity. This is because digital activities influence the way production is changing, impacting asset mix, asset lives and depreciation rates. One issue is that tablets and smartphones are now taking on the roles traditionally provided by computers, indicating higher replacement rates, especially in the more rapidly innovating industries. This implies higher depreciating assets requiring higher IT depreciation rates for user costs in productivity estimation. To improve visibility, this study separates the contribution of IT and non-IT capital in the productivity growth accounts.

2. Evaluating the Impact of Digital Activities on Australia's Productivity

As mentioned earlier, this paper follows the Syverson's approach to estimate the counterfactual real output growth for Australia. Syverson defined counterfactual output as output required to sustain average labour productivity growth in the pre-2004 period. This paper, instead, defines counterfactual output based on various MFP growth targets achieved in the pre-slowdown period.[2]

Figure I shows the evolution of the Australian market sector's MFP over the period 1994-95 to 2016-17. The slowdown phase lasted from 2003-04 to 2010-11 (averaging -0.3% annually). Since 2011-12, MFP recovered, averaging 1% annually.

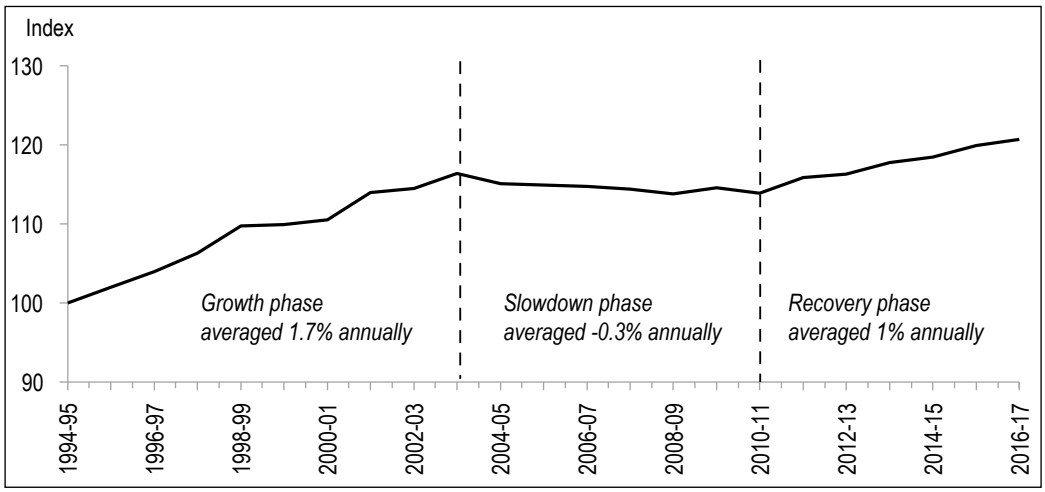
Using different average values of MFP growth over three different periods, the paper calculates the implied output growth required to explain each level of productivity growth. The following three scenarios are considered:

---

1. See ABS (2017) for more information on the recent ABS's methods and data sources in order to reflect the contemporary economy and consumer preferences in the Australian CPI.

2. While labour productivity measures output per hour of labour input, MFP measures output produced per unit of combined inputs of labour and capital.

Figure I – **Evolution of the 16 industry market sector MFP aggregate (1994-95 to 2016-17)**



Sources: ABS (Estimates of Industry Multifactor Productivity, 2016-17, Cat. no. 5260.0.55.002).

(i) the average annual long-term rate of MFP growth of the 12 selected industries aggregated over the period 1973-74 to 2003-04 (1%),[3]

(ii) the average annual rate of MFP growth of the 16 market sector industries aggregated over the period 1994-95 to 2003-04 (1.7%),[4] and

(iii) the average annual rate of MFP growth of the 16 market sector industries aggregated over the strongest growth cycle from 1993-94 to 1998-99 (2.6%).

Figure II highlights these three scenarios of MFP and the average annual output growth required to obtain each level of these MFP growth rates.

The analysis primarily focuses on potential mismeasurement in gross fixed capital formation (GFCF), household final consumption expenditure (HFCE), gross value added (GVA), and associated price deflators. Income measures, on the other hand, were found to be more robust. Although the ASNA adjust source data for under-statement of income (e.g. cash transactions in the construction industry), digital transactions are different as the relationship is typically three way via a facilitator. Transactions are more likely to be recorded in taxation data, since the facilitator has a registered Australian Business Number (ABN) to operate.[5] Accordingly, gross operating surplus or the income shares needed to aggregate capital services to the market sector, were not impacted by digital transactions. This assumes that any under-reporting of income would proportionally allocate to labour and capital income shares.

Detailed below are the empirical results for the digital activities identified as the most likely candidates for potential missing output: the sharing economy, telecommunication pricing, and IT enhancements.

2.1. Sharing Economy

Three general forms of intermediation of P2P services (sharing economy) are assessed in this paper: distribution services, ride-sharing services and financial intermediation services. For detailed discussion on these P2P services see Ahmad & Schreyer (2016).

2.1.1. P2P Retailing

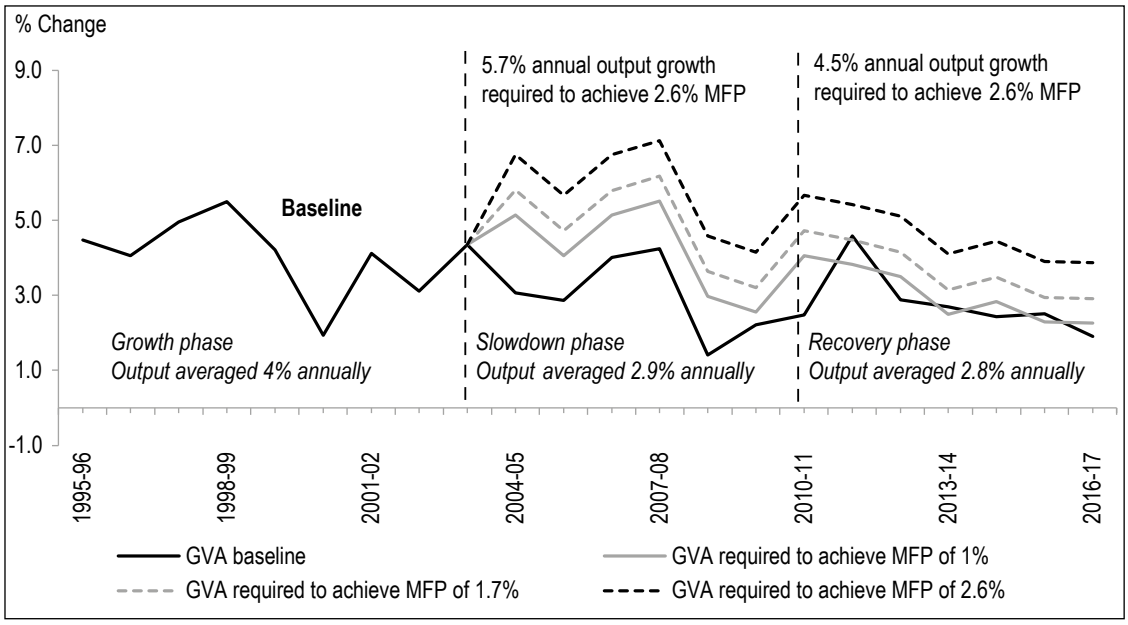
P2P or consumer-to-consumer retailing refers to transactions facilitated by a third party digital platform that brings buyers and sellers together. The underlying retail transactions are not new. Conceptually, all of these transactions and the

3. The 12 selected industries aggregate comprises the ANZSIC Divisions A to K and R, representing the ASNA definition of the market sector prior 2010-11. Estimates of MFP growth of this aggregate are published in the ABS cat. 5260.0.55.002.

4. The 16 market sector industries aggregate comprises the ANZSIC Divisions A to N, R and S (see Appendix). Estimates of MFP growth of this aggregate are published in the ABS cat. 5260.0.55.002.

5. In the digital activities, understatement of reported income is less likely to occur, as transactions between the facilitator and service provider are generally managed over a digital platform rather than cash. Moreover, they must have registered an ABN if their annual turnover exceeds AUD75,000.

Figure II – **The market sector output growth required post 2003-04 to achieve various MFP targets**



Sources: ABS (Estimates of Industry Multifactor Productivity, 2016-17, Cat. no. 5260.0.55.002) and authors' estimates.

GVA created are recorded in GDP. The main difference is that digital activities increase the scale of these transactions, since web-based intermediaries reduce entry barriers, and access to the internet facilitates consumer access (Ahmad & Schreyer, 2016). In Australia, many brick-and-mortar stores have extended their business to online retailer platforms via their own web portal or through a digital intermediary.

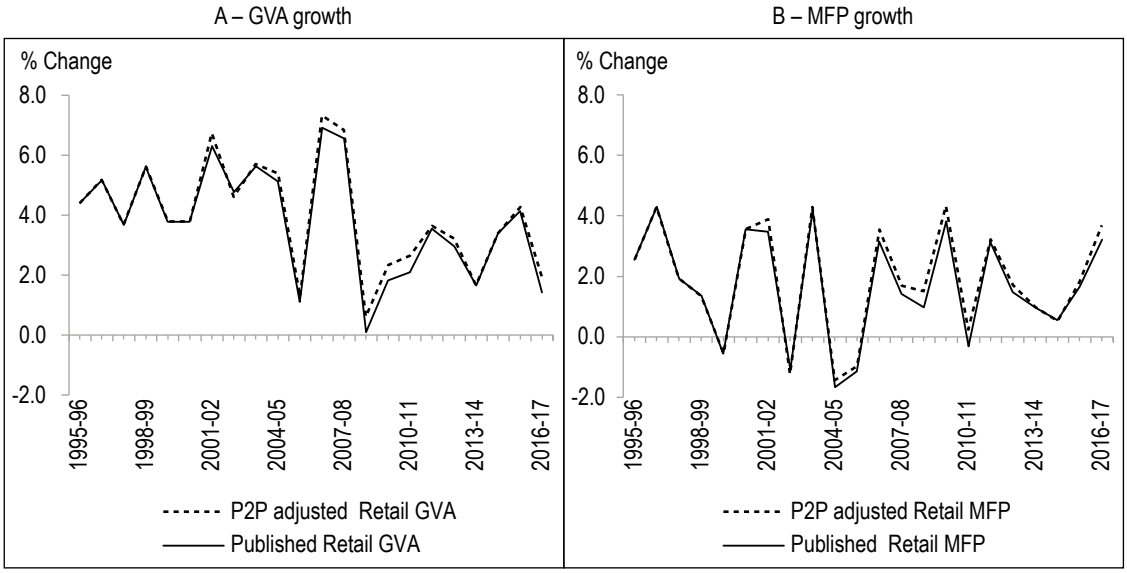
While investigations confirmed that products bought and sold through P2P retailing are captured through HFCE and imports, this study identified potential underestimation of Retail industry GVA caused by under-coverage of P2P intermediaries between 1999-2000 and 2013-14. This was mainly due to some online retailers not having the ABNs required for inclusion in the ABS surveys. To account for this, digital facilitator selling fees were modelled using an average fee per transaction method for the retailers without ABNs. The value of newly produced Australian goods (excluding second-hand goods and imports) that were sold by these online retailers was estimated to be approximately 4% of retail GVA chain volume measure in 2016-17. Figure III shows that the impact of the P2P adjustment on Retail industry GVA chain volume measures and MFP growth over the period 2001-02 to 2016-17 was immaterial.

2.2.2. Ride-Sharing Services

P2P ride-sharing refers to road transport services matching drivers and passengers digitally. These transactions are facilitated using tablets and smartphone applications. The taxi price index may not be a suitable deflator, because the price (depending on market forces of demand and supply) and ride quality are more dynamic. For example, cheaper rates are available relative to taxi fares when demand is subdued but when demand is strong, rates may exceed standard taxi fares. Accordingly, a separate price index for P2P ride sharing is applied in the ABS.

In Australia, ride-sharing services have grown substantially since 2014-15, with nearly 10 million rides facilitated since its inception. Potential missing output was identified before 2015-16, prior to a new tax law requiring ride sharers to register for an ABN. To quantify this, a 2% market share was assigned to ride sharing in 2012-13. From 2012-13, the market share progressively increased its share of transport services from 4% in 2013-14 to 6% in 2014-15, through to 10% since 2015-16. To reinforce potential missing volume growth, the analysis assumed prices were at the lower bound of 40% cheaper compared to taxis. Factoring into consideration the extent of ride-sharing discounting in the deflator not only reinforces the growth in real output through flexible

Figure III – **Retail trade, 1995-96 to 2016-17**



Sources: ABS (Estimates of Industry Multifactor Productivity, 2016-17, Cat. no. 5260.0.55.002 and Australian System of National Accounts, 2016-17, Cat. no. 5204.0) and authors' estimates.

pricing, real output growth is also reinforced by increased competition which has slowed the rate of price growth, or even resulted in some price decline over recent years. Moreover, since 2013-14, the modelled taxi and ride-sharing combined price deflator showed no significant price growth.

Modelling ride-sharing services is more complex than in the P2P retail scenario. Not only is pricing more dynamic, there is also the issue of reclassifying the HFCE component assigned to ride-sharing services as GFCF, which will impact upon capital services and hence productivity (Ahmad & Schreyer, 2016).[6] To account for this effect, a proportion of the stock of consumer durables attributable to ride-sharing was allocated to the capital services of road vehicles in the Transport, Postal and Warehousing (TPW) industry. Nonetheless, while significant in the context of transport consumption expenditure, this adjustment was small in the overall context of TPW industry GVA, at less than 0.1%. The impact of these adjustments on TPW real GVA and MFP growth, presented in Figure IV, are minimal.

2.2.3. Peer-to-Peer Lending

P2P lending facilitates the matching of borrowers and lenders via a digital platform, with intermediaries providing liquidity transformation services. Through a bidding process, these services potentially offer greater flexibility between interest rates and risk than traditional lending via financial institutions. P2P lending is an output of sub-division Finance in the Financial and Insurance Services (FIS) industry.

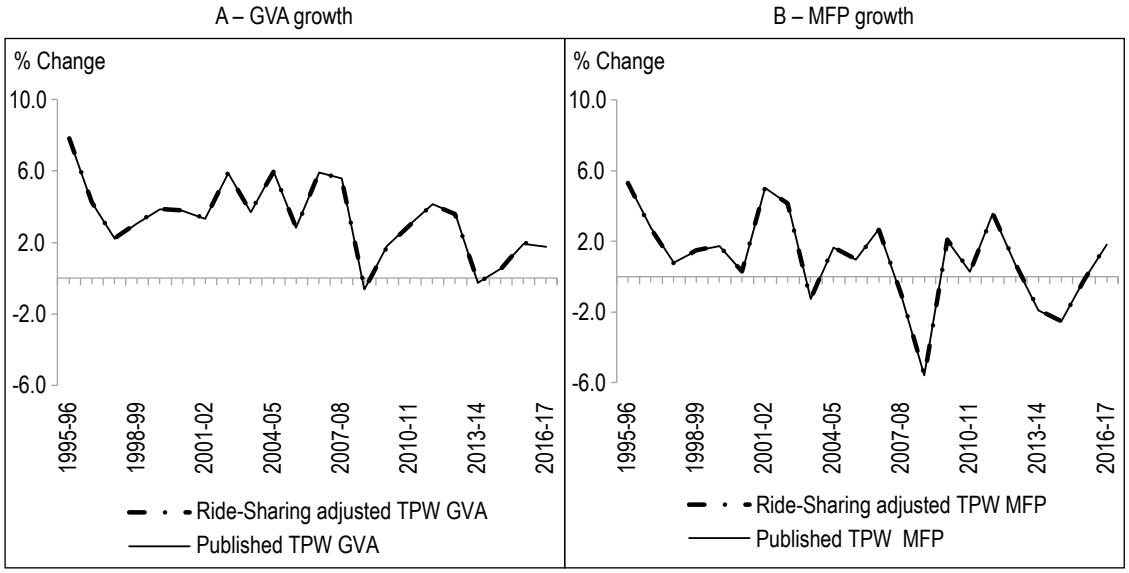
To capture emerging P2P lending impacts, FIS GVA was adjusted to capture missing loans. Industry analysis of the P2P lending market revealed that the potential missing FIS GVA for unrecorded loans from P2P lending, represented only 0.3% of total FIS GVA in chain volume terms in 2013-14, 0.6% in 2014-15, and 1% in 2015-16 and 2016-17. Therefore, the output adjustments had no material impact on FIS GVA and MFP growth (Figure V).

2.3. Quality-adjusted Internet Access Price

Quality change in digital products may also influence prices, and therefore volume estimates of GDP and productivity. One key area is the price of telecommunication services. Technological change has allowed internet service providers to offer customers progressively more generous download quotas at little or no additional cost, sometimes switching customers to an unlimited quota to maintain customer loyalty. In addition,

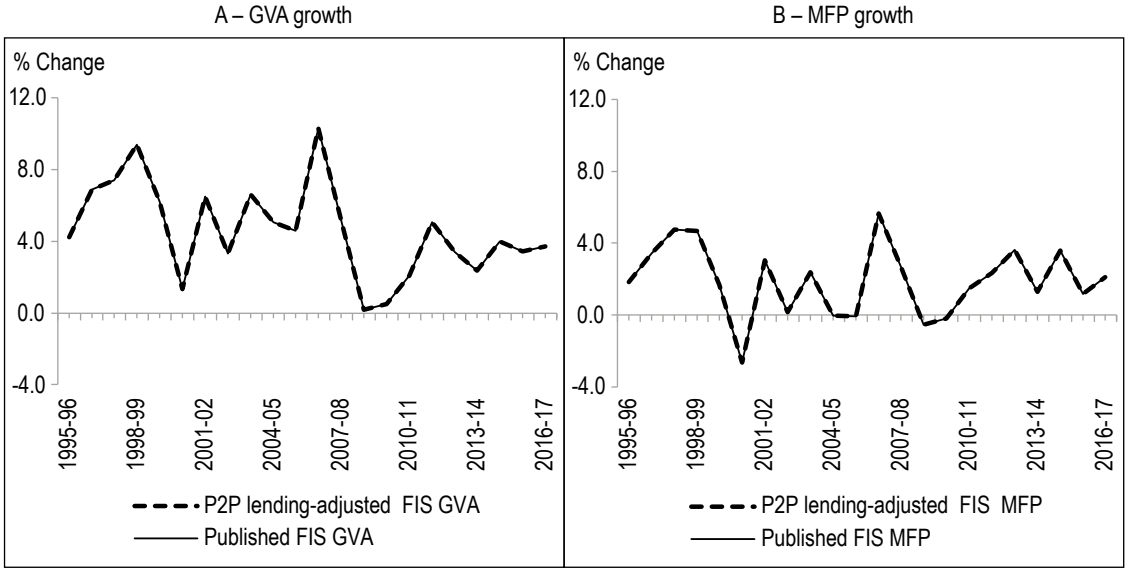
6. This reclassification has no impact on GDP growth so long as there is an equally offsetting adjustment to HFCE for the motor vehicles to be capitalised.

Figure IV – **Transport, postal and warehousing (TPW), 1995-96 to 2016-17**



Sources: ABS (Estimates of Industry Multifactor Productivity, 2016-17, Cat. no. 5260.0.55.002 and Australian System of National Accounts, 2016-17, Cat. no. 5204.0) and authors' estimates.

Figure V – **Financial and Insurance Services (FIS) (1995-96 to 2016-17)**



Sources: ABS (Estimates of Industry Multifactor Productivity, 2016-17, Cat. no. 5260.0.55.002 and Australian System of National Accounts, 2016-17, Cat. no. 5204.0) and authors' estimates.

with the emergence of tablets and smartphones, there has been a strong trend in the uptake of wireless internet services by customers. For example, the June 2018 issue of the Internet Activity ABS (2018b) reported that wireless data downloaded per customer tripled since 2010.

Commencing in 2013-14, the National Accounts quality-adjusted internet price index was based on the change in the Telecommunications equipment and services consumer price index (Telecommunications CPI). The Telecommunications CPI is quality-adjusted to reflect, for example, the progressively more generous download quotas, and fell around 20% since 2013-14.

From the March quarter 2014, the ABS significantly increased its use of transactions data to compile the Australian CPI, which included

transactions for telecommunications services.[7] The transactions data enabled replacing point-in-time prices for certain products (previously collected by field collectors) with a unit value (from transactions data). The unit value approach is described in ABS (2018a) and ILO (2004).[8]

The Telecommunications CPI generally grew in the pre-2014-15 period (prior to the unit value approach), and then fell steadily from 2014-15 as quality-adjusted prices steadily fell (Figure VI). However, information is available to model the quality change for certain subgroups of telecommunication prior to 2014-15. In particular, the volume of wireless data downloaded and the number of subscribers reflect that download volumes grew steadily since 2010, ABS (2018b). Therefore, an adjustment was applied to capture understated real output growth for the period 2008-09 to 2013-14, which was prior to the introduction of the unit value method. This was modelled by extending back the post 2013-14 relationship between the Telecommunications CPI and wireless downloads per subscriber, since the understated growth in GVA was determined to be mainly in the wireless telecommunication component. Prior to 2008-09, specific information on downloads per subscriber was not available to model the extent of quantity change, so the Telecommunications CPI was used.

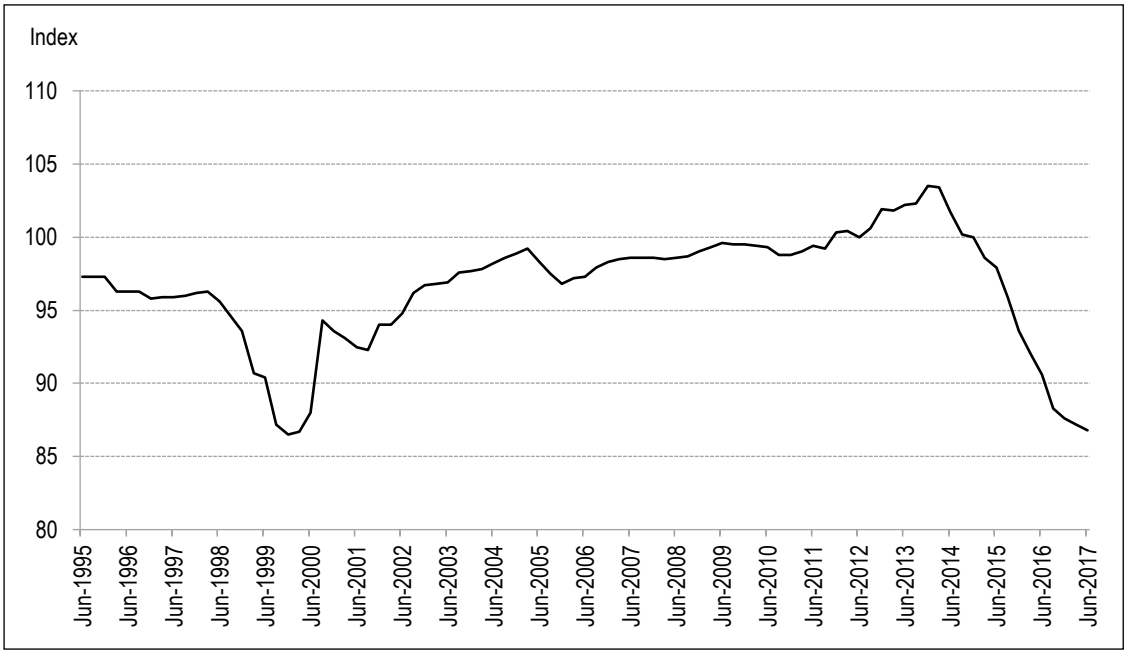
Using this approach highlights the impact of hidden quality change on Information, Media and Telecommunications (IMT) GVA and MFP between 2009-10 and 2013-14 (see Figure VII). Assuming the post 2013-14 relationship between wireless downloads per customer and Telecommunications CPI holds, the impact on GVA and MFP growth was positive, with adjusted GVA growth patterns aligning more closely with pre 2008-09 and post 2013-14 growth patterns.

2.4. Capital Services of Information Technology (IT)

Capital inputs are an important component for measuring MFP. This study evaluates the capital stock assumptions for computers and software, grouped as IT capital. Currently, the average asset lives (that is, average of the length of time they are used in production) for computers and software are fixed over time and across industries.[9] In addition, the same efficiency decay parameter, used to estimate the flow of capital

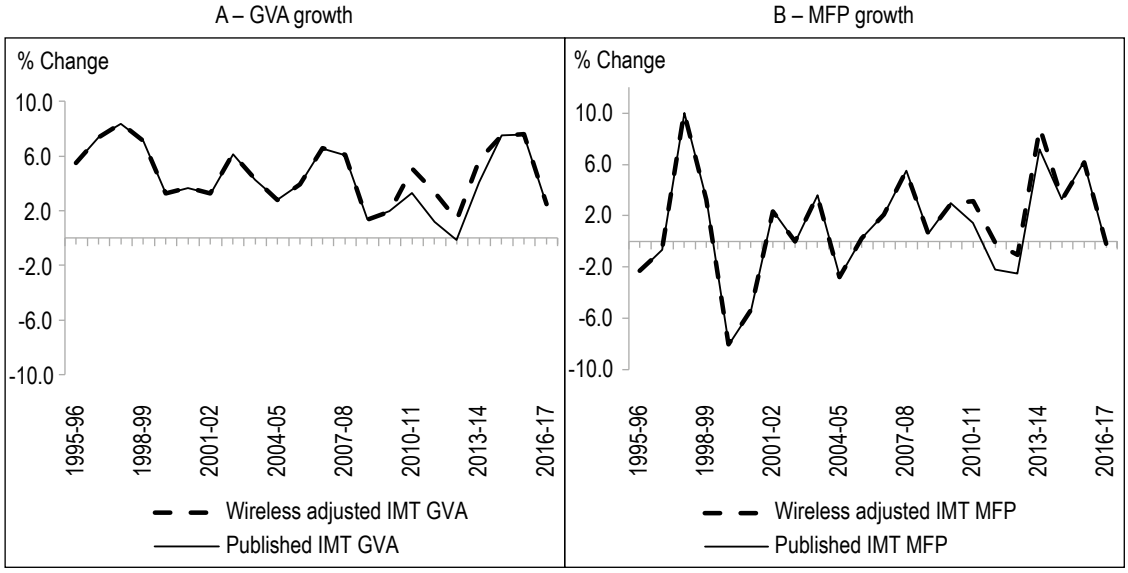
7. For more details, see the article "The Australian CPI: A Contemporary Measure of Household Inflation" in the September quarter 2017 issue of Consumer Price Index, Australia (Cat. 6401.0).
8. See sections 10.105 - 10.107 of (ILO, 2004).
9. SNA08 also recognises capitalisation of databases as intellectual property products, recommending the cost of production approach, or market price, if databases are sold. However, capitalised databases are not recorded in the ASNA due to lack of available data.

Figure VI – **Telecommunication equipment and services CPI**



Sources: ABS (Consumer Price Index, Australia, Jun 2017, Cat. no. 6401.0 and Internet Activity, Australia Cat. no. 8153.0) and authors' estimates.

Figure VII – **Information, Media and Telecommunications (IMT), 1995-96 to 2016-17**



Sources: ABS (Estimates of Industry Multifactor Productivity, 2016-17, Cat. no. 5260.0.55.002 and Australian System of National Accounts, 2016-17, Cat. no. 5204.0) and authors' estimates.

services, is applied to all types of machinery and equipment.[10]

The ABS currently applies the hyperbolic decay function with an efficiency reduction parameter of 0.5 for computers in all industries. Hyperbolic decay accelerates as computers age, due to the wear and tear effect.[11] However, Diewert & Wei (2017) argue that typically, the service flow that a computer generates over its useful life is roughly constant, implying a One-hoss shay age-efficiency function (a constant efficiency parameter of 1.0).[12]

The ABS's asset lives for computers and software currently are applied to all industries equally. However, the Bean review noted that efficiency and portability of IT capital may vary across industries (Bean, 2016). The ratio of IT GFCF to total GFCF indicates that certain industries (such as the FIS and Administrative and Support services) are more likely to intensively use computers as well as adopt new and improved computers more quickly. The faster replacement rates imply shorter IT asset lives in these industries (Bean, 2016).[13]

To account for these effects, a One-hoss shay age-efficiency function was applied to computers and software, and shorter asset lives for computers and software was applied for industries using IT capital more intensively. The existing and simulated new capital stock assumptions are presented in Tables 1 and 2 respectively. The adjustments to the efficiency reduction parameter and asset lives are made from 1999-2000 onwards.

Figure VIII-A shows that growth in capital services for the FIS industry slowed significantly due to the much shorter asset lives applied in the simulations. This is because shorter asset lives reduce the share of faster growing IT capital stock, relative to non-IT capital stock, in the industry. This more than offsets the increases in capital services growth due to applying the One-hoss shay age-efficiency function, resulting in MFP growth increasing in this industry. For most other industries, the One-hoss shay impact on capital services growth more than offset the reduction in capital services growth due to shortened assets lives, resulting in lower MFP growth in those industries. Figure VIII-B shows that the overall impact on the market sector is close to neutral.

10. The decline in the productive efficiency as an asset ages is described by an age-efficiency function. The age-efficiency function determines the loss in efficiency, mainly due to wear and tear as the asset ages.

11. The ABS uses a hyperbolic function in which the efficiency of the asset declines by small amounts at first and the rate of decline increases as the asset ages. See ABS (2015) for a discussion on capital stock measurement.

12. The One-hoss shay model assumes that the service flow of the asset is constant over the lifetime.

13. For example, the FIS industry embraces the use of digital technologies such as online banking services.

Table 1 – **Existing capital stock assumptions for IT assets**

Asset type	Age-efficiency slope (Beta)	Mean asset life		
		4 years	5 years	6 years
Computers	0.5		Divisions: All	
Computer software (purchased)	0.5	Divisions: All		
Computer software (in-house)	0.5			Divisions: All

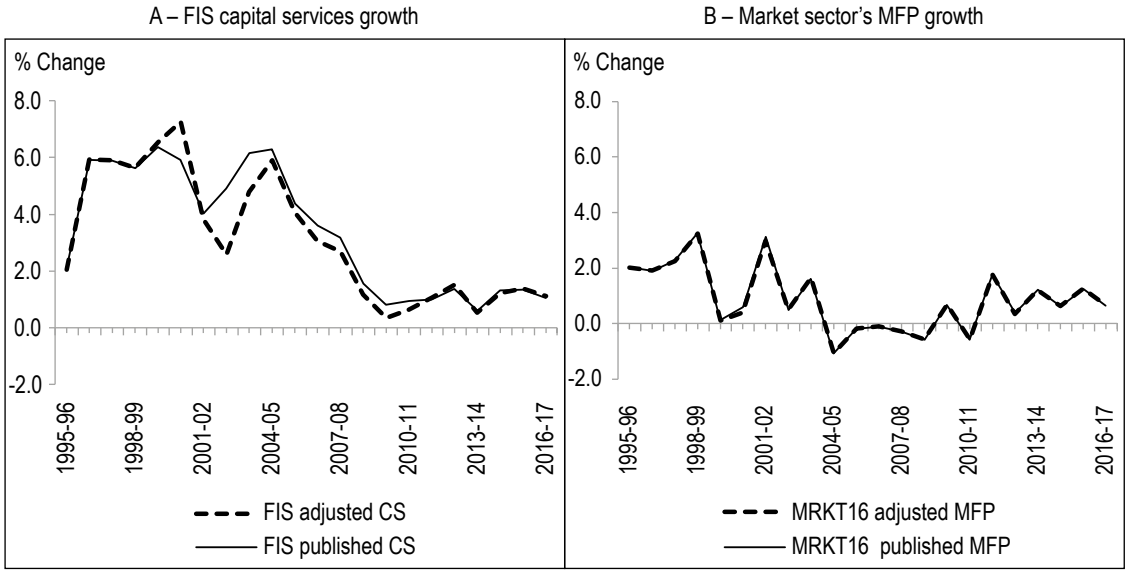
Note: A list of ANZSIC Divisions is provided in Appendix.

Table 2 – **Industry allocation of revised capital stock assumptions for IT capital services**

Asset type	Age-efficiency slope (Beta)	Mean asset life			
		2 years	3 years	4 years	5 years
Computers	1	Divisions K and N	Divisions G and M	Divisions A, B, C, D, E, F, H, I, J, L, R and S	
Computer software (purchased)	1	Divisions K and N	Divisions F, G, J, M and S	Divisions A, B, C, D, E, H, I, L and R	
Computer software (in-house)	1		Divisions K and N	Divisions F, G, J, M and S	Divisions A, B, C, D, E, H, I, L and R

Note: A list of ANZSIC Divisions is provided in Appendix.

Figure VIII – **The impact of ICT enhancement, 1995-96 to 2016-17**



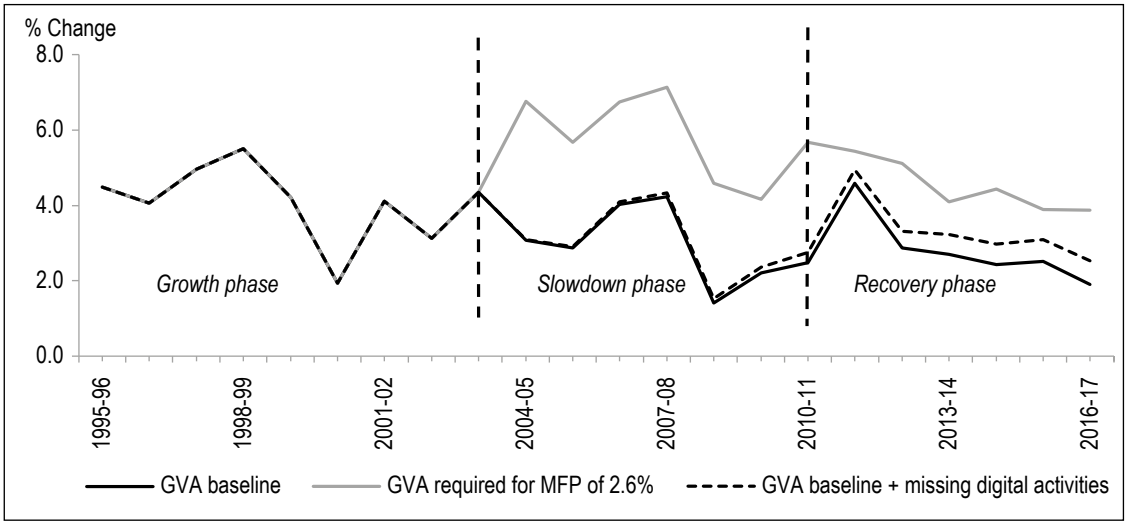
Sources: The authors' estimates based on data from ABS (Estimates of Industry Multifactor Productivity, 2016-17, Cat. no. 5260.0.55.002) and capital stock simulations under different age-efficiency assumptions.

2.5. The Overall Impact of Modelled Digital Activities

Each of the impacts discussed were combined into an overall impact of digital activities on Australia's market sector productivity. Figure IX presents the scenario of upper bound target output growth that corresponds to 2.6% MFP growth. The overall impact on the market sector GVA growth due to potential missing digital output is quite small during the productivity slowdown phase (2003-04 to 2010-11). Real GVA growth accelerates during the recovery phase (since 2011-12). By 2016-17, the GVA adjusted for digital output grew 0.6% per year stronger than the published baseline GVA.

The figure also shows that the upper bound target output was higher during the slowdown phase than the recovery phase. Going into the recovery

Figure IX – **The impact of digital activities on real GVA growth, 1995-96 to 2016-17**



Sources: ABS (Estimates of Industry Multifactor Productivity, 2016-17, Cat. no. 5260.0.55.002 and Australian System of National Accounts, 2016-17, Cat. no. 5204.0) and authors' estimates.

phase, the upper bound target output and GVA growth, including the potential missing digital output, start to converge. By 2016-17, the adjusted GVA is sitting almost midway between the baseline and upper bound output growth in 2016-17.

To put the counterfactual output into context, in 2016-17 market sector real GVA of AUD34,768 per capita would need to increase by an additional AUD12,278 per capita to maintain post-2004 market sector's MFP growth at 2.6%. However, potential missing digital output only accounted for AUD1,361 per capita, with the vast majority of AUD10,918 per capita of the counterfactual output gap unexplained. Similarly, Syverson estimated that to maintain labour productivity growth in the U.S. at the annual average growth recorded in the pre-slowdown period of 1994-95 to 2003-04 (of 2.8% per annum), would require a post 2004 counterfactual real output level around 17% higher in 2015, representing about USD9,300 per capita. The consumer surplus (outside of scope of production) from the new digital activities in 2015 was estimated to be approximately USD3,900 per capita, well short of the counterfactual per capita required to maintain labour productivity growth at 2.8%.[14] By comparison, the counterfactual output gap that can be explained by missing digital output in Australia is proportionally lower.

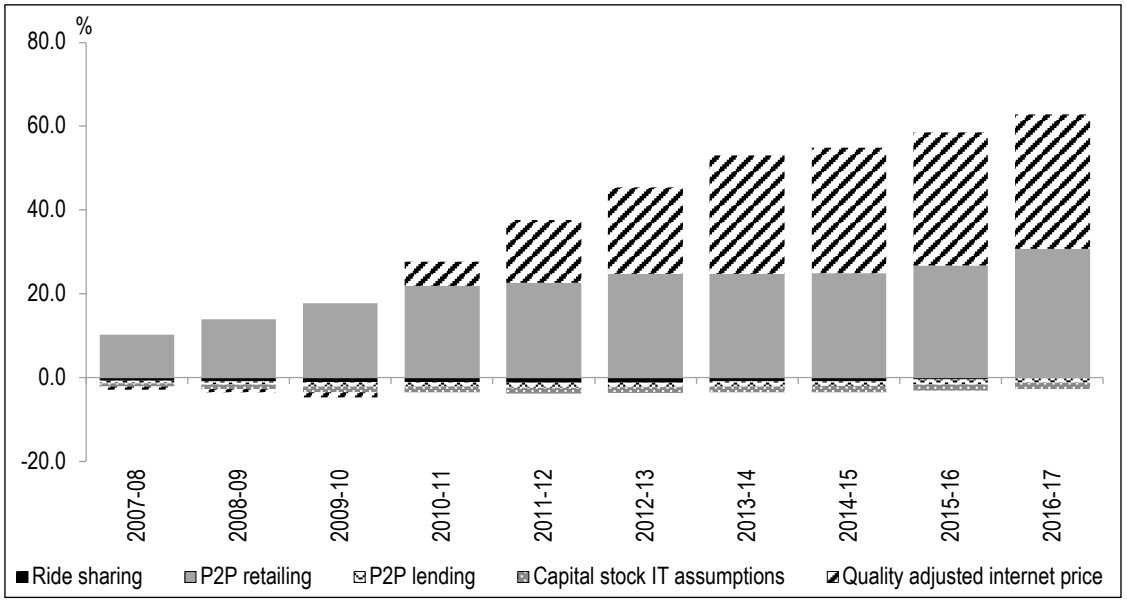
Figure X represents the breakdown of the contribution of digital activities to MFP growth by different digital platforms. The chart shows that P2P retailing and quality-adjusted internet price, were the main contributors to the potential digital mismeasurement. By comparison, ride-sharing, P2P lending and ICT enhancement impacts were negligible.

Another way to present the potential missing output lost due to digital activities, is to compare it to the remaining potential missing output required to achieve counterfactual output. Figure X plots this comparison for the upper bound MFP scenario. The figure highlights that during the productivity slowdown phase, the contribution of potential missing digital output to the counterfactual output required is small – although the share increases during the recovery phase. For example, in 2016-17 the potential missing output due to digital activities represented around 30% of counterfactual output. In part, this increased share can be attributable to the lower counterfactual output required during the recovery phase.

Figure XI suggests that, while the counterfactual output attributable to digital activities is small, its impact accumulates over time. This finding suggests that the ABS, and perhaps other statistical agencies, need to remain proactive to ensure data collection and pricing methods are adequately capturing the growing influence of digitally facilitated output. Figure XI also indicates non-digital, unattributed output,

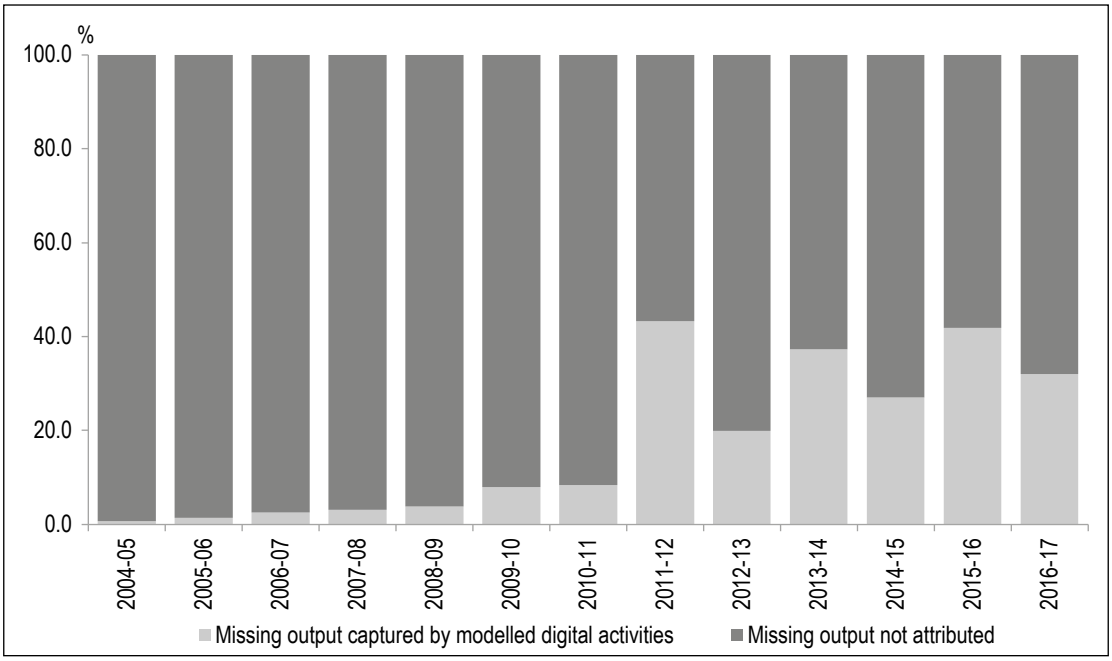
14. Estimates of consumer surplus can be wide ranging, depending on the model and approach. For example, Brynjolfsson & Oh (2012) estimated that the welfare gain from free digital goods and services averaged over USD100 billion per year during 2007-2011.

Figure X – **Contribution to MFP impact from different digital platforms, 2007-08 to 2016-17**



Sources: ABS (Estimates of Industry Multifactor Productivity, 2016-17, Cat. no. 5260.0.55.002 and Australian System of National Accounts, 2016-17, Cat. no. 5204.0) and authors' estimates.

Figure XI – **Overall impact of potential digital mismeasurement, 2004-05 to 2016-17**



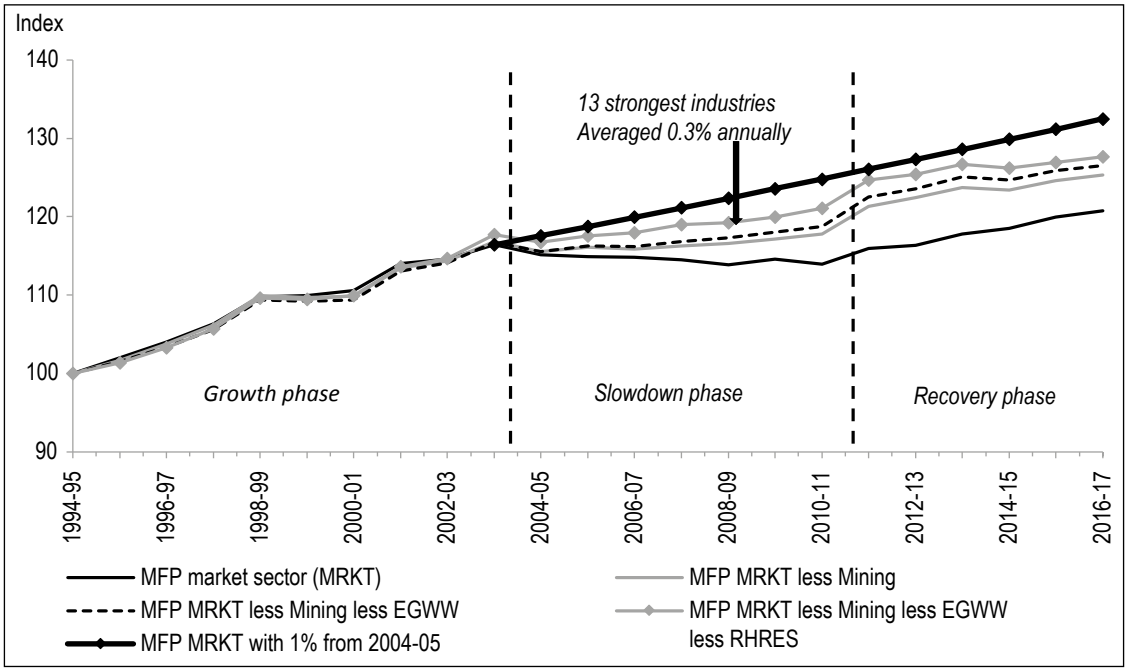
Sources: ABS (Estimates of Industry Multifactor Productivity, 2016-17, Cat. no. 5260.0.55.002 and Australian System of National Accounts, 2016-17, Cat. no. 5204.0) and authors' estimates.

represents the vast majority of counterfactual output, especially during the slowdown phase.

3. Industry Sources of Australia's MFP Slowdown

Since the earlier discussion established that digital mismeasurement is too small to play any significant role in Australia's post 2004 productivity slowdown, a crucial question then arises as to what are the sources of the productivity slowdown? One attempt to answer this question is to focus on the industries that have had the sharpest decline in productivity growth over this time. One key industry is Mining as it has generated a substantial proportion of the market

Figure XII – **Industry source of the market sector (MRKT) MFP slowdown, 1994-95 to 2016-17**



Sources: Authors' estimates based on data from ABS (Estimates of Industry Multifactor Productivity, 2016-17, Cat. no. 5260.0.55.002).

sector's output between 2000-01 and 2006-07 due to a resources boom. Nevertheless, it experienced a significant decline in its MFP growth, contributing substantially to the slowdown of the market sector's MFP growth. Besides Mining, the post 2004 slowdown in Australia's productivity was concentrated in a small number of industries, including Electricity, Gas, Water and Waste Services (EGWW) and Rental, Hiring and Real Estate Services (RHRES).

For the mining and EGWW industries, the digital intensity is too small to explain the slowdown. Rather, the slowdown in these industries can be attributed to other factors. For example, the Productivity Commission, noted timing lags between investment and output as well as using natural resources more intensively (e.g. mineral and energy resources in Mining, see Topp *et al*., 2008) and water resources in EGWW (Topp & Kulys, 2012) in the production process. Official productivity measures do not capture natural resources in capital services because producers do not exercise ownership over them. However, there have been several recent studies modelling the impact of mineral and energy resources in Mining. For example, the ABS (2013) found that Mining capital services growth slows when mineral and energy resource inputs are included, thus reducing the decline in measured MFP.

To understand the sources of Australia's productivity slowdown, it is useful to investigate the influence of each of these industries on the market sector's MFP. Figure XII shows that when Mining, EGWW and RHRES MFP are excluded from the estimation of the market sector's MFP, the remaining 13 industries showed positive growth averaging 0.3% per year during the slowdown phase.[15] This reinforces the results found above about the minimal role of digital activities in explaining the productivity slowdown.

* *
*

The empirical analysis of this paper found that mismeasurement of the digital activities is too small to account for the majority of the productivity slowdown. This finding confirms similar results by Syverson (2017) and Byrne *et al.* (2016) whom found that digital mismeasurement would need to have increased by several orders of magnitude to offset the U.S. productivity slowdown. Moreover, the slowdown in productivity appears to be largely unrelated to the penetration of information technologies across industries and countries. □

15. The ABS used data from Table 23 of Cat. 5260.0.55.002 to facilitate the estimation of several sub-aggregates, such as the non-mining market sector.

BIBLIOGRAPHY

ABS (2018a). Consumer Price Index: Concepts, Sources and Methods (Cat. no. 6461.0). Australian Bureau of Statistics. https://www.abs.gov.au/ausstats/abs@.nsf/mf/6461.0

ABS (2018b). Internet Activity, Australia (Cat. no. 8153.0). Australian Bureau of Statistics. https://www.abs.gov.au/AUSSTATS/abs@.nsf/allprimarymainfeatures/6342EA2E47A514E0CA25825F0014B06A?opendocument

ABS (2017). The Australian CPI: A Contemporary Measure of Household Inflation (Cat. no. 6401.0). Australian Bureau of Statistics. https://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/6401.0Feature+Article2Sep+2017

ABS (2015). The Australian System of National Accounts: Concepts, Sources and Methods (Cat. no. 5216.0). Australian Bureau of Statistics. https://www.abs.gov.au/ausstats/abs@.nsf/mf/5216.0

ABS (2013). Introduction of Mining Natural Resources into Australia's Productivity Measures (Cat. no. 5204.0.55.010). Australian Bureau of Statistics. https://www.abs.gov.au/ausstats/abs@.nsf/7d12b0f6763c78caca257061001cc588/db52709664089c9eca257d0a00118204!OpenDocument

Ahmad, N. & Schreyer, P. (2016). Measuring GDP in a digitalised economy. OECD Statistics *Working Papers* 2016/07. Paris: OECD Publishing. https://www.oecd-ilibrary.org/economics/measuring-gdp-in-a-digitalised-economy_5jlwqd81d09r-en

Bean, S. C. (2016). *Independent Review of UK Economic Statistics*. Available at: https://www.gov.uk/government/publications/independent-review-of-uk-economic-statistics-final-report

Brynjolfsson, E. & Oh, J. (2012). *The Attention Economy: Measuring the Value of Free Digital Services on the Internet.* Proceedings of the Thirty Third International Conference on Information Systems, Orlando 2012. https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1045&context=icis2012

Byrne, D., Fernald, J. & Reinsdorf, M. (2016). Does the United States have a productivity slowdown or a measurement problem? *Brookings Papers on Economic Activity*, 47(1), 109–182. https://www.brookings.edu/wp-content/uploads/2016/03/byrnetextspring16bpea.pdf

Cardarelli , R. & Lusinyan, L. (2015). U.S. Total Factor Productivity Slowdown: Evidence from the U.S. States. IMF, *Working Paper WP/15/116*. https://www.imf.org/external/pubs/ft/wp/2015/wp15116.pdf

Diewert, E. & Wei, H. (2017). Getting rental prices right for computers: reconciling different perspectives on depreciation. *Review of Income and Wealth,* 63(1), 149–68. https://onlinelibrary.wiley.com/doi/full/10.1111/roiw.12249

ILO (2004). *Consumer price index manual: Theory and practice.* International Labour Office, Geneva. https://www.ilo.org/public/english/bureau/stat/download/cpi/order.pdf

Nakamura, L. & Soloveichik, R. (2015). Capturing the Productivity Impact of the 'Free'Apps and Other Online Media. Federal Reserve Bank of Philadelphia *Working Paper,* N° 15-25. https://conference.nber.org/conf_papers/f84255.pdf

OECD (2017). Measuring GDP in a digital economy. *Global Conference on the G-20 Data Gaps Initiative.* Washington DC, June 14-15, 2017. https://www.imf.org/en/Publications/SPROLLs/G20-Data-Gaps-Initiative#sort=%40imfdate%20descending

Topp, V. & Kulys, T. (2012). Productivity in Electricity, Gas and Water: Measurement and Interpretation. Productivity Commission Staff *Working Paper*, Canberra. https://www.pc.gov.au/research/supporting/electricity-gas-water/electricity-gas-water.pdf

Topp, V., Soames, L., Parham, D. & Bloch, H. (2008). Productivity in the Mining Industry: Measurement and Interpretation, Productivity Commission Staff *Working Paper*, December. https://www.pc.gov.au/research/supporting/mining-productivity

Syverson, C. (2017). Challenges to Mismeasurement Explanations for the U.S. Productivity Slowdown. *Journal of Economic Perspectives,* 31(22), 165–186. https://ideas.repec.org/a/aea/jecper/v31y2017i2p165-86.html. An earlier draft of this paper is also available at https://www.nber.org/papers/w21974.pdf

APPENDIX

The 16 market sector industries for capital simulations

A	Agriculture, Forestry and Fishing	I	Transport, Postal and Warehousing Services
B	Mining	J	Information, Media and Telecommunication Services
C	Manufacturing	K	Financial and Insurance Services
D	Electricity, Gas, Water and Waste Services	L	Rental, Hiring and Real Estate Services
E	Construction	M	Professional, Scientific and Technical Services
F	Wholesale Trade	N	Administrative and Support Services
G	Retail Trade	R	Arts and Recreation Services
H	Accommodation and Food Services	S	Other Services

Does the Digital Economy Distort the Volume-Price Split of GDP? The French Experience

Lorraine Aeberhardt*, Florian Hatier, Marie Leclair*, Benoît Pentinat* and Jean-Denis Zafar*****

Abstract – The slowdown in economic growth over the past two decades is in contrast with the digitisation of the economy. As a result, certain economists are wondering about a possible problem in measuring GDP and, in particular, its volume-price split. The article reviews the methods used by statisticians, with a focus on France, to distinguish changes in price from changes in volume, with a particular attention to the particularities and difficulties linked with the digital economy: communication goods and services, the existence of forms of digital sales, the emergence of new digital services and the development of free services. While the methods put in place deserve to be questioned, a simulation shows that an error in the measurement of the prices of information and communication products is not likely to explain the slowdown in economic growth.

JEL Classification: E31, E01, O3
Keywords: volume-price split, GDP, consumer price indices, digital economy

**Insee (lorraine.aeberhardt@insee.fr; marie.leclair@insee.fr; benoit.pentinat@insee.fr); **DSED, Insee at the time of writing (florian.hatier@interieur.gouv.fr); ***AXA, Insee at the time of writing*

We would like to thank Didier Blanchet, Ronan Mahieu, Guillaume Hourriez, Chantal Cases and Pascal Chevalier, as well as two anonymous reviewers, for their reviews and support.

Received on 28 September 2018, accepted after revisions on 15 May 2019.
Translated from: L'économie numérique fausse-t-elle le partage volume-prix du PIB ? L'expérience française
Citation: Aeberhardt, L., Hatier, F., Leclair, M., Pentinat, B. & Zafar, J.-D. (2020). Does the Digital Economy Distort the Volume-Price Split of GDP? The French Experience. *Economie et Statistique / Economics and Statistics*, 517-518-519, 139–156. https://doi.org/10.24187/ecostat.2020.517t.2027

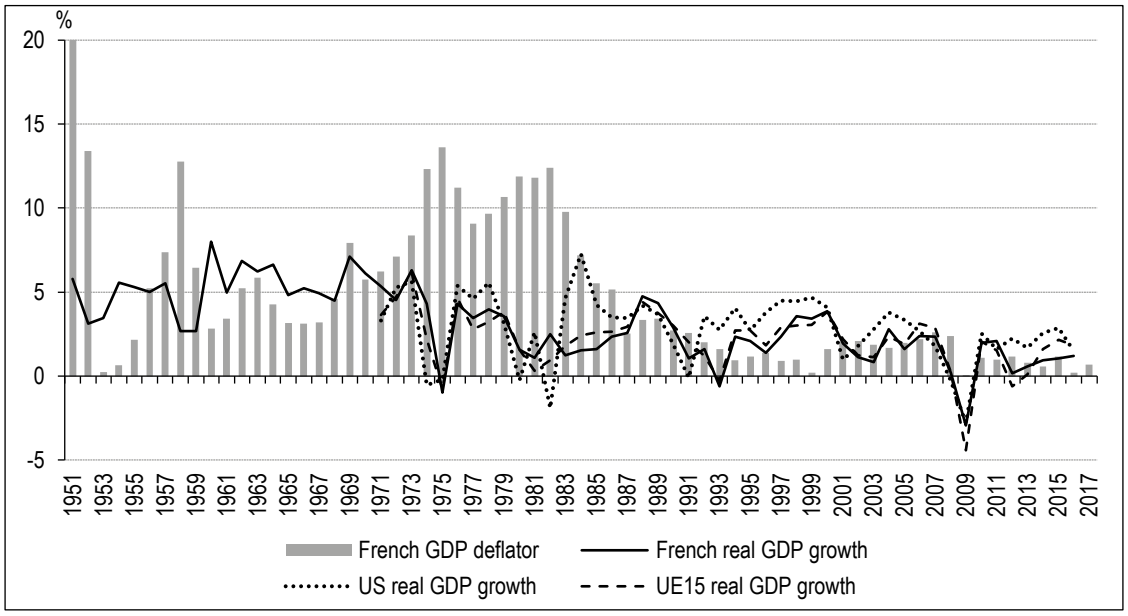
In many developed countries, economic growth has slowed considerably over the past two decades: with an initial decline in the early 1970s, some countries (but not France, Figure I; see also Cette *et al.*, 2016) saw a slight acceleration in productivity in the mid-1990s thanks to the development of new information technologies; since the mid-2000s, this productivity is thought to have slowed (Syverson, 2017). However, over this period, the economy has undergone major upheavals due to digital development: innovation in computer hardware and the integration of artificial intelligence in many goods, the development of communication services and *e*-commerce, as well as the digitisation of cultural content and traditional services and the emergence of new services, particularly intermediation services between private individuals. Therefore, according to the statistics, this digitisation of the economy would not have resulted in an increase in economic growth.

While some economists are looking for economic reasons for the slowdown in productivity, and in GDP more generally, others are wondering about a possible problem with the measurement of GDP (Feldstein, 2017), postulating that economic growth has not slowed but has evaded the traditional tools for measuring GDP. As summarised in Blanchet *et al.* (2018), this GDP measurement issue covers various dimensions. The first is that of the scope of GDP: GDP is not a measurement of well-being (Vanoli, 2002); it excludes a certain number of free productions that are at the origin of a long-standing debate concerning, in particular, the failure to take into account the non-market production of services by households. The explosion of free digital services (provided by companies financed by advertising or by households themselves) has revived this debate (Ahmad & Schreyer, 2016). As a result, there could be a mismatch between the measurement of economic growth and the perception by economic agents of an improvement in their well-being. A second issue is relating to the correct location of the production: globalisation, with the design and then production stages of the various components of a product in different countries, is thought to make it more difficult to locate the wealth created in a given country. A third and final issue is the measurement of the volume-price split and of the relevance of the traditional tools used to measure it.

This article focuses on this issue of the volume-price split. Reinsdorf & Schreyer (2018) believe there are three reasons why the digitisation of the economy would affect the volume-price split: the failure to adequately take into account changes in quality when a new generation of an existing product is introduced, a delay in taking into account completely new digital products and, finally, the failure to take into account many free digital products due to the lack of imputation of virtual prices.

Various articles have sought to measure the impact of these volume-price splitting problems

Figure I – **Annual growth rate for GDP and its deflator since 1951 (%)**



Sources: OECD, real GDP growth rate.

on the measurement of inflation or GDP. Aghion *et al.* (2019) thus estimate the underestimation of US growth, attributable to the failure on the part of the price indices to properly take into account the appearance of new products and the replacement of businesses by others, to be 0.7 point per year from 2006 to 2013. Applying the same model to French data, Aghion *et al.* (2018) obtain an underestimation of 0.4 point per year. However, this measurement problem already exists in previous periods and therefore cannot fully explain the economic slowdown. Groshen *et al.* (2017) estimate a measurement error bias of 0.4 point of growth per year that has been roughly stable since the mid-1990s. A bias of roughly the same extent was estimated by Reinsdorf & Schreyer (2017).

This problem and these criticisms are far from new. It should be remembered that, in the 1990s, the Boskin Commission Report (Boskin *et al.*, 1996) found that US inflation was overestimated by around 1.1 points per year, due to the failure to adequately take into account changes in consumer behaviour in the price indices. This report had given rise to questions from most price statisticians. In the case of France, the impact on inflation had been deemed to be much smaller (Lequiller, 1997). Since that report, a certain number of corrections have been made to better take into account replacements between products (faster updating of weights – which, in the case of France, had little impact in reality), but most of the questions raised at the time remain relevant today.

This article illustrates and discusses the difficulties of the volume-price split linked to the digitisation of the economy in the context of national accounts and the French consumer price index. The first section describes the methods used by the national accounts to carry out volume-price splitting of the GDP and the second then covers the specific difficulties of the volume-price split for products relating to information and communication technologies (ICTs). The third section looks at the more general changes to commercial offerings due to digitisation and the final section then provides a simulation of the impact that a measurement error concerning the prices of ICT products could have on the slowdown observed in the French GDP growth.

1. The Volume-Price Split in the French National Accounts

The national accounts measure all aggregates of supply and demand in value terms, i.e. in current euros. In order to determine whether one of these aggregates has increased between two periods, the accounts distinguish a price factor that reflects the movement in prices and a volume factor that measures the change in aggregates adjusted for the effects of inflation.

1.1. The Notion of Volume

Growth in volume, which is used in particular for macroeconomic and business steering, is a difficult notion to define precisely. In a "simple" economy, primarily composed of "physical" goods, this notion would be relatively easy to describe. In such an economy, the growth of GDP in volume terms would represent the change in the quantities consumed, invested and stored in the territory or exchanged with the rest of the world.

However, even in such a simple economy, measuring the volume of GDP faces several difficulties. First of all, adding up quantities of basic products makes no sense: these quantities must be commensurable and that is why the estimation of the volume of GDP is based on a Laspeyres formula[1] which results in the assignment to these basic quantities of the price they have at a given period (see Online Appendix C1 – link to Online Appendices at the end of the article). This accounting is based on the assumption that the relative prices of these basic products reflect the difference in the utility that can be derived from them.[2] This assumption is debatable and we will see that many questions relating to the measurement of the digital economy relate to this point.

Moreover, improving the quality of a good should result in a boost to real GDP growth: for example, a garment designed using a new, very high quality fabric, sold at the same price as a "traditional" garment, is likely to have a longer lifespan. Households that buy the very high quality garment then see its utility increase with the arrival of the new good. This increase in utility must be reflected in an increase in volume, as the role of the volume is to measure changes in both quantity and quality.

1. Using a Laspeyres formula is the most commonly used solution, mainly for practical reasons (simplicity of the formula and availability of information); however, there are many forms of indices that enable aggregation of these quantities, including superlative indices that make it possible to better take into account replacement effects.

2. The European System of Accounts (2010) specifies the various cases in which a difference in price at a given time cannot reflect a difference in product quality: lack of competition, imperfect consumer information, price discrimination, etc.

1.2. A Complex Measurement

To move from the notion of value to the notion of volume, in most cases, the national accounts rely on price indices; these make it possible to deflate aggregates in value terms by "pure" price changes (excluding any change in the structure of the aggregate or in the quality of the products that comprise it). To obtain such a measurement, the price indices are usually fixed basket indices, i.e. the prices of identical products are tracked over time, with their weight in the index also being fixed over time. This method is well suited for a stable economy without product renewal or changes in consumption. It is less well suited to an economy in a constant state of flux.

1.2.1. Taking Replacement Effects into Account

A first difficulty is that the behaviour of economic agents generally changes in accordance with prices. Let us take household consumption as an example: an increase in the price of a product will probably lead to the consumer substituting it for a less expensive similar product; the effect of this replacement by the consumer will make it possible to limit their loss of utility due to the increase in prices. If we wish to define inflation as the change in consumer income enabling the consumer to achieve the same level of utility as in the previous period and despite the increase in prices (the so-called "constant utility" price index, see Magnien & Pougnard, 2000 and Sillard, 2017), then it is clear that we wish to take account of these replacement effects. A volume index based on past fixed price weights will tend to give too much weight to the product for which the price decreases, while overlooking these replacement effects. In order to limit this phenomenon, which was criticised in the Boskin report for estimating US inflation in the 1990s, the French national accounts have been chaining these changes in volume annually since the 1995 base year.

The same difficulties arise with fixed-based price indices, which is why the consumer price index (CPI, see Online Appendix C2), the main index used by the national accounts, has also practised chaining since the 1970s: the basket of N products, the prices of which, p_i, are monitored each month and the associated weights q_i are fixed during year a, but are renewed each year; the CPI is thus an annual fixed-basket index.

More precisely, at the most aggregated level, the CPI is a Laspeyres index: in the course of a year, it weights the price ratio by the quantities observed in the past (period a-1). A CPI can thus be defined, during year a, as an index worth 100 in December of the previous year.

$$I_t^{12,a-1} = \frac{\sum_{i=1}^{N} p_i^t q_i^{a-1}}{\sum_{i=1}^{N} p_i^{12,a-1} q_i^{a-1}} \quad (1)$$

where $I_t^{12,a\text{-}1}$ is the price index observed in month t of year a, expressed with a reference of 100 in December of year a-1, p_i^t is the price for product i in month t and $q_i^{a\text{-}1}$ is the quantity of product i consumed in year a-1, while $p_i^{12,a\text{-}1}$ is the price of this product i in December of year a-1.

To obtain an index over a longer period ($I_{a,t}$), this index with a base of 100 in December of the previous year is chained to past indices:

$$I_{a,t} = \frac{I_{a-1,12} \times I_t^{12,a-1}}{100} \quad (2)$$

At the most disaggregated level, the CPI generally uses price aggregation formulas (Jevons formula) that, unlike Laspeyres formulas, take into account replacements between products (see Lequiller, 1997 for a discussion on this issue).

1.2.2. The Appearance of New Products and Discontinuations

A second difficulty relates to the renewal of products, which raises two questions: the estimation of the price of the new product, "cleaned" of any possible quality effects, and the date of inclusion of this new product in the basket of goods, insofar as the new product can replace an existing product. These issues are all the more important because it is sometimes through this renewal of products that a large part of growth or changes in prices occurs, making this issue a central question in the measurement of digital growth (Lequiller, 2000; Feldstein, 2017); indeed, products appear and are discontinued constantly while price indices follow a fixed basket of products.

In simplistic way, these appearances and discontinuations of products can correspond to two extreme cases. In the first one, the product is completely new/innovative and does not, even partially, replace an already existing product; in this case, the product must be taken into account in the measurement of inflation and GDP and in line with its economic weight and price. The annual chaining of the consumer price index makes it possible to revise the basket of goods and services for which the prices are tracked

each year and to add these new products: the products and weights tracked in *a* and in *a*-1 (equation 2) may indeed differ. Some critics claim this method fails to take into account the impact of the actual appearance of the new product on consumer well-being. From a theoretical point of view, they suggest estimating virtual reservation prices for these products, prices at which there is no longer any consumer demand, and quantifying the price drop linked to the appearance of this new product (the difference between the first price observed for the new product and this reservation price). This type of suggestion remains relatively theoretical and academic, especially in view of the estimation costs (see, for example, Diewert & Feenstra, 2018). Furthermore, it is based on the idea that if the new product does not exist, it is because there is no demand from the consumer, whereas very often it is because the innovative product has not been invented: there is therefore no reservation price. Finally, in general, new products have little impact on expenditure when they are introduced to the market and their omission, prior to the annual update of the index, is unlikely to cause a significant bias on inflation (and hence on GDP).

The second extreme case of product appearances and discontinuations corresponds to the appearance of new generations of an existing product that is already tracked in the basket of the price index and which they replace. In this case, in order to correctly calculate a price index, the old and new products will be matched and an adjustment[3] will be made to neutralise the difference in quality between the two products, so as to measure price evolution at constant quality. There are various methods to measure this quality adjustment (see IMF, 2004 for a review of all these methods): explicit methods seek to measure the difference in quality between the products and to determine a price difference that is justified by this difference in quality. These methods include option pricing and hedonic methods. The latter methods are based on the notion that the price of a product can be broken down in accordance with its main characteristics, which determine the differences in quality. The price of each of these characteristics can then be estimated by econometric regression. The pure change in the price will be measured by the change in the prices that cannot be explained by a change in these characteristics. Hedonic models seem very promising for measuring inflation in a context of frequent product renewal. However, in practice, their use remains limited: in the case of the French consumer price index, they are used for only a few durable goods.

Quality adjustments are most often estimated using implicit methods and, more specifically, overlap methods (in particular, the bridged overlap method). These methods are based on the assumption that price differences between two products at a given time reflect differences in the quality of those products. In the event that the prices of the discontinued product and the new product are not observed at the same time, the past price of the new product can be imputed on the basis of changes in the prices of similar products present in both periods.

The overlap method therefore assumes that prices are competitive prices, reflecting differences in marginal utility taken by the consumer, and that they adjust very quickly. However, pricing policies for new or ageing products may not respect these assumptions: new products may be offered at very low prices to gain market share or, on the contrary, at relatively high prices as producers rely on the attractiveness of novelty; conversely, end-of-life products may see their prices fall to be sold before the new product is fully introduced. Pragmatically, to avoid measuring these product life cycles, end-of-life products are excluded from the indices and new products are only introduced once they have been established on the market: the inclusion of these new generations of products in the price index before their inclusion in the calculation of the index is then done indirectly, through the change in the prices of existing competing products.

The fact that price differences observed at a given point in time reflect differences in utility for the consumer is crucial beyond the overlap methods. A similar assumption is found with hedonic models, as the price of the characteristics is estimated based on the price of different products at a given point in time, with the assumption that the price differential of these products reflects differences in characteristics.

This assumption, if correct, ensures that the effects of replacements between products are taken into account even when the new products are not included in the price index. Indeed, two extreme cases of product introduction have been presented above in an exaggerated manner: the cases of a completely innovative product and of

3. This adjustment involves equally, in equation (1), either making a correction to the observed current price p_i^t*, or changing the price of the reference period* $p_i^{12.a-1}$*.*

the new generation of an existing product that it replaces. In reality, there is a continuum between these two extreme cases, with innovative products fulfilling product functions that previously existed. Let us take the case of the very first smartphone as an example: it does not replace the traditional mobile phone and is introduced as a new product; however, it substitutes for the latter. Even without the introduction of the smartphone in the basket of the price index, the existence of competition from the smartphone is expected to be felt by a downward shift in the prices of the competing products, which are tracked in the price index. As indicated above, the impact of new products on prices would therefore be measured indirectly via the evolution of the prices of existing competing products.

1.3. A Variety of Sources and Methods

Before describing more precisely the issues raised by the digital economy in terms of price monitoring, let us reiterate, however, that the estimation of GDP in volume terms is not simply a blind and systematic application of a price index.

Far from being established "globally", by deflating GDP in value terms by a single price index, the measurement of GDP in volume terms is, on the contrary, carried out at a very fine level of nomenclature: for each product, the various components of the national accounts (household consumption, investment, foreign trade, production and intermediate consumption) are computed in terms of both value and volume, based on various pieces of information. It is all then added up, to measure each component in terms of volume, at an aggregate level over all products, which then makes it possible to determine real GDP.

For each product and each aggregate, the most appropriate index is selected: the aforementioned consumer price indices thus make it possible to measure consumption in terms of volume, the industrial producer price indices and service producer price indices make it possible to measure production in terms of volume and the industrial producer price indices for foreign markets make it possible to measure imports and exports of goods in volume terms, etc. (see Online Appendix C3).

In addition, national accountants carry out work to ensure the consistency of all this information, which may lead it to deviate from the price indices (see Online Appendix C4). Volume indices may be used in some cases: these are generally quantity indices. In this case, national accountants seek to determine the variation in quality by differentiating as many qualities of a product as possible. By way of example, the volume-price split of agricultural products is done using production quantity indices at a very fine level (durum wheat, soft wheat, barley, etc.).

In the end, the volume-price split in the national accounts cannot be summed up by simply taking into account a single price index: the methods used are varied and multiplied by the number of products on which the analysis is carried out (Aeberhardt & Bidault, 2018); making the different sources consistent (in terms of value, volumes, prices or quantities) makes it possible to go beyond the limitations associated with specific sources, as will be seen below when discussing communication services.

2. The Difficult Volume-Price Split for Information and Communication Technologies

Information and communication technologies (ICTs), as a vector of the digitisation of the economy, are the focus of major volume-price split difficulties. This is not a new issue: it was at the heart of the Boskin report and the questions about the low productivity growth in the 1990s in the midst of the IT revolution. While these technologies are no longer "new", the difficulty of measuring their prices, due to continuous innovation, remains a focus of the debate around measuring growth (Feldstein, 2017). International comparison work (Ahmad *et al.*, 2017, Reinsdorf & Schreyer, 2018) shows strong divergences in the prices of these products, even though the spread of these technologies and, in general, their importation (at least for goods) would suggest some price convergence within developed countries. Economists then point to differences in methods for measuring changes in the quality of these products and use the cross-country difference in price dynamics as a benchmark for the error in measuring the volume-price split.

2.1. Technological Goods, Frequent Innovations the Quality of Which is Difficult to Measure

2.1.1. Very Different Price Dynamics Depending on the Adjustment Methods Used

For information and communication technology goods alone, the difference between the French

and German harmonised consumer price indices (HCPIs) since the early 2000s has been more than six percentage points per year for telephony and fax equipment (including mobile phones, in particular) and almost three percentage points per year for audiovisual, photographic and data-processing equipment (including computers and tablets, in particular) (Figure II).

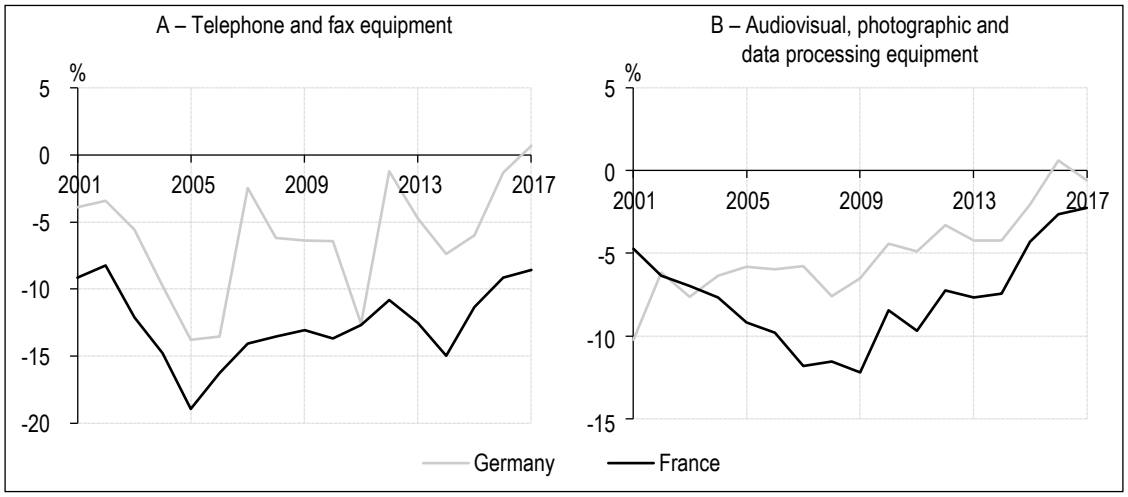
Given the turnover of these products, it is reasonable to assume that a large part of the change in their value takes place when new products are introduced. Therefore, quality adjustments are crucial. However, the methods used to make these adjustments are different in the cases of France (mainly an overlap method) and Germany (hedonic models). In the case of France, in relation to these highly technological products, almost all quality adjustments are made using an overlap method, considering that the difference in price observed between the new product and the discontinued product is a difference in quality. Hedonic models have been tested, but have been found to be of poor quality, either because the number of observations was insufficient for estimating the coefficients of the models in a robust manner, or because of the difficulty of modelling the price itself in accordance with observable characteristics. Hedonic models are based on the assumption that observable characteristics, which are stable over time, determine the quality and hence the price of products. In the event that these characteristics are themselves subject to major innovations, and are difficult to identify, hedonic models do not provide a solution to the problem of measuring the quality of new products.

The direction of the bias for each model is difficult to estimate. To illustrate the impact of the quality adjustments, a simulation is proposed for the French CPI from 2016 to 2018 with no quality adjustment made for the discontinued and replaced products belonging to the sector of telephony and fax equipment and audiovisual, photographic and data processing equipment, i.e. the new products are considered to be equivalent to previous generations in terms of quality. Without any quality adjustment, the overall index would have been 0.1 percentage point more dynamic per year (Figure III). For this sector, the new products are indeed more expensive, on average, than those they replace. Overlap methods neutralise the entirety of the price difference linked to the introduction of a new product, as a difference in quality. If the new product is offered at a price higher than the new quality it incorporates, relying on the attractiveness of the novelty, the overlap method will underestimate inflation. Hedonic models, on the other hand, only neutralise the price difference linked to changes in characteristics; however, if the economic model omits a characteristic (particularly a new characteristic specific to the new generation of products), it underestimates the change in the quality incorporated in the new product and overestimates inflation. It is therefore not surprising that adjustments made using overlap methods find less dynamic price changes than hedonic models.

2.2.2. Minimal Impact on the Measurement of GDP

The impact of these potential problems in measuring the prices of ICT goods on the

Figure II – **Annual average changes in the harmonised consumer price index in France and Germany (%)**



Sources and coverage: Eurostat, HICP database 2015; HICP for COICOP items 8.2 and 9.1.

measurement of the volume of GDP must be put into perspective. In France, as in many developed countries, household consumption of ICT goods is mainly based on imported products. French household consumption of computers and peripheral equipment (or communication equipment) thus represents, on average, 50% (or 40%) of the value of imports over the period 2000-2016.

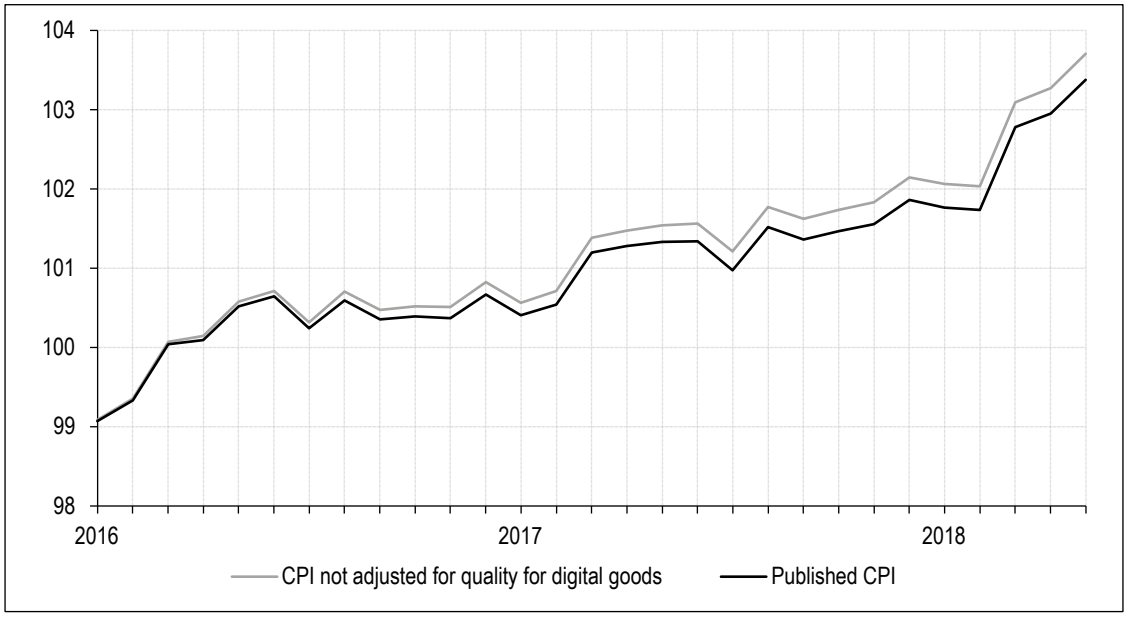
Consequently, and provided that the deflator for imports and the deflator for household final consumption have similar measurement problems, the impact of an inadequate volume-price split of consumption on the measurement of GDP is probably almost neutral, with an underestimation of consumption in volume terms resulting in an underestimation of imports on the same scale. National accountants carry out work to ensure the consistency of the deflators for the consumption and import of these products. In the event that the indices (CPI and import price indices) diverge, they proceed by arbitration, primarily in favour of the CPI, to bring the two deflators closer together: Figure IV presents the CPI and the producer and import price index in industry (*Indice des prix à la production et à l'importation dans l'industrie* – IPPI) as spontaneously measured, together with the consumption and import deflators used by the national accountants after arbitration.

2.3. Communication Services, Constantly Renewed Commercial Offers

The volume-price split for communication services also raises genuine difficulties. This difficulty is not only due to the innovations in this sector (development of the internet, mobile telephony, mobile phone data, 3G technology, 4G technology, etc.), but also to the extremely complex pricing of these services. Excluding innovation, the commercial offers proposed by operators generally cover more than one service (SMS, data, voice, domestic and mobile, national and international, etc.), with pricing that depends on consumption in a non-linear manner (a basic flat rate regardless of consumption, then specific pricing once beyond the included allowance). In addition, pricing changes are generally made by reviewing the scope of these commercial offers. Consequently, overlap methods are completely unsuitable since they would, by nature, mask any change in price, by neutralising them as a difference in quality. Finally, the characteristics of these commercial offers are often ill suited to hedonic models: how, for example, can we manage the transition to unlimited offers, knowing that ultimately the consumer will not have the use of them?

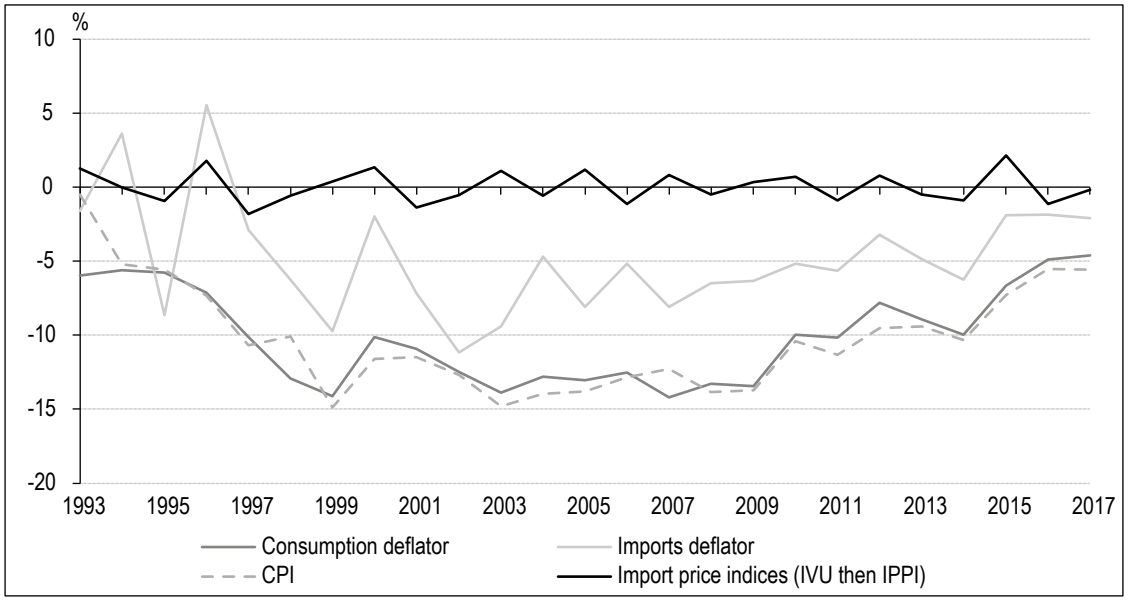
For all these reasons, European consumer price indices favour the so-called "constant use"

Figure III – **Overall consumer price index adjusted and not adjusted for the quality of digital goods (year 2015=100)**



Reading Note: The overall consumer price index was 103.4 in May 2018; if no quality adjustment had been made for COICOP items 8.2 and 9.1 between January 2016 and May 2018, then this index would have been 103.7 in May 2018.
Sources and coverage: CPI, database 2015 ; Metropolitan France.

Figure IV – **Annual change in price for computer, electronic and optical products (%)**



Sources and coverage: National accounts, database 2014; CPI, database 2014; France.

indices for communication services (Eurostat, 2017). These indices, which are an approximation of constant-utility indices, follow the minimum expenditure that a consumer must make to satisfy their specific use and are constant between two periods (Magnien, 2003). Thus, for example, the minimum expenditure of a consumer who usually sends 10 SMS messages per month will not be changed if all packages now offer unlimited sending of SMS messages for the same price: they will effectively not use it.

However, this method poses a certain number of difficulties. First of all, it is necessary to be able to describe consumer usage in a precise manner: it is not enough to simply track the minimum expenditure of a single rough consumer profile; the calculation of an index needs to be representative of all consumers. In the previous example, a consumer who usually sends 10 SMS messages for a package with a limit of 50 would indeed see their expenditure decrease. Fortunately, in the case of communication services, and contrary to other services for which one might be tempted to apply these constant use methods, the French regulator Arcep (*Autorité de régulation des communications électroniques, des postes et de la distribution de la presse* – an independent administrative authority), has very rich information on the operators' customers, making it possible to divide them into a set of consumer profiles.

A second difficulty in using this constant use method lies in the simplified modelling of consumer behaviour: in the case of the French CPI, it is assumed that consumers are aware of the various offers from operators and constantly adjust their packages to minimise their expenditure. In practice, there are a number of frictions (research costs, together with the costs linked to commitments) that are difficult to model without making the estimation of the communication services index too complex. In practice, the assumption made is that the consumer changes their package only within the offers provided by the same operator, thus disregarding mobility between operators, which amounts to treating each operator as providing a different product.

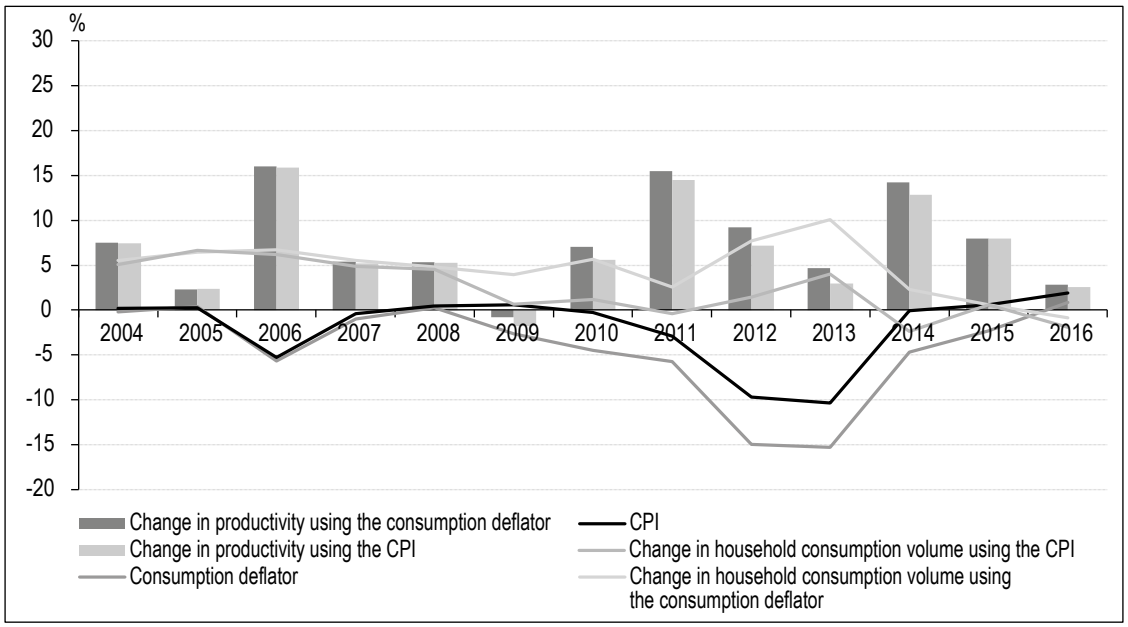
This realistic assumption in an initially highly segmented market has become less relevant with phone number portability. In particular, it posed a problem with the arrival in France of a fourth operator on the mobile phone market in 2012-2013: the arrival of this operator was accompanied by a massive transfer of subscribers from the old operators to this new one; as the prices charged by the latter were much lower, the revenue in telecommunication services fell while the number of minutes and SMS messages exploded. However, due to the modelling used, the consumer price index treated the new operator's packages as new products for the CPI and the price differential with the traditional packages as a quality differential. The CPI did fall significantly in 2012, but only due to the adaptation of the prices of the traditional operators, in response to the arrival of the new competitor. However, the

adaptation of the prices of the traditional operators was done on a gradual basis only and lagged behind the transfers to the new operator. The use of the CPI to measure the volume of telecommunication services, in this context, would have led to a sharp decline in consumption volumes, contrary to available information on consumption in terms of quantity. French national accountants have therefore preferred to estimate a volume index by calculating a weighted average of the basic volume indices (domestic telephone and internet and mobile telephone and internet), based on data on quantities of SMS/MMS messages or telecommunications minutes (Bessone *et al.*, 2014). This method makes it possible to avoid incorrectly measuring a decline in volumes in a highly competitive context: from 2011 to 2014, productivity in the telecommunications service sector thus grew by an average of 9.4% per year according to the national accounts, compared with 7.7% if the CPI had been used as the deflator (Figure V). As the market stabilises, this method is gradually being abandoned in favour of the CPI again.

2.4. The Difficulty in Measuring the Volume-Price Split for Investments in Software and Applications

In their international comparison, Ahmad *et al*. (2017) show that French price indices for investment in software and applications, generally services producer price indices (SPPIs), are around average for the countries described in their article. The disparity in the methods used for volume-price splitting for software is also a result of the lack of harmonisation of methods for measuring investment in software, which limits international comparability. Investment in software actually includes a variety of items, the number of which makes estimates more complex: expenditure on data processing and websites and expenditure on standard software, measured in France on the basis of business statistics, expenditure on specific software, measured though the use of computer engineering service companies and also estimated on the basis of business statistics but with the best possible extraction of anything relating to intermediate consumption. A significant portion of expenditure on software (over 30%) is also due to internal expenditure within companies to develop custom software, measured by French national accountants using data on remuneration by selecting the professions likely to be involved in these developments. In the absence of specific information on the real price of such expenditure, the market price for "external" expenditure is generally applied to them. Thus, the complexity and wide variety of types of software expenditure, together with relatively little international coordination in the area, makes it difficult to evaluate.

Figure V – **Change in apparent productivity of the work of the telecommunications service sector using different deflators (%)**



Reading Note: In 2012, the productivity of the telecommunication service sector grew by 9.2% according to the national accounts. It the latter had used the CPI as a deflator for these services, a 7.2% increase in productivity would have been measured. The CPI for telecommunications services in fact fell by 9.7% in 2012, compared with -14.9% for the telecommunications services consumption deflator.
Sources and coverage: National accounts, database 2010; CPI, database 2014; France.

3. The Digitisation of the Economy Changes the Existing Commercial Offer

Aside from the difficulties inherent in the volume-price split for ICTs, the digitisation of the economy is generating a certain number of phenomena for which a price and a volume must be identified: the emergence of a new form of sales, new services that shake up the traditional players, the production of free services and new price formation methods.

Before even addressing the question of their volume-price split, it should be noted that the emergence of the digital economy raises questions about the measurement of GDP in value terms. In the case of commercial products, taking into account this economy depends on its integration into the traditional data sources used by national accountants (in the case of France, panel data for household consumption and tax sources for production) and on its location (notably for the benefits of intermediation platforms). The case of free products, together with the economy based on sharing and the production of household-to-household services by intermediation platforms (*Airbnb*, *BlaBlaCar*, etc.), also raises questions about the scope of GDP (Bellégo & Mahieu, 2016; Blanchet *et al.*, 2018; Ahmad & Schreyer, 2016). This article is restricted solely to the issue of the volume-price split of this digital economy with a given scope of GDP.

3.1. The Appearance of a New Form of Sale: *E*-Commerce

The spread of the internet has allowed the emergence of a new form of sale, *e*-commerce. However, the quality of a good and its price do not depend solely on its intrinsic quality, they also depend on the commercial service associated with its sale: a single product may thus be sold at a higher price in a local business than in a hypermarket because the associated commercial service (in this case, proximity) is considered superior. Measuring the quality of the commercial service is probably even more complex (and less observable) than measuring the quality of the product actually sold. Faced with this difficulty, price statistics in France have adopted solutions that are sometimes contradictory, with ex-post work again required to ensure the consistency of the national accounts. In the case of the industrial producer price index, the prices measured are "ex-works" prices declared by the producers, regardless of the distribution channel chosen by the producers: the change in the method of sale for one of the producer's products is therefore neutral on the index measured. In contrast, the consumer price index is based on the prices in given outlets specified in its sample; the form of sale is an integral part of the quality of the product. Therefore, the appearance of a new form of sale is taken into account by the CPI as the appearance of an entirely new product and this new form of sale is only integrated through chaining at the time of the annual update of the basket of goods and services tracked by the CPI. The fact that prices are lower online (a finding that remains to be discussed, as shown in the review of the literature by Bellégo & Mahieu, 2016) would then not result in a fall in prices in the CPI but in a fall in quality. The assumption that the price difference reflects a difference in quality is of course debatable but, as with other quality issues, it is difficult to make an objective judgement about a difference in quality beyond the summary measurement of preferences that prices should reflect: online purchasing opens up the possibility of purchasing 24 hours a day, 7 days a week, without any cost of travel, but conversely, the buyer does not see the product or benefit from the seller's advice and product delivery is not immediate, etc. While the difference in price between online and physical retailers reflects, beyond a difference in quality, an improvement in competition through the arrival of new market players, the appearance of such new market players can also be expected to result in a fall of the prices charged by existing retailers. Therefore, the arrival of *e*-commerce will be measured in the CPI, but indirectly via the fall in prices measured in traditional outlets.

Finally, it should be noted that the question of a risk of a bias in the volume-price split due to the arrival of a new form of sale has already occurred in the past: similar debates about "purchasing channels" took place with the development of supermarkets and hypermarkets, then the hard-discounters (Lequiller, 1997). In the 1980s, the increase in the market shares of supermarkets and hypermarkets would have resulted in CPI growth around 0.2 percentage point lower per year (Saglio, 1995) if this had not been neutralised as a quality effect in the French CPI. American studies, in turn, estimated the maximum effect of purchasing channels to be around 0.25 percentage point during the 1980s (Reinsdorf, 1993).

3.2. The Appearance of New Services Online, Competing with Existing Services

The spread of the Internet has not only led to the emergence of a new form of sale, it has profoundly changed the services offered, with

the enrichment of existing services, the emergence of new market players and the appearance of new services that are entirely free of charge.

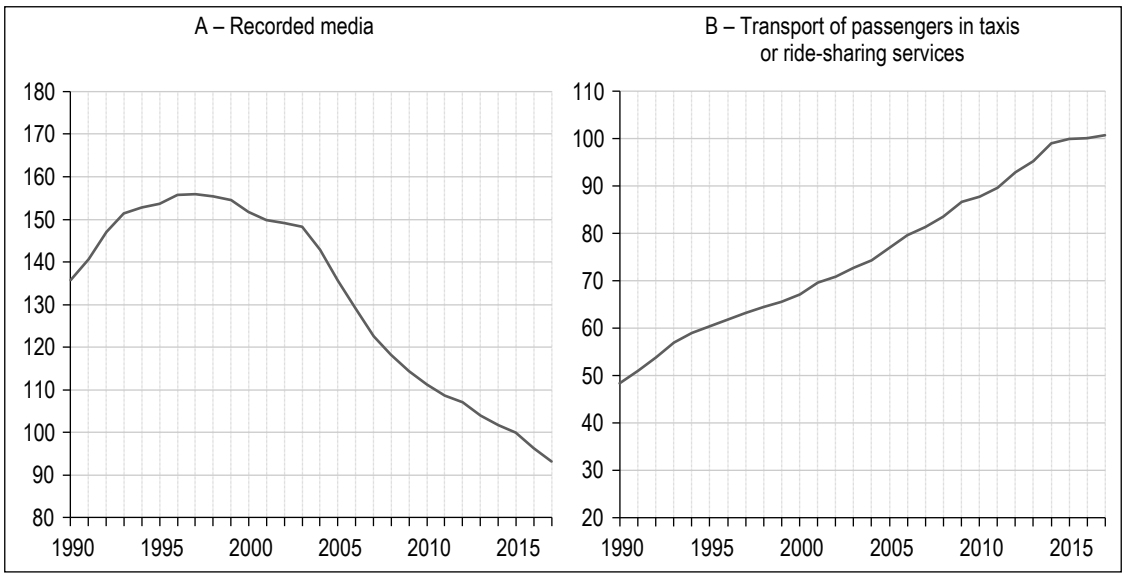
3.2.1. A Change in Existing Services, Without the Arrival of New Market Players

Many services have benefited from the digitisation of the economy. These include banking services and the ability to monitor one's bank account online, submitting an insurance claim online, receiving digital invoices (electricity, telecommunication services, etc.). The modification of these traditional services undoubtedly leads to a change in their quality, without it being possible to say authoritatively whether it is increasing or decreasing (the perception of quality will differ greatly depending on whether the consumer prefers paper or electronic billing, human contact or digital flexibility). Taking this change in quality into account in the volume-price split for these services will generally depend on the observation that can be made of it. In the absence of being able to do better, in the majority of cases, it will be considered that the modification does not substantially change the service, which generally remains elsewhere in reality (having electricity, holding a current account, etc.). In some cases, however, particularly when the internet was becoming widespread, online account access services could be one of the pricing parameters (for example, for banking services, in the case of France): in this case, this access was indeed taken into account as a change in quality.

3.2.2. New Market Players

However, the digitisation of the economy was able to bring about a more profound change in the supply of commercial services, with the emergence of new services replacing traditional ones. In this category, we can mention the development of streaming as a replacement for purchasing DVDs or CDs and the development of ride-hailing services or accommodation rentals through intermediation platforms (*Uber* and *Airbnb*). These products were introduced as entirely new products in the annual updates to the basket of the French CPI. Is GDP growth being underestimated in volume terms due to the inadequate taking into account of the fact that, by replacing existing products, these new services could make it possible to offer consumers a less costly alternative? For example, while a DVD is not equivalent to a streaming subscription, watching a film is now cheaper for consumers, on average. Again, statisticians have little choice but to rely on price differentials for measuring differentials in the utility or services provided by a particular product. Depending on the consumer, the relative utility of a streaming subscription or a DVD will be quite different. Faced with this difficulty, the price index will record the impact on streaming prices as a replacement for DVDs indirectly, via the price index for DVDs, which is expected to fall due to the competition from streaming. In fact, the price index for recorded media has fallen continuously in France since the early 2000s (Figure VI).

Figure VI – **Consumer price indices for recoded media and taxis and ride-hailing services (year 2015=100)**



Sources and coverage: CPI, database 2015; France.

However, this impact of substitution effects on prices can only be recorded in the CPI if the prices of competing products adjust and if the market is competitive. Ride-hailing services were also introduced into the consumer price index as new products (what ride-hailing services offer is not equivalent to that offered by taxis, particularly while on the go). However, the maximum fares that can be charged by taxis as established by prefectoral decrees (and tracked by the French CPI) have not fallen since the development of ride-hailing services. The CPI for passenger transport by taxi or private hire vehicle has thus been fairly stable since 2014. Should it have decreased following the arrival of competition from ride-hailing services? In view of the barriers to entry, this market was also characterised by a supply deficit and unmet demand at a given regulated price: it is therefore possible that the opening up of the market to competition may have made it possible to meet greater consumer demand without reviewing prices downwards. The measurement problem in this case may therefore have had a relatively small impact on the measurement of GDP.

3.2.3. The Development of Free Services

New services may also be free of charge, such as *Google Maps* or Wikipedia. The production of these services, financed by donations, online advertising or even the commercial exploitation of data gathered on their consumers, has no explicit counterpart in household consumption expenditure. In fact, as consumers can benefit from these services for a zero price, no household consumption expenditure in value terms is recorded in the accounts in relation to them, and no price is associated with them.[4] Therefore, and with even greater justification given that they are replacing old commercial services (a paper card or a dictionary), one might wish to record a drop in inflation or an increase in GDP when these free services appear on the market. From a conceptual point of view, the imputation of a virtual price, before the emergence of the new service, would make it possible to account for a price drop (switching from the reservation price to free) – see Reinsdorf & Schreyer (2018) for a discussion on this process. In view of the difficulty (and sometimes bias) in estimating such reservation prices, the fall in the price of the competing commercial services is recorded in the consumer price index only if it occurs: for example, the prices of publishing services have fallen by 1.2% since 2009, while consumption in volume terms has fallen by 3%.

3.3. New Price Formation Mechanisms

The existence of the internet as a source of information and/or place of purchase for the consumer, in theory, seems to bring price formation closer to the assumptions of perfect competition: the consumer would no longer have to pay any cost for information (they can simply search on the internet or use a price comparison site) or travel (to buy one product rather than another). Consequently, this would be more in line with the assumption that relative product prices equal the marginal utilities derived by consumers, an assumption generally required to measure quality differentials for price index calculation.

However, existing studies on various sectors (see the review of the literature on the impact of the internet on prices by Bellégo & Mahieu, 2016) do not make it possible to demonstrate that the internet offers systematically lower prices than in physical outlets; they also point to the persistence of high price dispersion on the internet. Research costs are thought to remain significant for the consumer, with information sometimes being limited on websites, particularly with regard to product quality.

While the internet does not seem to have revolutionised competitive price formation, it sometimes leads to new pricing practices, in particular by encouraging the formation of personalised prices that are differentiated according to the customer. Yield management policies have thus been developed, largely facilitated by the possibility of having a shared information system and allowing real-time price adjustments. They have gradually expanded from air transport to other services (tourist packages, hotels, rentals, etc.), which have gradually abandoned their catalogue prices. These policies consist of optimising prices in real time in accordance with demand, in a context in which the volume of the service produced is difficult to adapt and cannot be stored, but is generally reserved in advance.

For price statisticians, these yield management policies lead to a wide range of prices for a single service: within an aircraft and for the

4. On the issue of their valorisation, see Bourgeois (in this issue). It should be noted that these free services are not a new phenomenon (television programmes are an old example). However, they can be partially included in the GDP, for example as the production of an advertising service (see Bellégo & Mathieu, 2016). For a discussion on how to take them into account in GDP and on their financing see Ahmad & Shreyer (2016).

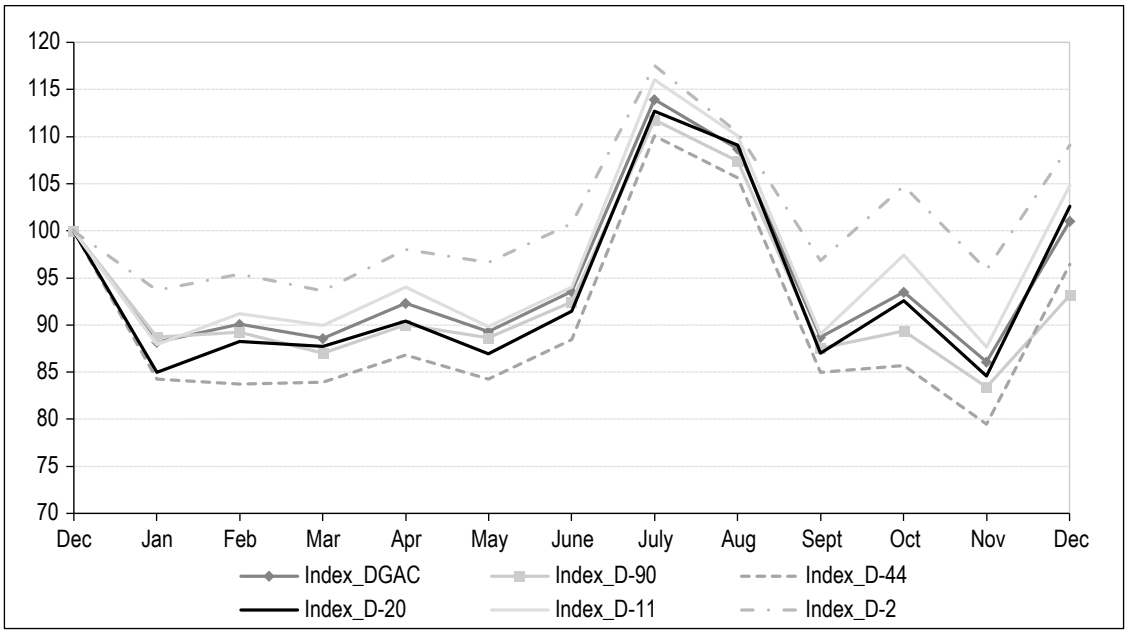
same journey and level of comfort, passengers will have paid different prices. What prices should be used in this context to measure inflation? First of all, the volatility and wide range of these prices make it necessary to increase price observations: the price of a single service will be observed at different intervals; the price of a plane ticket will thus be recorded the day before departure, two weeks before, one month before, three months before, six months before, etc. Webscraping techniques (robot-assisted data collection online) facilitate the mass observation of these multiple price collections (see Online Appendix C5). Then there is the issue of aggregating these multiple prices in a context in which pricing dynamics (and not only price level) are quite different, depending on how far in advance the plane ticket is reserved (Figure VII). Should an average price effectively paid by passengers on a flight be calculated in this way? Aside from the fact that information on the number of tickets sold at different intervals is generally not known (the prices on offer can be observed, but it is more difficult to observe the reality of sales), to what extent are constraints on purchase dates part of the quality of the service provided?

Specifically, the French CPI for air transport is based on a fixed sample of destinations, for which flight prices are collected almost daily, for various consumer profiles (flexible or non-flexible prices, in particular) and in accordance with the reservation interval. The prices are aggregated using fixed weights for each profile, destination and reservation interval. As a result, due to the fixed nature of these weights, any changes in the behaviour of consumers, who in exchange for work to optimise ticket selection could for example turn towards cheaper tickets, are not regarded as a price effect but as a quality effect. The national accounts, which use the CPI as a deflator for these services, therefore treat changes in consumer behaviour as an effect on volume: if all consumers prefer to buy tickets at reduced prices and are prepared to regularly monitor prices, then the volume of air transport services will decrease, taking into account the reduction in service quality resulting from the efforts to optimise the ticket price that the consumer has to make.

4. Is the Digitisation of the Economy Likely to Significantly Bias the Volume-Price Split?

The digitisation of the economy calls into question the relevance of traditional volume-price splitting tools, as it disturbs the offer of products,

Figure VII – **Monthly consumer price indices for air transport, calculated according to different reservation intervals, in 2016 (December 2015=100)**



Reading Note: The consumer price index for air transport published by Insee and calculated in collaboration with the DGAC (DGAC index) was 90 in May 2016 (December 2015=100); it has thus fallen by 10% compared with December 2015. If only the price of tickets purchased two days before departure had been tracked in the price index, this index (D-2_Index) was 97 in May 2016, representing a decrease of 3% compared with December 2015; if, on the contrary, only the price of tickets bought 44 days before departure had been tracked, the index (D-44_Index) would have been 85 in May 2016, representing a decrease of 15% compared with December 2015.
Sources and coverage: French Directorate-General for Civil Aviation (DGAC), Insee calculations; France.

whereas the measurement of a price index is based on the stability of that universe (with the notion of a fixed basket). However, the issue is not new and statisticians are not entirely at a loss when faced with the appearance and renewal of products, as we have sought to demonstrate above: methods exist and, moreover, the harmonisation of sources carried out by national accountants avoids many pitfalls. However, some assumptions are debatable.

In order to evaluate the importance of these assumptions, various studies have sought to quantify the uncertainty surrounding the volume-price split, particularly in view of the slowdown in growth. To do so, they generally rely on *ad hoc* maximum quantifications of biases for products likely to be affected by digitisation and their consequences for the measurement of GDP. As the weights of these products are generally quite low, their conclusion is most often that the problem of measuring the volume-price split does not call into question the finding of a real slowdown in real GDP (Reinsdorf & Schreyer, 2017; Ahmad *et al.*, 2017).

In the case of France, all computer, electronic and optical products associated with telecommunications services, programming and consultancy services and other computer activities only represent on average, over the period 1997-2016, 4.6% of GDP, which limits the impact of any error on the measurement of the consumer price index or other price indices. Two simulations were carried out to test the sensitivity of French growth in volume terms to the price indices used for these new products and, in particular, to the assumptions made to adjust quality.

In the first simulation, it is assumed that the quality of digital products (computer, electronic and optical products, telecommunication services and software) does not change despite the renewal of these products. To that end, there is traditionally a reliance on the CPI simulations in Figure III, which leads to an upward revision of consumer prices of around +7.5 percentage points per year for ICT goods alone. As indicated previously, the change in the CPI on digital goods theoretically has only a limited impact on GDP, as an error in the measurement of consumer prices is likely to result in an error in the measurement of the import prices of these products. However, in the context of their work to ensure consistency between the various indicators, if the consumer price index had been more dynamic by +7.5 percentage points, the national accountants would probably have had to raise also the investment and export prices of ICT goods, and this effect is therefore incorporated into the simulation. Furthermore, it is assumed that the +7.5 percentage point difference represents a general order of magnitude for the corrections made by Insee to deal with quality effects on products with high turnover, and that is why this difference is also traditionally applied to the deflator of GFCF for software. In contrast, the price indices for telecommunication services are not changed in this first simulation. Under these assumptions, without the correction for quality effects for technological products carried out by price statisticians, the GDP in volume terms would have grown by only 1.35% on average over the period 1997-2016, equating to an average annual growth rate of 0.26 percentage point less than that published in the national accounts.

In the second simulation, the assumption regarding the quality adjustment error is based on the work of Ahmad *et al.* (2017), which shows differences between the price indices used by US statisticians and the price indices proposed by other researchers, generally using hedonic models (Byrne *et al.*, 2016; Byrne & Corrado, 2017). These differences vary from product to product, but average around 7% over the period 1995-2014. Traditional assumptions were therefore applied to the French deflators based on the differences between official US deflators and the alternative deflators presented in this work[5] (cf. Table).

Assuming that the consumer price and investment indices for digital goods, software and telecommunications services have been greatly overestimated by Insee, French growth in volume terms is found to be underestimated by 0.23 percentage point over the period 1997-2016 (Figure VIII).

However, even the application of this rather extreme scenario in no way calls into question the diagnosis of a slowdown in French growth in volume terms, which would remain marked: the average annual rate of growth in GDP in volume terms would be 1.4% on average over the period 2010-2016 in this scenario, compared with published average growth of 1.2%, in comparison with a 1997-2008 growth rate (excluding the 2009 crisis year) of 2.5% in this scenario and 2.2% in the published accounts.

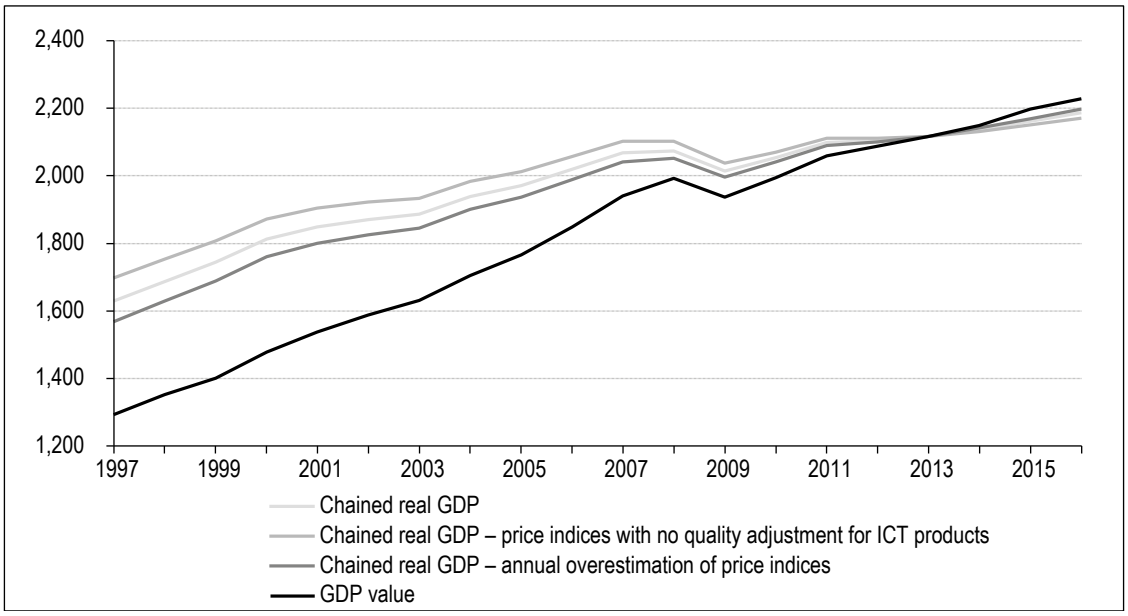
5. *Byrne* et al., *2016, Tables 2.2 and 2.5.*

Table – **Assumption used to correct the price indices for different technological products**

Difference in change rates between published deflators and the deflators used in the second simulation (%)	Mean 1995-2004	Mean 2004-2014
CI – Manufacture of computer, electronic and optical products	-7.4	-6.5
JB – Telecommunications	-6.8	-6.8
JC – IT activities and information services	-1.4	-0.9

Reading Note: It is assumed that the consumer price index for computer, electronic and optical products has been overestimated by Insee by 7.4 percentage points over the period 1995-2004 and by 6.5 percentage points over the period 2004-2014.

Figure VIII – **Impact on French GDP of the different price measurement scenarios (Billions of €)**



Sources and coverage: national accounts, database 2014; France.

* *
*

The low economic growth measured over the last two decades is challenging the perception that we have a digital and innovative economy. Is this a reflection of overestimated inflation? It should be noted, in passing, that inflation, which has already been very low in recent years, is regularly questioned by consumers who, in contrast, consider it to be underestimated (Accardo *et al.*, 2011; Leclair & Passeron, 2017).

This article has sought to demonstrate that the difficulties raised for the volume-price split by the digital economy are not ignored by statisticians. There are methods in place; effects are measured indirectly and significant attention is paid to the consistency of the various data sources (quantity, value, price, etc.). These issues are not new and have also affected previous measurements of GDP. However, the volume-price split is based on a certain number of assumptions (particularly that differences in price between products reflect differences in utility for the consumer) which may be questioned. In any event, the uncertainty surrounding these assumptions is not such as to explain the slowdown in the French economy over the recent period. □

Link to Online Appendices: https://www.insee.fr/en/statistiques/fichier/4770160/ES-517-518-519_Aeberhardt-etal_Online_Appendices.pdf

BIBLIOGRAPHY

Accardo, J., Célérier, C., Herpin, N., & Irac, D. (2011). L'inflation perçue. *Économie et Statistique,* 447, 3–31. https://doi.org/10.3406/estat.2011.9707

Aeberhardt, L. & Bidault, C. (2018). Le partage volume-prix base 2014. Insee, *Notes méthodologiques du système français de comptabilité nationale*. https://www.insee.fr/fr/metadonnees/source/s1375#documentation

Aghion, P., Bergeaud, A., Boppart, T., Klenow, P. J. & Li, H. (2019). Missing Growth from Creative Destruction. *American Economic Review,* 109 (8), 2795–2822. https://doi.org/10.1257/aer.20171745

Aghion, P., Bergeaud, A., Boppart, T. & Bunel, S. (2018). Firm dynamics and growth measurement in France. *Journal of the European Economic Association,* 16 (4), 933–956. https://doi.org/10.1093/jeea/jvy031

Ahmad, N. & Schreyer, P. (2016). Measuring GDP in a Digitalised Economy. *OECD Statistics Working Papers*, N°2016/07. https://doi.org/10.1787/5jlwqd81d09r-en

Ahmad, N., Ribarsky, J. & Reinsdorf, M. (2017). Can potential mismeasurement of the digital economy explain the post-crisis slowdown in GDP and productivity growth? *OECD Statistics Working Papers*, N°2017/09. https://doi.org/10.1787/a8e751b7-en

Bellégo, C. & Mahieu, R. (2016). La place d'Internet dans la description et l'analyse de l'économie. *Insee Références – L'économie française,* édition 2016.
https://www.insee.fr/fr/statistiques/fichier/2017586/ECOFRA16b_D1_internet.pdf

Bessone, A.-J., Broin, M., Hassan, M., Leclair, M. & Mahieu, R. (2014). Les comptes nationaux passent en base 2010. *Insee Références – L'économie française*, édition 2014.
https://www.insee.fr/fr/statistiques/fichier/3692679/ECOFRA14_b_D1_comptes.pdf

Blanchet, D., Khder, M.-B., Leclair, M., Lee, R., Poncet, H. & Ragache, N. (2018). La croissance est-elle sous-estimée ? *Insee Références – L'économie française,* édition 2018.
https://www.insee.fr/fr/statistiques/fichier/3614262/EcoFra2018.pdf

Boskin Commission (1996). *Toward a More Accurate Measure of the Cost of Living.* Final report to the Senate Finance Committee. http://www.ssa.gov/history/reports/boskinrpt.html

Bourgeois, A. (2020). Les services gratuits issus de l'économie numérique : faut-il, et comment, les valoriser ? *Economie et Statistique / Economics and Statistics*, ce numéro.

Byrne, D. & Corrado, C. (2017). ICT prices and ICT services: What do they tell us about productivity and technology? Board of Governors of the Federal Reserve. *Finance and Economics Discussion Series*, N°2017-015. https://doi.org/10.17016/FEDS.2017.015

Byrne, D., Fernald, J. & Reinsdorf, M. (2016). Does the United States have a productivity slowdown or a measurement problem? *Brookings Papers on Economic Activity,* Spring 2016, pp.109–157.
https://doi.org/10.17016/FEDS.2016.017

Cette, G., Fernald, J. & Mojon, B. (2016). The pre-great recession slowdown in productivity. *European Economic Review*, Elsevier, 88(C), pp. 3–20. https://doi.org/10.1016/j.euroecorev.2016.03.012

Diewert, E. & Feenstra, R. (2018). Estimating the Benefits and Costs of New and Disappearing Products. Communication au groupe des experts prix de l'UNECE-BIT.
www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.22/2018/University_of_British_Columbia.pdf

Eurostat (2010). *Système européen des comptes*, chapitre 10.
https://ec.europa.eu › eurostat › documents › KS-02-13-269-FR.PDF

Eurostat (2017). HICP methodological manual.
https://ec.europa.eu › eurostat › documents › KS-GQ-17-015-EN-N.pdf

FMI (2004). Manuel des prix à la consommation. Théorie et pratique.
www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/presentation/wcms_331155.pdf

Feldstein, M. (2017). Underestimating the Real Growth of GDP, Personal Income, and Productivity. *Journal of Economic Perspectives*, 31(2), 145–164. https://doi.org/10.1257/jep.31.2.145

Groshen, E. L., Moyer, B. C., Aizcorbe, A. M., Bradley, R. & Friedman D. M. (2017). How Government Statistics Adjust for Potential Biases from Quality Change and New Goods in an Age of Digital Technologies: A View from the Trenches. *Journal of Economic Perspectives*, 31(2), 187–210.
https://doi.org/10.1257/jep.31.2.187

Leclair, M. & Passeron, V. (2017). Une inflation modérée depuis le passage à l'euro. *Insee focus* N°87. https://www.insee.fr/fr/statistiques/2854085

Lequiller, F. (1997). L'indice des prix à la consommation surestime-t-il l'inflation ? *Économie et Statistique*, 303, 3–32. https://doi.org/10.3406/estat.1997.2542

Lequiller, F. (2000). La nouvelle économie et la mesure de la croissance. *Économie et Statistique*, 339-340, 45–71. https://doi.org/10.3406/estat.2000.7479

Magnien, F. & Pougnard, J. (2000). Les indices à utilité constante ; une référence pour mesurer l'évolution des prix. *Économie et Statistique*, 335, 81–94. https://doi.org/10.3406/estat.2000.7523

Magnien, F. (2003). Mesurer l'évolution des prix des services de téléphonie mobile ; une entreprise difficile. *Économie et Statistique*, 362, 3–31. https://doi.org/10.3406/estat.2003.7340

Reinsdorf, M. (1993). The Effect of Outlet Price Differentials in the U.S. Consumer Price Index. *In:* Foss, M., Manser, M. & Young, A. (Eds.). *Price Measurements and Their Uses,* pp. 227–258. Chicago: University of Chicago Press.

Reinsdorf, M. & Schreyer, P. (2018). Measuring inflation in a digital economy. OECD, *SDD Working Paper 101*. https://doi.org/10.1787/18152031

Saglio, A. (1995). Changements du tissu commercial et mesure de l'évolution des prix. *Économie et Statistique*, 285-256, 9–33. https://doi.org/10.3406/estat.1995.5975

Sillard, P. (2017). L'indice des prix à la consommation, Insee, *Document de travail n°F1706.* https://www.insee.fr/fr/statistiques/fichier/version-html/2964204/F1706.pdf

Syverson, C. (2017). Challenges to Mismeasurement Explanations for the US Productivity Slowdown. *Journal of Economic Perspectives*, 31(2), 165–186. https://doi.org/10.1257/jep.31.2.165

Vanoli A. (2002). *Une histoire de la comptabilité nationale*. Paris: La Découverte, Collection Repères.

Free Services from the Digital Economy: Do We Need to Measure Their Value and How?

Alexandre Bourgeois*

Abstract – For several years now, the economy has seen a change in consumption habits, driven by the development of digital technology. New products and services and new economic models have emerged, and value chains have changed. This raises many questions for national accounting, including how to take new forms of free services into account. There are three possible options for trying to assign these services a monetary value: *(a)* on the basis of the advertising income for those services that are indirectly financed through advertising, *(b)* direct valuation of the service provided to the users, by assessing their willingness to pay or by using the standard methods for valuing time spent on domestic tasks, and *(c)* valuing the data generated through the use of these services, which constitute another way of guaranteeing the immediate and long-term profitability of such services. Beyond practical difficulties of their implementation, another issue is that the three options do not all answer the same questions, which ultimately raises the issue of the purpose of the national accounts.

JEL Classification: D60, E01, E21, E22, L82, M37
Keywords: free, data, consumption, GDP, platform, intangible investment, advertising, brands

* *Insee (alexandre.bourgeois@insee.fr)*

I would like to thank Didier Blanchet, Jacques Magniez, André Vanoli, Ronan Mahieu as well as two anonymous reviewers for their careful reading and their precious help.

Received June 28, 2019, accepted after revisions le June 6, 2020.
Translated from: "Les services gratuits issus de l'économie numérique : faut-il, et comment, les valoriser ?"
Citation: Bourgeois, A. (2020). Free Services from the Digital Economy: Do We Need to Measure Their Value and How? *Economie et Statistique / Economics and Statistics*, 517-518-519, 157–172. https://doi.org/10.24187/ecostat.2020.517t.2021

The economy has been experiencing a radical transformation in recent years as a result of the increasing use of information technologies. These technologies are changing consumption habits. In 2018, 88% of the French population (57.3 million people) were internet users and 58% were active on social networks.[1] The average time spent on the internet in France in 2019 reportedly reached or even surpassed that spent in front of the television,[2] at 3 hours 20 minutes per day, including 1 hour 22 minutes on social media. This has been made possible with the emergence of new players and new economic models in which prices have seemingly lost a large part of their usual significance. Some services can be accessed at very low prices, while others are entirely free of charge, at least in appearance, since in practice, several free or pseudo-free models can be combined: fully free and funded by donations or open-source funding, different types of pseudo-free services funded by advertising and/or data collection, or even freemium models with chargeable add-on options. *Wikipedia*'s services are a perfect example of an entirely free service funded mainly by donors. *Facebook*'s services are indirectly funded by advertising and data collection. *YouTube* is funded through advertising and data collection, but also offers users the option of paying a fee for ad-free use and makes occasional payments to some content providers.

These phenomena are not entirely unprecedented. Renewing and diversifying goods and services have always been one of the main drivers of growth, and free or pseudo-free services are not new: television broadcasting, for example, has existed for a long time. However, the development of the internet and the capacity of networks to exchange large quantities of information on a massive scale (scalability) has created a technological breakthrough that has allowed these networks to reach unprecedented sizes, while also raising the question of how to include these in the national accounts. As there are no direct monetary transactions between the producer and end user, these free or pseudo-free services offered by the digital economy do indeed appear to be a glaring gap in the statistics of household consumption and, in the same way, seem to be "missing" from the GDP.

In reality, although the national accounts aim to describe the economic reality as accurately as possible, they do this primarily using market transactions. There are just two exceptions to this rule: public service production and the inclusion of imputed rental income for the housing services that owner-occupiers provide for themselves. These two exceptions are due to the need for international comparability: production and living standards must not appear lower in countries with strong public sector development or with high rates of home ownership. These two derogations can also be explained by the fact that it is possible to manage them with relatively natural imputation rules: assessment at production cost for public services and reference to market rents for imputed rents. There are no equivalents for the other types of free or non-market services. As it is, the treatment of these services follows the very minimal recording criteria that were agreed in the SNA 2008 (System of National Accounts) and implemented at European level in the ESA 2010 (European System of Accounts). Pseudo-free services funded by advertising are not included in household consumption: only advertising costs are accounted for as intermediate consumption by the advertising companies that rely on them. And those services that are truly free are only included up to their apparent production costs, ignoring any voluntary contributions from which they may benefit.

It is, therefore, true that the "digital replacements" currently in progress translate into a form of household consumption that is not included in the GDP, and which could, all other things being equal, even lower it. This calls us to question these criteria adopted by the SNA 2008 and ESA 2010 and the questions that we must address are many and complex. Do we need to assign a value for final consumption expenditure to households to reflect the free services from which they benefit? In the case of advertising companies that sponsor these services, do we need to consider that this is just one cost among others (recorded as intermediate consumption) or is this also, to some extent, an investment (gross fixed capital formation)? Where this financing of free services is supplemented by or fully reliant on data collection, do we need to consider that these data are a form of production by households leaving digital footprints or can we see this as an "inert raw material", the production of which does not merit valuation?

This article is going to show the complexity of these questions by considering the three potential ways of measuring the value of the free services currently at issue, which have, in

1. *Hootsuite and We Are Social 2018 Digital Report.*
2. *Sources: Zenith Media, 2019.*

some instances, been subject to initial attempts at implementation. These are: *(a)* measuring value on the basis of advertising income for those services that are indirectly financed through advertising, *(b)* directly assessing the service provided to the users, by assessing their willingness to pay or by using the standard methods for valuing the time spent on domestic tasks, and lastly *(c)* valuing the data generated through the use of these services, which constitute another way of guaranteeing the immediate and long-term profitability of such services. Our general guiding principle will be to ask what exactly it is that we want to measure. The fact that free services are not included in the GDP does not necessarily mean that they are underestimated or that we must systematically find ways of supplementing this with indirect valuations of these free services; this depends on how we want to use this indicator.

1. First Proposal: Using Advertising Income to Measure the Value of the Free Services That Are Indirectly Financed through Advertising

1.1. A Debate Dating Back to the 1980s

The first option for measuring the value of free services is based on measuring the advertising income that funds them, which corresponds to a traditional "input" approach used in the national accounts (cost-based valuation). Even if this gives the impression of putting things in a new light, this topic is nothing new for national accountants, as this question had already been raised under the same terms for measuring the value of recreational television programmes funded through this same advertising channel. This is what is known as a two-sided market (see Box 1), with one side aimed at the advertisers and the other at television viewers benefitting from free-of-charge broadcasting of recreational programmes funded by the advertising expenditure of these advertisers. If this form of recreational leisure is a substitute for purchasing tickets for performances, do we need to find a way to prevent this substitution from being considered in the national accounts as a drop in both consumption and production?

Until now, the predominant stance has been that such a correction was not necessary, at least in nominal terms, given that the consumer was already paying indirectly for this service through the heightened prices of the products "sponsored" by advertising, which cover these costs. This was, for example, the stance taken by Okun (1971). He recognised that valuing consumer expenditure on television broadcasting at zero was the "most puzzling single consequence" of the rules for processing the intermediate expenditure of companies. Although he accepted that these services are part of consumption, his view was that, as the services are not paid for on the market, they cannot be assessed and should therefore not be added to the GDP. He believed that "so long as radio and TV programs are free goods to the consumer, it is as meaningless to put a price tag on what comes over the airwaves as it is to put a price tag on air itself".

Several economists have, however, advocated for the opposite (in particular, Jaszi, 1973, Eisner, 1978 and Kendrick, 1971) as they believed it necessary to impute a value for these sponsored services from which consumers benefit. This option was developed in greater detail by Cremeans (1981), who proposed a scheme for extended measures of consumption and income by imputing flows so as to better account for household consumption of television programming that is, on the surface, free of charge. Cremeans underlines the specific nature of this type of exchange "in kind", which explains why it is not included in market flows.

Box 1 – **Two-Sided Markets**

A two-sided market refers to the activities of intermediation platforms that enhance positive externalities either one-way (audience platforms) from users to advertisers, or cross-cutting (marketplaces, application platforms, etc.) by targeting two or more markets at the same time and forming the two sides of the market. There are several examples of the latter phenomenon: intermediation between readers and advertisers by the media or between card-holding consumers and merchants *via* credit cards. In this model, the pricing on each side covers the interdependence between the two. The platform must have at least two distinct user groups or "sides" that produce mutually positive externalities. In terms of pricing, users on the side with the greatest price sensitivity (internet users) are charged a lower price than the marginal cost; the price may even be zero (free service) or negative (subsidised). Users on the other side of the market, who have a lower price elasticity (advertisers) are charged a price higher than the marginal cost. This creates cross-subsidies between the two sides.

In his scheme, which requires the creation of a new subcategory of services (listening to advertising services), advertising companies buy entertainment services from television companies in order to remunerate in kind the service of listening to advertising messages provided by households. Advertisers thus buy this listening service from households and remunerate them in kind through the benefit of entertainment programmes.

This scheme leads us to consider households that watch the television as a display enterprise selling time and entertainment space to television companies. This is described in Vanoli (2002, Box 28), who proposes a variation on the theory. Working from the initial idea that advertising expenditures cover the actual cost of both advertising programmes and entertainment programmes, he proposes a three-part scheme that implicitly represents a two-sided market:

- Advertisers consume a service (for which a value may be imputed) provided by households when they listen to advertising messages and a message broadcasting service (not imputed) provided by the television company.

- The television company provides an advertising programme broadcasting service to the advertisers and an (imputed) entertainment programme broadcasting service to households.

- Households consume the entertainment programme service provided by the television company in exchange for providing the service of listening to advertising messages from advertisers.

In the more recent context of digitalisation, and working from the observation that the "free services" and information from the internet that are funded by advertising income have a major impact on consumer behaviour, Nakamura & Soloveichik (2015) propose a new scheme for imputing such values that is relatively similar to those suggested by Crameans and Vanoli. That approach, however, is not based on costs, but on time spent. Cost-based approaches are often criticised for not taking markup-type margin behaviour into consideration; this is the type of pitfall that an approach based on time spent can avoid. By separating time spent viewing advertising from income per hour created by that same advertising, they give value to the time spent by households (distinguished from that spent by companies, which would be intermediate consumption) watching the adverts themselves, by considering this to be an act of production remunerated by the advertising company, which, in turn, pays for a right to broadcast with a broadcasting company. By taking all media into account, they identify a partial compensation (general equilibrium effect), but with a global effect that, ultimately, proves to be fairly negligible: between 1980 and 2013, they note an increase in the global growth rate of 0.018% per year.

This estimate does not include the consumer surplus and corresponds to a relatively conventional value imputation in the national accounts, in a similar way to imputed rents. The partial compensatory effect observed by the authors calls into question the "net" effect (in volume) on the advertising market made possible by the development of communication channels formerly owned by other operators (Tech Giants *vs.* historical operators).

1.2. Advertising Expenditure: Intermediate Consumption or Investment?

Of course, there are other possible ways of taking advertising expenditure into account. And even if this forces us to re-examine the issue of treating advertising costs as intermediate consumption for advertising companies, can we not instead choose to see these costs as intangible investments? After all, advertising influences consumers in various ways. It allows companies to influence the distinctive attributes of a brand, which are points of reference for consumers and often bring with them an ecosystem of values. Even in the background, advertising relies on people's memory structures, and its effects continue beyond the short term. Thanks to these characteristics, advertising helps to give value to a "brand" asset and a "produced" asset and contributes to advertisers' future return potential through a cluster of intangible assets that they hold or support. These intangible assets can act in various ways, for example by consolidating or strengthening a market position, or even increasing price positioning. Nakamura (2005), for example, observes that advertising increases company sales in the long term and should be considered as an investment in brand image.

Considering advertising in a new light, as an investment, is therefore another way of recording it in the GDP. There may, of course, be objections to this in the fact that "brand" assets are not considered as investments in the ESA 2010 on the basis that this is a zero-sum game in which the main impact of investments

in a brand is observed in terms of market shares between product classes that are close substitutes without net value creation, and therefore without the associated net revenue stream. However, this argument is contested by Corrado & Hao (2014), who deem that there is no proof supporting the zero-sum view, especially in an innovation economy with conveyors of market power between economic actors.

Therefore, if we consider that there is a net investment here, we could initially record an investment by the brands (*via* the advertisers), which would then lead us to record the production of a usage service, which, for the brands whose value has been increased through advertising, would consist in raising their prices in accordance with consumer trends (increased by the brand's intangible assets) in purchasing the sponsored product. The extra cost generated here would be equivalent to a form of royalty collected by the brands over time through the management of their intangible assets.

1.3. If There's No Such Thing as a Free Lunch, Who Pays at the End of the Day?

In order to decide between these various valuation options, we propose drawing up a typology of the situations, distinguishing between those that potentially fall into the categories of final consumption, intermediate consumption and intangible investments. Rather than describing all the relationships within the categories, we will simplify the relationship by considering a direct link between the upstream party (the advertiser issuing the advertising) and the downstream party (the consumer benefitting from access to a free, sponsored service). This simplification requires us to assume that there is intermediate consumption at the level of the companies playing an intermediary role (advertising agencies, content providers, etc.) at the point where the two sides of the market meet. Households benefit from a free service (recreational television programmes, for example) in the short-term (immediately). For the advertiser, advertising is a short-term cost; however, that cost can be recorded either as intermediate consumption or as an investment.[3] This is not a neutral choice; it has an impact on GDP as intermediate consumption is substracted from production when calculating the added value, whereas investment is added to the final expenditure. There is still uncertainty as to whether the advertiser will offset these costs through its sales prices or the quality of its service, depending on its production function. We can therefore define two stakeholders on this two-sided market:

- On the household side: recreational services that households access may potentially be recorded immediately as household consumption. The influence exerted by advertising may also lead to final consumption deferred in the medium or long term.

- On the advertiser side: they "sponsor" these recreational services, the expenses incurred immediately have expected benefits (in their revenues) that could be recorded in the short term as intermediate consumption, and in the medium-long term in the form of investments (gross fixed capital formation).

What are the advertisers' reasons for advertising? Either they expect financial benefits, or this is not the case, or in a very indirect way. The legitimacy of imputing a value for household consumption expenditure could depend on the way in which advertisers pass on their advertising costs in their prices:

- If this translates into a change only in price, whatever the term, the cost of the advertising is already included in the extra cost of the sponsored products or brands. This extra cost may be likened to a form of "tax", which of course generates more apparent consumption in terms of value, but not in terms of volume. It does not, therefore, seem legitimate[4] to assign an additional consumption expenditure to households for free sponsored services (as this risks counting these twice). Although there is ultimately no impact on volume, we may still be tempted to impute some value for volume in respect of the recreational service itself. This brings us back to the question of the objectives of measuring this.

- If this does not translate into a change in market prices, either this involves redistribution within a zero-sum market (constant volume, no extra cost of sponsored products), or an expansion of the market (increased total volume of sales, uncertain impact on the potential extra cost of sponsored products) through repositioning of the stakeholders. The consumers are the

3. Here, we do not believe it is possible to record this as individual company consumption expenditure, as the existence of this was rejected by the SNA 2008.

4. Legitimate in the sense that the users do not pay for this indirectly elsewhere and are the net beneficiaries (as the free service is of greater value to them than the cost). However, as these costs are difficult to measure due to the many factors involved and the fact they are sometimes difficult to identify (cost of viewing the advertising, the windfall effect of time spent in front of the television, etc.).

net beneficiaries in this situation as they access a service free of charge without this indirectly leading to an increase in the prices of the goods and services consumed. In this case, it would be legitimate to impute a value for part of the additional household consumption expenditure for free sponsored services; however, which part is still unknown.

This probably leads to segmentation by the types of goods sold by the advertisers: luxury goods should probably see more of a change in terms of prices and volumes, while budget goods are likely to see a greater change solely in terms of volumes. Where there is no expectation of financial benefits (for example, political influence, reputation management), we can assume that there may be future consumption connected with the broadcasting of the advertising, in which case this would only translate into potential profitability of the underlying intangible assets. We cannot, however, exclude the possibility that this translates into future consumption choices. In this case, this brings us back to the expectation of implicit financial benefits in the medium and long term.

Nevertheless, the approaches attempting to use the measured value of advertising as a basis suggest, at best, a prospect of profitability for the advertisers by attempting to identify who ultimately bears the cost (either the advertisers or the consumers in the form of a tax). These approaches therefore appear to miss the core subject: quantifying the benefit that the consumer gains from recreational services funded by this advertising.

1.4. Direct Imputation of Household Consumption Expenditure: Pros and Cons

There are several factual arguments in favour of imputing a value for expenditure for consumption of recreational services, so as to translate this benefit that households gain from these services into financial terms. An initial argument holds that if the consumer were to pay a fee to remove advertising, such as in *YouTube*'s "Premium" service, which offers users the option to pay a fee of around €10 per month to watch videos without adverts and offline on mobile devices, and gives access to exclusive content, we would record this amount as household consumption. Likewise, an ad-free newspaper would cost more for the consumer and would mechanically increase their consumption expenditure.

Furthermore, the methodological concepts that apply to the non-market sector consolidate this type of treatment, even if the transposition of this non-market treatment to the market sector does not always appear to be straightforward (Box 2 *infra*). In the case of free or almost-free public services, for example when cities make public transport services free of charge, the subsidised part of the expenditure is not included in household consumption expenditure but is, to a certain extent, reallocated as public administrations consumption expenditure *via* a transfer scheme. This mechanism makes the analogy with the principle of a tax quite clear. Indeed, where a public service stops charging for a service, there is still an instance of production recorded in the national accounts. The transition to a free service either has no impact on the production (if it is initially commercial) or changes it to a certain extent due to the difference between the market and non-market valuations (on the basis of costs). The production is therefore almost unchanged but, conversely, on the income and demand side, there is a drop in gross disposable income (GDI) and consumption in the amount of the price of the service. This is reconciled through adjusted gross disposable income by reallocating to households the public service made available to them. The transition to a free service does not, therefore, mean that the service production disappears; it simply changes the structure of the flows, specifically the distribution between consumption among households and consumption expenditure of government agencies.

Although, by analogy, we can consider that the rise in prices generated by advertising can, in some respects, be likened to an advertising "tax", it is nevertheless useful to highlight its specific nature: this would be a tax levied by companies on households (in return for their funding of free recreational services) and not the more traditional notion of tax that passes through public administrations (which play an intermediary role) and is ultimately redistributed among households.

Under the second hypothesis, which assumes that advertising increases the price of the advertised goods, any goods or services would include an immaterial "extra cost" associated specifically with reputation of the brand or the product itself, an extra cost acquired in part through advertising. In this case, this would purely affect the price, and have no impact on volumes. However, Nakamura & Soloveichik (2015) believe that the notion that the cost of the advertising is included in the products consumed supposes that consumers will automatically

Box 2 – **Free Services in the Market and Non-Market Sectors: A Smokescreen Analogy?**

The analogy between the rules of the non-market sector and the potentially equivalent treatment methods in the market sector can be called into question. Indeed, Robert Eisner (1988) notes that "Where we do not count production, we do not have a component of national income, with the curious exception of output related to government subsidies or losses in government enterprises." He believes that identifying the purchases of the various actors (households, companies, government agencies) as intermediate products results in anomalies regarding the sometimes changing identity of the purchasers. Hence, "the police services purchased by government are final products and are included in the GDP, while services of guards or watchmen purchased by a company are not; they are presumably resold as part of the business production in which they are used." Eisner believes that "in general, the difficulties arise from the inadequacy of market sales and purchases as measures of production". The government agency sector has specific features that, due to their nature, require conventional processing. It is therefore sensible to consider the analogy between the processing carried out in market and non-market sectors with caution.

purchase the sponsored products, yet, there is no legal obligation to buy, even if we were able to assume that this effect must be seen partly at the macroeconomic level, otherwise advertising would not be an economic incentive. Imputing a value for household consumption expenditure would, in this case, result in double counting. Cremeans (1981), on the other hand, does not fully reject the principle of potential double counting, drawing an analogy with the non-market example of VAT treatment (which suggests an analogy with an advertising "tax"). However, here once again, the parallel with processing in the non-market sector is not necessarily self-evident (cf. Box 2).

We can also conceive of other ways of measuring the value of free services, for example by calculating the time spent on these services, or even by trying to directly quantify the welfare gained from these services. The characteristic common to both of these ways of measuring value is an attempt to quantify the service provided to users.

2. Second Proposal: Direct Evaluations of the Service Provided to Users

Measuring the value of free services through, for example, an attention-based service raises the question of where to set the boundaries, and of what it is that we actually want to measure. A first option consists in valuing the time that a person dedicates to accessing content, in the form of opportunity cost, on the basis of an equivalent in terms of salary as a shadow price for the entertainment service, similar to the way in which the value of domestic activities is measured. The second approach consists in quantifying the "welfare" that consumers gain from the free service by evaluating their willingness to pay to access a service, or willingness to accept compensation to forgo access to a service.

2.1. Measuring the Value of Free Services on the Basis of the Time Spent in the Form of Opportunity Cost: At What Cost, and Where Do We Stop?

Brynjolfsson & Oh (2012) go beyond imputation in accounting terms by considering that even when people do not pay cash, they must still pay "attention", or spend time. They measure a consumer surplus through an opportunity cost of time spent on the internet by incorporating an annual quality effect of the digitalisation of the economy as well as substitution elasticities among media and between online and offline activity. They estimate the increase in consumer surplus created by free internet services to be over $100 billion per year in the United States (i.e. a 0.74% increase of the rate of annual GDP growth between 2007 and 2011). This estimate is significantly higher than that of Nakamura & Soloveichik (2015). These authors consider that most of the welfare gained from internet-based digital services would be overlooked by purely monetary approaches.

But where do we set the boundaries for measuring the value of attention time? The question of the accounting scope to be used for a potential imputation poses the risk of including non-digital contributions, for example widening the perimeter to include all leisure activities, or even the entire knowledge-based economy. For example, should we reclassify training as an intangible investment, or even education as a whole (Jorgenson & Fraumeni, 1992), and what about anything that can contribute to enhancing human capital?

In an approach similar to that of Brynjolfsson & Oh (2012) but with a wider scope, Goolsbee & Klenow (2006) suggest, for example, measuring the value of the internet on the basis of the time spent using it. Noting that the digitalisation of the economy has, in particular, led to an increase in the time spent by internet users on their computers, they use data on time spent online to estimate the welfare gains, which leads them to impute an additional consumption of between 2,500 and 3,800 dollars for the median household. They assess the opportunity cost of time spent on the internet and, using a simple utility model, calculate the consumer surplus associated with internet access, which differs from a more conventional approach based on costs (internet subscription and computer equipment, modem, etc.). This approach is based on the calculation of the consumer surplus, meaning it has a welfare focus and incorporates a measure of the value of leisure time in particular.

More specifically, this approach requires to consider that time spent on the internet is, in part, equivalent to productive activities (for example, booking a trip yourself rather than going to a travel agent) and that the measurement of its value is just one specific case of the more general valuation of domestic work. If we followed through with this approach, we would stumble upon the age-old problem of the value of all time spent performing domestic work. This question was recently addressed again by Roy (2013). Using the 1998 French Time Use Survey, she determines that the contribution of domestic work to household welfare makes a contribution to national production equivalent to 17.5% of GDP (and 27% of GDP under a broader definition of domestic work). This valuation is based on costs, by calculating the product of the number of hours of domestic work and a market price (the gross hourly minimum wage). It may seem preferable to use the opportunity cost, i.e. the wages of the person on the labour market, but, in reality, this approach would measure the wage differences between categories of individuals – and, across time, the change in wages – rather than the actual value of domestic work. In 2011, the OECD also highlighted the highly sensitive nature of assessing production from domestic services at the value attributed to the labour cost. Imputing an equivalent value in terms of time spent is itself, therefore, a complex issue.

2.2. Measuring the Value of Free Services on the Basis of Willingness to Pay or Accept Compensation: An Exhaustive Scope, Clear Objectives, but Tools Are Still Limited

More recently, and believing once again that the official statistics lack a growing portion of the real value created in our economy, Brynjolfsson *et al.* (2019) refined the previous approaches by suggesting a measure of new and free goods (which are generally insufficiently recorded in the national accounts) based on an empirical measurement framework of the willingness to accept compensation to forgo a service over a given period. This approach quantifies the benefits rather than the costs and defines an extension to GDP within a new indicator "GDP-B". This methodology is applied to several empirical examples and allows estimating the additional welfare gains associated with the use of *Facebook*, for example, between 0.05 and 0.11 percentage point of the US GDP-B growth per year in 2017. The benefit of this approach is that it allows us to include all types of free services: it is not limited to the advertising model, which ignores the value of digital services produced without compensation (e.g. *Wikipedia*), and neither does it arbitrarily impute an opportunity cost, which could be said to have questionable legitimacy.

While this method appears to be more substantive than previous works, it also highlights the limits of a macro assessment based on micro samples and questions the additivity of the utility. Indeed, one of the tests of willingness to accept compensation was carried out with a restricted sample of platforms (*Instagram*, *Snapchat*, *Skype*, *WhatsApp*, *Maps*, *LinkedIn*, *Twitter* and *Facebook*) and revealed that, by far, the most important value (five times greater than *Facebook* in second place) is that attributed to *WhatsApp* (€536 per month, compared with €97 for *Facebook* and €59 for *Maps*, and just €0.18 for *Skype*). The interviewees also stated that, for them, *WhatsApp* was an almost indispensable communication platform. We can assume that the smaller the number of substitutable services among those both within the suggested sample and beyond, the higher the suggested value; there need only be one service outside the survey that is fully capable of replacing a service within the scope of the study for the latter to lose a significant percentage of its value. For example, as *Skype* offers a similar service to *FaceTime*, which was not part of the sample, we can theoretically assume that this made a considerable

contribution to reducing the value of *Skype*. Even if the authors were to make corrections to manage this phenomenon, the resulting sampling effect seems to limit this type of evaluation, preventing us from extrapolating and then imputing a value for the entire economy.

2.3. Regarding the Questions that the Suggested Imputations Aim to Answer, the Proposed Approaches Face Several Pitfalls

The different imputations that we have just reviewed all aim to answer several questions, but these answers may be called into question. The first obstacle is the difficulty in establishing the volume-price splits. The approaches developed by Cremeans and Vanoli are methods for measuring the value of the attention service provided by households in viewing advertising, at a level close to the underlying advertising cost, which shifts from intermediate to final consumption and therefore increases GDP in line with the as yet unspecified volume-price split. The approach suggested by Nakamura & Soloveichik incorporates a deflator (the US deflator), which is based on a combination of "input" and "output" prices, but the compensation between valuations of the various media generally neutralises, by an effect of general equilibrium, the effects of the transition from paid financing and financing through advertising (and vice-versa) on real GDP (e.g. *YouTube*'s "Premium" service). This neutralising effect tends to strengthen one of the hypotheses mentioned above, that of a redistribution of market shares among market operators within a relatively stable perimeter, i.e. a relative positioning of the stakeholders on the market without increasing the size of the market. Finally, the consumer surplus approach suggested by Brynjolfsson & Oh (2012) seems instead to identify a quality effect for consumers, leaving aside any potential price effect, in the sense that the cost of the advertising is partly included in the extra cost of the sponsored products or brands. Here again, what do we actually want to measure? A "missing" consumption or the increased gain for users? The approaches put forward by Brynjolfsson & Oh (2012) and Brynjolfsson *et al.* (2019) seem to be the clearest in terms of the objectives pursued.

But why do we need to measure the value of time spent on domestic tasks, and where do we set the boundaries? Even though digital innovations are recorded as zero consumption in the national accounts, these services, even when they are free, can contribute to consumer welfare. While GDP is not intended to account directly for household welfare, consumption of free services appears for many observers to be a factor that is "missing" from GDP growth and assessments of productivity. However, it is difficult to measure the unobservable value created by activities on the internet, especially free activity, and the different types of proxy proposed do not allow us to clearly identify the question that this type of imputation is attempting to answer. Can we consider the time spent watching adverts to be a measure of the benefit gained from viewing the rest of the programmes? Or is this just, in a purely accounting-based logic, a potential way of accounting for this that neutralises the impact on real GDP of the transition from paid financing to advertising-based financing, without asking ourselves the question of what we are actually measuring in either case? Whatever the answer, we can see that the various answers put forward already include an approach for measuring the value of time spent doing something.

The different methods presented often combine welfare approaches with the more traditional output approach, which may contradict the conceptual framework for measuring GDP and income. Imputing a value for an attention service is not without problems: it increases household income (this service is produced by households) and therefore GDP; however, imputations of this kind tend to disrupt the added value and the underlying employment statistics. This is the case, for example, with imputed rents, which we may be tempted to take from the analysis of productivity. Coyle *et al.* (2018) also report that the imputed income could not truly have a monetary equivalent as it could not be saved or spent on anything elses, nor be fiscally taxed, which weakens the relevance of the GDP conceptual framework. This raises the additional question of the volume-price split, for which Bean (2016) suggests directly evaluating ("output" method) the volume growth in attention time, valued through the increase in the volume of data streams. This proxy, while interesting, does however raise the conceptual question of taking quality effects into account.

The issue of expanding the scope of GDP is a recurring one for national accountants and generally comes up against the risk of an uncontrollable expansion in all directions. This is why the international bodies have historically recommended expanding GDP within the specific framework of satellite accounts, outside of the main scope. To address the issue of what it is we actually want to measure, the Stiglitz-Sen-Fitoussi Commission (Stigltitz *et al.*, 2009) proposed a clarification

by differentiating between means and ends. The goods and services included in GDP or which are eligible to be included are classified as means. The notions of welfare and well-being are in the results category. The results are obtained by jointly implementing GDP components and other means, in particular time available, valuation of intangibles, factors of the social and natural environment, etc.

At this stage, we have examined two of the three conceivable solutions that attempt to give a monetary value to free services (direct assessments of the service provided to users, and valuation through advertising revenue). We will now look at the third option, which is relatively more recent and which, at the time of writing this article, is experiencing a marked increase in popularity.

3. A Third Solution in the Attempt to Attribute a Monetary Value to Free Services: Valuing the Data Generated Through the Use of These Services

According to Varian (2018), data have at least one characteristic in common with oil: they have to be refined to be useful. Without this, they have no intrinsic value. However, unlike oil, data are non-rival, even if they can be made partly excludable through general conditions of use or intellectual property regulations. Varian proposes a pyramid, which is a variation of the "data→ information→ knowledge→ wisdom" hierarchy set forth by Akerlof (1989). In Varian's scheme, data are collected and stored; they are then tidied up and analysed, which allows us to create information (stored in documents), which, once learned, creates knowledge for humans, which ultimately leads to action. In many cases, the data often existed well before we were interested in measuring their value (or in valuing them in new ways), in the form of a deposit of raw materials, of sorts. However, unlike deposits, fresh data can be generated, supplemented, increased and refined, as in the case of the *ImageNet* database, which has fostered recent advances in artificial intelligence. For Varian, data are not innately knowledge assets but can become information assets. This would make data a new production factor. According to Bean (2016), personal data become a production factor, in the same way as physical and intangible capital, which contributes to the productivity and competitiveness of market and non-market producers and creates a substantial surplus for the consumer.

In numerous cases, advertising and data collection both indirectly subsidise recreational services in respect of which we may want to impute a value for household consumption expenditure. Advertising and data collection are partly substitutable (Cecere *et al.*, 2018) and the analogy between the two models offers the prospect, by determining the value of the data, of suggesting an alternative method of measuring the value of free services. In the case of the financing of free services through data collection, the service provider is financed by reselling or immediately using these data or even by building an asset (databases containing the data produced by users, network of internet users, etc.), the value of which can then be increased in the production of services using these data. But is financing through data a simple substitute for financing through advertising? Indeed, in both cases, we may want to explicitly show household consumption expenditure in respect of the free (sponsored) services from which households benefit and intermediate consumption and investment expenditure on the part of the companies that subsidise these services. But these operations probably have quite different characteristics (see Table below).

3.1. Complementarity of Economic Models and Validity of Measuring the Value of Data in the Sphere of Free, Sponsored Services

Cecere *et al.* (2018) highlight the fact that 17.7% of *Google Play Store* applications in 2015 used personal data as a monetisation strategy, which is considerably less than the 32.4% that are financed through advertising, but still a very substantial figure. The data collected and valued include, in particular, user geolocation, contacts and access to SMS. However, in their sample, 53.3% of applications do not have a monetisation strategy, specifically because some are produced by non-market players (such as *Wikipedia*) or are used as "business cards" for the developers or may even be created directly by brands as a means of communication, advertising and influence. The authors are of the opinion that while advertising and integrated purchases are traditional commercial strategies in the internet economy, personal data can complement or even replace these commercial models. The notion of measuring the value of sponsored free services through the collection of data is largely the same as that of advertising-based funding; moreover, one platform monetisation strategy may often incorporate both models in an interlinked and complementary

Table – **Symmetry between financing through advertising and financing through data collection**

Two-sided market	Accounting classification	Indirect financing through advertising	Indirect financing through data and personal data collection
Household side	Potential classification as final consumption expenditure (FC)	Immediate consumption of free service + potential deferred effect of consumption of sponsored products and brand influence	Immediate consumption of free service
Company side	Potential classification of part of the costs as intermediate consumption (IC)	Short-term purchase of sponsored product	Better targeted advertising in the short-term, resale of personal data collected
	Potential classification of part of the costs as investment expenditure (GFCF)	Medium- and long-term effects linked to the brand, logic of influence	Later indirect value measurements, for example through artificial intelligence and algorithms and valuation of intangible assets (organisational, brands, etc.) linked to data

Notes : The companies may be advertisers in the case of an advertising model, or data collectors.

way. Indeed, 7.4% of applications use both advertising and data collection in their monetisation strategy. Consequently, there is no uniformity of economic models, which complicates the standardisation of data valuation methods.

The scalability of networks has increased the scope of free services, irrespective of whether they are financed through advertising or not, using data collection and processing. This new financing (or value measurement) channel is both an immediate profit-generating tool for companies (such as targeted marketing) and a potential future way of increasing value through new and future uses of big data and artificial intelligence. On the basis of this observation, Coyle *et al.* (2018) note that the acquisition of data and creation of free services with the aim of attracting and keeping users through network effects is characteristic of an investment, and recommend that there is a deeper consideration of the way in which the output generated through use of personal data is measured. Currently, however, as mentioned by Ahmad & Schreyer (2016), only the digitalisation of the data and not the inherent value of the data itself can be recorded as an investment in the national accounts, even though it is this intrinsic value that gives rise to the value measurement that justifies the provision of the free service. Imputing values for this type of asset in the national accounts would pave the way for a more general expansion of knowledge capitalisation, the scope of which seems to be difficult to pin down, which undoubtedly explains the reluctance of the community of national accountants to adopt this type of approach. But to claim to incorporate the national accounts framework, these assets would need to be compatible with the scope of "production".

3.2. Generating Data: Is This "Production" in the Conventional Sense of the Term?

The notion of production in the national accounts does not always overlap with the intuitive notion of production of value. This can be seen, for example, in the case of imputed rents, where we record a production of value solely on the basis of being an owner, which deviates from the intuitive idea of productive activity (see Blanchet, this issue). We can legitimately question the notion of "production" by users providing their personal data or leaving traces of their digital activity. This is not work *per se*, but instead a case of providing information, which can be used for multiple ends (targeted marketing, political targeting, various studies, algorithm training, etc.), some of which are still unknown at the time of collection, which could also justify its classification as an asset due to its potential production of future services.

The choice to integrate "data generation" into the scope of production could depend on the nature of the data and the way in which they have been generated. Some data would probably be classified as a produced intangible asset, for example data regarding a car journey generated through GPS tracking, or the construction of personal notoriety on social networks, while others would be a non-produced intangible asset (such as exchanges on social media). Seen in this light, we could consider, in the national accounts, that social networks (by way

of example) pay a rental fee to households for the right to use the non-produced asset created by their personal data, in the same way that we pay rent for the right to occupy undeveloped land, which is a non-produced asset. We should, nevertheless, note that income from the holding of non-produced assets is not considered as production in the national accounts. The digital economy is likely to lead to the registration of new intangible assets, but their production by households does not appear to be as straightforward. We could consider that household consumption of free services produces, whether voluntarily or as a by-product,[5] data that become a factor of production that underlies the intangible assets. Measuring the value of these assets does not, however, seem an easy task.

3.3. Measuring the Value of Data: A Delicate Exercise Due to the Nature of the Underlying Markets

Li *et al.* (2019) believe that data may have a very high and probably increasing value in the near future with 5G and the "Internet of Things" accelerating the accumulation of data, in terms of both type and volume. In order to make this observation, the authors examined several types of online platform that rely on commercial models of data valuation.

However, where this entails a specific shift to numerical values, measuring the value of the data can seem a complex task owing to the nature of the network effects that come with the "platformisation" of the economy, on both two-sided and multi-sided markets. The arrival of the internet and the capacity of networks to exchange large quantities of information on a massive scale (scalability) has created a technological breakthrough. A significant part of the collaborative economy is based on a model of two-sided markets (cf. Box 1), which has largely contributed to the democratisation of the "free service" model. The networks have unique economic properties; the higher the number of users of a networked service, the greater benefit of that service for a user. Hence, a user's decision to participate in a networked service depends on the number of service users and that decision increases the benefit for the existing users. The free nature and ease of use of the service are key factors in triggering these network effects. Once these have been triggered, growth becomes self-sustaining due to the snowball effect and the network reaches a key profitability threshold, which, once passed, increases the profits of the network organiser.

Yet, where the provision of free services allows operators to increase their network effects, the creation of value is not necessarily immediately visible and can sometimes take very indirect forms, which are difficult to quantify. Indeed, these operators increase the value of the intangible assets, but only disclose the value when transactions take place, which means they are often not accessible in real time and are difficult to transpose from one situation to another. Hence, as highlighted by the OECD (2013), the monetary, economic and social value of personal data is likely to be governed by non-linear principles, with increasing returns to scale. Furthermore, Brynjolfsson *et al.* (2019) believe that it would be helpful to further take the network effects into account when evaluating GDP-B.

Thus, the value of data is highly dependent on the context in which they are used. Indeed, while the value of a data record may be very low at an individual level, its value increases as more records are added to the database, as these data can be matched to other data, and information can be inferred from these data to further increase the amount of information available. Thus, for Li *et al.* (2019), unlike R&D activities which depreciate as they become obsolete, data can gain in value and generate new values by combining datasets. In this way, the aggregate value of a set of data is greater than the sum of its components (increasing returns). This is a unique functionality that changes the paradigm of asset depreciation over time and raises unprecedented challenges in terms of measuring value. In this regard, for Brynjolfsson *et al.* (2018), the fact that there is no depreciation means that the intangibles themselves generate intangibles, in a cumulative logic. This would, therefore, be a case of positive externalities. So, what consumption of fixed capital should we apply? This is a novel characteristic of the underlying intangible assets.

Since 2013, the OECD has examined several methodologies for measuring and estimating the monetary value of personal data. The first approach consists in examining the market capitalisations, income or net profit per individual record for the companies whose economic models are primarily based on personal data. However, the data on market capitalisation result in evaluations that can be highly variable. The OECD believes that the most direct way of determining the value of personal data is at the

5. A by-product is technologically fully linked to the production of another product.

intersection between supply and demand, and consists in evaluating the market price at which personal data are legitimately offered and sold. However, this market price is only visible at the time of purchase/sale, and often does not have a real equivalent as there is no true market. An alternative method, based on an observable counterfactual, consists in evaluating the economic costs of data theft. The costs associated with a loss of personal data can provide an assessment of their value; however, the OECD notes that the figures reported here still vary to a considerable extent. More recently, the OECD proposed other types of valuation, for example a cost-based approach (taking into account the costs of production together with a markup) or even an income-based approach. However, this last approach requires hypotheses regarding future cash flows and depreciation costs, which appear to be difficult to determine. In all of these cases, before moving towards an evaluation, it seems to be necessary to define the scope of produced assets and non-produced assets (and non-assets) in the light of the current SNA regulations, even if this means adjusting the boundaries in its future updated version. Assessing the potential value of data as a factor of production that underlies the intangible assets therefore seems to be a major challenge in measuring and understanding the future revolution in artificial intelligence and its implications for changes in productivity.

But if the cost of free services funded by data may, to some extent, be likened to an investment for the companies that rely on this economic model, how do we take purely free services into account?

3.4. Does Measuring the Value of Data Answer the Question of How to Measure the Value of Purely Free Services?

The model of "purely" free services is, on the surface, similar to the notion of volunteer work, but it can also help to increase the value of the intangible assets through data, sometimes without the initial explicit intention of doing so. The distinction between purely free models on the one hand, and free models supported by advertising and/or data collection, on the other, does not take into account the value of the underlying asset, which should appear as an additional wealth-management approach (but closely linked). Indeed, even though there are no market or even monetisation transactions, the content itself and the interaction with internet users may be valued. The platforms collect data produced or left by users (with the aid of cookies, for example) and infer information that improves the depth and quality of their understanding of the users. If the "purely" free model is, on the surface, similar to the idea of volunteering, with digital services, this creates a new economic component: the accumulation of knowledge, data, and their potential valuations as intangible assets.

We will take the example of *Microsoft*'s purchase of *GitHub* in 2018 for $7.5 billion. *GitHub* is a web-based hosting and code-sharing platform. The main service offered by *GitHub* is the provision of repositories accessible online, which is the reason for its enormous popularity for open source projects. *GitHub* is a flagship free-software platform with a culture that advocates the values of transparency, sharing and no-fee and operates mainly from voluntary contributions. Setting aside the pay-to-use section of the platform (which is limited), *GitHub* has no production nor consumption in the national accounts, yet its value was set at $7.5 billion. While *GitHub*'s (public and private) data, which are often free-to-access codes, do not have an intrinsic value, the site's volunteers have interacted and generated data, specifically programming code, which is a valuable resource in the field of big data and artificial intelligence that can be used, for example to automate code generation or improve automated bug detection. The purchase of the data on the site is also an aid that allows *Microsoft* to measure the value of different types of intangible assets (specifically marketing, brand and organisational assets). Indeed, *Microsoft* will benefit from potential synergies with the tools of its own community and will be able to strategically develop its brand image thanks to this shift towards the free software community. This goes back to the dichotomy in which the volunteers have assumed the roles of both consumers and producers on the site, at least in the sense that their consumption has generated, either voluntarily or as a by-product, data output that underlies intangible assets.

Several of the provisions laid down in the SNA 1993 (6.47 and 6.86), and recalled in Article 3.22(c) of the SNA 2008, could apply to this type of situation by allowing for the recording as gross fixed capital formation of the participation of informal unincorporated groups of households in a collective service (for example communal construction activities for the community). Here, this would be a case of own-account production by households. As the majority of inputs are provided free of charge

in such cases, the SNA 1993 recommended retaining a value estimate based on the wage levels for the remuneration of similar types of work, i.e. the use of professionals for the same work. This valuation could, however, be quite extensive.

The difficulties in taking "free services" into account in the central framework for the national accounts surely result, in part, from the hybrid nature of these services at the crossroads between the flow accounts and the balance sheets.

3.5. Generally Speaking, the Flow Approach Should Be More Systematically Linked to a Stock (Asset) Approach in Order to Assess the Reality of Some New Economic Models

For Li *et al.* (2019), the sheer disparity between the *Facebook* IPO in 2011 (value of its total assets: $6.3 billion) and its market valuation in 2013 ($104 billion) highlights the huge amount of its intangible assets, and, in particular, the value of the underlying data. These valuations can be very high, even where these companies are making heavy losses, as in the case of the "ubiquity now, revenue later" models (such as *Uber*), which use investments (venture capital) to rapidly gain market share by offering free or cut-price services. However, the central system of the national accounts records "transaction values" which have, since the SNA 1993, been distinguished from the conventional use of the expression "market values". This means that the properties attributed by neo-classical economic theory to market values cannot be transposed directly to the empirical transaction values.

According to Bean (2016), the digital economy has made it more difficult to measure economic production due to the transition from highly capital-intensive production to highly knowledge-intensive production, with a proliferation of intangible assets. The new characteristics of investments in intangibles are described by Haskel & Westlake (2019) as follows: "The intangible, knowledge-based assets that intangible investment builds have different properties relative to tangible assets: they are more likely to be scalable and have sunk costs; and their benefits are more likely to spill over and exhibit synergies with other intangibles".

The national accounts do, however, capture part of these intangibles. Notably, since the ESA 2010 was introduced, R&D expenditure has been recorded as an investment, rather than as intermediate consumption as before. Likewise, the datasets are already partly taken into account in the concepts laid down in the SNA 2008. With a revolution of artificial intelligence on the horizon and the changes it may bring, there seems to be an urgent need to build complementary reference indicators for intangibles in order to provide new insights into the way in which these phenomena, which are known but difficult to observe, are increasingly helping to shape economic dynamics.

* *
*

The issue of measuring the value of free services clearly does not have an obvious solution, but instead several types of potential solution depending on the objectives pursued. Yet the changes in the organisation of these pseudo-markets may, one day, come to the aid of the national accountants, whether in the form of spontaneous changes or changes resulting from regulatory policies to which these markets may be subject.

The platforms based on a free model are *de facto* in a "price-taker" position, which tends to standardise the value of the service provided for users; however, some users may be more active than others and contribute more actively to the reputation and the content value of some platforms. Users are often also consumers of the content on these sites, but there is nothing that guarantees a balance in this relationship; indeed, there are asymmetries as already mentioned, which are fairly characteristic of typical principal-agent problems. In some cases, the service provided may be significantly inferior to the contribution of the user, who can only request compensation from the platform in rare cases, even where that platform explicitly monetises the user's data "production", or even their personal data. For Li *et al.* (2019), online platform companies capture the majority of the benefits associated with data as they know how to exploit their value, whereas consumers lack the knowledge to give value to their own data. We would therefore be in a case of incomplete contracts. As these asymmetries are sometimes obvious, there are, despite everything, one-off cases of monetary compensation, whether direct or indirect. *YouTube*'s monetisation contracts are one example of this. These remunerate the most active YouTubers depending on their audience, or more indirectly *via* the recognition and notoriety that some influencers may acquire, which

may lead to free products and various invitations (trips, etc.). However, this article will not further examine the question of the value associated with acquired notoriety, which is more similar to an intangible asset for the household.

Can we rely on one proposal for measuring the value of free services, in particular, to steer economic policy? If we consider that measuring user willingness to accept compensation to maintain access to a free service (Brynjolfsson *et al.*, 2019) is probably the most accurate approach to assessing the value of the service provided (marginal cost), we can take a two-pronged approach by evaluating the "extra advertising cost"[6] and the benefits associated with data[7] on the basis of the amounts that platforms would be willing to pay not to part with the data generated by their users. This principle would be compatible with the logic of data portability and could be applied to flows as well as assets, which is important for buyback between platforms. An initial step has already been taken in this direction: the General Data Protection Regulation (GDPR) has, since May 2018, given companies the role of data "custodians" and not owners, and the responsibility of guaranteeing the portability of personal data. This theoretical extra cost cannot be quantified and would likely vary quickly on the basis of multiple parameters (notoriety, trends, etc.). However, this could be used as a basis for the public regulator to reflect on this. For example, the regulator could establish an option right granted to website users, who would then be able to individually or collectively request financial compensation in exchange for storing their data on the site. If this were refused, this would lead to the removal of the account and all information regarding that individual or produced by them. This "pay to keep" model could be applied to the private sector, but should not affect public data, as measuring the value of the latter is already of implicit benefit to the common good. The emergence of economic actors assuming the role of intermediaries on this market would lead to the definition and establishment of an option right approaching the marginal cost of the free service.

Generally speaking, while it is difficult to fully comprehend the notion of value of free services and to introduce a measure of it in the national accounts today, this may be partly because there are still significant asymmetries and the economic vectors underlying this market have yet to be fully formed. The economic policies of the future in this sector may offer statisticians the opportunity to evaluate such phenomena, even though it will not resolve the matter in its entirety. □

6. The extra advertising cost corresponds to collecting the willingness to pay the consumers who have been "influenced" by the advertising. This can, therefore, be likened to a sort of "produced" goodwill in terms of intangible investment (i.e. the gap between the acquisition price and the economic value of the goods in the absence of advertising).

7. The benefits associated with the data correspond to the revenue made from the data and their use, net of fees incurred in acquiring them (platform costs, service costs, etc.).

BIBLIOGRAPHY

Ahmad, N. & Schreyer, P. (2016). Measuring GDP in a Digitalised Economy. OECD*, Statistics Working Papers*, 2016/07. http://dx.doi.org/10.1787/5jlwqd81d09r-en

Akerlof, G.A. (1989). The economic of illusion. *Economics & Politics*, 1: 1–15. https://doi.org/10.1111/j.1468-0343.1989.tb00002.x

Bean, C. (2016). *Independent Review of UK Economic Statistics*. Cabinet Office, United Kingdom. https://www.gov.uk/government/publications/independent-review-of-uk-economic-statistics-interim-report

Blanchet, D. (2020). What Should the Concept of Domestic Production Mean in Globalized Economies? *Economie et Statistique / Economics and Statistics*, this issue.

Brynjolfsson, E. & Oh, J. (2012). The Attention Economy: Measuring the Value of Free Digital Services on the Internet. AIS Electronic Library. https://www.semanticscholar.org/paper/The-Attention-Economy%3A-Measuring-the-Value-of-Free-Brynjolfsson-Oh/9ff9bec84357dacc286b570937a955f358a9a8b5

Brynjolfsson, E., Eggers, F. & Guannamaneni, A. (2018). New Measures of the Economy: Measuring Welfare with Massive Online Choice Experiments - A Brief Introduction. *AEA Papers and Proceedings 2018, American Economic Review*, 108, 473–476. https://www.pnas.org/content/116/15/7250

Brynjolfsson, E., Diewert, E., Fox. K., Eggers, F. & Collis, A. (2019). GDP-B: Accounting for the Value of New and Free Goods in the Digital Economy. *NBER Working Paper*. https://www.nber.org/papers/w25695

Cecere, G., Le Guel, F. & Lefrere, V. (2018). Economics of free mobile applications: Personal data as a monetization strategy. Communication à la 16ème Conférence de IAOS OECD.
https://www.slideshare.net/StatsCommunications/iaos-2018-economics-of-free-mobile-applications-personal-data-g-cecere-f-le-guel-v-lefrere

Coyle, D., Riley, R. & Reinsdorf, M. (2018). Perspectives on Measuring the Digital Economy. Presentation at ESCoE. https://www.escoe.ac.uk/escoe-conference-economic-measurement-2018/

Corrado, C. & Hao, J. (2014). Brands as Productive Assets: Concepts, Measurement, and Global Trends. *World Intellectual Property Organization Economics & Statistics Series Economic Research, Working Paper* No. 13

Cremeans, J. (1980). Consumer Services Provided by Business through Advertising-Supported Media in the United States. *Review of Income and Wealth*, 26, 151–174. https://doi.org/10.1111/j.1475-4991.1980.tb00151.x

Eisner, R. (1978). Total Income in the United States, 1959 and 1969. *Review of Income and Wealth*, 24(1), p. 49. https://doi.org/10.1111/j.1475-4991.1978.tb00031.x

Eisner, R. (1988). Extended Accounts for National Income and Product. *Journal of Economic Literature*, 26(4), 1611–1684. https://www.jstor.org/stable/2726857

Goolsbee, A. & Klenow, P.J. (2006). Valuing Consumer Products by the Time Spent Using Them: An Application to the Internet. *American Economic Review*, 96(2), 108–113. https://doi.org/10.1257/000282806777212521

Haskel, J. & Westlake, S. (2017). *Capitalism without Capital: The Rise of the Intangible Economy.* Princeton University Press.

Jaszi, G. & Juster, T. (1973). A Framework for the Measurement of Economic and Social Performance. *In: The Measurement of Economic and Social Performance*. New York: Milton Moss, p. 50, and Jaszi, Comment, p. 89.

Jorgenson, D. & Fraumeni, B. (1992). Investment in Education and U.S. Economic Growth. *The Scandinavian Journal of Economics*, 94, S51-S70. https://doi.org/10.2307/3440246

Kendrick, J.W. (1971). Expanding Imputed Values in the National Income and Product Accounts. *Review of Income and Wealth*, 25(4), 349–364. https://doi.org/10.1111/j.1475-4991.1979.tb00112.x

Li, W., Nirei, M. & Yamana, K. (2019). Value of Data: There's No Such Thing as a Free Lunch in the Digital Economy. *U.S. Bureau of Economic Analysis Working Papers*.
https://www.bea.gov/research/papers/2018/value-data-theres-no-such-thing-free-lunch-digital-economy

Nakamura, L. & Soloveichik, R. (2015). Valuing "free" media across countries in GDP. *Federal Reserve Bank of Philadelphia* Working Papers 15-25. http://papers.ssrn.com/abstract=2631621

OECD (2013). Measuring the Internet Economy: A Contribution to the Research Agenda. *OECD Digital Economy Papers*, No. 226. http://dx.doi.org/10.1787/5k43gjg6r8jf-en

OECD (2013). Exploring the Economics of Personal Data: A Survey of Methodologies for Measuring Monetary Value. *OECD Digital Economy Papers*, No. 220. http://dx.doi.org/10.1787/5k486qtxldmq-en

Okun, A. M. (1971). The Mirage of Steady Inflation. Brookings Papers on Economic Activity*, Economic Studies Program, The Brookings Institution*, vol. 2(2), pp. 485–498.

Roy, D. (2013). La contribution du travail domestique au bien-être matériel des ménages : une quantification à partir de l'enquête Emploi du Temps. *INSEE, Document de travail* N° F1104.
https://www.insee.fr/fr/statistiques/1380932

Stiglitz, J., Sen A. & Fitoussi, J.-P. (2009). *Report by the Commission on the Measurement of Economic Performance and Social Progress*.
https://www.ladocumentationfrancaise.fr/var/storage/rapports-publics/094000427.pdf

Vanoli, A. (2002). *Une histoire de la comptabilité nationale*. Paris : La Découverte, Collection Grands Repères Manuels.

Varian, H. (2018). Artificial Intelligence and Antitrust: competition, regulation and pluralism in the online World. Presentation in Florence. https://fsr.eui.eu/event/competition-regulation-pluralism-online-world/

Irish GDP Growth in 2015: A Puzzle and Propositions for a Solution

Marie-Baïanne Khder*, Jérémi Montornès and Nicolas Ragache*****

Abstract – In July 2016, the Irish statistical institute significantly revised GDP annual growth in 2015 from 7% to 26%. This revision does not correspond to a similar increase in employment nor in the accumulation of new physical capital, but to the relocation of preexisting intangible assets by multinationals to Ireland. This article provides a comprehensive depiction of the effects of these relocations on the Irish GDP and balance of payments in 2015. We question the need to change the accounting standards defining the macroeconomic aggregates and the framework for economic analysis. We conclude that an effort to adapt and revamp the standards of national accounts is thus necessary to achieve a consistent recording of multinationals' transactions, crucially by clarifying the concept of economic ownership over production and intellectual property and then by facilitating its implementation.

JEL Classification: E01, F20, F40, F62
Keywords: national accounts, multinationals, globalization

** DGTrésor et Crest, Insee at the time of writing (marie-baianne.khder@dgtresor.gouv.fr); ** Banque de France, Insee at the time of writing (jeremi.montornes@banque-france.fr); *** AFEP, Insee at the time of writing (n.ragache@afep.com)*

*We thank D. Blanchet, R. Mahieu, S. Roux, N. Ahmad, O. Simon, T. Laurent and two anonymous referees for very helpful comments. We are also grateful to all participants to the National accounts Workshop "*Atelier de Comptabilité Nationale*", the internal Insee seminar (DEE) and the AFSE annual conference 2019. We remain responsible for errors and omissions.*

Received November 20, 2018, accepted after revisions July 17, 2020.
Citation: Khder, M.-B., Montornès, J. & Ragache, N. (2020). Irish GDP Growth in 2015: A Puzzle and Propositions for a Solution. *Economie et Statistique / Economics and Statistics*, 517-518-519, 173–190. https://doi.org/10.24187/ecostat.2020.517t.2026

In July 2016, the Central Statistics Office (CSO) significantly revised the Irish GDP growth in 2015 from 7% to 25.6% (CSO, 2016a). This upward revision has not been matched by a similar revision of employment nor of the accumulation of physical capital. Rather than stemming from new production capacities, it results mainly from the relocation of preexisting and intangible assets (research and development, software, etc.), worth €300 billion, by a small number of large multinationals enterprises (MNE) within their Irish legal units. This episode calls into question the principles and the rules according to which national accountants assign production to a territory, in a context of rapid movements of assets from one region of the world to another.

To begin with, the amount of intangible assets relocated in Ireland in 2015 was sufficiently sizable to have dramatic and indirect macroeconomic consequences on the Irish economy. Ireland, referred to as the "Celtic Tiger" in the 1990s, is a small and open economy. In 2008-2009, Irish banks were severely affected by the financial crisis in the wake of the collapse of the real-estate bubble. Since 2010, the Irish economy has recovered progressively driven by stronger exports. Foreign-owned subsidiaries accounted for approximately 70% of industrial production, 60% of exports of goods, 40% of exports of services and 60% of imports of services in 2015. Before 2015, employment was roughly in line with GDP. However, in 2015, compared to the sudden GDP growth, no paralleling shift in the employment level occurred.[1]

Another feature of Ireland is the tax and legal environment surrounding intangible assets. On one hand, the Irish tax system is the most favorable to companies in European Union (EU) and also, in 2015, compared to the US tax system.[2] This tax system is favorable to the establishment of MNE and to the relocation of intangible assets in Ireland. Since 2009, Ireland broadly extended the class of intangible assets eligible to a capital allowances scheme that enables companies to deduct their expenditure on the acquisition of eligible intangible assets from their taxable income, even when these assets are acquired from related parties (i.e. group subsidiaries). In 2015, to tackle tax avoidance schemes such as the "Double Irish with a Dutch sandwich" double non-taxation scheme, Ireland passed measures in the 2015 budget to close those loopholes.[3] However, the extension of the scope of the capital allowance for intangible assets has allowed a 0% effective tax rate on the associated income.[4] The Irish tax system is also particularly attractive for research and development (R&D), via a tax credit of 25%, and thanks to a patent box (the "Knowledge Development Box") that allows companies to deduct from their taxable income the product derived from their patents and then tax them at the rate of 6.25% (instead of the statutory rate of 12.5%).

On the other hand, the legal protection of income, the membership to the EU common market and the euro area, also confer Ireland advantages in terms of market access and regulation, per se and over offshore centres and tax havens (see e.g. Raspiller, 2005). In October 2015, personal information on European consumers was the subject of significant European regulatory attention. In particular, the European Court of Justice invalidated the automatic exchange procedures under the safe harbor regime, considering that European consumer data in the United States (US) were not sufficiently protected under current European standards. This trend has intensified, as highlighted by the 2016 General Data Protection Regulations. These developments in digital regulation are changing the relative attractiveness of each country for locating intangible assets in or outside the EU. This has in turn provided further incentives for US MNEs to strengthen their data processing activities within subsidiaries located in the Single Market. The same applies to other regulatory aspects specific to the EU market (pharmaceutical products, transport services, etc.). These elements add up so that Ireland is considered by some MNEs as an optimal place to register business in Europe, especially in case of intensive use of intangible assets.

The national accounting mechanisms at work in Ireland in 2015 are summarized below: a limited number of MNEs have transferred mainly intangible assets[5] and aircraft, from their balance sheets to resident units in Ireland. These transfers of assets and liabilities have deteriorated Ireland's external position. In return, these Irish resident units have become owners

---

1. Irish unemployment gradually decreased as the unemployment rate fell from 15% to 5% between 2010 and 2019 but no sharp decrease occurred in 2015.

2. The statutory corporate tax rate in Ireland is 12.5% compared to an average of around 22% for the European Union (European Commission, DG TAXUD, 2018).

3. By ending the use of this scheme for new tax plans, and implementing a staggered ban for established structures. Following the announcement, companies could still implement such a scheme during a three-month window.

4. In the 2015 budget, while the double Irish tax scheme ended, the 80% rule was abolished so firms could claim tax relief on up to 100% of profits from their Intellectual Property investment (Taylor, 2017).

5. These include R&D or commercial patents, trademarks, etc.

of MNEs' international production. They then receive payments generated from the production they own. This has led to a substantial increase in Irish exports and to a lesser extent in Irish imports, because these Irish units are remunerated directly from the proceeds of the sale of goods or services produced abroad traded under contract manufacturing. The sharp increase in the exports of goods by Irish resident units is recorded even though these goods are materially produced in the rest of the world and never crossed the Irish border.[6] As a result, the Irish resident units themselves have also been a source of income for the non-resident units, which ultimately own them and which had transferred the associated intangible assets to them in the first place. Property income payed to the rest of the world has also increased. Indeed, Ireland benefits on the one hand from the proceeds of the sales of products it now owns, and on the other hand pays dividends or reinvested profits to non-resident shareholders. These profits are also partly reused to reconstitute of the intangible capital newly recorded on the companies' balance sheets, leading to an increase in Gross Fixed Capital Formation (GFCF). All these changes could be observed because of the modest size of the Irish economy[7] and, conversely, the significant size of transfers. The first-rank counterparts of the asset transfers are not directly identifiable in the available statistical sources. To track "phantom" investments, these statistics need to be supplemented with data on global interconnections (Damgaard *et al*., 2019), including for tax havens.

The contribution of this paper is twofold. First, it provides a comprehensive and as detailed as possible picture of national accounts and balance of payments developments as well as GDP growth in Ireland in 2015. Its novelty is to trace the impact of the relocation of intangible assets on GDP and more generally on national accounts and the balance of payments. Second, without alternatives, we infer that a change in national accounting standards should occur to deal with the kind of episodes that happened in Ireland. This overhaul of accounting guidelines should be pursued, aimed at allocating multinationals' revenues from intellectual property to countries on the basis of economic considerations.

The paper is structured as follows. The first section presents GDP developments between 2014 and 2015 according to "expenditure" and "income" approaches and to what has triggered these unusual developments. The second section presents how national accounting principles assign economic activity to a territory and to which extent the Irish case challenges these underlying principles of national accounting. In particular, we highlight the role of economic property as a fundamental concept of national accounts. The third section reviews four solutions that have emerged so far: *i)* the release of complementary indicators such as the modified gross national income (GNI*) of the CSO; *ii) ex post* correction using formulary apportionment; *iii)* a change in national accounting rules; and finally *iv)* the enrichment of GDP modeling in the field of macroeconomics to better account for intangible capital as a production factor. In particular, we summarize the pros and cons of those four ways forward after the Irish case.

1. Investigating the Developments of the 2015 National Accounts and Balance of Payments

According to the CSO's publication July 2016, GDP in 2015 increased by 34.7% in value and 25.6% in volume terms. This development immediately seemed «abnormal» to observers (see e.g. Krugman, 2016). First, it contrasts with the pace of Irish GDP growth over the recent period. In comparison, growth was 1.6% between 2012 and 2013, then 8.3% between 2013 and 2014. Secondly, it is not caused by a positive shock in the domestic demand (higher public spending, higher consumption, etc.). On the financial side, the determinants of production (interest rates, oil prices, exchange rates) are close or identical to those observed in the euro area. Finally, this GDP increase does not translate into an increase in income for Irish households. This change in GDP reflects a sharp increase in trade balance that has no equivalent in other European countries.[8] The methodological notes gradually published by the CSO nevertheless highlighted the consistency of this evolution with that of the balance of payments without providing the full picture because of the rules of statistical secrecy. A detailed analysis of the developments in the components of GDP, GNI and Ireland's international investment position is thus presented.

6. For example, a smartphone or medicine are not necessarily owned by the industrial unit, which manufactures them. They are rather owned by the multinational company that immediately controls their marketing and that can allocate ownership right among its other units.
7. Ireland's GDP represents almost 2% of the euro area's GDP.
8. The public finance situation has been under surveillance since the 2007-08 crisis, with Ireland being one of the countries that received assistance from the European Union and the IMF due to the sharp increase of the deficit and public debt, which itself resulted from bank failures.

1.1. Demand Components Developments

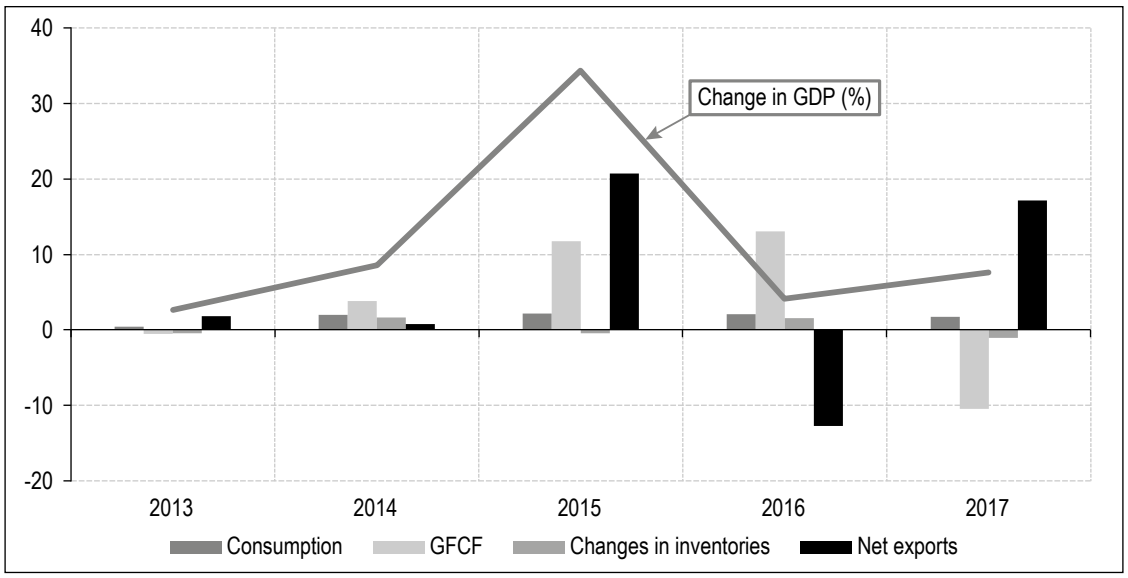
Two thirds of the increase in GDP in value between 2014 and 2015 is explained by Ireland's trade in goods and services, which contributes almost 21 points (Figure I). In addition, the GFCF's contribution to GDP growth amounts to 12 points. It also continued in 2016 and was followed by a backlash in 2017. Higher foreign trade and the increase in GFCF therefore substantially altered the composition of GDP in level and in terms of dynamics. In 2016 and 2017, the volatility of the demand components increased because their profile is more closely linked to the decisions of MNEs whose weight is now higher. The subsidiaries of resident MNEs in Ireland and hosting these balance sheet transfers received the export and import proceeds of goods whose production requires the assets they hold. Consequently, the expenses and income of the international production in which they participate were assigned to these units.

1.1.1. Exports and Imports without Border Crossings Due to Contract Manufacturing

This increase in imports and exports does not correspond to trade in physical goods but to margins made abroad and integrated into trade in goods. More specifically, most of the change in Ireland's trade balance in national accounts comes from the increase in trade adjustments, from cross border basis to ownership basis, including "goods for processing" and "contract manufacturing". Contract manufacturing occurs when a domestic company hires a company abroad to manufacture products on its behalf (and vice versa). These products could be either finished or semi-finished products, part of a value-chain. Crucially, the inputs and output in this production process remain in the ownership of the domestic entity and a change of economic ownership is not deemed to occur during this subcontracting process. Indeed, the foreign contract manufacturer supplies a manufacturing service to the Irish entity and never takes ownership of the product being product (CSO, 2016b).

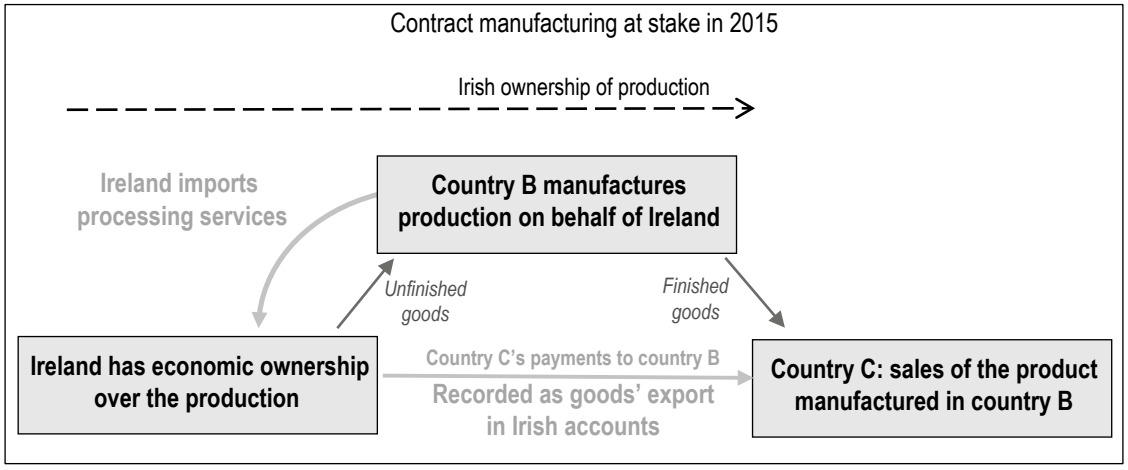
As an example, contract manufacturing occurs in the electronics sector when the originator provides the inputs of smartphones and the sub-contractor manufactures the finished goods. The flow chart in Figure II details the case of goods purchased, processed, and sold abroad underlined in the Economic Statistics Review Group (ESRG) report (2016). A resident unit in Ireland, part of a global production process, receives sales products of the goods manufactured abroad. The export is only recorded when a change of ownership occurs with the sale in Country C. In detail, inputs are purchased from abroad by the Irish company; the materials are sent and transformed into final goods by the processor (in country B), possibly using intangible assets or services that belongs to the Irish unit; the physical goods are then sold to the final consumer (in country C) without ever

Figure I – **Gross domestic product and main aggregates (variation in value and contributions in percentage points)**



Sources: CSO, National Accounts.

Figure II – **Exports of contract manufacture when the economic ownership is located in Ireland**



Notes: The diagram shows how, starting from trade data (in italics), balance of payments adjustments (in grey, bold): increase the goods' export value to the final sale value; may increase the goods' import value by the cost of material inputs delivered directly to country B for incorporation into production; and record imports of processing services in Ireland.

entering Ireland. The Irish company makes a profit as the owner of the product and possibly providing intangible assets (trademark, design, etc.) into the process. This profit enters the value added.

From a national accounts standpoint, production physically carried out abroad is considered as Irish production as long as an Irish resident unit has ownership, and the income from the sale of this production is thus recorded in the Irish GDP (see Section 2). Ireland's goods for processing and contract manufacturing exports increased by €60 billion between 2014 and 2015, accounting for more than two thirds of the increase in Irish exports in goods (+€86 billion in exports of goods). These levels were maintained in 2016 and 2017 (Table 1). Among these adjustments, €17 billion pertain to adjustments to goods exports to China in 2017.

Conversely, Ireland's trade balance in services is deteriorating sharply, with imports of services increasing by €53 billion while exports of services are growing by €20 billion (Table 2). In particular, imports of R&D services increase by €20 billion. This corresponds to the net acquisition of additional intellectual property products (IPP). Moreover, imports of royalties and license fees for the use of intellectual property rights are also increasing by more than €20 billion. Irish resident units therefore increased their payments of royalties and license fees to non-residents in return for permission to use intellectual property rights (patents, copyrights, trademarks, industrial

Table 1 – **From customs data to national accounts (in billions of euros)**

		2013	2014	2015	2016	2017
International trade (Cross-border basis)	Exports	89.2	92.6	112.4	117.6	122.5
	Imports	55.8	62.2	70.1	72.1	76.7
+ Goods for processing	Exports	7.1	18.6	78.6	67.6	64.7
	Imports	7.2	10.2	13.6	11.6	5.6
+ Net exports of goods under merchanting	Exports	3.7	3.5	6.4	5.3	7.6
+ Other conceptual adjustments	Exports	-1.3	-0.2	2.9	3.6	-1.9
	Imports	1.3	1.3	3.2	4.5	3.0
Merchandise (Ownership basis)	Exports	98.7	114.5	200.3	194.1	192.9
	Imports	64.2	73.7	86.9	88.2	85.2

Notes: This table breaks down the transition of customs data, which measures international trade in goods when crossing the Irish border, to imports and exports according to national accounts (i.e. based on the change in ownership criterion).
Sources: CSO, Trade statistics and National Accounts.

Table 2 – **Exports and imports of services (in billions of euros)**

		2014	2015	2016	2017
All services	Exports	99.9	120.2	135.1	161.8
	Imports	105.4	158.0	198.8	205.3
Insurance and Financial services	Exports	19.2	22.8	22.6	25.3
	Imports	13.3	17.4	18.2	19.9
Computer services	Exports	42.0	50.4	58.1	68.1
	Imports	0.6	1.1	1.4	3.2
Royalties and licence fees	Exports	5.2	7.3	8.1	9.1
	Imports	43.3	63.8	69.2	66.7
Research and developement services	Exports	2.3	1.8	4.0	6.7
	Imports	8.7	28.2	58.1	55.2
Other services not elsewhere stated	Exports	31.2	37.9	42.4	52.6
	Imports	39.5	47.5	51.9	60.4

Sources: CSO, Balance of Payments.

processes, etc.) or to use originals or prototypes produced (manuscripts, paintings, etc.) under licensing agreements.

The growth of imports of R&D services reflects that MNEs reshuffled the allocation of property and use of intangibles in 2015. R&D imports are recorded in case of outright results of different activities (patents, copyrights, etc.), which by the way renders royalties and license fees payments to use those intangible assets unnecessary. For example, cost-sharing agreements between a Irish resident unit and a US R&D centre, which are pointed to as a means to transfer intellectual property products quickly and at virtually no cost, enable Irish resident units to get ownership on IPP developed in the US provided the formers pay a fee to US units, covering for the R&D development costs. This fee is then also recorded as R&D import. This mechanism was at play in 2015 in Ireland, within MNEs (Richard Harvey, 2020; Coffey, 2018) and specifically *Apple* (Brehm Christensen & Clancy, 2018) contributing to the increase in R&D imports. R&D imports development in 2015 is then offset by investment developments, resulting in no effect on GDP. However, intangible assets that were relocated in 2015 by some MNEs were transferred to units before they became resident in Ireland and the relocation was in this case not recorded as R&D imports and matching investment, but as change of volume, see below. Royalties and license fees (which when imported, correspond to payments made by an Irish resident unit to overseas against the right to use IPP it does not own), also increased by €20 billion in 2015. This hints at the fact that some Irish firms might be continuing to use foreign intellectual property, but this increase is in line with the trend observed in the years prior to 2015.

1.1.2. The Current Account

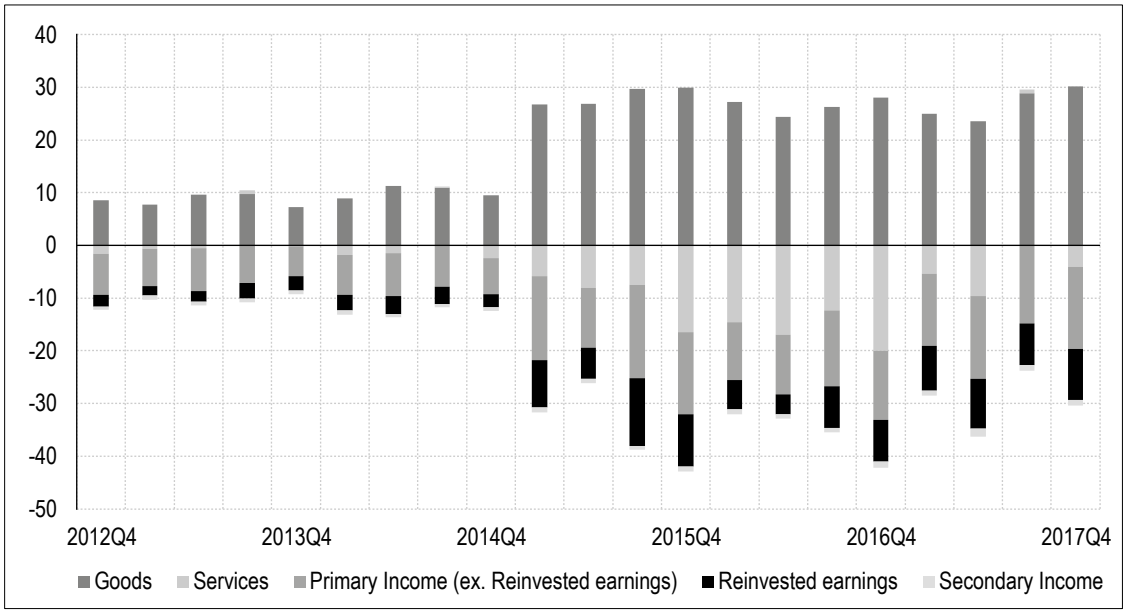
Figure III shows the current account of Ireland mirroring goods, services, primary and secondary income balances. As explained above, the balance of trade in goods improved in 2015 since Irish subsidiaries had become owners of goods traded under contract manufacturing. The increase in imports of R&D services, and, to a lesser extent, royalties, explains the degradation of the balance of trade in services. Besides, the net primary income balance,[9] which measures transfers of income between resident and non-resident institutional units remunerating the provision of labor or capital (wages of cross-border workers, flows of interest or dividends on securities held by non-resident agents, etc.), deteriorated by almost €30 billion between 2014 and 2015, as investigated below.

1.2. Income Developments

The analysis of the sequence of income also enables to identify the main mechanisms at work, and to better understand the deterioration of the net primary income balance. At current prices, Irish GDP rose from €195.3 billion in 2014 to €262.5 billion in 2015, a variation of €67 billion compared with €15 billion between 2013 and 2014 (Table 3).

9. This balance is defined as income received by resident institutional units in the rest of the world, minus income paid by resident institutional units to non-resident institutional units.

Figure III – **Ireland's current account (flows in billion euros)**



Sources: CSO, Balance of Payments.

Several observations can be drawn from this sequence of accounts. GNI is equal to GDP plus primary income flows received from abroad and minus primary income flows paid abroad (i.e. the net factor income, see Box). In Ireland, the increase in GNI growth is less pronounced than GDP, but still significant (€37 billion), and triggers a matching increase in gross national disposable income (GNDI). This means that almost 40% of the increase in GDP involves Irish resident units that are owned by the rest of the world, and that it is re-payed by those resident units to their final owner overseas. This payment does not need to effectively take place in the form of dividends: profits "reinvested in the subsidiaries" (i.e. undistributed) are also recorded as outflows. However, even if reinvested earnings mainly appear as outflows, a fraction of these reinvested earnings corresponding to depreciation (recorded as consumption of fixed capital) still remains parked within the Irish resident units by convention (see Online Appendix C1-A – link to Online appendices at the end of the article). The income outflows corresponding to the reinvested profits are reduced by the amount of fixed capital consumption. Provisions for depreciation indeed increased by €27 billion in 2015, explaining two thirds of the increase in GNI.

The analysis of income developments shed light on the distortion of usual links between macroeconomic aggregates. GNDI was not consumed in the usual proportions that existed before 2015.

Table 3 – **GDP and use of gross national disposable income (billions of euros at current prices)**

	2013	2014	2015	2016	2017
Gross domestic product (a)	179.9	195.3	262.5	273.2	294.1
Net factor income from the rest of the world (b)	-28.1	-30.4	-60.8	-50.1	-59.9
Gross national income (c =a+b)	151.8	164.9	201.7	223.2	234.2
Current transfers from the rest of the world (d)	-2.9	-2.7	-3.3	-3.6	-4.5
Gross national disposable income (e= c+d)	148.9	162.2	198.3	219.5	229.7
Total consumption expenditure (f)	111.3	114.9	119.2	124.7	129.5
Gross national savings (g= e-f)	37.6	47.3	79.2	94.8	100.3
Provision for depreciation (h)	26.7	28.8	56.5	63.9	72.0
Net national savings (i=g-h)	10.9	18.4	22.7	30.9	28.3

Sources: CSO, National Accounts.

Box - **Key identities of National accounts and Balance of payments**

In the following, the crucial transactions in the Irish case are marked with stars and a companion explanation is provided.

National accounts are based on the three approaches of GDP:

Income approach

GDP = compensation of employees + gross operating surplus* + gross mixed income + (taxes - subsidies) on production and imports

* Extra profits have been recorded by resident companies in Ireland

Expenditure approach

GDP = consumption + investment* + government spending + net exports of goods and services**

*Investment refers to gross fixed capital formation, which in particular includes depreciation on the capital stock, also known as consumption of fixed capital

** Exports of goods includes contract manufacturing. Import of services includes R&D services

Production approach

GDP = gross value added* + (taxes - subsidies) on production and imports

* The surge in value added is mainly recorded as manufactured production

The **Gross National Income (GNI)** is derived from the GDP. In Ireland, the GNI is inferior to the GDP by around €50 billion in 2015. The income outflows, mainly due to foreign MNEs which established subsidiaries in Ireland, far exceeds income that Irish resident units derive from investment abroad.

Gross national income

GNI = GDP + net primary incomes (interest, dividend, reinvested earnings and other primary income)*

* 'Primary income' less 'Other primary income' = 'Net Factor Income' mentioned in the paper

The Balance sheet account

The estimates of the stock of assets (K(t)) are usually computed with the Perpetual Inventory Method. We report here the law of motion of capital to clarify that relocations of assets would enter as an 'Other change in volume' in the sequence of accounts.

K(t) = K(t-1) - depreciations(t) + investment(t) + other change in volume(t)

The Balance of payments records all transactions made between entities in one country and the rest of the world. Balance of payments is consistent with the 'Rest of the world' sector in national accounts.

Current account (CA)

CA = net exports + net primary incomes* + net secondary incomes

CA = national savings - national investment

Financial account (FA)

FA= net acquisition of financial assets* - net acquisition of financial liabilities*

* Foreign direct investments, portfolio investments and other investments

Capital account (KA), defined such that:

KA + CA + FA =0

Net international position (NIPP)

NIPP(t) = NIPP(t-1) + current account (t) + other change in volumes* (t) + valuation effects (t)

*Relocation of assets recorded in the 'Other change in volumes' is, in the Irish case, similar to the asset side

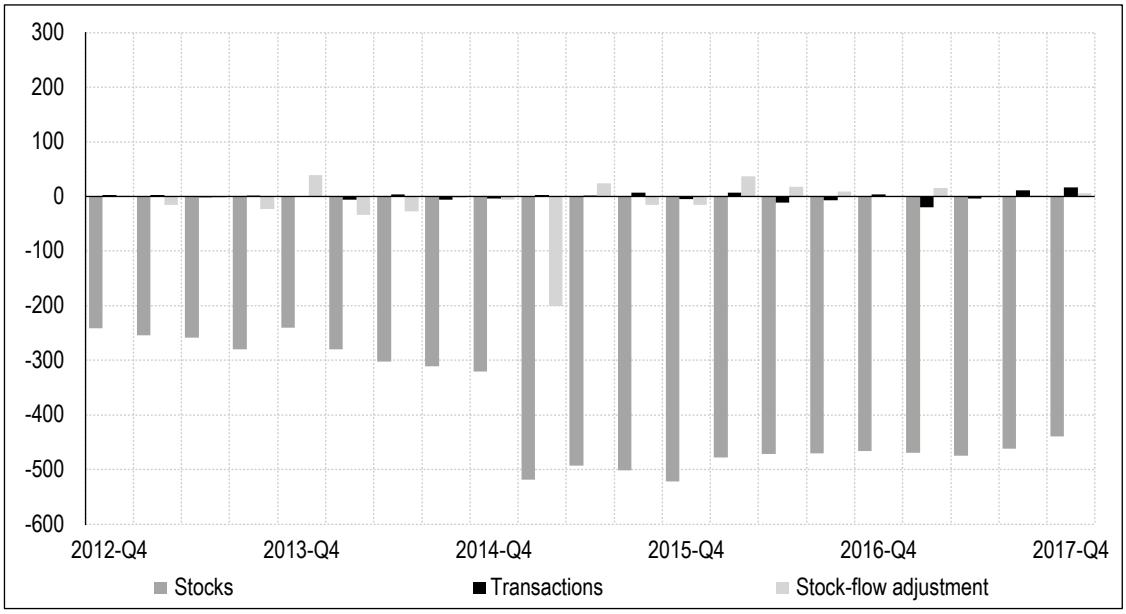
In 2014, final consumption represented 71% of the GNDI, this ratio was only 60% in 2015. The increase of disposable income by €36 billion in 2015 mainly led to an increase in national savings of €32 billion, including €27 billion hoarded by companies in the form of 'Provisions for depreciation', with virtually no impact on consumption. In total, the "new" value added gives rise essentially to two types of transactions involving Irish resident companies: repayment to foreign units and, above all, provisions for depreciation.

1.3. The Relocation of Assets as a Trigger

In total, €300 billion of intangible assets were transferred from the rest of the world to Ireland (ESRG, 2016, p. 8). Consistently, from 2014 to 2015, the external position recorded financial counterparts of net assets transfers as a stock-flow adjustment (more precisely a change in volume). Because those net assets were not newly produced, they are not recorded as an economic transaction (i.e. not as GFCF). Overall, these relocations led to a net decrease of €200 billion in Ireland's external position in the first quarter of 2015 (Figure IV).

The variations in Irish GDP in 2015 stemmed from the relocation of intangible assets from the rest of the world to Ireland. This may seem paradoxical because asset transfers via stock-flow adjustment do not constitute production as such. However, the relocation of intangible

Figure IV – **Ireland's international investment position (in billions of euros)**



Notes: The dark grey bars show the quarter stocks of foreign assets and liabilities. The difference between two end-quarter positions can be transactions (in black) or the impact of other changes in volumes and valuation (in light grey) occurring in the same period.
Sources: CSO, Balance of Payments.

assets indirectly affects GDP: with respect to the expenditure approach to GDP, those assets gave to Irish resident units economic ownership over some new goods, which led to an increase in net exports after accounting for contract manufacturing adjustments. To replenish the stock of intangible assets required, additional GFCF has to be carried out to offset the high share of depreciations.

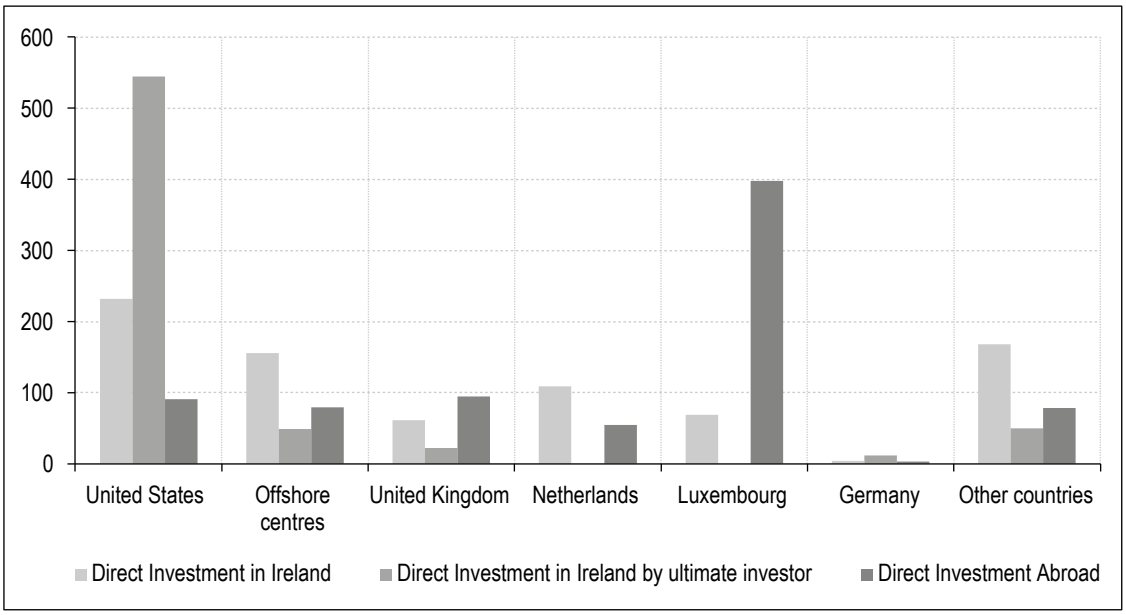
To understand the origin of these developments, we would want to access the geographical breakdown of the international investment position. The deterioration of net external position results mainly of portfolio investments[10] for which no geographical breakdown is available. At the same time, foreign direct investment assets and liabilities also increased dramatically, and the geographical origin of the direct investment stock is publicly available. The main direct counterpart countries are the US (€232 billion), Luxembourg (€69 billion), the Netherlands (€54.8 billion) and various offshore centres (€156 billion). These direct inward investments represent liabilities of Irish resident units that correspond to the financial first-rank counterparts of the intangible assets relocated to Ireland. Ownership of these intangible assets has been transferred to Irish resident units but they remain ultimately held by the rest of the world. Besides, direct investments by incoming offshore centres in Ireland are significant (€156 billion) when recorded under the immediate investor principle, as in figure V, but much smaller (€49 billion) under the ultimate investor principle (i.e. the ultimate country from which the investments originate, cf. Online Appendix C1-B).

In contrast, the US is an immediate investor in Ireland for €232 billion but is actually an ultimate investor for more than the double (€545 billion).[11] At the same time, Irish resident units increased their outward foreign direct investment by €325 billion between 2014 and 2015; their geographical distribution is different, reflecting the complexity of the reorganization at work. Assets are held for €397 billion by units resident in Luxembourg, €91 billion in the US and €11 billion in the United Kingdom (Figure V). Outward FDI statistics are calculated according to the direct beneficiary principle (and not the ultimate beneficiary), and Luxembourg's share in Irish outward FDI according to this principle reveals MNEs tax and jurisdiction arbitrage.

The change in the net FDI and portfolio investment position between 2014 and 2015 reflects

10. This is consistent also with the fact that the Irish resident unit sheltering Apple*'s relocated intellectual property (which is used for the sales outside the US that are themselves recorded as profits in the Irish resident unit) may have contracted loans towards other* Apple*'s subsidiaries outside Ireland (probably Jersey, see Brehm Christensen & Clancy, 2018).*
11. https://www.cso.ie/en/releasesandpublications/er/fdi/foreigndirectinvestmentannual2018/

Figure V – **Geographical distribution of direct investment in Ireland in 2016 (stocks in billions of euros)**



Sources: CSO, Balance of Payments.

a shift in balance sheets between the rest of the world and Ireland: new entries are simultaneously recorded in the liabilities and assets of resident units. For the most part, these assets do not result from new investments but from changes in the legal and/or geographical allocation of property rights between MNEs' units now located in Ireland.

2. The Measurement Problems Raised by the Irish Case

The extent of Irish GDP growth and the difficulty of rationalizing it as an evolution of national production have been much debated. Krugman (2016) and Fitzgerald (2018) expressed concerns about the source of growth and the economic relevance of such an accounting entry. Few companies were involved, so that the CSO could not give further explanations without infringing the rules of statistical confidentiality, which prevent access to the business statistics underlying the construction of the accounts.[12]

International institutions were immediately concerned with the accounting validity of this unprecedented GDP growth, to check whether it stemmed from misinterpretations of the rules (Stapel-Weber & Verrinder, 2016). The OECD (2016) pointed out in particular the difficulties of interpreting the concepts of resident unit, economic ownership and, overall, the implementation of the national accounting framework in the context of global production arrangements. In addition, the IMF (2016), in view of the assistance program received by Ireland, has paid close attention to Irish macroeconomic statistics.

2.1. Taking Globalization into Consideration in National Accounts

National accounts have gradually developed, taking into account the evolution of the concepts of territory, production and economic units since the second half of the 20th century. The uses of national accounts, the availability of data sources and the need for comparability have all played a role in its expansion (Vanoli, 2002). In particular, national accounts have developed with reference to the model of production by resident units on the national territory. In this framework, international trade is carried out between resident and non-resident companies. Imports and exports ensure that the economy's balance between resources and uses is achieved. Similarly, income transfers linked to the production process – dividends, wages – lead to the correction of GDP to GNI by tracing the balance of primary incomes with the rest of the world.

The Irish case shows that this framework is questioned by the globalization of production chains and situations where production is simultaneously carried out in several countries, organized

12. See https://www.cso.ie/en/aboutus/lgdp/csodatapolicies/statistical-confidentiality/

in a fragmented way, with the circulation of semi-finished products according to sophisticated contractual arrangements, involving off-market trade (at transfer prices between group units) and dissociating the commercial and financial aspects from physical production. The Irish case is made even more complex by the relocation of intangible assets. Therefore, when value added is created, where should it be located by national accounts (Avdjiev *et al*., 2018)?

The System of National Accounts (SNA) does not define the producer as one who physically participates in the production activity but as one who owns the product that is being processed (United Nations Statistics Division - UNSD, 2008). This principle is fundamental because it is the basis for consistency between the production and income approaches. However, this principle leads to the recording of the value added that results from physical production abroad in the country of residence of the owner of the product. For instance, a "factoryless" company that has designed a good but relocated its production uses subcontractors to produce the various elements and assemble them. The production of the various components and assembly can take place in several countries, all potentially different from the producer's country of residence. National accounts then allocate the value added to the "factoryless" producer's country of residence.

2.2. The Role of the Legal Unit

In national accounts, the definition of "domestic" production is based on that of resident[13] institutional units. Resident units are those that have a predominant centre of economic interest in the economic territory of the country. A "centre of economic interest" indicates that the unit carries out economic activities and large-scale operations on the economic territory for an either indefinite or fixed but relatively long period of at least one year (ESA 2010, 1.61). Some resident units may be re-domiciled (CSO, 2016c).[14]

The legal existence of a society does not automatically imply an economic «existence» from the national accounts perspective, the latter corresponding to the concept of institutional unit. According to the ESA (2010, 1.57), "institutional units are economic entities that are capable of owning goods and assets, of incurring liabilities and of engaging in economic activities and transactions with other units in their own right." This definition is detailed in ESA (2.12): "An institutional unit is an economic entity characterized by decision-making autonomy in the exercise of its principal function. A resident unit is regarded as constituting an institutional unit in the economic territory where it has its centre of predominant economic interest if it has decision-making autonomy and either keeps a complete set of accounts, or is able to compile a complete set of accounts."[15] Some subsidiaries within groups are "legal units", but may not be institutional units from national accounting standpoint.

The ESRG (2016) indicates the reasons why Ireland has become the predominant economic focus of subsidiaries receiving intangible assets: *(i)* the units in question are incorporated and registered in Ireland; *(ii)* the staff and in particular the senior management reside there; *(iii)* the units in question compile a complete set of accounts; and *(iv)* they have decision-making autonomy in economic matters. Eurostat's audit also agrees with the nature of resident institutional unit of the entities responsible for the increase in Irish GDP.

Nevertheless, the criterion of decision-making autonomy may remain difficult to establish within a group and sometimes, even in the relationship between a principal and a subcontractor. In the Irish case, whether foreign MNEs that have relocated intangible fixed capital (R&D, patents, etc.) to their Irish subsidiary have or not delegated operational decision-making process in Ireland has been questioned. The complexity of the organization of the MNEs concerned and the statistical secret also introduce uncertainty about the proper understanding of the classification of entities and the relations between them.

2.3. The Implications of the Economic Property Criterion

Since the SNA 2008, national accounts have used the criterion of change of economic ownership to record a transaction. This economic

13. Residence in the sense of national accounting slightly differs from tax residence.

14. Re-domiciliation is the relocation of the headquarters in Ireland of foreign multinationals that previously had only a subsidiary in Ireland, According to the CSO, the re-domiciliation of companies is not the main phenomenon underlying the Irish GDP growth in 2015.

15. "To enjoy decision-making autonomy in the exercise of its main function, an entity must: (a) *be entitled to own property and assets independently; it must be able to exchange ownership of property or assets in transactions with other institutional units;* (b) *have the capacity to make economic decisions and carry out economic activities for which it is held legally responsible;* (c) *have the capacity to enter into commitments, incur debts and other obligations and enter into contracts in its own name;* (d) *have the ability to establish a complete accounting system, i.e. a balance sheet of its assets and liabilities, and accounting documents showing all the transactions it has carried out during the accounting period."*

property is defined as the fact of bearing the benefits and risks associated with the use of an asset in a production. However, in the context of intra-group relations, determining whether a subsidiary enjoys economic ownership of a production is not straightforward (UNECE[16], 2015, 3.11). When economic ownership cannot be unequivocally defined, the legal ownership criterion is used by default.[17] According to the SNA 2008, legal ownership is characterized by the possibility for an institutional unit to "claim, as of right and under the law, the benefits associated with these entities" (UNSD, 2008, 3.21). Thus, while legal ownership corresponds to being able to claim an "advantage" by law, economic ownership consists of being able to claim an "advantage" *(1)* in the context of an economic activity and *(2)* by accepting the corresponding risks and *(3)* in the context of a use (see Online Appendix C1-C, for more details on the concept of legal ownership).

The difference between economic and legal ownership refers to a fundamental principle of national accounts: the distinction between production and income distribution operations. Indeed, production requires the economic ownership of the factors of production – capital and inputs – and of the product, but without all the criteria of legal ownership having to be met, since it may be sufficient to have the right to use the asset and enjoy its product. Conversely, income distribution operations refer to the ability to allocate the income received (related to exploitation, transfer, asset stripping) through legal ownership over an asset. The right of ownership therefore makes it possible to transfer income or risk as in the case of shares or bonds. Indeed, the ESRG (2016) highlights the fact that the relocation of intangible assets has reduced payments from Irish subsidiaries to non-resident units in return for the right to use intellectual property.[18]

Defining economic ownership is even more complex in the case of an intangible asset. Indeed, while the Irish resident unit may own a relocated intangible asset in the legal sense, it is difficult to decide on the origin of the relocation decision (Connolly, 2017). In the Irish case, the resident units receiving the intangible assets simultaneously saw their liabilities towards the rest of the world increase, which indicates that foreign entities keep ultimate control on the relocated assets. UNECE (2015, 3.56) warns in the general case that economic ownership may remain in the hands of a parent company and may never have been transferred to one of its subsidiaries even if the legal ownership of intellectual property has been transferred. In addition, the subsidiary may be a special purpose entity established to receive legal ownership of IPP and/or to centralize the associated income for tax optimization purposes. In this case, because economic ownership would be too difficult to determine without further guidance notably from tax authorities, UNECE (2015) recommends that national accountants record economic ownership in accordance with the legal declarations of the special purpose entity.[19] The distinction between economic and legal ownership therefore appears difficult in the era of transfers of intangible goods, as recognized by UNECE (2015). Frequent ownership stripping situations for intangible assets also contribute to blurring the notion of "economic ownership" based on "use". UNECE (2015) provides a decision tree of to define economic ownership over IPP (p. 50, Figure 4.1), but the criterions remain difficult to assess in the face of complex legal and contractual relationships within groups. Were the classification criterions too "blurry", the determination of economic ownership over IPP would be volatile and subject to disputes. The typical case would be the transfer of patent use rights from a parent company to a subsidiary, as for example under a cost-share agreement (Benshalom, 2006). Although the contractual situation is clear – every stakeholder knows what he can do with regard to the different contractual attributes of the partial transfer of ownership – the economic property as defined by UNECE (2015) needs further analysis. This therefore calls for clarifying and revamping the concept of economic ownership to make it more applicable.

In total, the changes in the Irish GDP in 2015 illustrate the difficulties in interpreting global production arrangements in accordance with national accounting rules, particularly with regard to the concepts of institutional units and economic ownership. By default, the legal

16. The national accountants grouped within the United Nations Economic Commission for Europe (UNECE), already alerted by the developments of cases of international production and their complexity in relation to the simple model of unified production on a single site, have addressed this subject in a guide on the effects of globalization.

17. For practical reasons, because the legal units are entitled to file financial statements. Consequently, it is often necessary to be a legal unit before being a resident economic unit (UNECE, 2012).

18. "In the past, the impact of contract manufacturing activities on exports of goods was largely offset by imports of Research & Development services, as Irish companies made payments to non-resident parts of the group for the use of intellectual property. However, when the intellectual property is located in Ireland, as seen in the results for 2015, these offsetting charges do not occur, and the full effect of contract manufacturing is attributed to GDP." (ESRG, 2016, p. 36)

19. "Applying the principles of economic ownership to such cases, in contrast to legal ownership, would be extremely difficult. National accountants usually have no alternative than to follow reality as reported by these SPEs i.e. recognize them as separate institutional units. Consulting the tax authorities may be a way to obtain a better understanding of the nature of these SPEs." (UNECE, 2015)

criteria of legal units and ownership overtook those of institutional units and economic ownership in the Irish case.

3. Four Ways Forward

GDP and GNI developments in Ireland challenge the economic analysis (sustainability, competitiveness, etc.), by substantially changing the debt and deficit ratios as a share of GDP or altering the computation of multi-factor productivity. They also led to operational uncertainties, for example regarding the increase in Ireland's contribution to the European budget. In this context, Eurostat conducted a methodological audit in 2016, concluding that the existing national accounting rules were respected, and validated the use of the revised Irish GDP in the context of the European excessive deficit or macroeconomic imbalance procedures (Eurostat, 2016a, 2016b). At the same time, the CSO also argued that there were no errors and that the accounting treatment was correct. The CSO mandated the ESRG[20] to consider these new phenomena. In this section, four ways forward are reviewed and discussed starting with the proposals made in the ESRG report.

3.1. New Complementary Indicators

The report contains thirteen recommendations (ESRG, 2016). The main conclusion was that the traditional indicators (GDP and GNI in particular) should be maintained, but that it was also necessary to add, at the same rate of publication, a modified gross national income neutralizing the effects of MNEs on GDP (the so-called GNI*) and a net national income. GNI* equals GNI minus the factor income of re-domiciled companies in Ireland and the consumption of fixed capital[21] on the imports of R&D services and trade in intellectual property[22] and on aircraft leasing in Ireland. GNI* is thus a hybrid concept (neither gross[23] nor net) that requires a separate national account of MNEs' subsidiaries. Indeed, GNI does not correct for all of MNEs' assets relocation, notably because the consumption of fixed capital on some foreign direct investments keeps being recorded in Ireland (Lane, 2017). The earnings reinvested in Ireland, which are removed from GDP to compute GNI, are computed net of consumption of fixed capital; the associated consumption of fixed capital therefore remains recorded in Irish GNI and GDP.

In 2015, the GNI* growth rate was 8.6% – compared to a rate of 26% for GDP. However, there are limits to this indicator. By nature, GNI* is an *ad hoc* aggregate, designed specifically for Ireland. At this stage, GNI* is mainly used by international organizations (IMF, European Commission, etc.) to compute Irish debt ratios for instance. The other users (academics, economic press, etc.) continue to refer to the GDP despite the level shift in 2015. This calls for alternative ways in addition to the publication of new complementary indicators.

3.2. Correcting *ex post* the Macroeconomic Aggregates for MNEs' Operations

A second approach lies in making an *ex post* correction of national accounts aggregates in order to single out the statistical distortion induced by MNEs, so that the resulting aggregates do not reflect the volatility of intangible capital location. Guvenen *et al*. (2017) and Bruner *et al*. (2018) take into account the US intra-group redistribution of income for tax optimization purposes and therefore seek to correct the US national accounts aggregates. Guvenen *et al*. (2017) based their analysis on the following hypothesis: US MNEs can decide to register, at no cost, a fraction of their income in foreign branches with more lenient taxation, by optimizing the registration of the legal ownership of intangible assets. The US shareholders, who ultimately hold these assets, financed and supported the R&D and innovation process, continue to be paid on the income recorded in foreign branches. Nevertheless, in this analysis, such profit shifting should entail a lower US GDP because part of national production is recorded in foreign affiliates and higher income on US direct investment abroad due to reinvested earnings.

To assess what the US GDP would be if profits of US MNEs currently recorded in foreign affiliates in low-tax jurisdictions were to be reallocated to the US rather than "repatriated" via returns on direct investments abroad, Guvenen *et al*. (2017) implement a "formulary apportionment" method similar to that used by tax experts. The

---

20. The ESRG, which brought together several stakeholders in these debates - academics, administrative and Eurostat and IMF experts - had the task of better assessing the effects of globalization on indicators derived from national accounts and the balance of payments. On this topic, see also Holton et al. *in this issue.*

21. In national accounts, consumption of fixed capital is roughly equivalent to depreciation in private general accounting.

22. The difference between GNI and GNI in 2015 almost exclusively stems from the correction for the consumption of fixed capital on the imports of R&D services and trade in intellectual property, that is consistent with a massive relocation of intangible assets. It amounts to roughly €30 billion.*

23. Gross refers to an aggregate that includes consumption of fixed capital (i.e. depreciation of assets), unlike net aggregates.

"formulary apportionment" consists in allocating the global profits of a MNE according to *(i)* the share of the wage bill represented by the country at stake in the MNE's total wage bill, and *(ii)* the share that the country represents in terms of sales to non-affiliated entities made by the MNE. The results obtained are not sensitive to the choice of criteria *(i)* or *(ii)* for ventilation. Although the study concluded that 65% of the returns on direct investment abroad are reassigned to US GDP, the final impact on GDP remains limited in the case of the US: over 2004-2014, this adjustment amounts to an average of $260 billion per year, roughly an annual 1.5% of 2014 GDP. However, a correction of the same magnitude would have substantial consequences for smaller economies, and in particular for Ireland. Guvenen *et al*. (2017) also estimate that $30 billion of the total amount reallocated to the US GDP based on their correction would come from Ireland, which represents about 13% of the Irish GDP in 2012. A recent statistical analysis estimates an even greater volume of profits shifted to Ireland, at around $117 billion in 2015 (Tørsløv *et al*. 2018)

Although promising, the *ex post* adjustment of macroeconomic aggregates is not without vulnerabilities. Suarez-Serrato (2018) shows that, following the repeal of the provisions in the Internal Revenue Code which enabled the US MNEs to shift profits to affiliates in Puerto Rico, the MNEs reacted to the increase in their overall tax burden by reducing employment and investment in the US, and increasing investment in their foreign subsidiaries. Changing the tax system therefore affects MNEs' organisation, and *ex post* correction of macroeconomic aggregates cannot sufficiently account for such feedback loops. Correcting *ex post* GDP also requires a review of the entire sequence of accounts to ensure consistency (Bruner *et al*., 2018). From a statistical point of view, the "formulary apportionment" method also requires detailed data on the activities and country-by-country financial statements of each entities of MNEs. Above all, Guvenen *et al*. (2017) or Bruner *et al*. (2018) aimed at providing an order of magnitude of profit shifting but did not suggest the release of adjusted statistics by national statistical institutes on a regular basis.

3.3. Changing the Rules of National Accounts

Lequiller (2019) paved a third way forward by urging a change in national accounts rules. He acknowledges that the 2015 Irish unprecedented growth, whose cause is a balance sheet movement rather than new production, illustrates that current rules can lead to a measure of GDP inconsistent with its use as an indicator of national production. Lequiller (2019) suggests to exclude R&D or software assets from capital and investment as it was the case in the previous manual of national accounts (SNA 1993). Lequiller (2019) also hints at the difficulty of distinguishing between production and financial operations and thus proposes to exclude "goods for processing" from transactions. In the same vein, Tedeschi (2018) advocates for a separation of the entire "offshore sector" of the Irish economy.

However, these proposals overlook intangibles as a source of economic growth. Recording intellectual property assets makes it possible to identify their contribution to value added in the analysis of productivity. Moreover, as pointed by Ahmad *et al*. (2018), production is generated through the use of R&D whether or not recognized and capitalized as such in the national accounts. Should national accountants stop capitalizing intangible assets, the income derived from these assets would still exist, but would not properly be explained by the traditional factors of production. In total, excluding intangible assets would have led to a more modest GDP growth in 2015 but does not address the economic issue of profit and revenue shifting at stake in Ireland.

Reviewing the current national accounting rules to provide a consistent recording of the global activity of MNEs resorting to profit shifting appears necessary: globalization is now one of the major topics on the agenda of the international System of National Accounts. However, we should focus on a clarification of the concept of economic ownership over production and intellectual property rather than removing/decapitalizing the latter.

3.4. Including Intangible and Mobile Capital in Production Functions

The Irish GDP growth in 2015 draws attention to an increasing difficulty in the analysis of aggregate production. Standard economic theory provides a guidance about the way to locate production only in an extreme case by assuming a linear production function (i.e. making the inputs separable or perfectly substitutable, see Online Appendix C2). How to deal with an increase in production that does not result from an increase in traditional factors of production such as employment, hours worked or physical capital? Moreover, in usual business

cycle analysis, the GDP level is explained by the demand fluctuations in relation to potential GDP. Traditionally, potential GDP depends of three components: the volume of hours worked – determined by demographic factors and the labor market –, the available capital – determined by investment – and multi-factor productivity – its determinants including levels of training, market organization and technological progress –. In the short term, these factors cannot change significantly. The novelty of the Irish case is that rapid and persistent movements in GDP are due to changes on the supply side in relation to the international mobility of intangible capital.

Integrating intangible capital, which is in essence more mobile, into business cycle models makes it possible to account for the rapid movements in an economy's supply, but not without difficulties (Corrado *et al*., 2009). First, intangible capital is not easy to define and measure (Thum-Thysen *et al.*, 2017). The problem of its valuation is acute, particularly because intangible assets are generally not traded on markets between independent players, but are the subject of intra-group transactions (Dischinger & Riedel, 2011). It depreciates more quickly than physical capital, and loses market value in the case of a patent that has fallen into the public domain. In addition, the inclusion of intangible capital poses a problem in estimating potential output, where the inclusion of physical capital already makes it difficult to estimate multi-factor productivity.

However, macroeconomic models incorporating intangible capital address issues such as changes in the labor share, corporate profitability differentials between home companies and foreign affiliates or gains from trade. For example, Koh *et al*. (2016) show that the decline in the labor share in the value added in the United States over the past 65 years is almost entirely due to the relative increase in the remuneration of intellectual property, while the share of traditional physical capital remuneration is stable. McGrattan & Prescott (2010) also develop a multi-country general equilibrium model that integrates an intangible capital called technological capital that is exclusive (i.e. cannot be used outside the MNE that owns it) but non-rival (i.e. can be used simultaneously by all entities belonging to the MNE). Their model allows them to explain 60% of the gap between the return on investment of American MNEs on their direct investments abroad and the return on investment of foreign MNEs in the United States. Using the same theoretical framework, Kapička (2012) explains and quantifies the movements of indirect investment abroad in the US the gains from trade.

Including intangible capital in an aggregate production function therefore makes it possible to understand movements in GDP but shifts the focus to the determinants of the allocation and accumulation of intangible capital within each country.

* *
*

The case of Ireland in 2015 is a telling illustration of the challenges posed by globalization to the measurement of economic activity, since part of the income and some factors of production are extremely mobile across jurisdictions. In particular, intangible assets such as patents or customer personal data play a major role in the new volatility of income. Besides, significant developments in GDP may result from restructuring within or between a few large groups.

So far, the national accounts rules face operational difficulties in their implementation, or in the availability of sources, but also in the interpretation of some of its core concepts such as economic ownership. In the case of Ireland, the GDP indicator has deviated from the measurement of production on the national territory. This step aside from the objectives traditionally assigned to GDP is all the more significant that the economy is "small" and "open", as is the case for Ireland. Consequently, national accounts are the subject of intense debate on how to take into account global value chains following UNECE (2015), for instance with respect to the definition of economic ownership, ownership over intellectual property products and control within multinationals. The alternative proposals that have emerged so far are either not fully satisfactory or not yet implemented and fully ripe. A continuous effort to adapt and revamp the standards of national accounts is thus necessary to achieve a consistent recording of transactions within MNEs, crucially by clarifying the concept of economic ownership over production and IPP, and by making it more applicable. This requires enhanced further exchange of information on MNEs between national accountants across countries. This effort should not aim at providing an *ad hoc* smoothing of macroeconomic aggregates in the manner of the GNI*, because the increased volatility of data series also inform on profound changes in economies,

for instance the increasing role of MNEs, and tax or legal competition across countries.

Even after the US tax reform of 2018, which aimed at reducing the tax base erosion and profit shifting in a context of international fiscal competition, the Irish statistical office continues to record new asset transfers. In the second and fourth quarter of 2019, Irish subsidiaries have become the new owners of intellectual property assets transferred from foreign subsidiaries within large groups. As a result, investment and imports surged in the Irish quarterly accounts. A phenomenon of a comparable order of magnitude on Irish investment and imports had already taken place in the quarter second 2017. The systematic approach of the 2015 Irish episode therefore paves the way for further research on the effect of localization of intellectual property products on GDP. □

Link to Online Appendices: https://www.insee.fr/en/statistiques/fichier/4770164/ES-517-518-519_Khder-etal_Online_Appendices.pdf

BIBLIOGRAPHY

Ahmad, N., Moulton, B., Richardson, J. D. & van de Ven, P. (2018). The Challenges of Globalization in the Measurement of National Accounts, Introduction. NBER conference March 2018. Forthcoming in *NBER Book Series Studies in Income and Wealth*. https://www.nber.org/books/ahma-1

Avdjiev, S., Everett, M., Lane, P. R. & Shin, H. S. (2018). Tracking the international footprints of global firms. *BIS Quarterly Review*, March, 49-66. https://www.bis.org/publ/qtrpdf/r_qt1803f.htm

Benshalom, I. (2006). Sourcing the Unsourceable: The Cost Sharing Regulations and the Sourcing of Affiliated Intangible-Related Transactions. *Virginia Tax Review*, 26, 631.

Blanchet, D., Khder, M.-B., Leclair, M., Lee, R., Poncet, H. & Ragache, N. (2018). La croissance est-elle sous-estimée ? Insee Références − L'économie française édition 2018, 59–79.
https://insee.fr/fr/statistiques/3614240?sommaire=3614262

Brehm Christensen, M. & Clancy, E. (2018). Exposed: Apple's golden delicious tax deals, Is Ireland helping Apple pay less than 1% tax in the EU? Report commissioned by the European Parliament GUENGL.

Bruner, J., Rassier, D. G. & Ruhl, K. J. (2018). Multinational profit shifting and measures throughout economic accounts, in: NBER conference March 2018, Chapter 12. Forthcoming in *NBER Book Series Studies in Income and Wealth*. https://www.nber.org/chapters/c14142.pdf

Coffey, S. (2018). What Apple did next? *Blog Economic Incentives*.
http://economic-incentives.blogspot.com/2018/01/what-apple-did-next.html

Connolly, M. (2017). The Expected and Unexpected Consequences of ESA 2010 – An Irish perspective. *Journal of the Statistical & Social Inquiry Society of Ireland*, 47, 39–70. http://hdl.handle.net/2262/86012

Corrado, C., Hulten, C. & Sichel, D. (2009). Intangible capital and US economic growth. *Review of Income and Wealth*, 55(3), 661-685. https://doi.org/10.1111/j.1475-4991.2009.00343.x

CSO (2016a). GDP increases significantly in 2015. Press release.

CSO (2016b). Contract Manufacturing. CSO Information Notice.

CSO (2016c). Redomicilied PLCs in the Irish balance of payments. CSO Information Notice.

Damgaard, J., Elkjaer, T. & Johannesen, N. (2019). What Is Real and What Is Not in the Global FDI Network? *IMF Working Paper*.
https://www.imf.org/external/pubs/ft/fandd/2019/09/the-rise-of-phantom-FDI-in-tax-havens-damgaard.htm

Dischinger, M. & Riedel, N. (2011). Corporate taxes and the location of intangible assets within multinational firms. *Journal of Public Economics*, 95(7-8), 691–707. http://dx.doi.org/10.1016/j.jpubeco.2010.12.002

ESRG (2016). Report of the Economic Statistics Review Group. ESRG, Dublin, December 2016.
http://www.cso.ie/en/csolatestnews/eventsconferencesseminars/resrg/

European Commission DGTAXUD (2018). Taxation trends in the European Union Eurostat (2010). *European System of Accounts – ESA 2010.*
http://ec.europa.eu/eurostat/fr/web/products-manuals-and-guidelines/-/KS-02-13-269

Eurostat (2016a). Irish GDP revision. Directorate C: National Accounts, Prices and Key Indicators, December 2016. http://ec.europa.eu/eurostat/documents/24987/6390465/Irish_GDP_communication.pdf

Eurostat (2016b). Impact of globalisation and multinational enterprise activities on GDP and other macroeconomic statistics. 63rd Meeting of the Partnership Group – Doc PG 2016/63/3.4, Ljubljana, Slovenia, 16 December 2016.

Eurostat (2010). European System of Accounts.
https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-02-13-269

Fitzgerald, J. (2018). National Accounts for a Global Economy: The Case of Ireland. In: NBER conference March 2018, Chapter 13, forthcoming in *NBER Book Series Studies in Income and Wealth*.
https://www.nber.org/books/ahma-1

Guvenen, F., Mataloni Jr, R. J., Rassier, D. G. & Ruhl, K. J. (2017). Offshore profit shifting and domestic productivity measurement. *NBER Working Paper* N° w23324. https://www.nber.org/papers/w23324

Holton, N., Kinsella, M., Mangan, O., McLaughlin, S. & Quill, P. (2020). Consistency in a Globalised Economy: Aligning the Treatment of R&D in the Irish National Accounts and Balance of Payments. *Economie et Statistique / Economics and Statistics*, 517-518-519

IMF (2016). Ireland Staff Report For The 2016 - Article IV Consultation and fifth Post-Program Monitoring Discussions – Supplementary Information. IMF July 21, 2016.
https://www.imf.org/en/Publications/CR/Issues/2016/12/31/Ireland-2016-Article-IV-Consultation-and-Fifth-Post-Program-Monitoring-Press-Release-and-44140

Kapička, M. (2012). How Important is Technology Capital for the United States? *American Economic Journal – Macroeconomics*, 4(2), 218–48. https://doi.org/10.1257/mac.4.2.218

Koh, D., Santaeulàlia-Llopis, R. & Zheng, Y. (2016). Labor share decline and intellectual property products capital. Queen Mary University of London, School of Economics and Finance *Working Papers* N° 873.

Krugman, P. (2016). Leprechaun economics. *Irish Times*, 31 July.

Lane, P. (2017). The Treatment of Global Firms in National Accounts. *Central Bank of Ireland Economic Letters* 01/EL/17.
https://www.centralbank.ie/docs/default-source/publications/economic-letters/economic-letter-vol-2017-no-1.pdf

Lequiller, F. (2019). Les multinationales rendent-elles le PIB obsolète ? *Variances* N° 14.
http://variances.eu/?p=3923

McGrattan, E. R. & Prescott, E. C. (2010). Technology capital and the US current account. *American Economic Review*, 100(4), 1493–1522. https://doi.org/10.1257/aer.100.4.1493

OECD (2016). Irish GDP up 26.3% in 2015? www.oecd.org/sdd/na/Irish-GDP-up-in-2015-OECD.pdf

OECD (2017). OECD/G20 Project on Tax Base Erosion and Profit Transfer - Addressing the Tax Challenges of the Digital Economy.
www.oecd.org/tax/beps/public-consultation-document-addressing-the-tax-challenges-of-the-digitalisation-of-the-economy.pdf

Raspiller, S. (2005). La concurrence fiscale : principaux enseignements de l'analyse économique. Insee DESE, *Document de travail* N° G2005/07. http://www.epsilon.insee.fr:80/jspui/handle/1/5858

Richard Harvey, J. (2017). Expert commentary on the impact of proposed tax reform.
https://www1.villanova.edu/villanova/law/newsroom/webstories/2017/1213.html

Stapel-Weber, S. & Verrinder, J. (2016). Globalisation at work in statistics - Questions arising from the 'Irish case'. *EURONA: Eurostat Review on National Accounts and Macroeconomic Indicators*, *2*(2016), 45–72.

Suarez Serrato, J. C. (2018). Unintended Consequences of Eliminating Tax Havens. *NBER Working Paper* N° w24850. https://doi.org/10.3386/w24850

Taylor, C. (2017). Noonan and Donohoe rule-switching could cost €1bn next year. *The Irish Times*. 8 November

Tedeschi, R. (2018). The Irish GDP in 2016. After the disaster comes a dilemma. *Questioni di Economia e Finanza.* Bank of Italy, Economic Research and International Relations Area Occasional Papers N° 471.
https://www.bancaditalia.it/pubblicazioni/qef/2018-0471/index.html?com.dotmarketing.htmlpage.language=1

Thum-Thysen, A., Voigt, P., Maier, C., Bilbao-Osorio, B. & Ognyanova, D. (2017). Unlocking investment in intangible assets in Europe. *Quarterly Report on the Euro Area (QREA)*, 16(1), 23–35.

Tørsløv, T. R., Wier, L. S. & Zucman, G. (2018). The Missing Profits of Nations. *NBER Working Paper* N° w24701. https://doi.org/10.3386/w24701

UNECE (2012). *Guide on Impact of globalization on national accounts*. https://www.unece.org/stats/na.html

UNECE (2015). *Guide to Measuring Global Production.* https://www.unece.org/publication/guide_mgp.html

UNSD (2008). *System of National Accounts.* https://unstats.un.org/unsd/nationalaccount/sna2008.asp

Vanoli, A. (2002). *Une histoire de la comptabilité nationale*. Paris : La Découverte, collection Repères.

Consistency in a Globalised Economy: Aligning the Treatment of R&D in the Irish National Accounts and Balance of Payments

Niamh Holton, Margaret Kinsella, Oisín Mangan, Shaun McLaughlin and Patrick Quill *

Abstract – Increasing globalisation of the world's business economy presents challenges to compilers of official statistics when producing consistent economic accounts. This paper examines two inconsistencies between national accounts and balance of payments – firstly, in the measurement of depreciation of intellectual property products, and secondly, in the recording of expenditure on research and development services. These inconsistencies can lead to sizeable discrepancies in net factor flows in the national accounts. Business data collected by the Central Statistics Office in Ireland is used to measure the divergence between the balance of payments and national accounts. The findings suggest that it is necessary to adjust the balance of payments data related to intellectual property products and R&D services to ensure consistency with national accounting methodology.

JEL Classification: F23, F62, O34
Keywords: globalisation, balance of payments, intellectual property, research and development, depreciation

*Central Statistics Office, Dublin (Niamh.Holton@cso.ie; Margaret.Kinsella@cso.ie; Oisin.Mangan@cso.ie; Shaun.McLaughlin@cso.ie; Patrick.Quill@cso.ie)

We are grateful to Jennifer Banim, Michael Connolly, other colleagues in the Central Statistics Office and two anonymous reviewers for helpful comments on an earlier draft of this paper.

Received on 20 July 2018, accepted after revisions on 6 May 2019.
Citation: Holton, N., Kinsella, M., Mangan, O., McLaughlin, S. & Quill, P. (2020). Consistency in a Globalised Economy: Aligning the Treatment of R&D in the Irish National Accounts and Balance of Payments. *Economie et Statistique / Economics and Statistics*, 517-518-519, 191–204. https://doi.org/10.24187/ecostat.2020.517t.2025

Balance of payments data are an essential input into the production of national accounts statistics. In particular, the net factor flows to/from abroad are subtracted from/added to gross domestic product (GDP) in the calculation of gross national income (GNI). In an economy like Ireland's, where the difference between GDP and GNI is large, these flows are particularly important. Differences in methodology between national accounts and balance of payments have the potential to affect the measurement of the factor flows and consequently GNI. In this paper, we focus on the challenges and possible inconsistencies that may arise when measuring a globalised economy. More specifically, we examine two examples – both of which concern differences in the treatment of research and development (R&D) by national accounts and balance of payments – and describe the approach taken to resolve the inconsistency.

In company accounts, and in the balance of payments data, expenditure on research and development appears as an expense and is deducted from profits. Expenditure on intellectual property products (i.e. patents) is treated as investment in intangible assets, which are added to the balance sheet and give rise to depreciation.

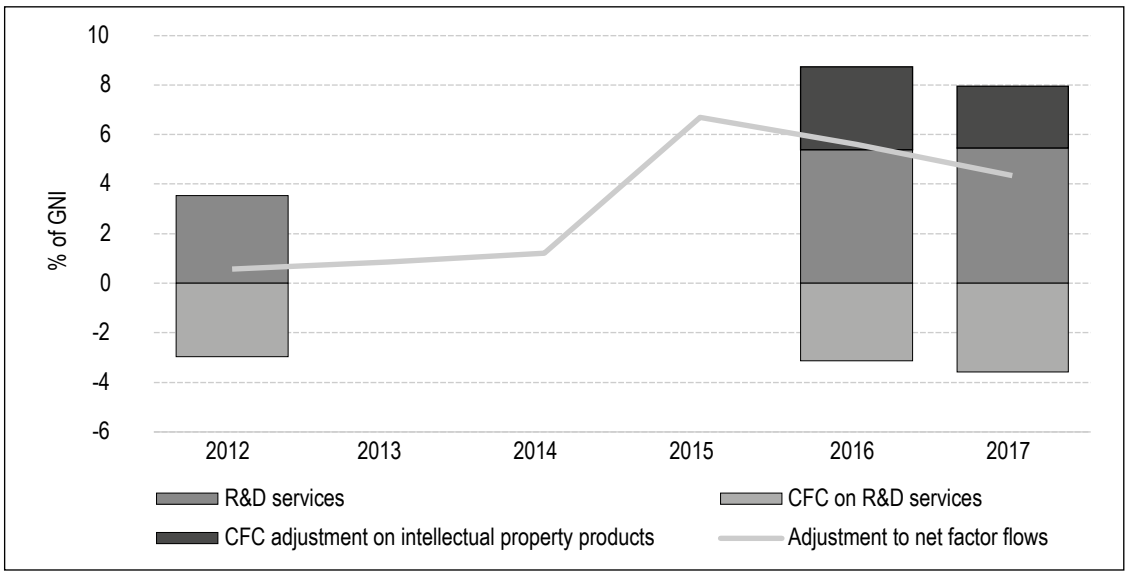
In national accounts, however, no distinction is made between expenditure on research and development services and expenditure on intellectual property products. Both are treated as investment in intangible assets. Both are added to the capital stock of fixed assets and give rise to depreciation. This difference in approach causes an inconsistency between the balance of payments factor flows and the national accounts net operating surplus, in that the net operating surplus figures have been adjusted to include the expenditure on R&D services, and to exclude any depreciation of those assets, while the equivalent figures in balance of payments have undergone no such adjustment.

A second inconsistency was also identified, where the depreciation of intellectual property assets recorded by companies was not in agreement with the depreciation recorded in the national accounts.

Since 2015, when trade in R&D services and intellectual property products in Ireland increased starkly, these have become pressing issues in the calculation of economic statistics, requiring adjustments to address discrepancies. Figure I illustrates the magnitude of the overall adjustment made to net factor flows. It shows the contribution made by the realignment of depreciation for large intellectual property assets, and that made by the treatment of R&D services as investment (and the depreciation which results from this).

We start by considering the literature and guidelines surrounding R&D activity and intellectual property products, before discussing the

Figure I – **Adjustment to net factor flows, Ireland, 2012-2017**



Notes: Data on the individual components of the adjustment for 2013-2015 are withheld for confidentiality reasons.
Sources: Authors' calculations.

challenges in recording these activities in a small globalised economy like Ireland. The potential for inconsistencies in economic statistics resulting from a misalignment in methodology are identified and the approach taken by the Central Statistics Office (CSO) to achieve a consistent treatment across the domains of national accounts and balance of payments is illustrated.

1. Literature Review

Haskel & Westlake (2018) state that there has been a long-term shift from tangible to intangible investment, and that "much of that shift does not appear in company balance sheets and national accounts because accountants and statisticians tend not to count intangible spending as an investment, but rather as day-to-day expenses". However, this has been changing in the case of national accounts based on the most recent updates to the standard systems of national accounts.

The decision to treat R&D as investment in The System of National Accounts, 2008 (2008 SNA) is the latest point in a progression towards expanding the asset boundary in the national accounts (UN *et al.,* 2008). The 1968 SNA defines gross fixed capital formation as the value of durable goods for civilian use; significant improvements to durable goods; reclamation of land; margins on transactions in land; and breeder stocks (UN, 1968). The System of National Accounts, 1993 (1993 SNA) extends gross fixed capital formation to include expenditure on mineral exploration, computer software and entertainment, and literary or artistic originals (UN *et al*., 1993). The current standard, 2008 SNA, extends the definition further. Expenditure on R&D and the outright transfer of the ownership rights of the outcome of R&D now fall within the asset boundary (UN *et al*., 2008). The Frascati Manual (OECD, 2015) sets out the guidance on collecting statistics on R&D. Ker & Galindo-Rueda (2017) describe the common background of the Frascati Manual and the 2008 SNA about this issue and how the fundamental change to the treatment of R&D in the 2008 SNA marks the convergence of the two frameworks.

The Balance of Payments and International Investment Positions Manual – Sixth Edition (BPM6) is harmonised with the 2008 SNA. Just as the SNA has evolved the asset boundary over time, the update to the balance of payments standard now records the provision of R&D services and the outright sales of intellectual property products under the service category 'Research and Development Services' (IMF, 2009). Previously, intellectual property products which are the result of R&D, such as patents and copyrights, were treated as non-produced assets and appeared in the capital account. By 2014, most OECD countries had implemented the new 2008 SNA standards. There was focus initially on the capitalisation of R&D, being the largest and most wide-reaching effect of the transition to the new standard. Expenditure on R&D is now treated as investment and not intermediate consumption and output is increased in the case of own-account R&D. Van de Ven (2015) shows an average increase of 2.2% of GDP due to the capitalisation of R&D (for OECD countries).

The recording of intellectual property products as assets in the accounting framework was not central to the discussion at the time of implementation of the new standards. When the CSO published the exceptional national accounts results for 2015, this aspect of the 2008 SNA took on a new significance. Commentators were baffled by these results when they were first published. *The Irish Times* said, "trying to interpret the official economic figures for 2015 is next to near impossible" (Taylor, 2016). *The Economist* (2016) called it a "virtual reality". What was illustrated was that the ability to measure domestic production in a meaningful way is tested when factors of production can be in different parts of the world, for instance through contract manufacturing. The robustness of the 2008 SNA standards in measuring activity in a globalised world was nevertheless demonstrated. The increase in capital stock of €262bn in 2015 helps to explain the increase in GDP in Ireland. The 2015 results would have been more difficult to comprehend without having intellectual property products within the asset boundary of the accounts, what Haskel & Westlake (2018, p. 5) refer to as "capitalism without capital".

Since the events in Ireland in 2015, the major role that intellectual property products play in modern production arrangements has been brought into sharp focus. The value of these assets can be extremely large, and they can transfer between multinational units in different countries relatively easily. These issues are discussed in *Globalisation, Intellectual Property Products and Measurement of GDP* (OECD, 2018, p.7) where it is noted that the change in treatment of R&D is often misunderstood to

be a driver of the problems in measurement. Stapel-Weber & Verrinder (2016, p. 36) observe that intellectual property assets do not behave like most other fixed assets as they are highly moveable and that the level of depreciation on the assets is very high. The latter aspect is focussed on in the current paper. De Haan & Haynes (2018) explore the economic ownership of intellectual property. They go beyond the 2008 SNA and suggest rerouting of transactions in these assets to the headquarters of the multinational enterprises (MNE) group (on these topics, see also Khder et al. in this issue).

Recent studies have attempted to approach measurement challenges in the national accounts due to globalisation effects by reallocating parts of the accounts across country boundaries. Tørsløv *et al*. (2018) adjust the profits of multinationals for activities of subsidiaries abroad. In describing the adjustment, they state: "We also subtract depreciation from profits, because depreciation is deductible from taxable profits". By subtracting the depreciation, the method attributes the depreciation part of the value added to the host country. In the case of intellectual property assets this is something that could be further examined. In a development of this approach Bruner *et al*. (2018) make a series of adjustments to compile a sequence of accounts adjusting the USA national accounts and balance of payments for effects of globalisation. Among other adjustments, is that for the relocation of the ownership of intellectual property. Through the redistribution of assets, charges for the use of intellectual property are reattributed in the USA Production Account. We would think that an entry for depreciation could also be considered in the Use of Disposable Income Account.

Following the dramatic level shift of Ireland's GDP for 2015 an expert group was set up to provide recommendations for the CSO to best meet the challenges for providing indicators that are more particular to the highly globalised nature of the Irish economy. The report of the Economic Statistics Review Group (CSO, 2016b) recommends the development of modified GNI, or GNI*, with a corresponding modified current account, or CA*. Recognising the exceptional situation of depreciation of capital stocks that are relocated to Ireland, these new indicators adjust for depreciation of foreign-owned IP assets. The discussion in the current paper is relevant to the work of the review group as GNI and GNI* are more consistently derived when the concepts of operating surplus and reinvested earnings are aligned.

Connolly (2017) examines many of the same issues discussed in this paper. Attention is drawn to the variation between the national accounts model of depreciation and the accounting measure used in companies' statutory accounts. In the context of the 2015 national accounts data for Ireland, he shows the need for balancing adjustments "to avoid introducing a distortion to the economic aggregates". Connolly mentions the need to make these adjustments. The current paper shows how these adjustments are applied in practice.

There is little direct reference in the literature to applying an adjustment to reinvested earnings in the balance of payments, due to the capitalisation of the provision of R&D services. This issue is dealt with in the current paper. The Czech Statistical Office (CZSO) outlined the challenges involved in aligning reinvested earnings based on business profit with the 2008 SNA concepts (Kermiet, 2017). Kermiet mentions R&D as a problematic aspect. Initial findings on the impact of R&D expenditure on the calculation of reinvested earnings was presented at the Balance of Payments Working Group of the EU Commission in November 2017 (Quill, 2017), and again at the Joint Eurostat – OECD Task Force on Land and other non-financial assets – intellectual property products (Mangan & Quill, 2018).

2. Background Information

Increasing globalisation of the world's business economy can pose challenges for the compilation of official statistics. Enterprises operate on a global playing field with complex international business models to maximise both productivity and profitability. The impact these global operations can have on official economic statistics is highlighted in the CSO's publication of national accounts indicators for 2015 (CSO, 2016a).

One aspect of globalisation and a subject of this paper is the importance of the provision of R&D services and intellectual property products in modern production arrangements. The know-how or technical specifications required to produce goods has become a central component of the production process. Under the 2008 SNA, intellectual property products that are the result of R&D are classified as produced assets and appear in the capital stock of the country of ownership. These assets do not have physical substance and are thus highly mobile.

This mobility can have a significant impact on the national accounts of the countries involved.

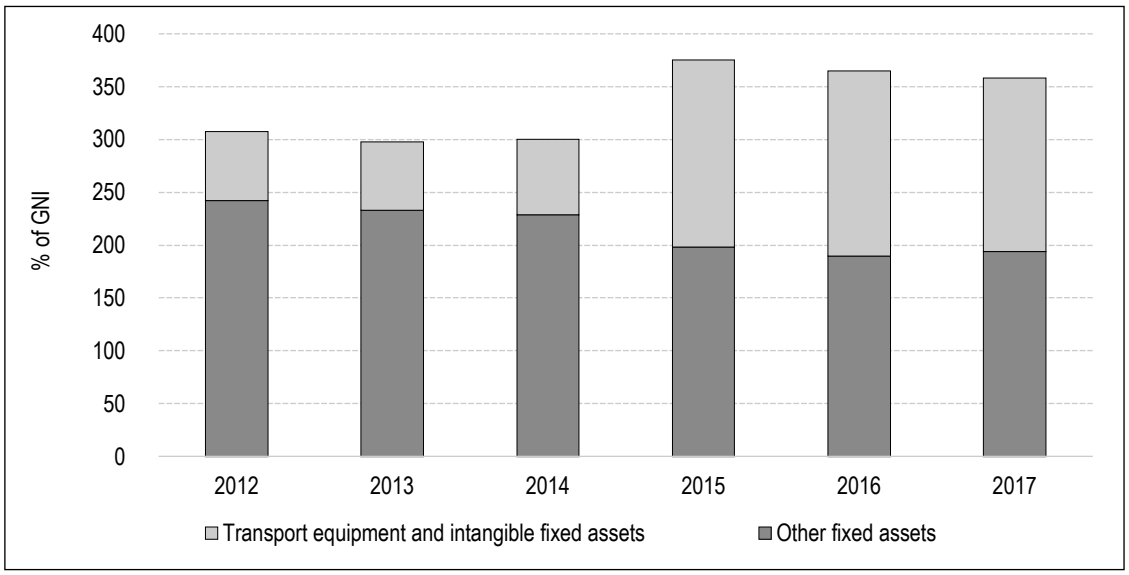
The provision of R&D services and the sale of proprietary rights arising from R&D are classified under the one heading 'Research and development services' in the balance of payments and are classified as investment in the national accounts. Both types of transactions are discussed in this paper. In general, the provision of R&D services tends to be a relatively smooth series in the accounts, whereas the acquisition and disposal of intellectual property products is much more erratic, characterised by one-off large transactions. In Ireland there has been a significant growth in the imports of R&D services and cross-border transactions in intellectual property products.

Figure II illustrates the value of net stock of fixed capital in Ireland from 2012 to 2017. There was steady growth in the capital stock before the level shift in 2015, and continuous strong growth since. The 2015 movement in the capital stock was driven by growth in the stock of Transport Equipment (which predominantly consists of aircraft) and Intangible Fixed Assets. These two asset types have been combined into a single category for confidentiality reasons. The pairing however is not unreasonable as they are both highly mobile assets which give rise to economic activity overseas that contributes to Ireland's GDP. These asset categories have soared in significance, equating to 46% of the total net capital stock of assets in 2017, in comparison to 24% in 2014, illustrating the changing composition of Ireland's balance sheet and business landscape. This increase in the capital stock is a combined result of the relocations of enterprises (and their entire balance sheets) to Ireland, the "onshoring" of IP assets by MNEs resident in Ireland, as well as the growth of the aircraft leasing industry.

The large increases in capital stock in Ireland coincide with two measures in the 2014 and 2015 Finance Acts. Firstly, the Finance Act (Government of Ireland, 2014) introduced measures to ensure that any enterprise incorporated in Ireland must also be tax resident, addressing the issue of so-called "stateless companies". Secondly, tax initiatives to promote R&D expenditure and the development of R&D assets by enterprises resident in Ireland were introduced. Coffey (2017, p. 124) discusses the 2015 level shift in capital assets alongside a consequential increase in capital allowances for intangible assets. He concludes that while income increases substantially, corporation tax receipts grow much more slowly due to the offsetting capital allowances.

Alongside the onshoring of intellectual property products, Ireland has seen a very strong growth in imports of R&D services. This service category has always been an important component of the balance of payments, illustrating the technical character of the MNE sector in Ireland.

Figure II – **Total net stock of fixed capital assets at current prices as a % of GNI, 2012-2017**



Sources: Central Statistics Office, Estimates of the Capital Stock of Fixed Assets 2017.
https://www.cso.ie/en/releasesandpublications/er/csfa/estimatesofthecapitalstockoffixedassets2017/

Table 1 shows the imports of R&D services since 2012.

Table 1 – **Imports of research and development services, 2012-2017**

Year	€Bn	% of GNI
2012	5.0	4
2013	:	
2014	:	
2015	10.4	5
2016	12.0	5
2017	12.8	5

Notes: Data for 2013 and 2014 are not available.
Sources: Eurostat, International trade in services (since 2010) (BPM6).

The imports of R&D services into Ireland have been significant and increasing. The particularly strong growth in recent years is seen in the 2015 figures which are double the 2012 value. Ireland is one of the biggest importers of R&D services in the EU with the values for Ireland in the years 2015 to 2017 comprising nearly one fifth of the total value of countries in the EU.

In most cases the increase in capital stock resulting from onshoring of intellectual property products or through the import of R&D services are shown in the current account of the balance of payments. However, where the movement of intellectual property products is the result of restructuring and reclassifications it is recorded as 'Other changes in volume'. In either case, these movements have a neutral impact on GDP. In the former case the imports of the R&D cancel the effect of the increase in investment. In the latter case, neither imports nor investment are recorded.

There is however a marked impact on measures of gross national income generated by investment in intellectual property products and expenditures on R&D. In 2016 the CSO published extraordinary national accounts data for Ireland showing a growth of 26.3% in annual GDP for 2015 measured at constant prices (CSO, 2016a). This pushed Ireland's economy into the spotlight, capturing the attention of economists and commentators throughout the world. Increasing globalisation and such dramatic changes in key economic indicators highlight a growing difficulty for official statistics to represent the domestic economy. The 2015 results in Ireland illustrate the possibility that GDP and GNI may no longer "provide useful insights into the economic activity that is physically taking place in the national territory, as such domestic production can be dwarfed by globalisation activities" (Stapel-Weber & Verrinder, 2016).

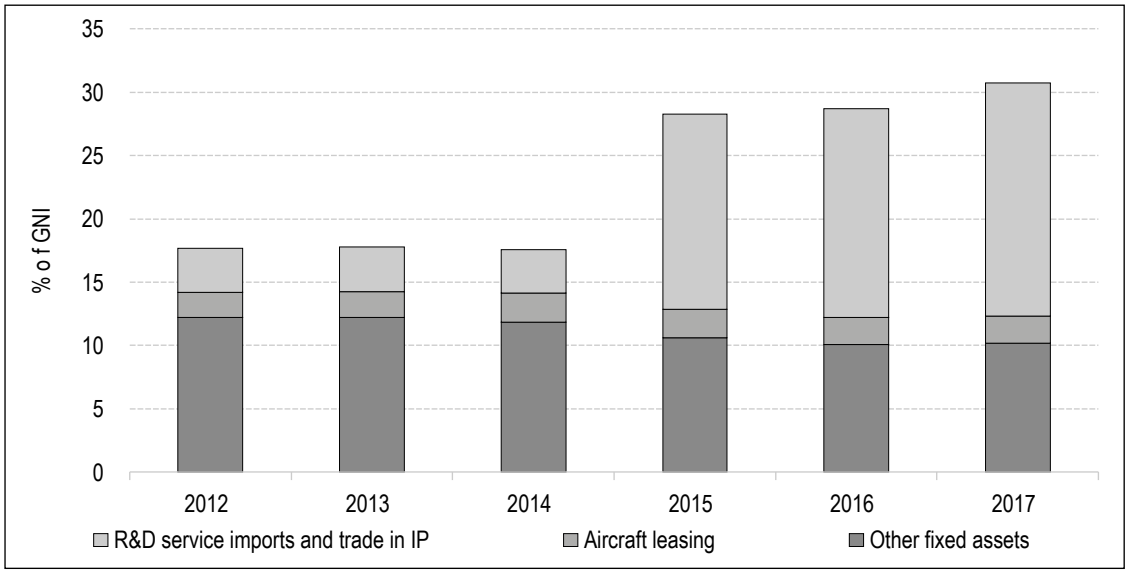
The level shift in the overall stock of assets in the Irish economy has a significant impact on consumption of fixed capital (depreciation), as evident from Figure III. Consumption of all fixed capital in the Irish economy increased by €47 billion between 2012 and 2017 (CSO, 2018b). Depreciation relating to R&D service imports and trade in intellectual property products, correspond to 81% of this growth, increasing from €5 billion in 2012 to €43 billion in 2017 (CSO, 2018c).

Connolly (2017) initiates a discussion of the impact such developments can have on key economic indicators in a small open economy, like Ireland. Once off increases in the capital stock are accompanied by corresponding growth in related economic activity. In the paper, Connolly (2017) identifies the potential consequence of an over/underestimate of GNI due to a misalignment in the depreciation estimates and the timing in which they are included at different stages across the accounting framework. This issue will be further examined in the body of this paper.

3. The Recording of Depreciation and R&D Expenditure in The Irish National Accounts and the Balance of Payments Statistics

The significance of R&D services expenditure and intellectual property products in a globalised economy was recognised in the most recent updates to the methodologies for compiling the national accounts and the balance of payments. Under the current balance of payments standard, BPM6 (IMF, 2009), the change of ownership of an intellectual property product is recorded in the current account under the Extended Balance of Payments Services classification heading: '10.1.1.2 - Sale of proprietary rights arising from research and development'. This item was previously recorded in the capital account of the balance of payments. Also, under the 2008 SNA, expenditure on R&D services is capitalised in the national accounts. This item is classified as '10.1.1.1 - Provision of customized and non-customized research and development services' in the BPM6 standard components. Indeed, both transactions, outright acquisition/ disposal of intellectual property products and

Figure III – **Consumption of total fixed capital at current prices as a % of GNI, 2012-2017**



Sources: Central Statistics Office, Estimates of the Capital Stock of Fixed Assets 2017.

expenditure on the provision of R&D services, are classified under component category '10.1-research and development services' in the balance of payments and are treated as capital formation in the national accounts. Previously, only expenditure on R&D services was treated as '9.3.3 Research and development' in the balance of payments BPM5 standard components (UN *et al*., 2002, pp. 82-84), and neither transaction was treated as investment in the national accounts.

Irish national accounts and balance of payments statistics are compiled in a well-integrated system within a single directorate in the CSO. Multiple data sources are used to compile the national accounts statistics including business surveys, company accounts from the tax authorities, household surveys and administrative data. The balance of payments statistics are based on comprehensive income and balance sheet surveys collected by the CSO with further surveys for parts of the financial sector collected by the Central Bank of Ireland.

The CSO benefits from access to extensive datasets, from both their own Large Cases Unit (LCU) and the national tax authority, The Revenue Commissioners, which may not be as freely available to other national statistical institutes (NSI). The LCU is a well-developed unit within the economics directorate of the CSO, that acts as a single collection point for all surveys issued to a selection of the largest enterprises resident in Ireland. The LCU is in regular contact with these key companies, developing an excellent understanding of globalisation issues that are present in a small open economy.

3.1. Depreciation

The perpetual inventory method (OECD, 2009) is used by national accounts in the CSO to estimate the capital stock of fixed assets (CSO, 2018b). Gross capital stock represents the value of stock still in use and is valued at the price of new capital goods. Net capital stock takes account of consumption of fixed capital (CFC, or depreciation) and is estimated by applying a suitable depreciation rate, which varies depending on the type of asset. CFC is calculated for each asset type and NACE A64 category, not at the individual company level. Net operating surplus is then derived at aggregate levels in the national accounts.

This differs from the compilation of balance of payments statistics, where the depreciation is reported by the individual companies. The information collected on the CSO's balance of payments survey form is generally consistent with the enterprises' end of year financial statements. Details of sales, expenses, depreciation, net interest, dividends, and other income flows provide the data required to derive reinvested or retained earnings by enterprise.

Although depreciation does not feature as an individual item in the balance of payments, it is indirectly included as an expense that reduces the income earned by an enterprise. If the enterprise is owned by a foreign shareholder this will result in a reduction in the related direct investment outflow. Section 11.43 of the BPM6 manual outlines the transition from net operating surplus to reinvested earnings in a series of steps adding and subtracting dividends, interest, taxes and other items (IMF, 2009). The manual goes on to say that "these items correspond exactly to SNA items". To achieve consistency, NSIs give guidance to respondents on how to record these items and closely monitor big items in the survey returns, such as the payment of dividends. It is not feasible to expect that the enterprise would adopt the national accounts PIM model to calculate depreciation. Blanchet *et al*. (2018) mention the difficulty in evaluating depreciation in intangible assets.

Each entity depreciates their stock of assets based on individual procedures that comply with national financial reporting standards. Enterprises resident in Ireland who participate in R&D are granted capital allowances which offset the total value of corporation tax they are liable for, possibly encouraging efficient planning in relation to the optimum depreciation strategies. There is little evidence that the methodologies and assumptions used at the enterprise level for depreciation of asset categories are in line with those used by national accounts. It is worth noting that any changes in the accounting rules and/or tax legislation relating to these assets could have an impact on the level of divergence between the depreciation reported in the balance of payments and the depreciation calculated for national accounts.

3.2. Expenditure on R&D Services

Under the 2008 SNA, R&D is treated as capital formation in the national accounts. This requires expenditure on in-house R&D activities, or the provision of R&D services by a third party, to be treated as investment, adding to the capital stock (UN *et al.*, 2008, p. 122). Expenditure on R&D activities in the domestic economy is collected in enterprise surveys for the compilation of statistics. The CSO's Business Expenditure on Research and Development (BERD) survey, is an example (CSO, 2017). Cross border expenditure on R&D services are captured in the international trade in services survey of domestic enterprises. In Ireland, the trade in services survey is part of the comprehensive balance of payments collection system.

In the national accounts, the sum of domestic R&D expenditure and trade in R&D services are the basis of the R&D services component of capital formation. R&D activities that are performed within the economy are either capitalised under the 2008 SNA or exported, with a positive impact on GDP in the period of activity. Imported R&D services have a neutral impact on GDP in the period when the expenditure occurs, as the increase in imports which would have a negative impact on GDP is balanced by a corresponding positive value in investment. While the collection of data relating to expenditure in R&D is relatively uncomplicated, difficulties arise when deriving data on direct investment flows arising from reinvested earnings for companies with R&D expenditure. As R&D is not regarded as intermediate consumption in the national accounts, it should not be subtracted in the calculation of operating surplus. Moreover, depreciation of R&D expenditure is necessary under the 2008 SNA.

This results in a challenge for compilers of statistics as reporting entities generally regard R&D as an expense, rather than a capital item. Specifically, companies consider R&D a business cost rather than an asset.

As a result, the usual accounting calculation for operating surplus will give a result which differs from the SNA calculation. This difference is equal to the value of R&D expenditure in the current period less the depreciation on any prior R&D expenditure. A further explanation is available in an information note on the CSO website (CSO, 2018a).

3.3. Misalignment of National Accounts and Balance of Payments Items

Connolly (2017) identifies a potential mismatch between national accounts' measurement of operating surplus and balance of payments measurement of primary income which is the main contributing factor in the calculation of net factor income. He suggests that if different estimates of depreciation are used at different stages in the accounting framework they may give rise to an over/under estimate in GNI.

Gross operating surplus in the national accounts corresponds to the income generated by enterprises based on production activities. It excludes

market price changes, interest payments and receipts but includes an estimate for financial intermediation service charges.

To arrive at GNI, which, for Ireland, was once considered the most appropriate measure of growth in the domestic economy, net factor income and EU subsidies less EU taxes are subtracted from GDP, this is displayed in Figure IV. Balance of payments' primary income or more precisely, its most significant component 'Investment income attributable to direct investors' (dividends plus retained earnings) is the driving force behind net factor income. Net factor income attributes the income of enterprises who are owned by foreign shareholders to the country of ownership as it is ultimately the income of the country of the direct investor rather than the income of Ireland.

The calculation of net factor income should be consistent with the national accounts' calculation of net operating surplus (gross operating surplus less depreciation). Therefore, inconsistencies in the recording of depreciation will result in an incorrect estimate of net factor income.

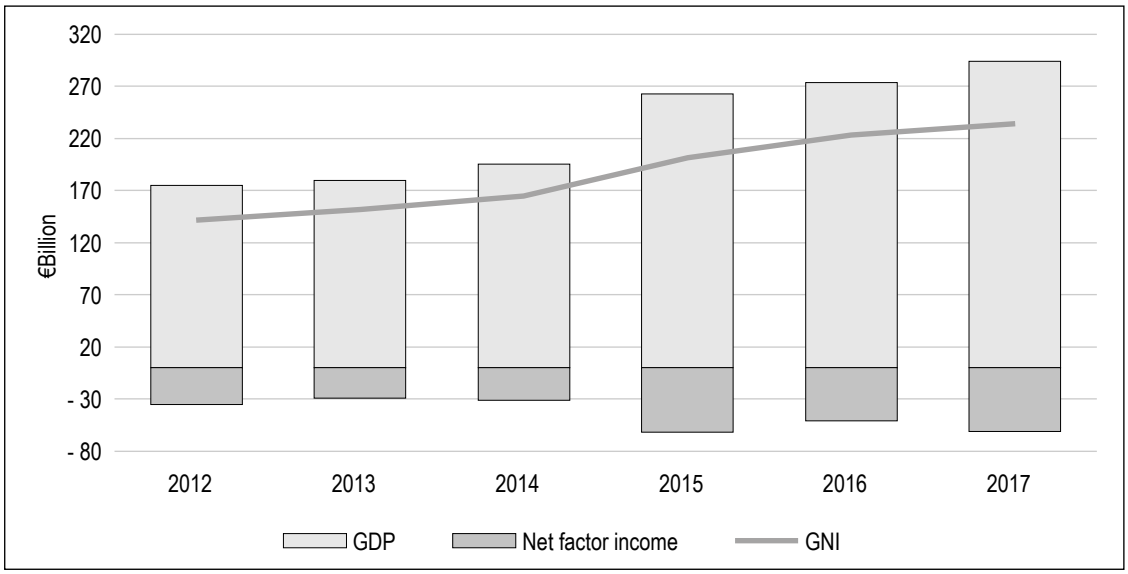
Recognising the potential for unintended variation between the two systems, the CSO established a group to monitor the consistency of depreciation of intellectual property (and other fixed assets), and to evaluate the impact of R&D on the calculation of retained earnings.

These consistency issues can be significant and challenging in economies with a concentration of R&D intensive industries. An examination of the consistency between national accounts and balance of payments for a selection of common indicators, uncovered a difference in the net operating surplus for a number of enterprises. The net operating surplus for a sample of multi-national enterprises, calculated using national accounting methodology was found to be higher than the net profit reported in their balance of payments survey return. We identify three reasons relating to R&D for these differences.

Firstly, enterprises who hold a portfolio containing high value intellectual property products generally report higher depreciation in the balance of payments survey than the value estimated using national accounts perpetual inventory method, as evident from Table 2.

The disparity between depreciation reported by enterprises and that calculated by national accounts suggests that enterprises depreciate intangible assets at a faster rate than used in official statistics, resulting in higher depreciation and lower net profits in the balance of payments. Due to the high percentage of MNEs in Ireland being owned by foreign shareholders, this asymmetry translates into a gap in the level of multinational profits recorded in the national accounts and the level of direct income outflows in the balance of payments, potentially inflating the indicators of growth in the domestic economy.

Figure IV – **GDP, GNI and net factor income at current market prices, 2012-2017**



Notes: The values for EU subsidies and EU taxes are not included separately in the graph as they make a very small contribution to the calculation of GNI.
Sources: Central Statistics Office, National Income and Expenditure 2017.

Table 2 – **Difference between the reported depreciation in the balance of payments and depreciation in national accounts for intellectual property products, 2013-2017**

Year	€Bn	% of GNI
2013	0.2	0
2014	1.1	1
2015	10.2	5
2016	7.8	3
2017	4.7	2

Sources: Authors' calculations.

Secondly, enterprises that have expenditure on the provision of R&D treat it as an expense in their calculation of profit. This conflicts with the national accounts convention in which it is classified as capital expenditure. Consequently, net operating surplus in the national accounts is higher than the corresponding net profit in the survey data.

A third accounting complication, which is a direct consequence of the second, relates to the capitalisation of R&D under the 2008 SNA. There is a requirement to record depreciation for the assets that arise from the capitalisation of the provision of R&D. As expenditure on R&D services is increasing (cf. Table 1), the adjustment to profits in the balance of payments for the non-inclusion of R&D services is not offset by the depreciation of R&D capital.

The analysis shows systematic differences between national accounts and the balance of payments. These differences may require an intervention by statisticians to ensure alignment of the two methods.

4. Achieving Consistency

4.1. Depreciation of Intellectual Property Products

The accountancy rules followed by enterprises when recording depreciation of intellectual property products are not directly linked to the national accounting methodology. A number of assumptions are made when depreciating assets. These include: the rate of depreciation, estimated average lifetime of the asset, and method of depreciation. There is variation in these assumptions across enterprises. There is also a divergence in the assumptions made by enterprises and the assumptions made by statisticians compiling the national accounts.

The PIM model produces a lower estimate of depreciation than that reported in the companies' accounts, suggesting that the economic life of the assets tends to be longer than the accounting life. The magnitude of these differences and their potential to cause inconsistencies is illustrated in Table 2. In the absence of any intervention by the CSO the differences are significant, ranging from 1% to 5% of GNI between 2014 and 2017.

To examine the consistency in the depreciation of intellectual property products across the accounting frameworks within the CSO the following approach was taken. Using a significant threshold, intangible assets are divided into 'Very large intellectual property products' and 'Other intellectual property products'. A different treatment is applied to each case. The former set of assets are examined individually and the depreciation is aligned between the two systems; the latter are monitored over time for consistency.

4.1.1. 'Very large Intellectual Property Products'

Purchases and imports of 'Very large intellectual property products' are identified by the compilers of balance of payments statistics in the CSO. These are jointly analysed with the compilers of national accounts. The national accounts division is responsible for the depreciation of these assets which is estimated using the perpetual inventory method. These depreciation values are then entered into the balance of payments accounting framework, replacing those used by the reporting entity. Further adjustments are subsequently required in the accounts of the entity, in particular to retained earnings in the current and financial accounts of the balance of payments. The alignment is supported in the IMF's BPM6 Compilation Guide: Consumption of fixed capital should be calculated on the basis of current replacement cost. However, company accounts may reflect a variety of bases and the balance of payments compiler may, in conjunction with the national accounts compiler, make an aggregate adjustment (IMF, 2014, para. 13.48).

The CSO procedure for addressing high value assets is aimed at ensuring that the reinvested earnings in balance of payments are consistent with the operating surplus in national accounts. The process results in full consistency between the two accounting systems for a small number of very large assets. It requires valuable effort from the LCU to maintain detailed information at the

company level relating to the assets and adjustments, along with the completion of updates to the balance of payments and national accounts data systems. This is time-consuming and requires manual input over an extended number of accounting periods, which is not ideal within what is otherwise a largely automated system.

4.1.2. 'Other Intellectual Property Products'

The smaller 'other intellectual property products' are first identified by the balance of payments compilers. Here the approach is to use the reporting entity's depreciation in the balance of payments. The assets are independently added to the capital stock of the economy and are depreciated in the usual way, at a macro level in the national accounts. This can lead to inconsistencies, as outlined previously. The aim here, however, is to achieve a broad agreement between the total depreciation of companies holding 'other intellectual property products' across the two accounting frameworks.

Each company's contribution to the macro-level depreciation can be estimated based on a profile of its assets and an application of the perpetual inventory method at a micro level. These data allow ongoing analysis, at the individual company level, of the consistency of depreciation in the two accounting systems. Thus, although the national accounts apply depreciation at an economy wide level, for this exercise, estimates of the contribution of individual companies to the total depreciation are made.

Experience of monitoring the intellectual property products has shown that there are features that compilers of statistics should be on alert for. For example, it is important to establish, from the beginning, if an intangible asset is a franchise or trademark; or if the asset is the outcome of research and development. Otherwise, there may be depreciation in the balance of payments company accounts but not in the national accounts, or vice versa. Another experience is that depreciation in the company accounts can behave erratically. Under a merger of two entities that hold intellectual property products, for example, the combined value of the assets may not equal the sum of the asset values prior to the merger. The perpetual inventory method is flexible enough to allow for this type of scenario (OECD, 2009, p. 49). If the asset is not too large, then the change in the value of the combined intellectual property products may not disturb the symmetry of the two accounting frameworks at the level of the whole economy.

The exercise in the CSO aims to measure and improve on discrepancies in the values of depreciation in the two accountancy frameworks. Currently we estimate close to full alignment for intellectual property products. For these assets the balance of payments depreciation differs from national accounts depreciation by around 2% on average over the years 2015 to 2017. Currently no adjustments have been implemented in this case, however, possible asymmetries will continue to be monitored and the need for adjustments will be reassessed.

4.2. Expenditure on R&D Services

As previously discussed, a further potential for divergence between net operating surplus in the national accounts and retained earnings in balance of payments, is the conceptual difference between the treatments of R&D expenditure. In national accounts R&D is treated as investment whereas in company accounts it is generally considered an expense.

The different treatment can lead to significantly different calculations of retained earnings in the balance of payments in economies with a concentration of R&D intensive multi-national enterprises, such as Ireland. For this reason, the CSO finds it necessary to apply an adjustment to the retained earnings in the balance of payments for certain companies, who have purchases of R&D services. The modification is twofold: in the first instance, the cost incurred for expenditure on R&D must be removed in the calculation of the net profits of the enterprise, resulting in higher net profits. Secondly, depreciation of prior years' R&D has to be considered as an expense item, reducing the net profits. The CSO has found that combining these two adjustments results in a higher net operating surplus for companies with expenditure on R&D than would have been the case, for example, under the 1993 SNA (UN *et al*., 1993).

The CSO is aware that this treatment may be necessary for companies whose reinvested earnings are recorded as inflows to Ireland. However, this data is not so readily available, and no adjustments have been implemented to inflows in the balance of payments. The effect of the modification to retained earnings has been to increase factor income outflows by between 1 and 2 percentage points of GDP in recent years.

The actual implementation of this process is not entirely straightforward. Removing current expenditure on R&D from the profit calculation

involves handling the company data. The second adjustment requires a calculation of depreciation on R&D expenditures, up to the current period, estimated in accordance with the national accounts perpetual inventory method. The process also needs manual input over all accounting periods.

* *
*

Increasing globalisation of the world's business economy presents challenges to compilers of official statistics when producing consistent economic accounts.

The extraordinary growth in the Irish economy experienced in 2015 shows how sensitive small open economies, like Ireland, are to globalisation measurement standards. The CSO became acutely aware of the possible inconsistencies in the measurement of activity related to globalisation with the onshoring of intellectual property products in 2015 and put in place a group of practitioners to monitor the alignment of depreciation and the treatment of R&D across the two accounting frameworks of national accounts and balance of payments. The scale of events relating to R&D required intervention by the CSO to align company reporting with concepts of national accounts. The result of this project is consistency in R&D related data, so that the factor flows, calculated in the balance of payments and used in the national accounts, are based on national accounting rules and are therefore consistent with other concepts in the national accounts. An indication of the magnitude of the overall adjustment is given in Figure I in the introduction. This shows separately the realignment of depreciation for large intellectual property assets, the alignment of income through the treatment of R&D services as investment, and the further adjustment needed to allow for depreciation on the R&D services.

Due to the magnitude of differences and the need for an adjustment in Ireland we recommend that compilers of balance of payment statistics and international organisations take note of the CSO's experience, and that any future manuals or guidance could contain a discussion of these items. □

BIBLIOGRAPHY

Blanchet, T., Khder, M., Leclair, M., Lee, R., Poncet, H. & Ragache, N. (2018). Is growth underestimated? Insee Références, *L'économie française - comptes et dossiers, édition 2018*, pp. 59–79. https://www.insee.fr/en/statistiques/fichier/3614242/ECOFRA18b_D1_croissance_EN.pdf

Bruner, J., Rassier, D. & Ruhl, K. (2018). Multinational Profit Shifting and Measures throughout Economic Accounts. *Bureau of Economic Analysis Working Papers* N°0150. https://ideas.repec.org/p/bea/wpaper/0150.html

Central Statistics Office (2016a). National Income and Expenditure 2016. https://www.cso.ie/en/media/csoie/methods/nationalincomeandexpenditureannualresults/NIE2016MethodologyNote.pdf–

Central Statistics Office (2016b). Report of the Economic Statistics Review Group (ESRG), December 2016. https://www.cso.ie/en/media/csoie/newsevents/documents/reportoftheeconomicstatisticsreviewgroup/Economic_Statistics_Review_(ESRG)_Report_Dec_2016.pdf

Central Statistics Office (2017). Business Expenditure on Research and Development 2015-2016. https://www.cso.ie/en/releasesandpublications/er/berd/businessexpenditureonresearchdevelopment2015-2016/

Central Statistics Office (2018a). Consistency of Treatment of R&D in the National Accounts and the Balance of Payments. https://www.cso.ie/en/releasesandpublications/in/rdnabop/consistencyoftreatmentofrdinthenationalaccountsandthebalanceofpayments/

Central Statistics Office (2018b). Estimates of the Capital Stock of Fixed Assets 2017. https://www.cso.ie/en/releasesandpublications/er/csfa/estimatesofthecapitalstockoffixedassets2017/

Central Statistics Office (2018c). National Income and Expenditure 2017.
https://www.cso.ie/en/releasesandpublications/ep/p-nie/nie2017/

Coffey, S. (2017). Review of Ireland's Corporation Tax Code. Report for Minister for Finance and Public Expenditure and Reform, 12 September 2017.
https://www.gov.ie/en/publication/fa53d4-review-of-irelands-corporation-tax-code/

Connolly, M. (2017). The expected and unexpected consequences of ESA 2010 - an Irish perspective. Read before the second meeting of the 2017/2018 session of the Statistical and Social Inquiry of Ireland on 30-November 2017. http://www.ssisi.ie/CSO_Connolly_SSISI_v4.pdf

De Haan, M. & Haynes, J. (2018). R&D capitalization: where did we go wrong? *Eurostat Review on National Accounts and Microeconomic Indicators.* Massachusetts: University of Chicago Press.
https://ec.europa.eu/eurostat/cros/system/files/euronaissue1-2018-article1.pdf

Eurostat (2013). *European System of Accounts (ESA 2010)*. Luxembourg: Publication Office of the European Union.
http://ec.europa.eu/eurostat/documents/3859598/5925693/KS-02-13-269-EN.PDF/44cd9d01-bc64-40e5-bd40-d17df0c69334

Eurostat (2017). International trade in services (since 2004).
http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=bop_its_det&lang=en

Eurostat (2019). International trade in services (since 2010) (BPM6).
http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=bop_its6_det&lang=en

Government of Ireland. *Finance Act 2014.* Number 37 of 2014. Ss.26, 43. Dublin: Houses of the Oireachtas.
http://www.irishstatutebook.ie/eli/2014/act/37/section/43/enacted/en/html#sec43

Government of Ireland. *Finance Act 2015.* Number 52 of 2015. Ss.31, 32. Dublin: Houses of the Oireachtas.
https://data.oireachtas.ie/ie/oireachtas/act/2015/52/eng/enacted/a5215.pdf

Haskel, J. & Westlake, S. (2018). *Capitalism without Capital: The Rise of the Intangible Economy*. New Jersey: Princeton University Press.

IMF (2009). *Balance of Payments and International Investment Position Manual – Sixth Edition (BPM6).* Washington D.C.: International Monetary Fund.
https://www.imf.org/external/pubs/ft/bop/2007/pdf/bpm6.pdf

IMF (2014). *BPM6 Compilation Guide – Companion Document to the Sixth Edition of The Balance of Payments and International Investment Position Manual.* Washington D.C.: International Monetary Fund.
http://www.imf.org/external/pubs/ft/bop/2014/pdf/guide.pdf

Kermiet, V. (2017). Reinvested earnings on FDI. Presented at Balance of Payments Working Group in November 2017.

Ker, D. & Galindo-Rueda, F. (2017). Frascati Manual R&D and the System of National Accounts. OECD *Science, Technology and Industry Working Papers*, 2017/06. http://dx.doi.org/10.1787/edb6e020-en

Khder, M.-B., Montornes, J. & Ragache, N. (2020). Irish GDP Growth in 2015: A Puzzle and Propositions for a Solution. *Economie et Statistique / Economics and Statistics*, this issue.

Mangan, O. & Quill, P. (2018). The impact on reinvested earnings of expenditure in R&D. Presented at Joint Eurostat - OECD *Task Force on Land and other non-financial assets – Intellectual Property Products* (February 2018).

OECD (2009). *Measuring Capital - OECD Manual 2009: Second edition*. Paris: OECD Publishing.
https://doi.org/10.1787/9789264068476-en

OECD (2015). *Frascati Manual 2015 guidelines for collecting and reporting data on research and experimental development*. Paris: OECD publishing.
https://www.oecd-ilibrary.org/docserver/9789264239012-en.pdf?expires=1554216827&id=id&accname=guest&checksum=786A11CD6334CE55744F9106B00BE2D5

OECD (2018). *Globalisation, Intellectual Property Products and Measurement of GDP: Issues and Proposals.* Committee on Statistics and Statistical Policy Informal Reflection Group, 15th Meeting of the Committee on Statistics and Statistical Policy 20-21 June 2018, Geneva.
http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=SDD/CSSP(2018)6&docLanguage=En

Quill, P. (2017*).* Treatment of research and development and investment income in balance of payments and national accounts. Presented at Balance of Payments Working Group in November 2017.

Stapel-Weber, S. & Verrinder, J. (2016). Globalisation at work in statistics – Questions arising from the 'Irish case'. *Eurostat Review on National Accounts and Macroeconomic Indicators,* 2/2016, 29-44. http://ec.europa.eu/eurostat/documents/3217494/7784358/KS-GP-16-002-EN-N.pdf/af4b1474-cc3a-4453-9814-bfbcb74e31d0

Taylor, C. (2016). Ireland's GDP figures: Why 26% economic growth is a problem. *Irish Times*, 15 July 2016. https://www.irishtimes.com/business/economy/ireland-s-gdp-figures-why-26-economic-growth-is-a-problem-1.2722170

The Economist (2016, July 16). Not the full shilling. Why GDP growth of 26% a year is mad. Europe, Jul 16th 201. https://www.economist.com/europe/2016/07/16/not-the-full-shilling

Tørsløv, T., Wier, L. & Zucman, G. (2018). The Missing Profit of Nations. NBER, *Working Paper* N° 24701. https://gabriel-zucman.eu/files/TWZ2018.pdf

UN (1968). *A System of National Accounts.* New York: United Nations Publications. https://unstats.un.org/unsd/nationalaccount/docs/1968SNA.pdf

UN, Eurostat, IMF, OECD, World Bank (1993). *System of National Accounts 1993*. Brussels/Luxembourg, New York, Paris, Washington, D.C. https://unstats.un.org/unsd/nationalaccount/docs/1993sna.pdf

UN, EC, IMF, OECD, UNCTAD, WTO (2002). *Manual on Statistics of International Trade in Services*, pp. 82–84. New York: United Nations Publications. https://www.oecd.org/trade/its/2404428.pdf

UN, Eurostat, IMF, OECD, World Bank (2008). *System of National Accounts 2008*. New York. https://unstats.un.org/unsd/nationalaccount/docs/sna2008.pdf

UN, IMF, OECD, Eurostat, UNCTAD, World Tourism Organisation, WTO (2011). *Manual on Statistics of International Trade in Services 2010*, pp. 145–148. New York: United Nations Publications. https://unstats.un.org/unsd/publication/seriesm/seriesm_86rev1e.pdf

van de Ven, P. (2015). New standards for compiling national accounts: what's the impact on GDP and other macro-economic indicators? OECD *Statistics Brief* N° 20. http://www.oecd.org/sdd/na/new-standards-for-compiling-national-accounts-SNA2008-OECDSB20.pdf

What Should the Concept of Domestic Production Mean in Globalized Economies?

Didier Blanchet*

Abstract – Traditional criticisms of GDP generally focus on its 'P' and 'G', the fact it is only a measure of gross output, without offering any insight into well-being and sustainability. Globalization adds in the 'D' problem, with the increasing difficulty of determining the location of major segments of production by multinational companies. When distinct factors contribute to production from several sites, there is effectively no analytical way of characterizing what each of these factors produces on its own in each of these sites, *a fortiori* for intangible factors that are located in a purely conventional way. An interpretation of GDP in terms of income avoids this problem; it invites us to distinguish between income associated with mobile or volatile factors and income attributed to factors that can be deemed purely domestic. It also clarifies the links with the issue of measuring well-being.

JEL Classification: E01, E23, F6, O3
Keywords: globalization, production, income, national accounts

* *Insee (didier.blanchet@insee.fr)*

This article draws on a paper presented at the 16th IAOS Conference "Better Statistics for Better Lives", held at the OECD, Paris, 19-21 September 2018. I would like to thank participants at this conference for their remarks, more specifically N. Ahmad, A. Braakmann, P. Schreyer and P. van de Ven. I also benefited from comments by V. Biausque, G. Houriez and R. Mahieu. The usual disclaimer applies.

Received January 4, 2019, accepted after revisions February 2, 2020.
Citation: Blanchet, D. (2020). What Should the Concept of Domestic Production Mean in Globalized Economies? *Economie et Statistique / Economics and Statistics*, 517-518-519, 205–214. https://doi.org/10.24187/ecostat.2020.517t.2019

Criticisms of GDP traditionally focus on what is meant by its 'G' and 'P'. The well-known problem with the 'G' is that a concept of "gross" domestic production disregards everything destroyed in the production process, including environmental assets: a measure of net production would be more appropriate to characterize how well or badly our economies are performing. Concerning the 'P', the problem is that "production" cannot be taken as the ultimate aim of socio-economic policies: instead of comparing GDP, performance comparisons between countries and time periods should be in terms of economic or general well-being rather than production. All these criticisms are well known. National accountants are accustomed to responding to them by asserting that, whatever its limits, production is a notion that is still worth measuring for its own sake. This favours a simple division of labour: national accountants measure production, without reference to the notion of well-being, leaving others to tackle the more ambitious attempt to quantify well-being and its sustainability.

This pragmatic position looks sensible, but leaves many problems unresolved. The possibility of completely separating measures of production from well-being considerations is particularly questionable. Measuring aggregate production requires aggregation of quantities for a huge variety of goods and services, and it is hard to see how this can be done without referring in some way to their utility content (Blanchet & Fleurbaey, this issue). National accountants realise this: their rationale for using prices as keys for aggregation (Lequiller & Blades, 2014), is the fact that prices are representative of the relative utility of the different goods, at the margin at least. This is not to say that GDP can be regarded as a direct measure of well-being or even a measure of economic well-being, nor that it should aim for that, as it will always lack too many things to enable it to fulfil that role. But it implies that the characterization of GDP has to endorse its connection with well-being (Schreyer, 2016); national accountants cannot disregard this connection.

The ongoing "mismeasurement debate" offers good illustrations of this difficulty of thinking about GDP without reference to the notion of well-being. This debate is about the ability of accounts to give a proper view of how the renewal of goods and services contributes to economic progress, in particular those provided by the digital economy (see e.g. Feldstein, 2017; Syverson, 2017; Blanchet *et al.*, 2018; Byrne *et al.*, 2018; Aghion *et al.*, 2019). This is typically about the marginal gains in utility generated by those goods. A particularly interesting aspect of this debate concerns free digital services for which no explicit price is observable, and whose utility must be measured in other ways. This reignites the very old issue of the conventional market or quasi-market boundary that accountants use for distinguishing between what they do or do not regard as production (Coyle, 2017; Bourgeois, this issue). Should we move this boundary to make room for some of these new free goods, as suggested recently for instance by Brynjolfsson *et al*. (2019) proposal of a "GDP-B". If we start broadening the scope of accounts, where should we stop?

Our conclusion will return to these matters, but in relation to another category of questions about how GDP can be interpreted, which will form the main focus of this paper. In addition to the debate about GDP's conceptual boundaries there is now the question of its geographical boundaries. It is the "D" of GDP that is the issue here. National accounts have been historically developed to monitor economic trends, country by country, with GDP aiming to measure what is produced domestically in each country. This notion of domestic production does not raise any problem when countries are autarkic or when international trade is limited to finished products. But this is no longer the case when trade also includes intermediate products, *a fortiori* when the factors of production are in different locations.

Until recently, these problems have been deemed sufficiently under control and without any major impact on the credibility of accounts: discussion of them has been limited to specialist circles (UNECE, 2015). They have come under much greater scrutiny since the reporting of Irish GDP figures for 2015, indicating a one-year increase of over 25% due to the relocation to Ireland of intangible assets and related income for one large multinational enterprise or a small number of such enterprises. This emblematic example is discussed at greater length by ESRG (2016), Stapel-Weber & Verrinder (2016) and by Holton *et al.* and Khder *et al.* (in this issue). Since 2015, Irish accounts have been submitted to particular scrutiny and they continue to display movements that are difficult to interpret in terms of the amount of production actually taking place on Irish territory.

Some references will be made here to this Irish case, but the approach will be both more general and more basic. It will consist of just

expressing the problem of output location in the context of basic production functions, and this will suffice to show there is indeed a major conceptual difficulty. The problem we face with globalization is not only a matter of international coordination in collecting business data, nor of just clarifying the national accountants' formal notions of economic or legal property, two matters on which the debate has tended to focus until now. The fundamental point is that, for standard realistic forms of production processes, there is no established way, in theory, of breaking down production according to the differing contributions from factors located in different places.

Yet this situation is less damaging for national accounts than it may seem, as there is still some economic significance for another production-related concept, which is not production *stricto sensu*, but the income derived from factors' contribution to production. This is indeed what GDP can and does try to measure, it is just out of convenience that we tend to view it incorrectly as a direct measure of production. We should therefore start afresh by interpreting GDP in terms of income so as to clarify what it has to say about what is happening in the economies participating in the globalization process. This would, at the same time, shed light on our opening question about its connection with the measurement of well-being, as it is easier to move from the notion of income to well-being than from the notion of production. In short, on this issue of globalization, either in addition to or before there is a need to develop new concepts, an important preliminary step may simply be to ensure we are more precise in our communication about what the existing concepts actually capture.

This paper will expand on this point in progressive steps. We shall first recall how the notion of domestic value-added already faces some frequently ignored difficulties in a world where only intermediate goods are traded across countries. We shall then show how the problem becomes clearly unsolvable when cross-border cooperation of the factors themselves is involved, *a fortiori* with intangible factors for which the place of operation cannot be determined. The conclusion will summarize how this can legitimize the shift of emphasis from a production-based interpretation of accounts to an income-based interpretation and how this could lead to a better expression of how accounts are linked with the measurement of economic well-being.

1. Defining Domestic Production: Already Some Difficulties with Standard Forms of International Trade

Throughout this paper, we will disregard the difficulty raised by the heterogeneity of finished products, which is a subject in itself (Blanchet & Fleurbaey, this issue). We will therefore consider production Q of a single representative finished product. Subscripts will be used to identify countries. When a country i autarkically produces Q_i with local amounts of capital K_i and labor L_i, without any outside contribution, our ability to develop a volume indicator of what is produced in country i is not in any doubt, this is what its GDP is expected to do.

Intuitively, the same should hold true when production Q_i also relies on intermediate consumption IC_j of goods and services imported from another country j. This was predominantly the case when the main concepts of national accounting were worked out. In this case, identifying what is produced in the two countries should also be a no-brainer. Flows of IC_j's are observable when they cross the border and these intermediate goods have clearly been produced in country j. All the statistician has to do in country i is measure the value-added generated by local factors K_i and L_i.

Yet already, even in this case, the notion of domestic real value-added for country i proves more contentious than might be expected. Its interpretation was debated in the 1960s and 1970s, as recalled by Vanoli (2002). What national accounts observe in country i is only value-added, expressed in monetary terms by subtracting the price paid for the intermediate products from the value of total output Q. This defines the amount of money that accrues to production factors K_i and L_i in compensation for their contributions to production. Conveying this in terms of "volume added" is less straightforward. The recommended method for estimating changes in real value-added is double deflation: nominal changes in real output and real imports of *IC*s are deflated using their own specific indices, and changes in real value-added are estimated by difference.

One well-known problem with this technique of double deflation is the risk of estimating negative levels of real value-added, when using base prices that are too distant, but this problem is practically avoided by moving to chained prices.

The more fundamental issue is to determine the meaning of what is calculated in that way, since there is nothing intuitive about the notion of "volume added" and it is *de facto* never used by accountants. As noted by Sims (1969) and Arrow (1974), it implicitly requires a separable form of production function.

$$Q_i=F(K_i,L_i,IC_j)=G(H_1(K_i,L_i),IC_j) \quad (1)$$

with H_1 the quantity that is expected to be measured. There is nothing self-evident about such a specification. One case where it would have fully observable counterparts is when K_i and L_i separately produce "something" – i.e. another form of identifiable intermediate product – and where this thing would then be combined with the *IC*s to produce finished product *Q*. But the nature of this "thing" may be quite abstract: what is a "volume" of cars netted out of everything included in the manufacturing of these cars: the car body, engine, tyres, etc.? More formally, as explained by Arrow (1974), H_1 is an unobservable latent variable that can be replaced by any arbitrary transformation of it, compensated by a suitable adaptation of the form of function *G*. And this specification (1) implies a pattern of substitutability that is not always appropriate. For instance, if workers produce final output *Q* using machines of varying types, requiring a variable quantity of imported energy, the right specification is rather:

$$Q_i=F(K_i,L_i,IC_j)=G(L_i,H_2(K_i,IC_j)) \quad (2)$$

in which case it appears impossible to say exactly what will be measured by double deflation, for this specification does not entail any equivalent of the subcomponent H_1 in the first formulation.

This does not mean double deflation is completely irrelevant, but that we need a proper understanding of what it measures. The interpretation proposed by Sato (1976) consists of saying that double deflation applied to specifications (1) or (2) delivers an overall index of how changes in quantities of *K* and *L,* combined with technical progress and economies of scale, contribute to the total growth of real final output *Q*, i.e. everything that cannot be attributed to changes in IC_j. This solves the interpretation problem but only if we are interested in growth rates. What we thus construct is just an index, i.e. a measure of relative changes. To convert this to levels, national accountants apply these real changes to base year amounts that are purely nominal, offering no indication of how much of its own product each country generates in real amounts, compared to what is produced in other countries.

This lack of insight into production levels is, in fact, completely intuitive. What would be the meaning of a comparison between the volume of intermediate goods IC_j produced by country *j* and the pseudo-volume added by country *i*? The only thing of any meaning in terms of international comparisons is the money income generated by the activities carried out in each country, and how much of a given representative finished product these incomes can purchase in each of these countries. This is exactly what we do when we rank economies by size or when we compare their labour force productivity in terms of GDP adjusted for purchasing power parity.

In short, double deflation is a method that, at best, informs us about changes. Such a limitation might be acceptable if national accounts were only used to measure change, but this is not the case. From a "level" perspective, there is no way of comparing "real" levels of production for countries with different production specializations, as there cannot be any common unit for measuring them. If one country specializes in the production of airplanes, a second in the production of clothes and a third in oil mining, we are unable to compare their production of these three goods, nor are we interested in doing so; what we actually do is compare the standards of living that these countries are able to achieve thanks to their respective specializations. This is easier to characterize in terms of income rather than production.

2. The Case of Production Factors with Defined Locations, but Cooperating From Different Places

Even in the apparently simple case of well-identified products moving across borders, the definition of what is produced in each country is not that straightforward: we feel more comfortable if we focus on associated income flows and what this income can buy, rather than on the amount of physical production they represent.

These difficulties become still more evident when the production factors themselves are located in different places, and this is increasingly the situation we face in the context of digital economies with the increasing possibility of factors cooperating remotely.

To illustrate this, let's simplify the modelling still further and disregard, from now on, the role of intermediate products, i.e. a production function with only K and L, but now located respectively in countries i and j, hence $Q=F(K_i,L_j)$. One can think for instance of K_i representing the infrastructure of a data centre, with workers from country j using this data centre. Or, more generally, one can think of any setting where multiple production factors $K_1,\ldots,K_m$, $L_1,\ldots,L_n$ can cooperate from different places.

In such a case, how does one define what is produced in i and j respectively? In general, there is no answer to this question. The only situation where it would be possible to determine this is of little empirical relevance, concerning the case of an additively separable function of the form $F(K,L)=G(K)+H(L)$ where K and L would be able to produce the same category of output separately, without the help of the other factor. It goes without saying that this specification is of no practical interest; the general concept of a production function has been precisely developed to account for the fact that the production of Q generally combines K and L in more complex ways than mere addition. Once additivity is no longer applicable, the contributions of the different factors are so interwoven that it is impossible to say how much of Q has been produced by each of them.

In such a case, what we can answer is another question. Once again, the thing that is observable is the income that factors K and L derive from their contributions to production. Of course, this income is, in some way, representative of these contributions. In a competitive context, we know this income corresponds to the marginal productivity of these factors. On the added assumption of constant returns to scale, this allows total output to be broken down as $Q= F'_K K_i + F'_L L_j$ hence a breakdown of total output that reflects some properties of the production function, but only marginal properties, i.e. contributions of changes in factor quantities to the variation of production, exactly as with the Sato (1976) interpretation of double deflation, and only insofar as the functioning of markets can be deemed competitive. All this cannot be likened to measures of what K and L separately produce in the two countries.

A correlate of this result is the impossibility of offering any sound theoretical basis for any of the empirical *ad hoc* suggestions that can be or have been envisaged for calculating "how much is produced where". One could, for instance, consider K to produce a service that is exported from i to j and constitutes an intermediate consumption for workers operating in j. But we return to the problem we had of defining real value-added generated in j. The only measurable thing is, at best, the dynamic contribution of ICs and other factors to real growth, not levels of contribution.

One could also argue that, if GDP is essentially calculated in order to carry out employment policies, one could choose to locate all production where labour is located, this being a particular case of the "formulary apportionment" methodology which consists of splitting multinational enterprises' results according to conventional keys for determining apportionment, the key being, in this case, the wages paid by these MNEs in the different countries where they operate. But this is conventional. In our stylized example, can we argue that location i plays no role in production? And what if different kinds of labour are cooperating from different workplaces in different countries? Allocating this production according to these worker's relative wages raises the same problem of reflecting, at best, marginal contributions. This can result in biased messages that are clearly identified at infra-national level when one attempts to calculate regional GDP by taking the nationwide production of national companies with multiple sites and apportioning it according to the wage bills of their local units. This results in the finding that productivity is concentrated in regions where headquarters and top wages are concentrated, an observation that is not very helpful if the purpose is to identify productivity-enhancing policies (Bouba-Olga & Grossetti, 2015).

All this is not to say that reliance on this kind of methodology is not an answer, it just means it is only an answer to the question of the distribution of earnings, and not the distribution of production. For instance, a recent application of the method to an evaluation of "missing growth" caused by fiscal optimization in the US (Guvenen *et al.*, 2017) can be interpreted in this way. Such a study provides a proxy of the income impact of fiscal optimization and its implications for the tax base of the US economy, i.e. an "income-based" use of the approach, rather than information about missing production or productivity.

3. Intangible Factors

In cases similar to the Irish one, the problem is further complicated by the intangible nature

of the capital whose contribution needs to be evaluated. We can formulate the problem by shifting to a $Q=A_i.F(K_j,L_j)$ specification, where K_j and L_j now stand for physical capital and labour supposed to be located in the same place j where products (e.g. smartphones) are actually produced, and A_i the intangible capital embedded in the finished product, supposed to be "located" in country i, whether for ordinary reasons – such as having an R&D centre actually located in I – or purely for fiscal optimization purposes.

The way this A factor is treated in national accounts has paradoxical aspects that, once again, can be understood only if we shift the accent from a "production" to an "income-based" interpretation of accounts.

To show how, let's put location considerations aside for a moment. From the $A.F(K,L)$ perspective, the contribution of this intangible capital should be the same, irrespective of whether or not it is covered by an intellectual property right. But this is not the case. If A is privately owned, national accounts will deem related "production" to correspond to the income generated by this property right. The "production" generated by K and L will be reduced accordingly. On the other hand, if A is a free public resource, it will be considered as "unproductive", leading to higher levels of apparent production attributable to K and L. The problem is similar to the one we have with free natural resources, in that they are a production factor in the $F(\ldots)$ sense of the term but their contribution to production is ignored in the absence of any property right.

This can be illustrated through the production of a drug based on a formula temporarily protected by a patent. As long as the patent is active, it is deemed to produce something, thus lowering the income of other production factors (or increasing prices) and generating apparent production where the property right is located. Then, when the formula falls into the public domain, it will no longer be deemed productive, despite the fact that nothing at all has changed in the production process. The $Q=A.F(K,L)$ function remains the same; the only thing that has changed is the fact that A is not remunerated anymore for its contribution to Q. Such a convention only makes sense in terms of production if the formula becomes obsolete when the patent expires, with the associated Q also dropping to zero, which is obviously not the general rule. In other words, the problem with intellectual property rights is not just their facility to cross borders, but also the difficulty of reflecting their contribution to production on the sole basis of the money they generate.

Here again, the consequence is that it is more appropriate to talk about incomes generated by contributions to production rather than of production itself. When it is possible for the factor A to move across countries, the issue is not about knowing what is produced in this factor's host country, but what income accrues to this country as a result of being the factor's official location. As long as A is protected by a patent, we will observe an income for A that will accompany any moves it makes across borders. It will then drop to zero once the patent expires. This is a story about how incomes are generated by A, it cannot be a story about what A is actually producing, given there is no principle for quantifying A, and still less a story about where such a productive activity takes place. Being intangible, A's production location is nowhere. For intangible factors whose location is unavoidably conventional, we just observe incomes and where they end up.

Such an income-oriented approach is not only more consistent with what is theoretically measurable, it also offers a simple and understandable justification for the fact that some segments of GDP may behave in a much more volatile way than others. Volatility of the magnitude observed in Ireland is hard to accept in a measure of production. The same volatility is much easier to accept once it is clear we are measuring income, for which it is easy to conceive a high rate of transferability between places, irrespective of how production is actually organised, geographically. This kind of volatility is not necessarily pure noise that has to be neutralized; it can and must be measured for its own sake. The recent changes in Irish GDP are troubling for anyone interested in production, but they did more for the debate on fiscal optimization than if Irish accountants had immediately found a way of hiding or smoothing over the revelations made by their business data sources.

Of course, this does not mean we can feel satisfied with such volatile series as our only source of information on how small, open economies are evolving. What we need is a combination of (a) series that render this volatile behaviour, and (b) series that would more adequately reflect the domestic fundamentals of each country. This means splitting the issue of income measurement into separate sub-questions, depending on

users' detailed needs. One of these uses is to know the income accruing to Irish households or "stable" Irish production units. One can expect this indicator to move relatively slowly. GNI is a first basic step in this direction. The difference between GDP and this GNI is not that we have a production concept on one side and an income concept on the other. Both are income concepts but with different scopes, the second being more appropriate for capturing truly domestic income. Yet the 2015 Irish shock has also shown that moving from GDP to GNI is not enough, on its own, to purge volatility in its entirety, as a proportion of the profits made by multinational enterprises is not necessarily redistributed to foreign owners of the capital stock of those MNEs. The additional correction proposed by Irish accountants has been to subtract from GNI an evaluation of intangible asset depreciation, but other possibilities might be considered.

The distinction between (a) and (b)-type series may then be useful for one other major use of GDP or GNI data, namely the evaluation of fiscal bases for assessing the sustainability of budgetary policies, since volatile and more fundamental components of income have, by nature, different rates of reaction to their level of taxation.

4. At What Level Does Production Remain Measurable and with What Object?

So, what role might still be attributed to the concept of production?

First, a preliminary clarification. One objection to the above line of argument is that it apparently contradicts the well-established accounting principle of equivalence between the income and production approaches to GDP: distinguishing between the two concepts is meaningless if they are equivalent by construction. But this objection overstates the scope of this principle of equivalence. This equivalence is only an equivalence between different ways of constructing the same aggregate, and it only holds because it is not a substantive notion of "production" that national accounts capture, but simply incomes derived from participation in the production process. Moreover, this equivalence only holds in nominal terms. In real terms, we have seen that production is a non-measurable object as far as levels are concerned. At best, it is its growth rate that is measurable, but through the application of deflators that are not the same as those applicable to the income approach: output price indices in one case, *vs.* consumer price indices in the other, thus eliminating symmetry between the two approaches.

If we thus accept that these production and income perspectives are not strictly superimposable, what is the role of the production side? A limited notion of domestic production remains manageable for fully domestic activities, including services in particular and public services especially. These public services are indeed easier to think of in terms of production than in terms of income, for, in that case, it is levies on income that permit public production, as opposed to public production generating market income.

But for production with a high level of international integration, we need to accept the idea that only transnational production functions make sense. A comparison of the relative productive performances of two multinational companies producing the same kind of finished product can only be done by examining their global production functions. Looking at the domestic traces of this production does not help, except to inform us of the extent to which different countries draw monetary benefits from the presence of segments of these multinational companies on their territory. This is not a negation of "national" accounting, but a clarification of what still makes sense at national or domestic levels. Income does. Some parts of production can also continue to do so. But not all production: some can be measured and analysed only at global level. As stated in OECD (2018), "nominal GDP maintains its interpretation as the income generated in a particular territory through the use of the factors of production, including intellectual property" but "from a production-perspective, the productivity of MNEs can only be properly measured at the level of the MNE, i.e. across national borders".

5. Production, Income and Economic Well-Being

To summarize, globalization clearly provides some obvious reasons for revisiting the importance we place on the concept of production in our reporting of national accounts. Production and income are the two main keywords available to characterize what is measured by accounts, with the first of these still evidently strongly pre-eminent, since it is the one that qualifies their headline indicator, GDP. It will

always be necessary to have some measures of production, but we have to address some of the difficulties with the concept that are too frequently skirted.

The first set of difficulties is independent of the geographical organization of production. These difficulties were mentioned in the introduction: the issue of the conceptual boundaries of production. How do you define the boundary between what is said to be produced and what is not? We know the conventional nature of the definition that national accounts have adopted for this boundary. The emergence of new categories of free goods has reignited tensions that have always existed around this boundary. Even within this boundary, there is then the matter of how do you calculate the total volume produced? Can it be defined free from any reference to the amount of well-being or utility that this production is supposed to generate? GDP is unquestionably not well-being, nor even a measure of economic well-being: even this more modest objective would need to account for a wide set of things that cannot be summed up in a simple aggregate, such as the way resources are distributed across the population, the economic risks to which people are exposed, the way their assets do or do not protect them against these risks, etc. But, at the same time, contributions to well-being are the only reference metric one can have in mind for aggregating everything included in GDP. This is what makes GDP a welfare-related concept.

There is then this additional difficulty of defining "domestic" production, with an important difference in nature. As far as well-being or related notions are concerned, even if we are using concepts that are not easily observable, there is no theoretical impossibility of detailing them at local level. Well-being relates to people and, to some extent, we know where these people are or where they spend most of their time: hence it makes sense to refer to a concept of "domestic" well-being, which could take the form of an aggregate of individual utility levels $U_{i,l}$ of the individuals l=1 to N_i living in country i, if these utilities were observable. Doing the same with production is possible only when everything is locally produced, which is less and less often the case. Once production results from the interaction of factors located in different places, it is no longer possible to conceptually isolate what is produced in each of them.

Faced with this theoretical impossibility, a pragmatic position might be to agree to fall back on conceptual rigour: all that matters should be the production of indicators that meet users' needs, and some of these users still expect measures of domestic production. If such is the position, one possibility would be to rely on variants of the formulary apportionment method that we briefly outlined above. Problems of interpretation will remain however, and emphasizing them is not excessive rigour; they reflect a very basic and intuitive fact of life. The question of splitting total production by factors of production does not make any more sense than the question of knowing how much of a cake is produced by the ingredients, the oven, the recipe and the pastry chef's time and know-how respectively. What we can say, at best, is how much these different inputs have been paid (or not) for these contributions, i.e. their different incomes, which is not a measurement of how much of the cake each of them has individually produced.

One response to this could, in turn, be that speaking of "production" is just a convention, to be accepted as such, with the realisation that this word is just shorthand for "incomes accruing from contributions to production". But the question remains of whether or not to maintain the use of a vocabulary that globalization renders utterly problematic *vis-à-vis* the general public: the Irish shock of 2015 has severely affected the credibility of GDP as a production concept; it would not have had the same consequences if GDP had been presented as an income concept.

Globalization is thus a strong element in favour of increasing the emphasis on the "income" compared to the "production"-based interpretation of accounts. How far would such a reorientation help alleviate the other questions raised about national accounts?

The first clear advantage is that it makes it much more natural to limit accounts to elements with monetary counterparts. There is nothing natural about this monetary trade boundary for what is presented as a concept of production, because we have things on both sides of this boundary that can all be said to be "produced". Conversely, this monetary boundary goes without saying once it is made clear that the primary purpose is to measure incomes. And this makes it much easier to explain in what sense and to what extent GDP is linked to a measurement of well-being: everyone knows money cannot buy happiness, but everybody is also aware that it contributes to well-being.

It is on such a basis that the issue of other non-monetary dimensions of well-being can be explored, with the provision of new free services being a special case in this regard, in the spirit of the recommendations of the Stiglitz Report eleven years ago (Stiglitz *et al*., 2009). The assessment of real incomes, i.e. what money can buy, is a starting point. Broadening the scope to other dimensions of well-being can be the task of specialist satellite accounts, extending the central core of these accounts. □

BIBLIOGRAPHY

Aghion, P., Bergeaud, A., Boppart, T., Klenow, P. J. & Li, H. (2019). Missing Growth from Creative Destruction. *American Economic Review,* 109(8), 2795–2822. https://doi.org/10.1257/aer.20171745

Arrow, K. J. (1974). The measurement of real value-added. *In:* P.A. David and M.W. Reder (eds), *Nations and households in economic growth*. New York: Academic Press.

Blanchet, D. & Fleurbaey, M. (2020). Building Indicators for Inclusive Growth and its Sustainability: What Can the National Accounts Offer and How Can They Be Supplemented? *Economie et Statistique / Economics and Statistics*, this issue.

Blanchet, D., Khder M.-B., Leclair M., Lee, R. & Ragache N. (2018). La croissance est-elle sous-estimée ? Insee Références, *L'Economie française -comptes et dossiers, édition 2018,* pp. 59–79. https://www.insee.fr/fr/statistiques/fichier/3614240/ECOFRA18b_D1_croissance.pdf

Bouba-Olga, O. & Grossetti, M. (2015). La métropolisation, horizon indépassable de la croissance économique ? *Revue de l'OFCE*, 143, 117–144. https://doi.org/10.3917/reof.143.0117

Bourgeois, A. (2020). Free Services from the Digital Economy: Do We Need to Measure Their Value and How? *Economie et Statistique / Economics and Statistics*, this issue.

Brynjolfsson, E., Collis, A., Diewert, W. E., Eggers, F. & Fox, K. J. (2019). GDP-B: Accounting for the Value of New and Free Goods in the Digital Economy. NBER Working Paper N° 25695. https://doi.org/10.3386/w25695

Byrne, D. M., Fernald, J. G. & Reinsdorf, M. B. (2016). Does the United States Have a Productivity Slowdown or a Measurement Problem? *Brooking Papers on Economic Activity, 2016* (spring), pp. 109–157. https://www.brookings.edu/bpea-articles/does-the-united-states-have-a-productivity-slowdown-or-a-measure

Coyle, D. (2017). Do it yourself digital: the production boundary and the productivity puzzle. *ESCoE Discussion paper* N° 2017-01. http://dx.doi.org/10.2139/ssrn.2986725

ESRG (2016). *Report of the economic statistics review group.* https://www.cso.ie/en/csolatestnews/eventsconferencesseminars/resrg/

Feldstein, M. (2017). Underestimating the Real Growth of GDP, Personal Income, and Productivity. *Journal of Economic Perspectives*, 31(2), 145–164. https://doi.org/10.1257/jep.31.2.145

Guvenen, F., Mataloni, R. J., Rassier, D. J. & Ruhl, K. J. (2017). Offshore Profit Shifting and Domestic Productivity Measurement. *NBER and BEA Working Paper* N° 23324. https://doi.org/10.3386/w23324

Khder, M.-B., Montornes, J. & Ragache, N. (2020). Irish GDP Growth in 2015: A Puzzle and Propositions for a Solution. *Economie et Statistique / Economics and Statistics*, this issue.

Lequiller, F. & Blades, D. (2014). *Understanding National Accounts* (Second Edition). OECD Publishing. https://www.oecd.org/sdd/understanding-national-accounts-9789264214637-en.htm

OECD (2018). Globalization, Intellectual Property Products and Measurement of GDP: Issues and Proposals. Note SDD/CSSP(2018)6.

Sato, K. (1976). The meaning and measurement of the real value-added index. *The Review of Economics and Statistics*, 58(4), 434-442. https://doi.org/10.2307/1935875

Schreyer, P. (2016). GDP. *In:* M. Fleurbaey & M. Adler (eds). *The Oxford Handbook of Well-Being and Public Policy.* Oxford: Oxford University Press.

Sims, C. (1969). Theoretical basis for a double deflated index of real value-added. *The Review of Economics and Statistics*, 71, 470–471. https://doi.org/10.2307/1935875

Stapel-Weber, S. & Verrinder, J. (2016). Globalisation at work in statistics - Questions arising from the 'Irish case'. *Eurostat Review on National Accounts and Macroeconomic Indicators*, N° 2, pp. 29–44. https://ec.europa.eu/eurostat/cros/content/globalisation-work-statistics-%E2%80%94-questions-arising-%E2%80%98irish-case%E2%80%99-silke-stapel-weber-and-john_en

Stiglitz, J., Sen A. & Fitoussi, J.-P. (2009). *Report by the Commission on the Measurement of Economic Performance and Social Progress*. https://www.ladocumentationfrancaise.fr/var/storage/rapports-publics/094000427.pdf

Syverson, C. (2017). Challenges to Mismeasurement Explanations for the US Productivity Slowdown. *Journal of Economic Perspectives*, 31(2), 165–186. https://www.jstor.org/stable/44235004

UNECE (2015). *Guide to Measuring Global Production*. New York and Geneva: United Nations. https://www.unece.org/fileadmin/DAM/stats/publications/2015/Guide_to_Measuring_Global_Production__2015_.pdf

Vanoli, A. (2002). *Une histoire de la comptabilité nationale*. Paris: La Découverte, Collection Repères.

• Quarante ans d'évolution de l'offre et de la demande de travail par qualification – Progrès technique, coût du travail et transformation sociale / *Forty Years of Change in Labour Supply and Demand by Skill Level – Technical Progress, Labour Costs and Social Change* – Dominique Goux & Éric Maurin

• Intelligence artificielle, croissance et emploi : le rôle des politiques / *Artificial Intelligence, Growth and Employment: The Role of Policy* – Philippe Aghion, Céline Antonin & Simon Bunel

• Quelle valeur donner à l'action pour le climat ? / *What Value Do We Attach to Climate Action?* – Alain Quinet

• Inférence causale et évaluation d'impact / *Causal Inference and Impact Evaluation* – Denis Fougère & Nicolas Jacquemet

• L'émergence et la consolidation des modèles de microsimulation en France / *The Emergence and Consolidation of Microsimulation Methods in France* – François Legendre

N° 509 (2019) – **BIG DATA ET STATISTIQUES 2ème PARTIE / *BIG DATA AND STATISTICS PART 2***

◦ Introduction – La chaîne de valeur des données de caisse et des données moissonnées sur le Web / *Introduction – The Value Chain of Scanner and Web Scraped Data* – Jens Mehrhoff

• Les données de caisse : avancées méthodologiques et nouveaux enjeux pour le calcul d'un indice des prix à la consommation / *Scanner Data: Advances in Methodology and New Challenges for Computing Consumer Price Indices* – Marie Leclair, Isabelle Léonard, Guillaume Rateau, Patrick Sillard, Gaëtan Varlet & Pierre Vernédal

• Mesure de l'inflation avec des données de caisse et un panier fixe évolutif / *Inflation Measurement with Scanner Data and an Ever Changing Fixed Basket* – Can Tongur

• Comparaison des indices de prix des vêtements et des chaussures à partir de données de caisse et de données moissonnées sur le Web / *Comparing Price Indices of Clothing and Footwear for Scanner Data and Web Scraped Data* – Antonio G. Chessa & Robert Griffioen

• Écarts spatiaux de niveaux de prix entre régions et villes françaises avec des données de caisse / *Spatial Differences in Price Levels between French Regions* – Isabelle Léonard, Patrick Sillard, Gaëtan Varlet & Jean-Paul-Zoyem

N° 507-508 (2019) – **VARIA**

• Financer sa perte d'autonomie : rôle potentiel du revenu, du patrimoine et des prêts viagers hypothécaires / *Private Financing of Long Term Care: Income, Savings and Reverse Mortgages* – Carole Bonnet, Sandrine Juin & Anne Laferrère

◦ Commentaire – L'auto assurance du risque dépendance est elle une solution ? / *Comment – Is Self Insurance for Long Term Care Risk a Solution?* – Jérôme Wittwer

• L'impact distributif de la fiscalité locale sur les ménages en France / *The Distributional Impact of Local Taxation on Households in France* – Clément Carbonnier

• Les allocations logement ne peuvent à elles seules empêcher les arriérés de loyer / *Housing Allowances Alone Cannot Prevent Rent Arrears* – Véronique Flambard

• Le sentiment d'insécurité de l'emploi en France : entre déterminants individuels et pratiques managériales / *The Perception of Job Insecurity in France: Between Individual Determinants and Managerial Practices* – Stéphanie Moullet & Zinaïda Salibekyan

• L'impact du dispositif Scellier sur les prix des terrains à bâtir / *The Impact of the 'Scellier' Income Tax Relief on Building Land Prices in France* – Pierre Henri Bono & Alain Trannoy

• Croissance de la productivité et réallocation des ressources en France : le processus de destruction création / *Productivity Growth and Resource Reallocation in France: The Process of Creative Destruction* – Haithem Ben Hassine

N° 505-506 (2018) – **BIG DATA ET STATISTIQUES 1ère PARTIE / *BIG DATA AND STATISTICS PART 1***

◦ Introduction – Les apports des Big Data / *Introduction – The Contributions of Big Data* – Philippe Tassi

• Prévoir la croissance du PIB en lisant le journal / *Nowcasting GDP Growth by Reading Newspapers* – Clément Bortoli, Stéphanie Combes & Thomas Renault

• Utilisation de Google Trends dans les enquêtes mensuelles sur le commerce de détail de la Banque de France / *Use of Google Trends Data in Banque de France Monthly Retail Trade Surveys* – François Robin

• L'apport des Big Data pour les prévisions macroéconomiques à court terme et en « temps réel » : une revue critique / *Nowcasting and the Use of Big Data in Short Term Macroeconomic Forecasting: A Critical Review* – Pete Richardson

• Les données de téléphonie mobile peuvent-elles améliorer la mesure du tourisme international en France ? / *Can Mobile Phone Data Improve the Measurement of International Tourism in France?* – Guillaume Cousin & Fabrice Hillaireau

• Estimer la population résidente à partir de données de téléphonie mobile, une première exploration / *Estimating the Residential Population from Mobile Phone Data, an Initial Exploration* – Benjamin Sakarovitch, Marie-Pierre de Bellefon, Pauline Givord & Maarten Vanhoof

• Big Data et mesure d'audience : un mariage de raison ? / *Big Data and Audience Measurement: A Marriage of Convenience?* – Lorie Dudoignon, Fabienne Le Sager & Aurélie Vanheuverzwyn

• Économétrie et Machine Learning / *Econometrics and Machine Learning* – Arthur Charpentier, Emmanuel Flachaire & Antoine Ly

• Données numériques de masse, « données citoyennes » et confiance dans la statistique publique / *Citizen Data and Trust in Official Statistics* – Evelyn Ruppert, Francisca Grommé, Funda Ustek-Spilda & Baki Cakici

Economie et Statistique / Economics and Statistics

Objectifs généraux de la revue

Economie et Statistique / Economics and Statistics publie des articles traitant de tous les phénomènes économiques et sociaux, au niveau micro ou macro, s'appuyant sur les données de la statistique publique ou d'autres sources. Une attention particulière est portée à la qualité de la démarche statistique et à la rigueur des concepts mobilisés dans l'analyse. Pour répondre aux objectifs de la revue, les principaux messages des articles et leurs limites éventuelles doivent être formulés dans des termes accessibles à un public qui n'est pas nécessairement spécialiste du sujet de l'article.

Soumissions

Les propositions d'articles, en français ou en anglais, doivent être adressées à la rédaction de la revue (redaction-ecostat@insee.fr), en format MS-Word. Il doit s'agir de travaux originaux, qui ne sont pas soumis en parallèle à une autre revue. Un article standard fait environ 11 000 mots (y compris encadrés, tableaux, figures, annexes et bibliographie, non compris éventuels compléments en ligne). Aucune proposition initiale de plus de 12 500 mots ne sera examinée.

La soumission doit comporter deux fichiers distincts :

- Un fichier d'une page indiquant : le titre de l'article ; le prénom et nom, les affiliations (maximum deux), l'adresse e-mail et postale de chaque auteur ; un résumé de 160 mots maximum (soit environ 1 050 signes espaces compris) qui doit présenter très brièvement la problématique, indiquer la source et donner les principaux axes et conclusions de la recherche ; les codes JEL et quelques mots-clés ; d'éventuels remerciements.
- Un fichier anonymisé du manuscrit complet (texte, illustrations, bibliographie, éventuelles annexes) indiquant en première page uniquement le titre, le résumé, les codes JEL et les mots-clés.

Les propositions retenues sont évaluées par deux à trois rapporteurs (procédure en « double-aveugle »). Les articles acceptés pour publication devront être mis en forme suivant les consignes aux auteurs (accessibles sur https://www.insee.fr/fr/information/2410168). Ils pourront faire l'objet d'un travail éditorial visant à améliorer leur lisibilité et leur présentation formelle.

Publication

Les articles sont publiés en français dans l'édition papier et simultanément en français et en anglais dans l'édition électronique. Celle-ci est disponible, en accès libre, sur le site de l'Insee, le jour même de la publication ; cette mise en ligne immédiate et gratuite donne aux articles une grande visibilité. La revue est par ailleurs accessible sur le portail francophone Persée, et référencée sur le site international Repec et dans la base EconLit.

Main objectives of the journal

Economie et Statistique / Economics and Statistics publishes articles covering any micro- or macro- economic or sociological topic, either using data from public statistics or other sources. Particular attention is paid to rigor in the statistical approach and clarity in the concepts and analyses. In order to meet the journal aims, the main conclusions of the articles, as well as possible limitations, should be written to be accessible to an audience not necessarily specialist of the topic.

Submissions

Manuscripts can be submitted either in French or in English; they should be sent to the editorial team (redaction-ecostat@insee.fr), in MS-Word format. The manuscript must be original work and not submitted at the same time to any other journal. The standard length of an article is of about 11,000 words (including boxes if needed, tables and figures, appendices, list of references, but not counting online complements if any). Manuscripts of more than 12,500 words will not be considered.

Submissions must include two separate files:

- A one-page file providing: the title of the article; the first name, name, affiliation-s (at most two), e-mail et postal addresses of each author; an abstract of maximum 160 words (about 1050 characters including spaces), briefly presenting the question(s), data and methodology, and the main conclusions; JEL codes and a few keywords; acknowledgements.
- An anonymised manuscript (including the main text, illustrations, bibliography and appendices if any), mentioning only the title, abstract, JEL codes and keywords on the front page.

Proposals that meet the journal objectives are reviewed by two to three referees ("double-blind" review). The articles accepted for publication will have to be presented according to the guidelines for authors (available at https://www.insee.fr/en/information/2591257). They may be subject to editorial work aimed at improving their readability and formal presentation.

Publication

The articles are published in French in the printed edition, and simultaneously in French and in English in the online edition. The online issue is available, in open access, on the Insee website the day of its publication; this immediate and free online availability gives the articles a high visibility. The journal is also available online on the French portal Persée, and indexed in Repec and EconLit.

N° 517-518-519 - 2020

Economie ET Statistique

Economics AND Statistics

Prochain numéro / Next issue

Varia

ISBN 978-2-11-151273-3 - ISSN 0336-1454 - ECO 517-518-519
Parution octobre 2020 - PRIX : 20,90 €