Universitat Autònoma de Barcelona
Departamento de Economía Aplicada

Essays on Equality of Opportunity

Christelle Sapata

Tesis presentada para la obtención
del grado de "doctor en ciencias económicas"

Supervisor :

Professor Xavier Ramos (Universitat Autònoma de Barcelona)

Barcelona, April 2012

Acknowledgments

I would like to express my gratitude to all those who contributed to the completion of this thesis.

My upmost gratitude goes to my supervisor, Xavier Ramos, for his encouragement and continued valuable suggestions during this research and, last but not least, for his great patience.

I am also extremely grateful to Dirk Van de gaer who accepted to revise my work on many occasions. Our discussions in Louvain-La-Neuve and Ghent, his availability and his kindness provided me an essential support when I needed it most. He also suggested to me to go to Oxford to present my work in front of such an outstanding audience.

I also thank Chico Ferreira for his comments, especially for the first and third chapters. Meeting him in conferences in Buenos Aires, Louvain-la-Neuve and Oxford were a great opportunity not only to receive his precious comments but also to share great moments. I would like to thank him also for accepting me for a stay at the World Bank, which allowed me to fulfill kind of a dream.

I thank Juan Gabriel Rodriguez for his comments on the first chapter during the ECINEQ conference. I will always remember when, walking to the dinner conference, he told me to talk to Chico about a dataset. Thanks to his advice, I was invited to work with Chico.

My extreme gratitude goes also to Marc Fleurbaey. In Canazei, he accepted to arrange for me a 6-month stay, initially in Paris, eventually at CORE in Louvain-la-Neuve. There, I received his very helpful comments and support with an extreme sense of pedagogy and a great dose of patience. This invitation has changed the course of my Ph.D and I am really indebted for this for the rest of my life.

Spending six months at CORE has been the opportunity to work under the supervision of François Maniquet, whom I want to thank for his availability,

his commitment with rigor and his invitation to conferences and seminars, especially these black board presentations in Louvain-la-Neuve. I want to thank all the participants to these meetings for their beneficial comments during the last two years.

The CORE will always be a special place for me. I could benefit from the regular suggestions and sense of humor of Erik Schokkaert, to whom I am also grateful. I could meet on a regular basis Dirk Van de gaer who was always extremely patient regarding my questions.

I thank all the friends and colleagues from France, Barcelona, Louvain-La-Neuve and Brussels, who have always been there for me during good times and hard times of the Ph.D. Their warmth, affection and great sense of humor have been essential ingredients of this thesis. I am especially indebted to Josep Lluís Raymond-Bara, for his answers on econometric topics, and to the CEREC members who provided the best environment for completing this thesis.

Especially, I want to thank my parents for their unconditional love and support, continued encouragement and great cooking that have been essential for the completion of this work.

My profound gratitude goes to my husband Jean-François. Without him, I would not have finished this thesis. He was my strongest support and my best coach. He beard the tears in hard times and found the words to renew my motivation. He accepted my bad mood when I faced problems with my loops and encouraged me until I found the right solutions. He even managed to make me work on Sunday mornings between two fresh made croissants. Finally, he found the best means to complete this thesis: Louis-Léon, our sweet little boy has been the best deadline I have ever had. His moves in my growing belly were a constant reminder for working harder and harder. Thank you lovely Louis-Léon for being born just in time.

Contents

List of Figures

List of Tables

General Introduction

The theory of equality of opportunity has developed a responsibility-sensitive egalitarianism that leads to legitimate some inequalities: inequalities due to factors for which the individual can be held responsible are fair. As a consequence, there is no room for redistribution when inequalities arise from the exercise of individual's responsibility. On the contrary, when inequalities are due to factors that are beyond individual's responsibility, these inequalities are unfair and should be removed.

The division between responsibility factors and non-responsibility factors is far from straightforward. Nevertheless, in the literature, some variables such as family background are always used as a non-responsibility factor. Thus, according to this theory, in a fair society everybody should be able to reach the same set of outcomes whatever his family background. But this does not mean perfect equality among individuals, because individuals bear the consequences of their preferences, choices, effort, all these factors that can be included into the responsibility factors. Therefore, if individuals make different choices, the inequality resulting from these choices leads to fair inequalities.

This illustrates the two principles of equality of opportunity: the compensation principle, according to which we should compensate for unfair inequalities, and the reward principle that consists in accepting inequalities explained by responsibility factors. These two principles explain why this concept of equality of opportunity has a strong appeal.

On the one hand, it manages to justify redistribution schemes by relying on fairness principles that are hardly refutable. Indeed, only few people would not accept removing inequalities due to the family background. So, even when restricting to the minimum the factors for which we should compensate, redistribution may find a justification.

On the other hand, this theory is also worth being studied because it puts emphasis on individuals' responsibility and this corresponds to current political concerns. In the developing world, the World Bank is promoting equality of

opportunity as a means to achieve development. In the developed world as well, the fact that many countries make unemployment subsidies conditional to demonstrable search for a job shows how policies are increasingly inspired by the principles of equality of opportunity.

However, even though a consensus emerges about the concept itself of equal opportunities, the same is not true for the measurement methods. Very distinct measurement strategies have emerged and may lead to distinct conclusions. This thesis studies three aspects concerning the measurement of equality of opportunity. We aim to study the impact of the choice of the techniques used to measure inequality of opportunity on the conclusions about unfair inequalities. To the extent that techniques might matter, examining distinct measurement strategies could be a first necessary step before designing public policies that would reduce inequality of opportunity. Therefore, every chapter of the thesis deals with specific methods for measuring inequality of opportunity and investigates how the methods of measurement impact on our understanding of the magnitude of unfairness in distinct economies.

The first chapter measures ex-ante inequality of opportunity in Spain. In the literature, the ex-ante view of equality of opportunity refers to a situation where non-responsibility factors do not impact on individuals' outcomes. There is ex-ante equality of opportunity when, whatever the endowments of individuals in non-responsibility factors, individuals can obtain the same opportunity sets, that is to say, the same set of outcomes. It remains that the final distribution of outcome may be unequal if individual differ in their responsibility factors.

To measure ex-ante (in)equality of opportunity, we can measure the impact of one or several non-responsibility factors on individuals' outcomes. Also, we can decide to account or not for the indirect impact of non-responsibility factors on responsibility factors and finally we can move the cut that separates both types of factors. Our study investigates to which extent the magnitude of inequality of opportunity is sensitive to the change in one of these three decisions. To this end, we use the model proposed by Bourguignon et al. [13, 14] to measure the direct and indirect contribution of each non-responsibility factor on individuals' income and use the 2005 EU-SILC dataset to perform the analysis for Spain.

We find that the magnitude of ex-ante inequality of opportunity is very sensitive to the inclusion of several non-responsibility factors. As expected, family background is found to explain inequalities in income, but other non-responsibility variables such as country of birth or gender are also found to be important determinants of inequalities. Therefore, when enlarging the set of non-responsibility factors, we better estimate by how much each of these

variables impact on unfair inequalities. The indirect effect of non-responsibility factors is also found to be of main importance as it may explain around half the overall inequality of opportunity. Finally, the cut between responsibility factors and non-responsibility factors is not found to be of main impact as the variables that are obviously beyond individuals' responsibility are the ones that impact the most on inequality.

The second chapter adopts an ex-post view of equality of opportunity to study inequality of opportunity across regions in France. The ex-post approach considers that there is equality of opportunity if two individuals with the same responsibility factors (called effort) achieve the same outcome. To measure ex-post (in)equality of opportunity, we need to define a measure of effort exerted by the individual. Because the responsibility factors are hardly observable, one alternative has consisted in identifying them indirectly as being "all what is not a non-responsibility factor (called circumstance) and impacts on outcomes". Here, we take advantage of the 2005 EU-SILC dataset for France that includes many indicators of effort to measure ex-post inequality of opportunity with a direct measure of effort. The purpose is to address two original questions: Is effort equally rewarded across regions of France and is ex-post inequality of opportunity distributed in the same way as income?

To this end, we estimate for each region a wage equation as a function of circumstance and effort. In this way, every circumstance and effort variables may impact unequally on individuals' income across regions. Then, we calculate the fair income as defined by Almas et al. [4]. This is the income that depends on effort only. Finally, we take the distance between the actual earnings of an individual and the fair income to obtain a measure of ex-post inequality of opportunity.

We exhibit the presence of inequality of opportunity in France as we find that non-responsibility variables significantly impact on individuals' incomes. In addition, the way responsibility factors are rewarded differ across regions. On the one hand, the distribution of sectors across regions cannot explain these differences in any region, which indicates that the issue of equality of opportunity should be addressed at a decentralized level. On the other hand, the rankings of the regions in terms of equality of opportunity and inequality in income change slightly, and this is consistent with other studies on the correlation between income inequality and inequality of opportunity (Checchi, Peragine [20], Lefranc et al. [42])

The third chapter applies criteria of equality of opportunity when individuals have heterogeneous preferences on consumption and leisure. In the first two chapters, we considered that people are only interested in their earnings, the consequence of this hypothesis of homogeneous preferences is that we can

aggregate individuals' outcome and obtain aggregate measures of inequality of opportunity. In this chapter, we assume individuals' preferences are heterogeneous. As a consequence of this, interpersonal comparisons are restricted and the criteria of equality of opportunity have to be adapted. In fact, criteria of fairness can still be proposed as Fleurbaey and Maniquet [31, 34] do. They achieve rankings of individuals according to the compensation and reward principle in a framework where individuals' preferences are identifiable through their actual choices on consumption and leisure in an ordinal framework. Our aim is to propose a model to apply these criteria and make them comparable with some cardinal criteria.

To this end, we start from a model given by Decoster and Haan [25] that identifies groups' preferences through the observed choices made by individuals on consumption and leisure. Then we make use of the information on groups' preferences and on the individuals' choices to approximate individuals' preferences. This allows us to apply Fleurbaey and Maniquet criteria with an ordinal measure of well-being. Finally, we propose two ways of cardinalizing our measure such as to compare our results with the criteria proposed by Roemer [59, 58] and Van de gaer [61].

The empirical application uses the Cross National Equivalent File for the US in 2005 and identify the worst-off according to each criterion. We show very little matching among the criteria. Firstly, the target of redistribution policy changes as we move from the egalitarian equivalence criterion to the conditional equality criterion. Secondly, the differences we observe are consistent with the theoretical predictions. Then, we find strong similarities between Van de gaer's criterion and the conditional equality criterion. Roemer's criterion appears to give very distinct conclusions with respect to the other three criteria. As a conclusion, the difference in the results given by each criterion shows that the strategies of measurement are at least as important as the criterion used.

As a conclusion, the results obtained in this thesis show how much the technics of measurement drive the results on inequality of opportunity. Far from discrediting these methods, this thesis would tend to show that making explicit all the assumptions that are included in any empirical research is extremely useful to make reliable conclusions. Finally, these results may constitute a reason to encourage a political debate on the scope of individuals' responsibility. Indeed, neither sensitivity analysis nor comparison between methods could ever replace a normative debate among citizens to determine what should be compensented and what should be rewarded.

Chapter 1

Comparing Frameworks for Measuring Inequality of Opportunity

Abstract

The aim of the paper is to assess by how much competing strategies of measurement of inequality of opportunity differ in terms of the magnitude of measured inequality of opportunity. Because techniques that measure (in)equality of opportunity reflect particular views of equality of opportunity, comparisons are not straightforward. We choose a model and propose an application that relies on distinct ideas of fair and unfair inequalities. It allows us to (1) measure inequality of opportunity with one or several circumstances, (2) distinguish between the direct and indirect effect of circumstances and (3)draw up a flexible responsibility-cut. We apply the model to the acquisition of labour income in Spain and compare the results from each perspective. We find that including only one circumstance implies a significant underestimation of the magnitude of inequality of opportunity. In addition, measuring the indirect effect of circumstances appears to have a large effect on the magnitude of inequality of opportunity. Indeed, it accounts for half the overall observed inequality of opportunity. Finally, we find that this model offers a close link between the measurement of inequality of opportunity and the design of the redistribution policies.

1.1 Introduction

The concept of inequality of opportunity has emerged as the new benchmark for assessing justice issues [31, 58]. Put simply, under the equality of opportunity approach, inequalities are fair if they are caused by factors for which the individual can be held responsible; these factors are called effort variables or responsibility factors. On the contrary, inequalities are unfair if they come from determinants that are beyond individual's responsibility; these are called circumstances. To the extent that the individual is not responsible for them, maintaining these inequalities is unfair and therefore a compensation is required.

This theoretical approach has been widely accepted [31, 58] but measuring the magnitude of inequality of opportunity has lead to competing strategies such that the task of comparing empirical results remains difficult. Indeed, distinguishing between the fair and unfair components of inequality of outcome requires paying attention to the determinants of the individual outcome and dividing them according to the realm of individual responsibility, as first exposed by Roemer (1993, 1998). However, this classification between effort and circumstances is far from straightforward and has raised three types of questions[1].

Firstly, empirical studies do not use the same number of effort and circumstance variables. A body of the literature [42, 43, 57] estimates inequality of opportunity based on one circumstance variable only. In general, an indicator of the family background is selected, since it allows for international comparison (to the extent that this data is quite homogeneous among countries). As explained by Lefranc et al. [43], circumstances are often highly correlated. As a consequence, the estimation of the observed circumstance often captures the unobserved circumstances as well. Though just accounting for one circumstance prevents to examine the channels causing inequality of opportunity, this approach offers a good estimate of the impact of the circumstances.

An alternative is to perform multi-circumstance analysis. The added value of this option for particular countries is not clear since measures of inequality of opportunity depend on which variables are selected. For instance, Nilsson [49] has found little inequality of opportunity for income in Sweden, mainly because his circumstance variables (number of siblings, fathers employed, fathers divorced, fathers' age) do very little to explain the overall inequality. By contrast, Bourguignon et al. [13] have used other variables (gender, race, birth place, fathers' labour market status) for Brazil and they have concluded that inequality of opportunities accounts for approximately 10 percentage

[1] Here we focus on the Roemer's approach and we do not consider other views of inequality of opportunity, for instance those exposed in Fleurbaey [31].

points of the Gini coefficient for individual income. The difference in the overall inequality between Brazil and Sweden could be a natural explanation for these diverging findings. Nevertheless, it also illustrates the potential impact of including several variables in the set of circumstances.

Secondly, empirical studies make distinct choices about the realm of individual's responsibility. This issue is also of main interest because conclusions substantially depend on the way the responsibility-cut is defined. In fact, if we assume independence between the relevant circumstance variables, including one more circumstance variable into an income equation would generate a commensurate increase in inequality of opportunity for income. Surprisingly, for Belgium, Devooght found that the magnitude of unfair inequalities conditional to changes in the responsibility-cut varies only by 5%. Nevertheless, this kind of study remains useful to justify (or disapprove) the importance of discussing the proper realm of individual responsibility.

Finally, models differ on the techniques used to measure the effect of circumstances on outcomes. Roemer [58] has convincingly explained how circumstances shape outcomes through two channels: A direct channel captures the effect of circumstances on outcome and an indirect one measures the same phenomenon through the effect on effort. His famous example is as follows: It has been observed that Asian children tend to work harder at school because culture at home promotes this. On a scale of 0 to 10, an average European child exerts an absolute level of effort of 5 and an average Asian child exerts an absolute level of effort of 7. This absolute level of effort mixes the individualistic component of effort with the effect of parental influence. It would be more convenient to work with the relative level of effort, that is to say, the distribution of the effort conditional to their circumstances. Here, the average European child with an absolute level of effort of 5 would be exerting a similar relative level of effort to an average Asian child that exerts an absolute level of effort of 7. Using this method, the relative level of effort would be cleaned from the effect of circumstances on effort. As a matter of fact, the indirect effect of circumstance in Brazil has been found to explain a substantial part of the overall inequality of opportunity [13]. On the contrary, Jones et al. [40] obtain that the indirect effect of circumstance almost do not contribute to inequality of opportunity for health in Sweden.

To summarize, many studies that measure inequality of opportunity have been focusing on three distinct questions: the importance of including many circumstance variables to approximate the magnitude of inequality of opportunity, the relevance of including flexible responsibility-cuts and the way of disentangling the double effect of circumstances on effort and outcomes. As each study proposes its own method of measurement to answer one particular question, comparison is difficult. In addition, the conflictive results are not

conclusive due to the heterogeneity of the countries.

In this paper, we seek to assess by how much the competing strategies of measurement differ in terms of the magnitude of measured inequality of opportunity. To this end, we propose a model that embraces the different perspectives explained above. It allows us to (1) measure inequality of opportunity with one or several circumstances, (2) distinguish between the direct and indirect effect of circumstances and (3) draw up a flexible responsibility-cut. As a result, we are able to work out the magnitude of inequality of opportunity obtained with different views about this concept.

To compare the differences in the amplitude of inequality of opportunity, we implement this model for the acquisition of labour income in Spain. Roemer et al. [57] as well as Rodriguez [56] have already studied inequality of opportunity in Spain by showing how parents' economic and academic situations impact on children's outcome. They also find that Spain displays an intermediate level of inequality of opportunity, which has been decreasing in the last decades [56]. As a consequence, our results will have the advantage of relying on a country that does not display extreme levels of inequality. Also, this analysis will complement Rodriguez's study by giving additional information on the sources of inequality of opportunity.

We propose a baseline specification where earnings depend on one circumstance variable and several effort variables. We equalize the circumstances among individuals and compute the reduction of total inequality we would obtain. Then, we enlarge the circumstance set to other circumstances which are obviously beyond individuals' responsibility and compare with the baseline specification. We also use Bourguignon et al.[13] study to measure the direct and indirect share of unfair inqualities in the total income inequality. Finally we propose different responsibility-cuts and compare the magnitude of the unfair inequalities according to different views of individual's responsibility.

The rest of this paper is organized as follows: Section 1.2 presents the model. Section 1.3 discusses the choice of the data and variables. In Section 1.4 are grouped the main results. Finally, Section 1.5 provides a conclusion.

1.2 Comparing the measures of inequality of opportunity: the model

To compare the frameworks for measuring inequality of opportunity we propose to work with Mincerian equations adapted to a responsibility-sensitive perspective, a strategy that has gained an increasing support [4, 13, 27, 49] since we can identify how effort and circumstances shape overall inequality.

Firstly, we estimate an earning equation as a function of effort variables and one circumstance variable only, then we measure by how much the inequality would decrease if we equalized this circumstance among individuals. This is the baseline specification. Then, we enlarge the set of circumstances and perform the same analysis to check if we can improve substantially our understanding of the sources of inequality of opportunity. In this first step, we only consider the direct impact of circumstances on income. Then, we account for the indirect effect of circumstances on income through their impact on effort variables. This is the point made by Roemer [58] and it has been widely used in empirical studies when measuring the true impact of circumstances [40, 41]. In order to measure the importance of the indirect effect, we measure by how much the overall inequality would be reduced if we account for both the direct and indirect effect of circumstances. Finally, we equalize each circumstance separately among indivuals, and measure the indirect and direct effect of this equalization on the overall inequality. We use these results to establish a ranking of the circumstances as suggested by Bourguignon et al. [13]. The methodology is precisely described as follows.

1.2.1 The baseline specification

In the baseline specification, we estimate an equation where individual's earnings is a function of one circumstance variable and other control variables, which are, by default, effort variables. The unique circumstance we use is the father's education since it is the most used variable in the literature when a single circumstance variable is selected [19, 41, 42, 43]. The other variables are effort variables as they cannot be considered automatically as circumstances. These are the region of residence, the degree or urbanisation of the region of residence, the sector where the individual works and the own education.

Formally, the earnings equation is specified as follows:

$$\ln(w) = A_1 + \alpha_1 E + \beta_1 C + u_1 \tag{1.1}$$

Where A_1 is the constant, $\ln(w)$ is the logarithm of the labour income, E is the vector of effort variables, C the circumstance variable and u_1 is the error term. Individual i subscripts are omitted for convenience in the rest of the Chapter[2].

Equality of opportunity requires that the circumstance variable does not cause significant differences in the acquisition of labour income among individuals. As a consequence, the estimator associated with the circumstance variable should not be significantly different from zero. We thus test the significance of β

[2]The effort and circumstance variables are listed in further detail in the next Section.

to check the relevance of father's education in causing inequality of opportunity.

Since one of the effort variables, education, might be subject to endogeneity, we perform Bound et al. tests [12] to identify potential instruments. As we could find in our dataset no variable at the same time correleated with education and not correlated with income, we had to estimate the baseline equation with OLS. Even if the estimator can be biased in presence of endogeneity, it does not prevent to capture the existence of inequality of opportunity as long as β_2 remains significantly different from zero.

Then, we assign to all individuals the same circumstance (here it is the same level of father's education for all) and build the distribution of income in this hypothetical case. Therefore, the difference between the actual income inequality and the inequality of this counter-factual distribution is interpreted as the direct contribution of father's education to overall inequality. The possible bias of the estimator does not matter to perform this analysis[3].

This method has been first developed by Bourguignon et al. [13] to rank the circumstances variables that cause inequality of opportunity in Brazil. It gives a good intuition about the importance of a variable at causing inequalities. This method assumes that more a circumstance generates inequality of income, more it contributes to inequality of opportunity. In this Chapter, this assumption will allow us to compare on a same basis different frameworks on inequality of opportunity.

We use the estimated parameters of the baseline specification and the estimated residuals, we equalize the value of the circumstance for all the individuals and take the observed values of effort, so as to obtain the counter-factual distribution of income as follows:

$$\ln(\tilde{w}_1) = \hat{A}_1 + \hat{\beta}_1 \bar{C} + \hat{\alpha}_1 E + \hat{u}_1 \tag{1.2}$$

Where $\ln(\tilde{w}_1)$ is the counter-factual individual's income, E corresponds to the vector of observed effort variables, $\bar{C}$ is the equalized circumstance variable, $\hat{\beta}_1$, $\hat{\alpha}_1$, and $\hat{u}_1$ are the parameters estimated in Equation 1.1.

We measure the inequality of this counter-factual distribution of income and compare it with the inequality of the actual distribution of income as proposed by Bourguignon et al.:

$$\Theta_d = 1 - \frac{I_{\tilde{w}_1}}{I_w} \tag{1.3}$$

Where I_w and $I_{\tilde{w}_1}$ are inequality indices of the observed and counter-factual distribution of income.

[3]See the Appendix A for a discussion about the estimation with instrumental variables.

1.2.2 Enlarging the set of circumstances

We enlarge then the set of circumstances to all the available variables in the dataset that are obviously out of the real of individual's responsibility. As long as the variables ar found to impact significantly on incomes, taking them into account allows to better understand the cause of unfair inequalities. From a policy perspective, it seems to be a first crucial step to move forward to the design of policies aiming at reducing these inequalities.

To the previous framework we now insert additional circumstance variables into the estimation. We obtain:

$$\ln(w) = A_2 + \alpha_2 E + \beta_2 C + u_2 \tag{1.4}$$

where E and C are both vectors of effort and circumstance variables.

Due to potential endogeneity of individual's education we reiterate tests to detect potential instruments. We now identify one satisfactory instrument: the number of siblings. We use it and perform an IV estimation[4]. Based on Hausman test comparing OLS and IV estimates, we conclude that we cannot reject the null hypothesis of equality of the parameters. So, our analysis remains based on OLS results.

Then, we equalize all the circumstance variables among the individuals and measure by how much the overall inequality will be reduced in this case. We proceed as in the baseline specification. It gives:

$$ln(\tilde{w}_2) = \hat{A}_2 + \hat{\beta}_2 \bar{C} + \hat{\alpha}_2 E + \hat{u}_2 \tag{1.5}$$

Where the estimated coefficients and residuals come from Equation 1.4 and $ln(\tilde{w}_2)$ is the resulting counter-factual distribution of income. By comparing the inequality in this distribution with the actual income inequality, we obtain the corresponding share of unfair inequalities:

$$\Theta_d = 1 - \frac{I_{\tilde{w}_2}}{I_w} \tag{1.6}$$

These two procedures allow to assess by how much widening the set of circumstances improves our understanding of the component of unfair inequalities.

1.2.3 Distinguishing the direct and indirect effect of circumstances

In order to measure the importance of accounting for the indirect effect of circumstances, we estimate the impact of circumstances on income through their effect on effort variables. To this end, we estimate the effort variables as

[4]See Appendix B for detail on the estimations and results.

a function of circumstance variables. This estimation is then transposed into the previous framework as follows. We build a counter-factual distribution of effort where we equalize the circumstance variable among individuals. Then, we generate another counter-factual distribution of income by using the counter-factual distribution of effort and the same circumstance for all. As a result, the difference between the overall inequality and the inequality in this counter-factual distribution measures the total effect of circumstances on income.

Formally, we estimate each effort variable j on the circumstances:

$$E_j = A_3 + \gamma C + \eta \tag{1.7}$$

Removing the effect of circumstances on effort by keeping the residuals only of this equation makes sense for continuous variables but gives rise to problems in case of categorical variables. Indeed, the region of residence, an effort variable, takes no ordered values. Taking the residuals would not be meaningful. So alternative approaches are in this case required (as the one used in Chapter 2).

Before using sophisticated methods, we first check if circumstances explain effort. Based on logit and multilogit estimations[5], we find that circumstance variables have a very low power of explanation for the categorical effort variables. On this basis, since these are not subject to circumstances' influence, categorical effort variables are considered as exogeneous and not transformed in our analysis.

For years of education and experience, that we consider as approximatively continuous, we can obtain the residuals from OLS regressions. We thus estimate the counter factual distribution of effort when equalizing the circumstances only. We estimate these variables as a function of the father's education only and then we repeat the estimation for the complete set of circumstance in order to compare both approaches.

After removing the inequality in circumstances, we have:

$$\hat{E} = \hat{A}_3 + \hat{\gamma}\bar{C} + \hat{\eta} \tag{1.8}$$

where $\hat{E}$, the counter-factual effort, is the part of the effort that is explained by residual factors only. With this term, we generate the counter-factual distribution of income after equalizing the circumstance and replacing the actual effort by the counter-factual effort.

$ln(\ddot{w1})$, the second counter-factual income distribution is defined as:

$$ln(\ddot{w1}) = \hat{A}_1 + \hat{\beta}_1\bar{C} + \hat{\alpha}_1\hat{E}_1 + \hat{u}_1 \tag{1.9}$$

[5]Results are available on request.

Where father's education is the unique circumstance, where $\bar{C}$ is the equalized circumstance, $\hat{E}_1$ is the counter-factual effort when we use the father's education as a unique regressor in the education equation, and $\hat{\beta}_1$, $\hat{\alpha}_1$, $\hat{u}_1$ are estimated in Equations 1.8 and 1.9.

When we use the complete set of circumstance (and not only father's aducation), the counter-factual income distribution is:

$$ln(\ddot{w}2) = \hat{A}_2 + \hat{\beta}_2\bar{C} + \hat{\alpha}_2\hat{E}_2 + \hat{u}_2 \tag{1.10}$$

Where $\hat{E}_2$ is the counter-factual effort when we use the complete set of circumstances as regressors in the effort equations.

Hence, the total opportunity share of earnings inequality is alternatively

$$\Theta_t = 1 - \frac{I_{\ddot{w}1}}{I_w} \tag{1.11}$$

or

$$\Theta_t = 1 - \frac{I_{\ddot{w}2}}{I_w} \tag{1.12}$$

Finally, the indirect effect is the difference between the total effect and the direct effect of circumstances:

$$\Theta_i = \Theta_t - \Theta_d \tag{1.13}$$

1.2.4 Changing the responsibility-cut

Our last objective is to assess the impact of each circumstance separately. Since we already take into account the indrect effect of circumstance, we do not enlarge the set of circumstance with previous effort variables. Moreover, it would go too far to consider that people are not responsible at all for their education or their region of residence.

Instead, it seems relevant to measure by how much each circumstance shape inequalites in income in order to understand which are the circumstances of main impact on inequality of opportunity. For policy purpose, it may help to determine which kind of policy would be more urgent for reducing inequality of opportunity.

This is why we measure the impact of equalizing each circumstance separetely and then their respective impact by grouping them by topic: family background, geographical origin, genetics, and luck. Luck is assumed to be captured by the residuals of the equation. It is an important aspect of inequality of opportunity. As a matter of fact, some papers interpret the residuals as unfair inequalities and other do not, and this may lead to different conclusions. Here we will be able to disentangle some components of unfair inequalities.

1.3 Data

Models that have been employed to measure inequality of opportunity are difficult to compare because they use very heterogeneous countries or variables. On the one hand, we can compare a large set of countries [57, 43, 19] if we use a unique simple circumstance (national surveys rarely provide homogenized data across countries). On the other hand, if we focus on one country, we can extend the number of variables that explain inequality of opportunity by examining the variables provided by national surveys. The implication is the impossibility of comparing different studies because the variables are often too different or because the countries are very heterogeneous.

Our proposal is to focus on one country only: Spain because it presents some characteristics that make an analysis of inequality of opportunity in this country particularly interesting. It has been shown that inequality of opportunity has been reduced through the last decade in Spain [56] but this estimation is based on the inclusion of only one circumstance variable. Thus, a detailed analysis of the sources of unfair inequality appears of special relevancy.

Moreover, Roemer et al. reveal that Spain present an intermediate level of inequality of opportunity in comparison with other European countries which is convenient so as to obtain reliable results.

Finally the EU-SILC (European Survey on Income and Labour conditions) completed in 2005 provides a dataset with complete information on individuals' outcome, socio-economic characteristics but also geographical indicators and detailed family background that permit the inclusion of a variety of variables which are often present in the literature.

Over a sample of 30,375 individuals, we select those between 25 and 65 years old in order to avoid people who combine education or pensions with work. We also restrict the sample to wage earners because the reported wages of the self-employed are not reliable: reported earnings are on average 45% lower for self-employed than for employees who declare themselves as wage earners. This is a problem often discussed by the Spanish Institute of Statistics but still unsolved.

In addition, we select those who worked full time the whole year. Indeed, we only have 400 individuals over 10,000 who work part time. Our results could be biased in case we include both kinds of workers (because of the lack of representativity of the part-time workers who are very numerous in Spain). In the end, we obtain a sample of 9635 individuals.

Concerning the dependent variables, we select the net labour income. Capital

income is excluded to avoid mixing two dimensions of inequality of opportunity. Rodriguez [56] shows that in Spain capital income is more unequally distributed than labour income and positively contributes to inequality of opportunity. In this way, we are under-estimating the extent of total inequality of opportunity in Spain. The net income is supposed to represent in a larger extent the income people care about, this is why we do not study the gross income. The net income is composed of the monetary and non monetary labour income the individuals declared for the year 2004.

The effort variables are the region of residence (which corresponds to the autonomous communities in Spain), the degree of urbanization (which is a categorical variable that indicates if the individual lives in a small, medium or large urban area) and the civil status of the individual (which indicates if the individual declares himself as single, married or other).

We also have the sector in which the individual works, his labour experience, we also take the experience squared to check for descreasing return of experience and a dummy variable that indicates if the firm in which the individual works has at least 10 employees.

Finally, years of schooling are considered as effort variables. Perhaps, up to a certain age, years of schooling are not under the individual's control but they are at least partly a question of choices and we can interpret them as the individual's investment in academic human capital. This variable is a linearized variable. We use the level of education and follow the instructions of the Spanish Statistical Institute to transform the distinct levels into years of schooling[6].

The circumstance variables, in the baseline specification, is father's education, a categorical variable which has been linearized following the same method as mentionned for individual's education.

When enlarging the set of circumstances, we include the gender and the presence of chronic disease, a categorical variable with two values indicating if the individual has or not a permanent disease. These two variables represent the genetic individual's characteristics.

We also have one geographical circumstance: the individual's birth country that indicates if the individual was born in Spain, the rest of Europe or out of Europe.

Finally, we complete the family background by using the mother's education (we use the same method as the one used for father's education), the fact

[6]See Table 1.7 for further detail.

the father worked when the individual was between 12 and 16 years old, the same for the mother. We also have information about the presence or not of financial problems when the individual was between 12 and 16 years old (a categorical variable with two values, 1 if the individal declares he had no financial problems, 0 if he had at least few financial problems).

Table 1.7 summarizes the information about each of the variables. Next Section presents the results.

1.4 Results

1.4.1 The importance of widening the set of circumstances

The results concerning the baseline specification (table 1.8) show that the explanatory variables explain about 30% of the variation in income and most of the variables impact significantly on income, which is quite satisfactory. As expected, father's schooling has a significant effect on earnings. This is an evidence of inequality of opportunity, but father's schooling has a lower impact than some effort variables. For instance, in terms of elasticity, one additional year of father's schooling increases by 1.1% the yearly individual's income, instead, one additional year of own education increases by 4.1% the yearly income. In consequence, the extent of inequality of opportunity appears to be quite reduced. Besides, when assuming orthogonality between effort and circumstances and equalizing the father's schooling, the income inequality would be reduced by 5.3% when using the Theil index or by 2.5% when using Gini index. Other variables are clearly significant: region, firm size, education, experience with very small decreasing return of labour experience.

Then, when enlarging the set of circumstances (table 1.10): the explanatory variables explain almost 36% of the variation in individual's income. It could indicate two phenomena: The additional variables allow us to increase the explanatory power, but these new variables seem to be strongly correlated with father's education since, when adding eight variables, this increases by 6% only the explanatory power of the model.

The correlation between the variables is clear when we pay attention to the parameter for father's schooling. The value of the parameters is twice as small as in the baseline specification. It shows that father's schooling captures other family characteristics but it cannot explain on its own all the causes of inequality of opportunity.

Moreover, the direct share of inequality of opportunity brings a new light on this debate. When we equalize all the variables related to the family back-

ground, the Gini and the Theil indexes would decrease by 50% more than before (table 1.13). The Gini coefficient would decrease by 4% and the Theil index by 8.2%. These results show the relevance of enlarging the set of circumstances.

Precisely, we observe some interesting additional phenomena concerning the family background: As expected, mother's schooling has a positive and significant effect on individual's income. The parameter has almost the same value as the one for father's schooling, which means both are equally important. On the contrary, this contrasts with the fact that mothers who worked when the individual was between 12 and 16 years old impact negatively on individual's income. This is a common result in the literature that we confirm, the role of the mother at home seems to be crucial for off-springs' income.

This is not true for fathers as we obtain a positive effect on income when the father worked when the individual was a teenager.The correlation between this variable and the presence of financial problem explain why the presence of financial problems is significant for a confidence interval of 10% only.

In comparison with the baseline specification, these first results reveal how informative is the distinction between family academic background (the effect of parental schooling) and family economic position. Neverheless, an analysis that only includes parental schooling as a circumstance is expected to yield similar results on the presence of inequality of opportunity because academic and economic backgrounds drive inequality of opportunity in the same direction. But it will lead to a large under estimation of the magnitude of inequality of opportunity.

The other cirumstances give very clear conclusions as well, the wage equation indicates a high level of inequality across genders, which is an obvious component of inequality of opportunity. Since some studies have restricted their samples to men, these measures ignore one key source of unfair inequalities. And, when equalizing the gender among individuals, the income inequality would decrease by 1.5% using the Gini index or by 3.145% using the Theil index (table 1.13). This decrease is not so large since women perform much better at school (as revealed by the education equation). This illustrates the fact that the return of schooling contributes to reduce income inequality of opportunity across genders.

Extending the number of circumstances is also instructive for the case of the country of birth. Indeed, net earnings equations put into evidence equality of opportunity between people born in the European Union (EU) and people born in Spain. This result demonstrates how the European policies are being successful in setting up a single labour market for Europeans. On the contrary, people who were born outside the EU suffer inequality of opportunities for

income acquisition. Lack of more detailed information about the country of birth prevents us from more precise conclusions although the variety of immigrant origins would require more information. In terms of share of unfair inequalities, the direct effect of the country of birth is a little smaller than the effect of family background, indeed, the income inequality would decrease by 2% using the Gini index or by 4.2% using the Theil index . It shows the large inequality of opportunity between immigrants and people born in Spain, which may be explained by the types of jobs the immigrants take in comparison to natives. Indeed a large proportion of immigrants take low skilled jobs whatever their qualifications.

In a nutshell, the decomposition of the wage inequalities into its different components illustrates the added value of adding other variables to the traditional family background variables. An equalization of all the circumstance variables would generate a decrease in the Gini index by 6.2% and by 12.5% for the Theil index (table 1.13), which represents almost half the reduction that would be observed if father's education were equalized only. To conclude, this result strongly advocates for giving priority to multi-circumstance analysis. In addition, the difference between the Gini index and the Theil index shows that inequality as well as inequality of opportunity is larger for the poorest, so that an analysis of inequality of opportunity by quantile could be of main interest[7].

1.4.2 The indirect effect of circumstances matters

We have assessed the indirect impact of circumstances on income through two variables: the education and the experience (and its squarred value) (as reported in Tables 1.9, 1.11 and 1.12).

Clearly, the schooling determinants highlight the role played by circumstances on effort variables and permit a better understanding of the role of circumstances.

Firstly, the explanatory power of the education equation increases when we include the whole set of circumstances, which confirms that the indirect effect is better captured with a multi-circumstances analysis. Education does not depend only on parents' education but also on economic factors. Again, it is worth noting that mothers' work, when the individual is a teenager, impacts negatively on the education. This explains why the indirect effect of this variable turns out to be negative.

Another interesting finding is that birth country is never significant, whatever

[7]See the analysis made by Aaberge, Mogstad and Peragine [3] on the inequality of opportunity by quantile in the long run which gives extremely interesting results on the changes of inequality of opportunity when comparing incomes by type.

the country of origin. This strengthens our conclusion on the inequality of opportunity between Spaniards and the individual born outside the European Union, to the extent earn less with a average equivalent education.

Regarding education, the presence of chronic disease impacts negatively on education but, on the contrary, has no impact on income. As a result, though the variable has an indirect effect only, it should be considered in the measure of inequality of opportunity for policy purpose.

The results for experience are almost the reverse of the ones obtained for education. The significativity of the parameters remain the same but the signs are, in general, the opposite (table 1.12). Experience appears to depend negatively on fathers' education and positively on financial problems. As experience has a positive effect on earnings, the indirect effect goes in opposite directions if we look at education or experience. In such a case, it appears to be crucial disentangling the distinct forces that compose the indirect effect.

This finding is also instructive from the perspective of human capital. Assuming experience represents professional human capital and education represents academic human capital, the wage equation indicates that the academic human capital impacts more positively on earnings than experience (table 1.10). In terms of inequality of opportunity, the experience permits to reduce the impact of circumstance on education. Moreover, we observe that experience more than compensates the impact of circumstances on education. The market appears in Spain to neutralize part of the inequality of opportunity due to family background as it rewards the academic and the professional human capital and as the two variables are correlated in opposite ways with circumstances.

To summarize, this exercise shows the relevance of accounting for the indirect effect of circumstances. Taking into account the role played by circumstance on children's schooling and experience is essential to approximate the true level of inequality of opportunity within a country. Whether a compensation is required for the indirect effect is not clear. As discussed in Jusot, Tubeuf and Trannoy [41], different philosophical views exist: On the one hand, we can consider that parents exert an effort to help their children to achieve better earnings and we should be neutral with respect to this effort. This is Swift's point of view. On the other hand, if we give priority to children's effort, they are not responsible for receiving unequal support from their parents, so we should compensate for it. But in any case, when the task is to measure the magnitude of inequality of opportunity, it turns out to be fundamental to make explicit and systematic the measure of the indirect effect of circumstance through the children's effort.

1.4.3 Compensating for what?

Inequality of opportunity is a clear concept. We should compensate for variables for which we are not responsible at all. This is precisely the added value of this concept. Therefore we assume that education or urbanization degree belong to effort variables because, in Spain, people are partly responsible for this choice.

These last results have the following objective. Due to the fact we may not be able to compensate for everything, one question remains: for what should we compensate first? For instance, should we compensate for inequality due to gender or to the indirect effect of gender on inequality in education? We want to show here how decomposing inequality of opportunity between its different components, first, allows to rank the circumstances, and, secondly, makes closer the link between the measurement of equality of opportunity and some policy recommendations. We use a multi-circumstance analysis and assess to which policy government should give priority to promote equality of opportunity.

Here, counter-factual distribution of income allows us to establish a ranking among the unfair sources of unfair inequalities to tackle. We can affirm that the more a circumstance generates inequality the more the government should aim at reducing the effect of this circumstance. Here, we found that the variables that are obviously beyond the individual's responsibility account for around 8% of the overall inequality among net earnings so that there remains plenty of room for governments to improve equality of opportunity.

Modern societies have scarce resources to allocate to redistribution and the proper redistribution task is often criticized because it generates changes in incentives that damage efficiency. In this context, it could be very instructive to employ the precise measurement of the inequality of opportunity so that the government could design policies according to these results. These policies would be easier to defend because they would not erase all inequalities but only the unfair component of inequalities.

Our results show that family background and genetics are the largest unfair sources of inequalities. In particular, the family background plays an important role through the effect on children's schooling. Therefore, equalizing educational opportunities appears to be the first determinant step to neutralize the transmission of advantages across generations.

By contrast, to decrease inequality of opportunity between genders, it would be more convenient to implement policies in the labour market because women already perform better at school. Regarding natives and migrants, we can deduce that policies in the labour market should be recommended because there

is evidence of inequality of opportunity for the acquisition of labour income even when differences in education are removed.

1.5 Conclusion

As shown in this chapter, the concept of equality of opportunity has been increasingly used to analyze within-country inequalities in income. Its definition relies on the separation between the factors that are under or beyond the individual's responsibility called the effort and circumstance variables. Even though this definition is widely accepted among economists, this concept remains compatible with very distinct frameworks of measurements, so that comparison of empirical results is far from straightforward. Mainly three conflictive perspectives have been adopted. Firstly, models differ in the number of variables they use. Secondly, models do not always account for the indirect effect of circumstances on effort and, lastly, models do not agree on the responsibility-cut.

We therefore have proposed a model that encompasses these three issues to measure the differences in the magnitude of inequality of opportunity reported by each framework. The objective is to be able to discriminate between these perspectives. We also seek to determine which issue is most critical regarding the accuracy of inequality of opportunity measurement.

We found that including several circumstances is crucial in measuring inequality of opportunity. We also distinguish between the direct and indirect effect of circumstances on outcome and found that the indirect effect should be measured to identify the different channels through which circumstances shaped inequalities. Finally, the ranking we obtain shows that family background is still the circumstance most influential on unfair inequalities.

Appendix A: Instrumental Variable Estimation

Estimating a wage equation to assess equality of opportunity requires some preliminary checks. Specially, two major problems should be clearly defined. These are the endogeneity of the education variable, and the selection bias introduced by estimating a wage equation.

The endogeneity issue is of interest here because we want to disentangle between the direct and the indirect effect of circumstances. Here, the income equation does not include any data on individual ability; as a result the estimates may be biased because ability affects educational attainment and outcomes. If ability is not observable, the education variable can be endogenous. And, as explained by Card [18], a wage equation estimated with OLS (Ordinary Least Square) may lead to under-estimating or over-estimating the return to schooling.

To tackle this issue, distinct options have been used in the literature. A first strategy is to restrict the sample to twins who are supposed to have the same ability [8] . But this information is not available for Spain and we can also question whether twins don't significantly differ in ability. Another solution consists in using a proxy for ability. Woessman [62] uses scores obtained in mathematics tests; Roemer [57] uses IQ tests as a way to measure individual ability. Lack of such data prevents us from choosing this option. Another solution is to use the Instrumental Variable (IV) method.

Where ability and education are positively correlated, IV estimates of the return to schooling will be lower than OLS estimates. This IV method is however not fully satisfactory.

Bound et al. [12] find that IV method with weak instruments produces more biased estimates than those obtained by OLS. To prevent this, based on the data available for Spain, among the instruments used in the literature and available for Spain, we collect the quarter of birth [5], the mother's educational attainment [17] and the number of siblings [24]. We finally only select the last one since it alone fulfills all the tests proposed by Bound et al. [12][8].

However, we find a higher return to schooling by using the IV rather than the OLS method. This is precisely in the line with Card's explanation where the OLS estimates are interpreted as a "conservative estimate of average causal effect of education for groups typically affected by supply-side reform" [18]. OLS results may reflect the return to schooling of those who face a higher cost

[8]Right instruments are highly correlated to the education variable, but their coefficients are not significantly different from zero into the wage equation, the R2 of the partial regression is quite high and the Sargan test does not reject the hypothesis of exogeneity of the instruments.

of investing in human capital. This is why when we use IV, we will obtain higher estimates for education.

Finally, to the extent that we are interested in the evaluation of inequality of opportunities, we assume that people with a higher cost of investment are also more affected by unequal opportunities so this can justify using the OLS method.

As a consequence, we proceed as follows: we first select the instruments which fulfill Bound et al. [12] tests. Then, we run the estimation with IV and OLS. We test for the equality between the coefficients and use OLS coefficients because we cannot reject the null hypothesis.

Table 1.1: Education equation with instrument and one circumstance

Years of schooling	Coef.	Std. Err.	t	P-value
Number of siblings	-0.210	0.043	-4.95	0.000
Father's years of schooling	0.199	0.019	10.42	0.000
More than 10 employees	0.320	0.182	1.76	0.079
Experience	-0.097	0.031	-3.16	0.002
Experience squared	0.000	0.001	0.31	0.755
Agriculture	0.122	0.528	0.23	0.818
Construction	-0.979	0.307	-3.19	0.001
Retail	0.165	0.296	0.56	0.577
Hotel	0.050	0.446	0.11	0.911
Transport	0.377	0.340	1.11	0.268
Finance	2.246	0.454	4.95	0.000
Real estate	1.069	0.350	3.05	0.002
Public adm.	1.158	0.293	3.95	0.000
Education	2.838	0.358	7.92	0.000
Other sector	0.933	0.293	3.18	0.002
High density area	0.916	0.230	3.97	0.000
Low density area	0.027	0.294	0.09	0.927
Galicia	0.921	0.676	1.36	0.174
Principados de Asturias	2.009	1.286	1.56	0.118
Cantabria	-0.206	1.543	-0.13	0.894
Pais Vasco	1.085	0.626	1.73	0.083
Navarra	4.256	2.945	1.45	0.149
Aragon	0.257	0.980	0.26	0.793
Madrid	0.615	0.525	1.17	0.242
Castilla y León	0.324	0.726	0.45	0.655
Castilla la Mancha	0.781	0.687	1.14	0.256
Extremadura	3.664	2.126	1.72	0.085
Cataluna	0.001	0.525	0.00	0.998
Valencia	0.149	0.558	0.27	0.789
Islas Baleares	0.138	1.007	0.14	0.891
Andalucia	0.082	0.536	0.15	0.878
Murcia	-0.096	0.889	-0.11	0.914
Married	0.504	0.194	2.60	0.009
Other civil status	-0.990	0.403	-2.46	0.014
Constant	12.272	0.657	18.69	0.000

Table 1.2: Wage equation with one instrument and one circumstance

ln(income)	Coef.	Std. Err.	t	P-value
Number of siblings	-0.017	0.005	-3.23	0.001
Father's years of schooling	0.012	0.002	5.00	0.000
Years of schooling	0.040	0.003	11.61	0.000
More than 10 employees	0.139	0.023	6.15	0.000
Experience	0.018	0.004	4.81	0.000
Experience squared	0.000	0.000	-2.57	0.010
Agriculture	-0.113	0.065	-1.72	0.085
Construction	0.066	0.038	1.73	0.084
Retail	-0.112	0.037	-3.05	0.002
Hotel	-0.218	0.055	-3.95	0.000
Transport	0.036	0.042	0.86	0.387
Finance	0.099	0.057	1.74	0.082
Real estate	-0.050	0.043	-1.15	0.249
Public adm.	-0.024	0.036	-0.67	0.503
Education	0.007	0.045	0.15	0.877
Other sector	-0.144	0.036	-3.96	0.000
High density area	0.022	0.029	0.77	0.440
Low density area	0.005	0.036	0.13	0.894
Galicia	0.086	0.084	1.03	0.305
Principados de Asturias	0.094	0.159	0.59	0.556
Cantabria	0.006	0.191	0.03	0.975
Pais Vasco	0.268	0.078	3.45	0.001
Navarra	0.242	0.365	0.66	0.508
Aragon	-0.009	0.121	-0.08	0.938
Madrid	0.189	0.065	2.90	0.004
Castilla y León	0.171	0.090	1.91	0.057
Castilla la Mancha	0.147	0.085	1.73	0.084
Extremadura	0.069	0.263	0.26	0.794
Cataluna	0.207	0.065	3.18	0.002
Valencia	0.090	0.069	1.31	0.190
Islas Baleares	0.377	0.125	3.03	0.003
Andalucia	0.133	0.066	2.01	0.045
Murcia	0.176	0.110	1.60	0.111
Married	0.078	0.024	3.24	0.001
Other civil status	0.075	0.050	1.50	0.135
Constant	8.461	0.092	92.41	0.000
R squared	0.292			

Table 1.3: Education equation with one instrument only

Years of schooling	Coef.	Std. Err.	t	P-value
Number of siblings	-0.412	0.013	-31.07	0.000
Constant	13.120	0.046	287.71	0.000
R squared	0.0485			

Table 1.4: Education equation with instruments and multiple circumstances

Years of schooling	Coef.	Std. Err.	t	P-value
Number of siblings	-0.193	0.045	-4.300	0.000
Father's years of schooling	0.155	0.024	6.440	0.000
Mother's years of schooling	0.063	0.028	2.240	0.026
Financial problem	-0.326	0.171	-1.910	0.056
Father worked	0.467	0.491	0.950	0.342
Mother worked	0.022	0.186	0.120	0.907
Women	0.360	0.183	1.970	0.049
Born in Europe	0.257	0.555	0.460	0.643
Born in another country	0.386	0.437	0.880	0.377
Chronic disease	-0.204	0.248	-0.820	0.410
More than 10 employees	0.367	0.184	1.990	0.046
Experience	-0.086	0.031	-2.780	0.005
Experience squared	0.000	0.001	0.220	0.826
Agriculture	0.162	0.529	0.310	0.759
Construction	-0.893	0.312	-2.860	0.004
Retail	0.074	0.299	0.250	0.804
Hotel	-0.020	0.450	-0.040	0.965
Hransport	0.389	0.343	1.140	0.256
Finance	2.100	0.454	4.620	0.000
Real estate	1.008	0.351	2.870	0.004
Public adm.	1.106	0.298	3.710	0.000
Education	2.676	0.367	7.290	0.000
Other sector	0.871	0.301	2.900	0.004
High density area	1.020	0.234	4.360	0.000
Low density area	0.120	0.296	0.400	0.686
Galicia	1.037	0.675	1.540	0.124
Principados de Asturias	1.609	1.387	1.160	0.246
Cantabria	-0.440	1.540	-0.290	0.775
Pais Vasco	1.312	0.628	2.090	0.037
Navarra	4.117	2.931	1.400	0.160
Aragon	0.234	0.983	0.240	0.812
Madrid	0.638	0.525	1.220	0.224
Castilla y León	0.341	0.736	0.460	0.643
Castilla la Mancha	0.903	0.687	1.310	0.189
Extremadura	4.067	2.121	1.920	0.055
Cataluna	0.012	0.525	0.020	0.981
Valencia	0.194	0.556	0.350	0.727
Islas Baleares	0.268	1.009	0.270	0.791
Andalucia	0.210	0.535	0.390	0.695
Murcia	0.116	0.911	0.130	0.899
Married	0.545	0.196	2.790	0.005
Other civil status	-0.958	0.408	-2.350	0.019
Constant	11.439	0.823	13.890	0.000
R squared	0.370			

Table 1.5: Wage equation with instrument and multiple circumstances

ln(income)	Coef.	Std. Err.	t	P-value
number of siblings	-0,008	0,005	-1,520	0,129
Father's years of schooling	0,008	0,003	2,710	0,007
Mother's years of schooling	0,008	0,003	2,410	0,016
Financial problem	-0,039	0,020	-1,890	0,059
Father worked	0,118	0,059	2,020	0,043
Mother worked	-0,047	0,022	-2,110	0,035
Women	-0,195	0,022	-8,960	0,000
Born in Europe	-0,002	0,066	-0,030	0,973
Born in another country	-0,203	0,052	-3,910	0,000
Chronic disease	-0,016	0,030	-0,540	0,587
Years education	0,041	0,003	12,200	0,000
More than 10 employees	0,130	0,022	5,940	0,000
Experience	0,016	0,004	4,370	0,000
Experience squared	0,000	0,000	-2,400	0,017
Agriculture	-0,129	0,063	-2,050	0,041
Construction	0,062	0,037	1,660	0,096
Retail	-0,076	0,036	-2,140	0,032
Hotel	-0,146	0,054	-2,730	0,006
Transport	0,044	0,041	1,070	0,283
Finance	0,139	0,055	2,540	0,011
Real estate	-0,019	0,042	-0,450	0,651
Public adm.	0,022	0,036	0,610	0,542
Education	0,084	0,045	1,880	0,060
Other sector	-0,073	0,036	-2,030	0,042
High density area	0,012	0,028	0,440	0,661
Low density area	0,000	0,035	0,010	0,990
Galicia	0,128	0,081	1,590	0,111
Principados de Asturias	0,192	0,165	1,160	0,245
Cantabria	-0,066	0,184	-0,360	0,717
Pais Vasco	0,282	0,075	3,760	0,000
Navarra	0,194	0,350	0,550	0,579
Aragon	0,050	0,117	0,420	0,671
Madrid	0,220	0,063	3,520	0,000
Castilla y León	0,190	0,088	2,170	0,030
Castilla la Mancha	0,147	0,082	1,790	0,074
Extremadura	-0,044	0,253	-0,170	0,862
Cataluna	0,232	0,063	3,710	0,000
Valencia	0,108	0,066	1,630	0,103
Islas Baleares	0,381	0,120	3,170	0,002
Andalucia	0,136	0,064	2,140	0,033
Murcia	0,206	0,109	1,900	0,058
Married	0,082	0,023	3,510	0,000
Other civil status	0,080	0,049	1,630	0,103
Constant	8,395	0,105	79,680	0,000
R squared	0,365			

Table 1.6: Instrumental variables vs OLS

ln(income)	IV coeff.	OLS coeff.	Difference	Std. error
Years education	0.083	0.041	0.042	0.029
Father's years of schooling	0.001	0.008	-0.006	0.004
Mother's years of schooling	0.005	0.008	-0.002	0.002
Financial problem	-0.025	-0.037	0.012	0.013
Father worked	0.099	0.124	-0.026	0.025
Mother worked	-0.048	-0.048	0.000	0.008
Women	-0.211	-0.194	-0.017	0.012
Born in Europe	-0.013	0.002	-0.015	0.023
Born in another country	-0.220	-0.222	0.002	0.018
Chronic disease	-0.007	-0.016	0.008	0.013
More than 10 employees	0.115	0.131	-0.016	0.014
Experience	0.020	0.015	0.005	0.003
Experience squared	0.000	0.000	0.000	0.000
Agriculture	-0.136	-0.136	0.000	0.021
Construction	0.100	0.057	0.043	0.031
Retail	-0.079	-0.082	0.002	0.012
Hotel	-0.146	-0.151	0.006	0.017
Transport	0.027	0.036	-0.008	0.017
Finance	0.050	0.138	-0.088	0.066
Real estate	-0.062	-0.023	-0.038	0.032
Public adm.	-0.025	0.020	-0.045	0.034
Education	-0.029	0.059	-0.088	0.080
Other sector	-0.110	-0.082	-0.028	0.027
High density area	-0.031	0.012	-0.043	0.032
Low density area	-0.005	0.002	-0.007	0.013
Galicia	0.084	0.132	-0.047	0.044
Principados de Asturias	0.124	0.187	-0.063	0.072
Cantabria	-0.048	-0.066	0.019	0.061
Pais Vasco	0.227	0.286	-0.059	0.049
Navarra	0.019	0.203	-0.184	0.172
Aragon	0.040	0.059	-0.019	0.041
Madrid	0.193	0.224	-0.031	0.031
Castilla y León	0.176	0.191	-0.015	0.031
Castilla la Mancha	0.108	0.149	-0.040	0.041
Extremadura	-0.216	-0.049	-0.167	0.204
Cataluna	0.232	0.228	0.004	0.023
Valencia	0.100	0.110	-0.010	0.025
Islas Baleares	0.369	0.386	-0.017	0.042
Andalucia	0.127	0.135	-0.008	0.024
Murcia	0.201	0.205	-0.004	0.037
Married	0.059	0.080	-0.022	0.018
Other civil status	0.120	0.075	0.046	0.033

Appendix B: Main Tables

Table 1.7: Description of the variables

Variables	Description
Total labour earnings	Sum of monetary and non monetary labour earnings
Years of schooling	Linearized variable in function of the educationnal attainment declared by the individual (2 years of pre-1ary, 7 years if complete 1ary, 11 years if lower 2ary, 14 years if upper 2ary, 16 if higher not 3ary, 17 if tertiary, 19 if PhD
Years of experience	Years the individual declares having worked in paid job
Years of experience squarred	Squarre of the years the individual declares having worked in paid job
Region of residence	They are autonomous communities in Spain
Urbanization degree	The reference value is “living in a large area", the other two values are “living in a medium size area" and “living in a small size area"
Father's years of schooling	Same method as the individual's years of education
Mother's years of schooling	Same method as the individual's years of education
Financial problems	It takes the value 1 if the indivual's family, when he was between 12 and 16 years old, had at least some financial problems
Country of birth	“Born in Europe" indicates the individual was born in the rest of the European Union (except Spain). “Born in another country" indicates the individual was born out of the European Union
Civil status	“married" indicates the individual is married, “single" indicates the individual is single and “other civil status" indicates the individual is neither married nor single (reference value is being single)
Number of siblings	The variable goes from zero to seven
Sector	It is the sector where the individual works. The reference value is the category “other" in the original survey
Firm Size	The reference value is “working in a firm having less than or 10 employees"
Father worked	It takes the value 1 if the father worked when the individual was between 12 and 16 years old
Mother worked	It takes the value 1 if the mother worked when the individual was between 12 and 16 years old
Chronic disease	It takes the value 1 if the individuals has a chronic disease

Table 1.8: Wage equation with one circumstance

ln(income)	Coef.	Std. Err.	t	P-value
Father's years of schooling	0.012	0.002	4.75	0.000
Own years of schooling	0.041	0.003	12.11	0.000
More than 10 employees	0.142	0.023	6.24	0.000
Experience	0.017	0.004	4.35	0.000
Experience squared	0.000	0.000	-2.28	0.023
Agriculture	-0.126	0.066	-1.91	0.057
Construction	0.053	0.038	1.40	0.162
Retail	-0.119	0.037	-3.22	0.001
Hotel	-0.229	0.056	-4.11	0.000
Transport	0.027	0.042	0.64	0.525
Finance	0.098	0.057	1.72	0.086
Real estate	-0.056	0.044	-1.28	0.202
Public adm.	-0.027	0.037	-0.72	0.469
Education	-0.020	0.045	-0.44	0.661
Other sector	-0.161	0.036	-4.43	0.000
High density area	0.024	0.029	0.82	0.411
Low density area	0.009	0.037	0.24	0.809
Galicia	0.101	0.084	1.20	0.231
Principados de Asturias	0.097	0.161	0.60	0.547
Cantabria	0.013	0.193	0.07	0.948
Pais Vasco	0.281	0.077	3.62	0.000
Navarra	0.267	0.369	0.73	0.468
Aragon	0.014	0.122	0.12	0.908
Madrid	0.200	0.065	3.08	0.002
Castilla y León	0.177	0.090	1.96	0.051
Castilla la Mancha	0.159	0.085	1.86	0.063
Extremadura	0.086	0.221	0.39	0.698
Cataluna	0.212	0.065	3.27	0.001
Valencia	0.102	0.069	1.47	0.141
Islas Baleares	0.382	0.126	3.04	0.002
Andalucia	0.138	0.066	2.08	0.038
Murcia	0.175	0.111	1.58	0.115
Married	0.077	0.024	3.17	0.002
Other civil status	0.073	0.050	1.47	0.141
Constant	8.419	0.091	92.98	0.000
R-squared	0.2923			

Table 1.9: Education and experience equations with one circumstance

Years of schooling	Coef.	Std. Err.	t	P-value
Father's years of schooling	0.326	0.019	17.07	0.000
Constant	11.116	0.175	63.49	0.000
R-squared	0.176			
Experience	Coef.	Std. Err.	t	p-value
Father's years of schooling	-0.780	0.060	-13.06	0.000
Constant	24.351	0.547	44.52	0.000
R-squared	0.112			
Experience squarred	Coef.	Std. Err.	t	p-value
Father's years of schooling	-32.345	2.692	-12.01	0.000
Constant	702.665	24.645	28.51	0.000
R-squared	0.096			

Table 1.10: Wage equation with multiple circumstances

ln(income)	Coef.	Std. Err.	t	P-value
Father's years of schooling	0.008	0.003	2.620	0.009
Mother's years of schooling	0.008	0.003	2.340	0.019
Financial problem	-0.037	0.020	-1.800	0.073
Father worked	0.124	0.059	2.100	0.036
Mother worked	-0.048	0.022	-2.160	0.031
Women	-0.194	0.022	-8.830	0.000
Born in Europe	0.002	0.067	0.030	0.975
Born in another country	-0.222	0.050	-4.410	0.000
Chronic disease	-0.016	0.030	-0.530	0.594
Years education	0.041	0.003	12.410	0.000
More than 10 employees	0.131	0.022	5.900	0.000
Experience	0.015	0.004	4.080	0.000
Experience squared	0.000	0.000	-2.240	0.025
Agriculture	-0.136	0.064	-2.140	0.032
Construction	0.057	0.038	1.520	0.129
Retail	-0.082	0.036	-2.270	0.023
Hotel	-0.151	0.054	-2.790	0.005
Transport	0.036	0.041	0.870	0.384
Finance	0.138	0.055	2.490	0.013
Real estate	-0.023	0.042	-0.550	0.582
Public adm.	0.020	0.036	0.550	0.581
Education	0.059	0.044	1.340	0.180
Other sector	-0.082	0.036	-2.280	0.023
High density area	0.012	0.028	0.440	0.662
Low density area	0.002	0.036	0.060	0.952
Galicia	0.132	0.081	1.630	0.103
Principados de Asturias	0.187	0.167	1.120	0.262
Cantabria	-0.066	0.185	-0.360	0.720
Pais Vasco	0.286	0.075	3.810	0.000
Navarra	0.203	0.353	0.580	0.565
Aragon	0.059	0.118	0.500	0.617
Madrid	0.224	0.062	3.590	0.000
Castilla y León	0.191	0.088	2.170	0.030
Castilla la Mancha	0.149	0.082	1.810	0.071
Extremadura	-0.049	0.212	-0.230	0.816
Cataluna	0.228	0.062	3.650	0.000
Valencia	0.110	0.066	1.660	0.097
Islas Baleares	0.386	0.121	3.190	0.001
Andalucia	0.135	0.064	2.120	0.034
Murcia	0.205	0.109	1.880	0.060
Married	0.080	0.024	3.420	0.001
Other civil status	0.075	0.049	1.530	0.126
Constant	8.3810	0.104	80.320	0.000
R-squared	0.3544			

Table 1.11: Education equation with multiple circumstances

Years of education	Coef.	Std. Err.	t	P-value
Father's years of schooling	0.231	0.025	9.17	0.000
Mother's years of schooling	0.117	0.030	3.94	0.000
Financial problem	-0.504	0.179	-2.81	0.005
Father worked	0.151	0.521	0.29	0.772
Mother worked	-0.036	0.197	-0.18	0.856
Women	1.068	0.183	5.85	0.000
Born in Europe	0.248	0.602	0.41	0.680
Born in another country	-0.565	0.440	-1.28	0.199
Chronic disease	-0.564	0.262	-2.16	0.031
Constant	10.944	0.559	19.59	0.000
R-squared	0.2156			

Table 1.12: Experience equations with multiple circumstances

Experience	Coef.	Std. Err.	t	
Father's years of schooling	-0,339	0,078	-4,360	
Mother's years of schooling	-0,570	0,092	-6,210	
Financial problem	1,735	0,555	3,130	
Father worked	4,459	1,598	2,790	
Mother worked	-2,153	0,608	-3,540	
Women	-2,601	0,564	-4,610	
Born in Europe	-0,314	1,848	-0,170	
Born in another country	-0,823	1,350	-0,610	
Chronic disease	4,070	0,803	5,070	
Constant	20,499	1,717	11,940	
R-squared	0,193			
Experience squared	Coef.	Std. Err.	t	P-value
Father's years of schooling	-14.196	3.517	-4.04	0.000
Mother's years of schooling	-22.633	4.156	-5.45	0.000
Financial problem	77.962	25.115	3.10	0.002
Father worked	164.879	72.359	2.28	0.023
Mother worked	-82.092	27.557	-2.98	0.003
Women	-117.301	25.547	-4.59	0.000
Born in Europe	-33.007	83.689	-0.39	0.693
Born in another country	-49.856	61.131	-0.82	0.415
Chronic disease	205.411	36.385	5.65	0.000
Constant	551.334	77.750	7.09	0.000
R-squared	0.1730			

Table 1.13: Share of inequality of opportunity

	Gini	Theil	Share (Gini)	Share (Theil)
Total Inequality	0.232	0.089		
Direct Effect of Equalizing:				
Groups of Circumstances				
All Circumstances	0.217	0.078	6.234	12.741
All Circumstances and Residuals	0.116	0.021	49.983	76.347
All Genetics	0.2289	0.0857	1.445	4.227
All Family Background	0.223	0.082	4.027	8.219
Each circumstance separately				
Father's education	0.226	0.084	2.681	5.535
Mother's education	0.227	0.085	1.986	4.567
Financial Problems	0.228	0.086	1.460	3.462
Father worked	0.229	0.084	1.267	5.268
Mother worked	0.229	0.085	1.095	4.942
Birth Country	0.227	0.084	2.049	6.613
Chronic disease	0.229	0.086	1.381	3.476
Gender	0.229	0.085	1.445	5.533
Total effect of Equalizing:				
Groups of Circumstances				
All Genetics	0.228	0.085	1.781	4.567
All Family Background	0.223	0.082	3.795	7.790
Each circumstance separately				
Father's education	0.224	0.083	3.269	6.931
Mother's education	0.227	0.085	1.986	4.567
Financial Problems	0.228	0.086	1.553	3.611
Father worked	0.2292	0.086	1.320	3.204
Mother worked	0.2297	0.085	1.095	4.942
Birth Country	0.227	0.084	2.049	6.613
Chronic disease	0.229	0.086	1.381	3.476
Gender	0.228	0.085	1.651	4.561

Chapter 2

Unfair Inequalities in France: A Regional Comparison

Abstract

This paper proposes a regional comparison of ex-post inequality of opportunity in France by measuring, within each region, the inequality between individuals exerting the same effort. According to the concept of equality of opportunity, inequalities due to factors for which the individual is not responsible are unfair and should be removed but inequalities due to factors for which the individual can be held responsible (effort) are fair and should be preserved. Therefore, our analysis defends a responsibility-sensitive egalitarianism whose aim is to show whether (1) regions reward equally effort, (2) inequality of opportunity is equally distributed among regions and (3) it is correlated with income inequality. In this paper, we use a direct measure of effort to measure ex-post inequality of opportunity across regions in France. To this end, we follow Almas et al. methodology [4] to define a fair income that fulfills ex-post equality of opportunity requirements. Unfairness is measured by an unfair Gini based on the distance between the actual income and the fair income. Our findings reveal that regions display differences in the magnitude of ex-post inequality of opportunity and this is due to differences in reward schemes and differences in the impact of the non-responsibility factors on income. As a consequence, this papers motivates decentralized policies to solve the problem of unequal opportunities in France. Finally, the positive correlation between income inequality and inequality of opportunity confirms previous results given in the literature.

2.1 Research Question

In the literature on redistributive justice, growing attention has been paid to the concept of equality of opportunity. This defends a responsibility-sensitive egalitarianism. In an nutshell, individuals' advantage is decomposed into two kinds of determinants and these ones determine if inequality among individuals' advantages is fair or unfair. Unfair inequalities are due to factors for which the individual is not responsible, called circumstances. As a consequence we should remove these inequalities. On the contrary, the factors for which the individual is responsible, called effort, generate fair inequalities and this does not give room to redistribution.

The division between responsibility and non-responsibility factors is far from straightforward. Nevertheless, a consensus prevails concerning certain variables. For instance in the literature, it is often accepted that education is a responsibility factor but father's education is a non-responsibility factor. But, to the extent that father's education may impact on the education of their offspring, should we consider that education is an effort variable? This example raises the question of the right measure of effort which is related to the issue about the frontier between both determinants.

More educated parents often incentive their children to get high education. If we consider that this corresponds to parents' effort, we may not want to remove the effect of parents' education on their offsrings' education. In this case, the effort of a child should be measured by his level of education. We can say that the true effort should be the gross effort.

On the contrary, we may consider that the impact of father's education on offsprings' education should be removed because this is beyond the child's responsibility. Therefore the true effort is the effort once it has been cleaned from its correlation with non-responsibility factors. This can be called the net effort.

Another question concerns the measurement of inequality of opportunity. Even though the concept of equality of opportunity is unambiguous, distinct strategies of measurement have emerged. On the one hand, the ex-ante approach consists in studying the contribution of non-responsibility factors in shaping inequalities. Effort is not a key variable in this kind of analysis because we do not need to observe it to measure inequality of opportunity. On the other hand, the ex-post approach consists in determining if individuals who exert the same effort actually obtain the same outcome, in other words, if outcome is a function of effort only. These two measures are equally valid but may lead to distinct conclusions because they measure inequality among distinct groups. The ex-ante approach measures the inequality between individuals who share distinct circumstances and the second

perspective focuses on inequalities between individuals exerting the same effort.

These two measures of inequality of opportunity have been conducted in many developed (Italy [20], Belgium [27], France [43, 44], Norway [4], US [53], Spain (Chapter 1) among others), and developing countries (for instance in Brazil [14], Latin America [30], India [48] , Africa [21]). In France, inequality of opportunity has been recently studied for income and for health [26, 41, 44] with the ex-ante approach.

In the case of income, Lefranc et al. [44] have applied first and second stochastic dominance tools to compare the income distribution conditional on the social origin. They establish a clear hierarchy between groups of distinct social origins, this ranking has been quite constant over time (1979-2000) but the income distributions have become closer to each other, this indicates the presence of inequality of opportunity but a reduction of its magnitude. They also find that, when using the variance of incomes as an indicator of inequality, social origin explains only 5% of the overall income inequality. This reveals how difficult it is to explain income and this is true for most developed countries.

Regarding health unfair inequalities, Jusot et al. [41] and Devaux et al. [26] analyze the channels through which health inequalities are transmitted across generations. The first study investigates by how much the correlation between circumstance and effort impacts on the magnitude of inequality of opportunity. They show that circumstances affect health inequalities through their impact on effort measured by obesity, vegetable consumption, smoking habits. However, whatever we consider we should or not compensate for this indirect effect, the results do not change in a large extent and in any case, parental background explain the largest part of inequalities of opportunities. The second study scrutinizes more on the mechanisms that are at the origin of inequality of opportunity in health. They show that mother's education has a direct impact on offsprings' health whereas father's education has only a indirect impact through the education and profession of the offsprings. In total, unfair health inequalities could be considerably reduced when assigning the best circumstances to all, as Gini index would decrease by 57%.

The studies on income and health cited above [26, 41, 44] correspond to the ex-ante approach for inequality of opportunity to the extent that they study the contribution of circumstances in shaping income and health inequalities. Nevertheless, it could be of main interest to change the perspective such as to include indicators of effort in order to implement an ex-post aproach. To our knowledge, this has not been studied for France. Instead this approach has been implemented in Norway [4] where direct indicators of effort have been used (education, profession, working hours). We believe that an ex-post

approach with a direct measure of effort may be relevant for France:

First, disparities across regions of France are found to be determined more by individuals' characteristics [23, 45], mainly human capital, rather than by spacial determinants such as density and potential markets. More precisely, differences in human capital and in the sector of activity have played a large role in explaining inequalities between individuals: Godechot [38] found that the increasing inequalities in the last years are due to a rise in the salary of workers working in the financial and entertainment sectors. Variables related to effort such as human capital and the sector are mainly used in the literature on equality of opportunity [4, 14, 26] and they are related to income acquisition. Thus, we have effort variables that explain inequalities between individuals but no study has adopted the theory of equality of opportunity to explore the relationship between effort and income.

Secondly, when we turn to the perceptions of individuals with respect to inequality in France [37], we observe that individuals, whatever their professions, legitimate income inequalities when they are due to differences in the number of working hours, the type of job, the dedication of the employees. This gives a strong appeal to the analysis of ex-post inequality of opportunity because this theory justifies inequalities due to effort variables and precisely states that income inequalities are fair if income is a function of effort only.

Last, we have access to the 2005 EU-SILC dataset that contains indicators of income, effort and circumstance that allow such an analysis.

Our second purpose is to measure ex-post inequality of opportunity across regions in France for the following reasons:

A regional analysis provides a more precise understanding of the unfair inequalities in France and, because there is no reason to accept unfair inequalities in any region, we should not accept distinct levels of unfairness across regions. In a context where income inequalities has been found to be rather low between regions except Ile de France[1] [23], we ask whether the same is true for unfair income inequalities. This is important for policy design. Indeed, policies should be adapted if unfair inequalities are not distributed as income inequalities.

In addition, even when there is no correlation a priori between income inequality and inequality of opportunity, empirical studies have shown a positive correlation between both terms in European countries and the US [43], in Italy [20] and in India [48]. In this way, our study may give new evidence on this correlation by comparing the ex-post inequality of opportunity and the

[1]See also the Table 2.9 for the levels of wage and GDP per region in 2004.

income inequality across regions.

Then, we choose to explore unfair inequalities across regions because regions are old and homogeneous entities in terms of culture and this may impact on the way effort and circumstances may impact on income. Also, regions of France have displayed distinct economic histories and dynamics and this may lead to differences with respect to equality of opportunity. And regarding the size of the regions, they are sufficiently big in terms of population to allow reliable conclusions and sufficiently distint to allow differences in terms of ex-post inequality of opportunity[2].

As a result, this paper will provide a first measure of ex-post inequality of opportunity across regions in France. We aim to study to which extent regions reward equally effort and only effort. We also want to determine if ex-post inequality of opportunity is equally distributed among regions and to compare the rankings of regions in terms of income inequality and inequality of opportunity. To this end, we use the 2005 EU-SILC dataset for France. We regress the gross labour income on responsibility and non-responsibility factors for each region separately and use Almas et al. [4] methodology to define ex-post inequality of opportunity.

The rest of the paper is organized as follows: Section 2.2 presents the methodology for measuring ex-post inequality of opportunity. Section 2.3 introduces the data, Section 2.4 summarizes the results and Section 2.5 is the conclusion.

2.2 Methodology

The ex-post approach checks whether effort is the only determinant of income, therefore effort is the core input in this view. A first method proposed by Roemer [58] following the ex-post approach considers that effort can be captured indirectly through the Roemer's Identification Axiom: assuming the outcome is a monotonous increasing function of effort, once we partition the population into types, that is to say into groups of individuals sharing the same non-responsibility factors, the individuals who locate at the same quantile of their income distribution per type are supposed to exert the same effort. The fact we first partition individuals into types permits to correct for the effect of the type on effort. This means that we measure net effort.

In this case, because everything that is not a circumstance is effort, this method is often said to provide a lower bound for inequality of opportunity. However, though difficult to delimit, effort could be partially captured by some

[2]An analysis per department could be also valuable but we would have too few observations.

responsibility factors. This restricts the power of residuals to proxy effort. In addition, there is no guarantee that residual distributions conditional on types are equal, and no guarantee that residuals precisely capture effort. Residuals capture effort but also omitted circumstances.

For all these reasons, we propose to take into account effort variables explicitely, rather than implicitely through the residuals. The cut is not always clear between circumstances and effort. Following Arneson [6], Cohen [22], Roemer [58, 59], our cut is strictly based on control. If the agent has, even partially, control on the determination of the variable at some moment of his life, then it is a responsibility factor. The residual still includes unobserved effort and unobserved circumstances because we are still far from capturing all the determinants of income. Nevertheless, as shown in Section 2.4, the responsibility and non-responsibility factors explain about half of income inequality.

We thus estimate a classical Mincerian wage (y_i) equation, where the regressors are split into two categories, X_i^E and X_i^C, where E and C stand for effort (responsibility factors) and circumstances (non responsibility factors), respectively:

$$log(y_i) = \beta_0 + \beta^E X_i^E + \beta^C X_i^C + \epsilon_i \tag{2.1}$$

We do not consider residuals as part of the effort variables set. Since the residuals could include some omitted circumstances we follow the method proposed by Bjorklund et al. [10] which consists in splitting the residuals in a term which is standardized by standard error conditional on circumstances, and an heteroskedastic term depending on circumstances[3]. Since the dependence of the heteroskedastic term on circumstances was close to zero, we do not include this additional term in our specification for the rest of the paper.

As regards the potential correlation of effort and circumstance variables, it is common to distinguish the direct effect of circumstances on the income and the indirect effect of circumstances, exerted through the effort variables. We thus propose to measure ex-post inequality of opportunity according to the two views, one where effort is included as such (gross effort) and the other

[3]Bjorklund et al. decomposition is based on the following formulas:

$$\begin{aligned} Var(\epsilon_i|X^C) &= \sigma_c^2 \\ \epsilon_i &= \epsilon_i - \epsilon_i/\sigma_c + \epsilon_i/\sigma_c \\ \epsilon_i &= \tilde{\epsilon}_i^c + u_i \\ u_i &= \epsilon_i/\sigma_c \\ log(y_i) &= f(X_i^E, X_i^C) + \tilde{\epsilon}_i^c + u_i \end{aligned}$$

We find no significant effet on our data.

where effort is cleaned from its correlation with the circumstances (net effort).

To this end, we regress each effort variable on the set of circumstances and the residual of each of these equations corresponds to the net effort. For continuous variables, we proceed as follows:

$$X_i^E = f(X_i^C) + \tilde{X}_i^E \tag{2.2}$$

$$\tag{2.3}$$

The net effort is simply $\tilde{X}_i^E$. For binary variables such as education or gender[4], the latent variables underlying the binary outcomes are unobservable. By consequence, the residuals of the probit cannot be obtained. We thus replace it by their best predictions, called the generalised residuals, whose formula (see Jusot et al. [41] for details) is the following:

$$E(\tilde{X}^E|E) = \frac{\phi(bX^C)}{\Phi(bX^C)(1-\Phi(bX^C)}(E - \Phi(bX^C)) \tag{2.4}$$

where $E(.|E)$ is the expectation conditional on gross effort, $\phi(.)$ and $\Phi(.)$ are the normal density and cumulative density function of a $N(0,1)$ (the residuals are assumed to be normally distributed). Therefore, we estimate two wage equations. The first wage equation is the the one given by Equation 2.1 and the second one replaces the gross effort by the net effort as follows:

$$log(y_i) = \beta_0^{'} + \beta_E^{'}\tilde{X}_i^E + \beta_C^{'}X_i^C + \epsilon_i \tag{2.5}$$

Where $\tilde{X}_i^E$ is the net effort.

Once we obtain the estimates, we define a situation of ex-post equality of opportunity by following the methodology developed by Almas et al. [4]. There is ex-post equality of opportunity if the individual receives a fair income defined as the income due to the sole responsibility factors (and not to non-responsibility factors). This definition fulfills the conditions of ex-post equality of opportunity as income is a function of effort only.

We replicate all our results based on gross effort for net effort. Just replace X^E by $\tilde{X}^E$ in the next formulas to find the equivalent. So, the fair income based on gross effort is:

$$log(y_i^F) = f(X_i^E) \tag{2.6}$$

$$y_i^F = exp(\beta_0 + \beta X_i^E) \tag{2.7}$$

We replicate all our results for net effort by replacing the coming formulas E by $\tilde{E}$. The individual's fair income y_i^F is then standardized so as to have the

[4]All the categorical variables are binary in this research paper, since we transformed the multicategorial ones into dummy variables.

same average as the actual income.

$$y_i^F = \frac{exp(\beta_0 + \beta X_i^E)}{\sum_j exp(\beta_0 + \beta X_j^E)} \sum_j y_j \tag{2.8}$$

As a result, the standardized fair income (onwards it stands for the fair income) depends on the non-responsibility factors of the whole population and on the own individual's responsibility factors. The fair income is proportional to the effort exerted by the individual. In this sense, Almas et al. state that this fair income based on the general proportionality principle.

Many statistics, such as the Gini, Theil and General Entropy indexes, are quite standard to summarize and measure inequality of opportunity. Here, we now use a new version of the Gini Index, as generalized by Almas et al. [4] to capture unfair income inequality. The standard Gini, based on Lorenz curves implicitely compare actual incomes to average income. Here, the reference income is not the average income but the fair income and this fair income is not necessarily an equal income because differences in effort justify income inequality. Almas et al. [4] thus propose an unfair Gini where the actual incomes are compared to fair incomes [5]. The distance between the actual income and the fair income is a measure of unfairness in the distribution of the individuals' income. They formulate the unfair Gini as follows:

$$G^u(A) = \frac{2}{n(n-1)\mu(A)} \sum_i i(y_i - y_i^F) \tag{2.9}$$

$$\mu(A) = n^{-1} \sum_i y_i^A \tag{2.10}$$

$$A = \left[(y_1, y_1^F), ..., (y_n, y_n^F)\right] \tag{2.11}$$

The unfair Gini is not derived from the standard Gini (based on actual incomes). Indeed, incomes are sorted according to an ascending order to compute the standard Gini. Instead, the unfair Gini orders the individuals according to the distance between the own actual income and the own fair income (from negative values to positive values). This gives no guarantee that individuals are ordered identically under both statistics. As a consequence, we do not obtain here a decomposition of the standard Gini between an unfair Gini and a residual part. Instead, we have an original measure of inequality where the reference income to asses inequality is not anymore the mean income but the fair income.

The measures of fairness are obtained for each region separately. To this end, we regress the wage equation on a same set of circumstances and effort variables for each region. There is equality of opportunity in each region if the circumstances are found not to explain income inequality. Also, we can check

[5] The properties of the index are available in the paper by Almas et al. [4]

by how much effort is rewarded across regions.

Firstly, we regress wage on effort and circumstance for each region. Then, we obtain a distribution of fair income per region and an unfair Gini per region based on the regional regressions. We do not define the fair income nationally in order to maintain the heterogeneity among regions regarding the impact of circumstances and the way each region rewards effort. This heterogeneity may be observed because regions have distinct economic histories and this may impact on the effect of circumstances. Because people across regions may have distinct preferences, the reward schemes may vary. Also, the impact of effort variables on income may be driven by market forces. This may reflect the relative abundance/scarcity of human capital, some specific sectorial needs and so on. If this is true, the state should take it into account when designing policies aiming at reducing unfair inequalities. An alternative would be to perform an analysis per department. As the departments that belong to the same regions are homogeneous in terms of cultures, we rather focus on regions in order to have more data for each region. For these reasons, we consider that we should not assume a priori that the national fair income should be unique, instead we prefer to be agnostic regarding the national fair income.

Another consequence of this choice is that the regional fair income may be unequal and this is not unfair. We think that people within a country such as France are responsible for their decision of moving from one region to another one in the sense that people have, at least parlially, control on this variable. Given the fair income depend only on effort variables, the differences in fair income across regions are explained by differences in reward schemes. Therefore, if people move to regions where their effort are more rewarded, according to the theory of equality of opportunty, inequalities in fair income across regions are fair. On the contrary, whatever the region, the effect of circumstances on income is unfair, thus it should be removed in all the regions. Moreover, the effect of circumstances may differ across regions, thus this analysis may help to design redistribution policies that would take these differences into account.

Defining a distribution of fair income per region has a cost: we have less observations per region and this may alter the significance of the estimates. In addition, we do not have a unique fair income distribution to display but 21 distributions.

Finally, the residual in this earning equation explains about half the income inequality. Obviously, this residual includes effort as well as circumstance variables and it is impossible to disentangle between both determinants. As other papers that use earnings equations, we include the residual into the circumstance set [4, 13]. We think this is also more appropriate in the context of our study because we study to which extent individuals who exert the same

effort obtain the same outcome with a direct measure of effort. Because we cannot capture any direct effort through the residual, it is included into the circumstance set. Then, we will not obtain a lower bound for inequality of opportunity. Fair inequalities will be due to differences in education, experience, sector of activity and working hours only.

2.3 Data

The data come from the 2005 EU-SILC (European Survey on Income and Life Conditions) dataset from France. This dataset has been designed at the European level and implemented by INSEE for France. We take the year 2005 because this module contains detailed information about indicators of effort, family background and income.

The dataset contains two parts, the first one contains information on households, the second one concerns individuals. For our purpose, we take the module of the survey that is addressed to individuals, since inequality of opportunity refers to inequality between individuals and not between groups of individuals. Moreover, it would be much more challenging to define the concept of effort and circumstance for households.

Concerning the sample, we select individuals who are between 25 and 65 years-old, who were wage earners, worked full-time during the 12 months in 2004 and did not change their jobs between 2004 and 2005. These restrictions are partly imposed by the dataset: questions about the sector of activity refer to the year 2005 but incomes refer to 2004. Therefore, we select people who did not change their profesional situation to keep the information about the sector of activity[6]. We restrict our sample to individuals between 25 and 65 years-old to focus on individuals who are more likely to be active and we avoid people being at the very begining and very end of their profesional career. Finally, because the income of self-employed is not clearly identifiable, we select wage earners only. We obtain in total 4279 observations. Details on the variables and on the sample are provided in Table 2.1 and in Figure 2.1.

We split our sample into 21 regions. The actual number of regions in France is 27 but the survey does not separate Corse and Provence-Alpes-Cotes-d'Azur (PACA) such that we have to proceed in the same way and no data is available for the 5 overseas regions.

[6]This constraint could bias our result if job mobility is correlated with effort but the sign of the correlation is not clear. On the one hand, individuals, the youngest in particular often progress by changing jobs. On the other hand, those who exert less effort are more likely to change jobs in order to benefit from uemployment fees and are more likely to be fired. Therefore, to the extent that the sign of the bias is not clearly identifiable, we maintain this choice.

The dependent variable is the gross labour income. We do not include capital income in order not to mix two dimensions of equality of opportunity and we take the gross earnings to measure by how much the labour market on its own generates ex-post inequality of opportunity. As a consequence, our analysis is restricted to ex-post inequality of opportunity among employed individuals before the state operates any redistribution. Thus it is not an analysis concerning the whole adult population.

In Table 2.9, we indicate the mean gross labour earnings given by the sample. We can notice that income inequality is rather low between regions except for Ile de France where the level of income is much higher than in all the other regions.

Table 2.1: Description of the variables

Name	Format	Values	Category
Gross labour income	Numerical	[3,055;271,962]	Dep. variable
Age	Numerical	[25;65]	Circumstance
Gender	Dummy	0: Woman 1: Man	Circumstance
Father's education	Categorical	1: lower than 1ary education 2: 1ary education 3: 2ary education 4: 3ary education	Circumstance
Financial problems during adolescence	Categorical	0: Never or few 1: More than few	Circumstance
Weekly working hours	Numerical	[11,97]	Effort
Years of experience	Numerical	[0,49]	Effort
Individual's education	Categorical	1: Uncomplete 1ary 2: Complete 1ary 3: Complete lower 2ary 4: Complete upper 2ary 5: Complete lower 3ary 6: Ms.C, Ph.D.	Effort
Sector of activity	Categorical	1: Agriculture 2: Energy and Industry 3: Construction 4: Services 5: Public sector	Effort

For the circumstance variables, we take father's education, financial problems during adolescence, gender and age. Parental education is often used in empirical analysis as a circumstance [14, 30, 43, 20]. Here, father's education takes 4 values according to the highest diploma obtained by the father (see Table 2.1). Financial problems during adolescence is a dummy that indicates whether or not the individual's family faced financial difficulties when he was between 12 and 16 years old. We include this variable in order to complement the information on family background as we consider this may measure the economic capital whereas father's education indicates human capital. Gender

is included as a circumstance because we control for working hours. So, even if women may tend to prefer more leisure [25] and so tend to work less hours, because we control for this, gender should not be a determinant of inequalities. We adopt the same view for age: we consider two individuals with the same working experience should not obtain a distinct income whatever their age, because we control for experience, age is included as a circumstance variable.

Effort is defined by a set of variables: individual's education, years of experience, sector of activity and weekly working hours (see Table 2.1 for complementary information). Individual's education is a categorical variable. It takes 6 values that depend on the highest level of education attained by the individual. This variable, although correlated with father's education is in general considered as an effort variable [4, 14] because the individual is at least partly responsible for it. Moreover, we clean this correlation by using the net measure of effort.

Years of experience corresponds to the years the individual spent in paid jobs. This is an indicator of generic professional capital as the sector of activity where the individuals obtained their experience is not specifically reported.

Sector of activity takes 5 values and corresponds to the sector in which the individual works. Considering this variable as being an effort variable is disputable: labour market restrictions may impose constraint on individual's choice concerning the sector of activity. But to the extent that individuals have at least a partial control on their sector of activity, we consider this is an effort variable.

Lastly, the weekly working hours correspond to the hours per week an individual worked in his job. In the survey, the individuals are not asked to declare the official weekly working hours but the number of effective weekly working hours. This variable should be an indicator of dedication at work. But in this way, we do not account for the fact that working hours, and specially part-time jobs can be imposed.

More descriptive statistics concerning the frequencies of the variable per region for categorical variables, the mean and variance for continous variables are reported in Tables 2.1 and 2.2. In a nutshell, all the variables cited above have been commonly used in the literature on equality of opportunity. Almas et al. [4] as well Bourguigion et al. [13] for instance also use indicators of working hours, level of education, gender and sector of activity (public vs private). For France, mainly Lefranc et al. [42, 43, 44] have worked on inequality of opportunity from the ex-ante perspective using stochatsic dominance tools. Therefore they use only one circumstance, the occupation of the father. Also, because they adopt a long term perspective, they use the Household Budget Survey that is available from the seventies whereas the

EU-SILC has been launched since 2004. This is a reason why our data are not directly comparable. The advantage of the EU-SILC survey is the fact that many indicators of effort are available and this is not the case for the Household Budget Survey. Moreover, Lefranc et al. have focused on the gross and disposable income, by including labour and asset income, and we focus on the gross labour income only.

Table 2.2: Summary statistics

Variable	Mean	Std. Dev.
Gross labour income	29,852	18,357
Working hours	39.7	8.1
Age	42.6	9.3
Years of experience	21.5	10.5
Years of experience squared	573.7	470.5
Observations	4279	

2.4 Results

2.4.1 Preliminary results

An overview of wage determinants in France

Wage equations show that the effect of circumstances is mostly significant. It is not our purpose to develop these results and we refer to the table 2.4 for more details. Nevertheless, we notice some interesting points[7].

Father's education has a significant impact on offspring's education when the offspring has achieved at least upper secondary school. In this case, more educated fathers grow more educated children. Also, education depends on financial problems: if the family faced financial problems during the adolescence, it increases the probablility of having a low level of education. Instead, having financial problems during adolescence has a positive effect on experience. An explanation is that the years of experience is negatively correlated with education, since individuals stop studying and enter the labour market. Concerning working hours, it depends significatively on father's education and on gender. As expected, men tend to work more hours than women do and more educated people tend to work more.

[7]We make a national regression to purge the effort variables from their correlation with circumstances.This may appear as inconsistent with our explanations to motivate a regional analysis. Nevertheless, we choose this option due to data constraints. Regional regression of effort variables often give very unprecise coefficient because the number of observations is limited in some regions. Thus, by doing a national regression, we get a national measure of net effort but this is a more precise measure.

Regarding the national regression, coefficients vary when taking the gross or net measure of effort. As expected with the results of the effort equations, when taking the net effort (the one that is cleaned from the correlation with cirumstances), the significance and the magnitude of the parameters of circumstance and effort respectively increase and decrease. Actually, we provide the wage regression at a national level in order to check for the significance of the parameters when we have a large number of observations (4279). Table 2.3 presents the results.

We obtain that the explanatory variables explain around 41% of income inequality. And if we decompose the R squared, we obtain that 34% of income inequality is explained by the effort variables and 7% by circumstances. Our indicators of effort provide a good explanation of income inequality. Also, almost all the coefficients are significant and of expected sign. Only the fact of having accomplished primary school or not having any diploma does not make any difference and financial problem becomes significant only when using the net measure of effort. When turning to the rest of the variables, we observe that father's education and age have a positive impact on income as well as being a men and working in a sector that is not agriculture. Also education is highly significant and has a increasing impact on income as the level of education increases. Experience does not present non linearity but still has a positive impact on earnings as well as the number of working hours. This national regression does not present any unexpected results, on the contrary, it is consistent with classical results in the literature and confirms the hypothesis of the presence of inequality of opportunity in France. At a national level, inequality of opportunity is mainly due to the impact of gender and father's education and, to a lower extent, to age and financial problems during teenhood.

Table 2.3: Wage equation - national level

Variables	**Gross effort**		**Net effort**	
Father's education: 1ary	0.083***	[3.187]	0.132***	[5.066]
Father's education: 2ary	0.131***	[4.796]	0.242***	[8.646]
Father's education: 3ary	0.198***	[6.226]	0.464***	[14.569]
Financial problems during adolescence	0.001	[0.109]	-0.034***	[2.982]
Gender	0.165***	[13.747]	0.195***	[17.525]
Age	0.008***	[5.294]	0.006***	[4.224]
Industry	0.258***	[5.150]	0.126***	[4.580]
Construction	0.158***	[3.000]	0.059**	[2.359]
Services	0.233***	[4.680]	0.114***	[4.017]
Public sector	0.124**	[2.481]	0.049*	[1.704]
Education: 1ary	0.003	[0.092]	0.006	[0.399]
Education: low 2ary	0.169***	[6.019]	0.088***	[6.442]
Education: up 2ary	0.197***	[9.491]	0.128***	[10.145]
Education: low 3ary	0.566***	[25.141]	0.339***	[25.594]
Education: Ph.D	0.867***	[23.622]	0.394***	[23.628]
Experience	0.028***	[11.137]	0.007***	[5.197]
Experience squared	-0.000***	[9.709]	-0.000***	[6.169]
Weekly working hours	0.011***	[16.140]	0.011***	[15.854]
Constant	8.381***	[110.046]	9.665***	[136.676]
Observations	4279		4279	
R-squared	0.415		0.408	

Notes. Absolute value of t statistics in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%

Figure 2.1: Data frequencies

Region	Education						Father's education				Gender		Sector					Fin. prob		Tot.
	No	1ary	Low 2ary	Up 2ary	Low 3ary	PhD	No	1ary	2ary	3ary	Women	Men	Agr	Ind	Cons	Serv	Publ	No	Yes	
Ile de France	82	49	78	309	386	54	75	402	304	177	445	513	2	137	58	385	376	520	438	958
	8.6%	5.1%	8.1%	32.3%	40.3%	5.6%	7.8%	42.0%	31.7%	18.5%	46.5%	53.5%	0.2%	14.3%	6.1%	40.2%	39.2%	54.3%	45.7%	100%
Champ-Ardenne	9	6	11	48	25	2	2	59	34	6	39	62	1	33	8	24	35	57	44	101
	8.9%	5.9%	10.9%	47.5%	24.8%	2.0%	2.0%	58.4%	33.7%	5.9%	38.6%	61.4%	1.0%	32.7%	7.9%	23.8%	34.7%	56.4%	43.6%	100%
Picardie	26	8	12	71	17	1	3	104	24	4	49	86	3	38	19	41	34	63	72	135
	19.3%	5.9%	8.9%	52.6%	12.6%	0.7%	2.2%	77.0%	17.8%	3.0%	36.3%	63.7%	2.2%	28.1%	14.1%	30.4%	25.2%	46.7%	53.3%	100%
H.-Normandie	9	3	4	36	37	3	2	60	23	7	31	61	2	21	4	27	38	51	41	92
	9.8%	3.3%	4.3%	39.1%	40.2%	3.3%	2.2%	65.2%	25.0%	7.6%	33.7%	66.3%	2.2%	22.8%	4.3%	29.3%	41.3%	55.4%	44.6%	100%
Centre	19	25	13	78	51	2	7	131	43	7	72	116	3	45	10	45	85	101	87	188
	10.1%	13.3%	6.9%	41.5%	27.1%	1.1%	3.7%	69.7%	22.9%	3.7%	38.3%	61.7%	1.6%	23.9%	5.3%	23.9%	45.2%	53.7%	46.3%	100%
B.-Normandie	10	13	6	42	21	1	0	68	23	2	42	51	3	35	4	14	37	52	41	93
	10.8%	14.0%	6.5%	45.2%	22.6%	1.1%	0.0%	73.1%	24.7%	2.2%	45.2%	54.8%	3.2%	37.6%	4.3%	15.1%	39.8%	55.9%	44.1%	100%
Bourgogne	11	4	5	55	38	3	5	61	44	6	47	69	3	31	10	33	39	68	48	116
	9.5%	3.4%	4.3%	47.4%	32.8%	2.6%	4.3%	52.6%	37.9%	5.2%	40.5%	59.5%	2.6%	26.7%	8.6%	28.4%	33.6%	58.6%	41.4%	100%
Nord-P-d-Calais	24	18	24	141	92	12	9	199	84	19	107	204	3	87	19	89	113	173	138	311
	7.7%	5.8%	7.7%	45.3%	29.6%	3.9%	2.9%	64.0%	27.0%	6.1%	34.4%	65.6%	1.0%	28.0%	6.1%	28.6%	36.3%	55.6%	44.4%	100%
Lorraine	15	11	7	76	40	4	12	83	42	16	42	111	2	40	13	41	57	85	68	153
	9.8%	7.2%	4.6%	49.7%	26.1%	2.6%	7.8%	54.2%	27.5%	10.5%	27.5%	72.5%	1.3%	26.1%	8.5%	26.8%	37.3%	55.6%	44.4%	100%
Alsace	16	4	4	76	46	2	7	56	75	10	51	97	0	56	18	36	38	90	58	148
	10.8%	2.7%	2.7%	51.4%	31.1%	1.4%	4.7%	37.8%	50.7%	6.8%	34.5%	65.5%	0.0%	37.8%	12.2%	24.3%	25.7%	60.8%	39.2%	100%
Franche-Comté	15	8	3	59	28	0	4	80	26	3	42	71	2	47	8	23	33	53	60	113
	13.3%	7.1%	2.7%	52.2%	24.8%	0.0%	3.5%	70.8%	23.0%	2.7%	37.2%	62.8%	1.8%	41.6%	7.1%	20.4%	29.2%	46.9%	53.1%	100%
Pays de la Loire	19	17	20	141	81	7	6	184	77	18	114	171	5	78	32	75	95	152	133	285
	6.7%	6.0%	7.0%	49.5%	28.4%	2.5%	2.1%	64.6%	27.0%	6.3%	40.0%	60.0%	1.8%	27.4%	11.2%	26.3%	33.3%	53.3%	46.7%	100%
Bretagne	2	5	10	92	68	2	2	89	77	11	73	106	1	45	11	48	74	103	76	179
	1.1%	2.8%	5.6%	51.4%	38.0%	1.1%	1.1%	49.7%	43.0%	6.1%	40.8%	59.2%	0.6%	25.1%	6.1%	26.8%	41.3%	57.5%	42.5%	100%
Poit-Charentes	12	3	13	75	43	2	0	97	45	6	65	83	1	22	17	37	71	84	64	148
	8.1%	2.0%	8.8%	50.7%	29.1%	1.4%	0.0%	65.5%	30.4%	4.1%	43.9%	56.1%	0.7%	14.9%	11.5%	25.0%	48.0%	56.8%	43.2%	100%
Aquitaine	7	6	15	99	56	3	14	90	65	17	72	114	6	39	8	39	94	109	77	186
	3.8%	3.2%	8.1%	53.2%	30.1%	1.6%	7.5%	48.4%	34.9%	9.1%	38.7%	61.3%	3.2%	21.0%	4.3%	21.0%	50.5%	58.6%	41.4%	100%
Midi-Pyrénées	4	9	15	65	56	8	11	75	57	14	67	90	2	25	9	40	81	90	67	157
	2.5%	5.7%	9.6%	41.4%	35.7%	5.1%	7.0%	47.8%	36.3%	8.9%	42.7%	57.3%	1.3%	15.9%	5.7%	25.5%	51.6%	57.3%	42.7%	100%
Limousin	5	2	6	26	16	1	3	34	14	5	27	29	2	11	3	12	28	27	29	56
	8.9%	3.6%	10.7%	46.4%	28.6%	1.8%	5.4%	60.7%	25.0%	8.9%	48.2%	51.8%	3.6%	19.6%	5.4%	21.4%	50.0%	48.2%	51.8%	100%
Rhône-Alpes	42	11	20	177	105	20	18	212	108	37	145	230	0	110	27	94	144	202	173	375
	11.2%	2.9%	5.3%	47.2%	28.0%	5.3%	4.8%	56.5%	28.8%	9.9%	38.7%	61.3%	0.0%	29.3%	7.2%	25.1%	38.4%	53.9%	46.1%	100%
Auvergne	3	5	4	44	26	6	1	54	30	3	32	56	4	17	9	27	31	38	50	88
	3.4%	5.7%	4.5%	50.0%	29.5%	6.8%	1.1%	61.4%	34.1%	3.4%	36.4%	63.6%	4.5%	19.3%	10.2%	30.7%	35.2%	43.2%	56.8%	100%
Lang-Roussillon	9	8	6	56	31	7	7	62	36	12	53	64	3	15	13	28	58	69	48	117
	7.7%	6.8%	5.1%	47.9%	26.5%	6.0%	6.0%	53.0%	30.8%	10.3%	45.3%	54.7%	2.6%	12.8%	11.1%	23.9%	49.6%	59.0%	41.0%	100%
PACA Corse	24	11	20	117	99	9	20	141	84	35	122	158	5	38	16	93	128	158	122	280
	8.6%	3.9%	7.1%	41.8%	35.4%	3.2%	7.1%	50.4%	30.0%	12.5%	43.6%	56.4%	1.8%	13.6%	5.7%	33.2%	45.7%	56.4%	43.6%	100%
France	363	226	296	1,883	1,362	149	208	2,341	1,315	415	1,737	2,542	53	970	316	1,251	1,689	2,345	1,934	4,279
	8.5%	5.3%	6.9%	44.0%	31.8%	3.5%	4.9%	54.7%	30.7%	9.7%	40.6%	59.4%	1.2%	22.7%	7.4%	29.2%	39.5%	54.8%	45.2%	100%

Table 2.4: Regression of effort on circumstances

Circ	Probit	Probit	Probit	Probit	Probit	Probit	Probit	Probit	Probit	OLS	OLS
	Ind	Constr	Serv	Public	Edu1ary	Low2ary	Up2ary	Low3ary	PhD	Exprc	Wrk hrs
Fathed1ary	0.088	-0.196	-0.109	0.125	0.402***	0.266*	0.44***	0.042	-0.221	-0.411	1.063*
	[0.869]	[1.632]	[1.135]	[1.267]	[3.410]	[1.784]	[4.716]	[0.395]	[1.091]	[0.563]	[1.848]
Fathed2ary	-0.095	-0.36***	-0.013	0.261**	-0.031	0.104	0.26***	0.50***	0.243	-5.50***	1.78***
	[0.893]	[2.755]	[0.132]	[2.553]	[0.219]	[0.665]	[2.601]	[4.524]	[1.196]	[7.268]	[2.962]
Fathed3ary	-0.21*	-0.61***	-0.033	0.436***		-0.133	-0.44***	1.00***	1.00***	-7.87***	4.31***
	[1.660]	[3.594]	[0.290]	[3.768]		[0.698]	[3.791]	[8.184]	[4.806]	[9.082]	[6.275]
Fin prob	-0.006	0.111*	-0.081*	0.023	0.187***	0.098	0.10***	-0.25***	-0.081	1.42***	0.131
	[0.136]	[1.808]	[1.942]	[0.547]	[2.642]	[1.634]	[2.603]	[5.871]	[0.972]	[4.503]	[0.527]
Gender	0.51***	0.875***	0.120***	-0.77***	0.079	-0.083	0.19***	-0.26***	-0.021	1.15***	3.22***
	[11.077]	[10.948]	[2.873]	[18.879]	[1.098]	[1.391]	[4.864]	[6.244]	[0.268]	[3.675]	[13.012]
Age	-0.001*	-0.004	-0.01***	0.02***	0.06***	0.01***	-0.01***	-0.02***	0.01***		0.04***
	[1.732]	[1.287]	[5.073]	[7.197]	[12.196]	[4.294]	[2.904]	[6.686]	[3.124]		[3.192]
Constant	-0.90***	-1.70***	-0.033	-0.72***	-4.76***	-2.29***	-0.33**	0.135	-2.5***	22.9***	34.4***
	[5.863]	[8.255]	[0.233]	[5.007]	[17.279]	[10.354]	[2.401]	[0.909]	[8.894]	[29.913]	[40.252]
Obs.	4279	4279	4279	4279	4279	4279	4279	4279	4279	4279	4279
R-squared										0.089	0.05

Notes. Absolute value of t statistics in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%

A regional analysis is relevant

Our point to compare regions is related to the possible heterogeneity characterizing the French regions. A first satisfactory result is that regional dummies added to the wage equation are strongly significant (see Equation 2.1). Moreover, we also find that these regional differences are not due to different sectorial endowments of the regions because interaction terms between sectors and regions do not remove the significance of region specific dummies[8]. This confirms a first intuition according to which there is heterogeneity in income inequality across regions in France.

When turning to the regional equation results displayed in the Tables 2.5, 2.6, 2.7 and 2.8, we first observe that coefficients have lower significativity compared to the national regression. A possible explanation is the sample size, smaller for the regions.

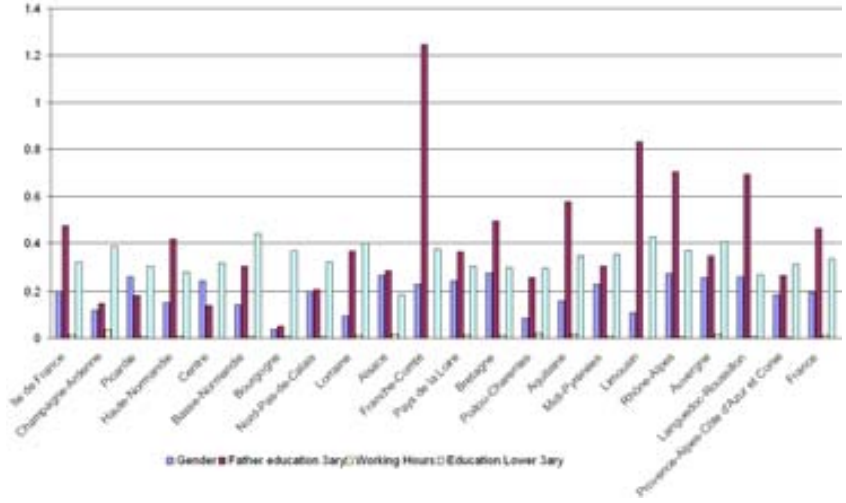
Regarding the circumstances, we show that they do not affect income inequality equally across regions. Firstly, despite of the small sample size, the variables "age", "gender" and "having a father who went to the university" are circumstances that are almost always significant and have a positive impact on earnings. But their impact is very different from one region to another. Indeed, when taking into account only the significant coefficients, the coefficient "gender" varies between 0.094 for Lorraine and 0.285 for Auvergne. The effect of gender in Auvergne is three times as much as in Lorraine once we controlled for all the other sources of income inequality. The Figure 2.2 shows the variation in the coefficients across regions. The same difference is observed for the effect of father's education. By comparing only significant coefficients, the effect of having a father who went to the university with respect to a father with no diploma is the smallest in PACA-Corse with a coefficient equal to 0.366 and the effect is the largest for Limousin with a coefficinet equal to 0.832. The Tables 2.5, 2.6, 2.7 and 2.8 rank the regions in increasing order of income level (the first table contains the poorest regions), and the smallest and largest effects have been found for the half poorest regions, therefore we cannot conclude that income level might explain the variation in the impact of unfair sources of inequalities. Finally, financial problem during teenhood has a very small impact on earnings, even in Ile de France. The explanation cannot be the small sample size. As a conclusion, the transmission of advantages seem to be better captured by the human capital rather than the economic capital of the family.

When turning to the effect of effort variables, we first observe that regions do not reward equally effort. Working in a sector or another does not appear to have a significant impact on earnings[9] as for half the regions, not even one

[8]Results are not reported but available upon request.

[9]We reproduce the results by changing the reference sector and obtain similar results,

Figure 2.2: Impact of gender, father's education, education and working hours on (log) earnings across regions



coefficient is significantly different from zero. On the contrary, the other effort variables are often significant, especially "working hours", "experience" and "education" (precisely the dummies that indicate that the individual achieved at least upper secondary education). This shows how much our indicators of effort are powerful in explaining income inequality and this justifies our direct measure of effort. In addition, we also obtain a great variation in the reward schemes across regions. The Figure 2.2 displays the effect of education and working hours on the earnings. The impact of having lower terciary education varies between 0.073 in Nord-Pas-de-Calais and 0.314 in Limousin. The same large difference is observed for "working hours" whose impact goes from 0.07 in Nord-Pas-de-Calais to 0.036 in Champagne-Ardenne.

As a conclusion for these first results, we obtain that circumstances have in France a direct and also an indirect effect on effort. Even though the sources of unfair inequalities are in general the same across regions, the impact of each source is very distinct from one region to another one. This would justify a regional approach for the design of redistribution policies. Concerning fair inequalities, the reward schemes are not the same across regions but this is not a concern in this analysis. In order to analyse to which extent regions differ in terms of ex-post inequality of opportunity, we turn to the comparison of the unfair Gini.

---

they are available upon request.

Table 2.5: Wage equation - regions 1/4

	Langued-Rouss		**Nrd-P-d-Calais**		**Picardie**		**Limousin**		**Lorraine**		**Basse-Norm.**	
	Gross	Net	Gross	Net	Gross	Net	Gross	Net	Gross	Net	Gross	Net
Fathed1ary	0.263	0.324	-0.052	0.001	-0.284	-0.234	-0.059	0.092	0.158*	0.217**	-0.122	-0.183**
	[1.237]	[1.551]	[0.490]	[0.006]	[1.364]	[1.126]	[0.301]	[0.500]	[1.918]	[2.607]	[1.391]	[2.020]
Fathed2ary	0.285	0.368*	-0.02	0.11	-0.194	-0.126	0.061	0.367*	0.129	0.243**		
	[1.279]	[1.743]	[0.182]	[0.998]	[0.888]	[0.561]	[0.301]	[1.763]	[1.390]	[2.287]		
Fathed3ary	0.546**	0.693***	-0.123	0.208	-0.012	0.182	0.383	0.832***	0.142	0.366***	0.085	0.308
	[2.158]	[3.000]	[0.967]	[1.619]	[0.044]	[0.663]	[1.609]	[3.352]	[1.303]	[3.161]	[0.330]	[1.161]
Fin prb ad	0.001	-0.032	-0.009	-0.044	0.001	-0.015	0.016	-0.05	0.065	0.03	-0.056	-0.091
	[0.008]	[0.352]	[0.253]	[1.227]	[0.023]	[0.234]	[0.202]	[0.592]	[1.547]	[0.690]	[0.756]	[1.185]
Gender	0.227**	0.262***	0.176***	0.200***	0.234***	0.262***	0.055	0.108	0.058	0.094*	0.161**	0.141*
	[2.552]	[3.057]	[4.217]	[5.404]	[3.290]	[4.048]	[0.712]	[1.406]	[1.077]	[1.948]	[1.998]	[1.905]
Age	0.002	0.003	0.007	0.006	-0.001	-0.002	0.031*	0.038**	0.01	0.005	0.013*	0.009
	[0.251]	[0.393]	[1.508]	[1.295]	[0.092]	[0.243]	[1.986]	[2.305]	[1.127]	[0.613]	[1.677]	[1.063]
Industry	0.548*	0.231	0.141	0.022	0.29	0.15	0.370*	0.227*	0.118	0.12	0.097	0.037
	[1.834]	[1.310]	[0.789]	[0.226]	[1.419]	[1.280]	[1.759]	[1.841]	[0.646]	[1.192]	[0.444]	[0.288]
Constr	0.107	0.009	0.065	-0.019	0.065	0.023	0.787**	0.384**	-0.093	-0.001	-0.076	-0.059
	[0.349]	[0.055]	[0.347]	[0.215]	[0.303]	[0.217]	[2.533]	[2.423]	[0.491]	[0.015]	[0.291]	[0.428]
Services	0.16	0.018	0.144	0.026	0.355*	0.198	0.316	0.182	0.108	0.115	0.213	0.098
	[0.548]	[0.099]	[0.796]	[0.252]	[1.698]	[1.613]	[1.483]	[1.404]	[0.593]	[1.124]	[0.950]	[0.718]
Public	0.205	0.049	0.084	-0.011	0.187	0.098	0.277	0.182	-0.064	0.013	0.037	-0.005
	[0.700]	[0.261]	[0.461]	[0.110]	[0.875]	[0.767]	[1.343]	[1.432]	[0.356]	[0.127]	[0.168]	[0.033]
Edu1ary	-0.036	0.031	-0.026	-0.024	0.159	0.112	0.371	0.165	-0.045	-0.035	-0.164	-0.09
	[0.150]	[0.248]	[0.255]	[0.484]	[1.117]	[1.628]	[1.523]	[1.396]	[0.407]	[0.700]	[1.082]	[1.141]
Lower2ary	0.089	0.084	0.163*	0.075*	0.164	0.091	0.501**	0.240**	0.299**	0.134**	0.163	0.088
	[0.356]	[0.638]	[1.805]	[1.765]	[1.338]	[1.539]	[2.502]	[2.449]	[2.527]	[2.364]	[0.917]	[0.961]
Upper2ary	0.178	0.16	0.123*	0.073*	0.192**	0.132**	0.534***	0.314***	0.291***	0.171***	0.286**	0.185**
	[1.005]	[1.445]	[1.765]	[1.719]	[2.324]	[2.602]	[3.099]	[2.840]	[4.011]	[3.855]	[2.309]	[2.324]
Lower3ary	0.393**	0.269**	0.556***	0.324***	0.479***	0.306***	0.772***	0.425***	0.676***	0.400***	0.723***	0.441***
	[2.017]	[2.343]	[7.139]	[7.184]	[3.609]	[3.877]	[4.013]	[3.667]	[7.815]	[7.716]	[5.283]	[5.228]
PhD	0.799***	0.424***	0.911***	0.401***	1.202***	0.449***	1.991***	0.707***	0.462***	0.199***	1.197***	0.506***
	[3.151]	[3.551]	[7.600]	[7.832]	[3.396]	[3.445]	[6.048]	[5.636]	[2.904]	[2.781]	[3.257]	[3.119]
Experience	0.051***	-0.002	0.011	0.017***	0.032**	0.012	0.037	0.015**	0.005	0.008	0.041**	0.009
	[3.105]	[0.540]	[1.282]	[6.112]	[2.140]	[2.013]	[1.659]	[1.175]	[0.500]	[0.971]	[2.498]	[1.306]
Experience2	-0.001***	-0.001*	0	0	0	0	-0.001***	-0.001**	0	0	-0.001**	0
	[2.898]	[1.686]	[0.502]	[0.419]	[1.630]	[1.615]	[3.008]	[2.029]	[0.171]	[0.116]	[2.103]	[1.065]
Wrkng hours	-0.001	0.006	0.016***	0.007*	0.007*	0.007**	0.012	-0.017	0.012***	0.012***	0.007**	0.007*
	[0.145]	[0.710]	[5.965]	[1.684]	[1.968]	[1.382]	[1.657]	[2.036]	[3.123]	[3.126]	[2.014]	[1.864]
Constant	8.605***	9.520***	8.596***	9.755***	9.034***	10.333***	7.207***	8.368***	8.561***	9.678***	8.372***	9.897***
	[20.560]	[22.699]	[31.455]	[40.292]	[21.024]	[20.934]	[12.030]	[11.291]	[24.553]	[23.753]	[20.029]	[29.396]
Observ.	117	117	311	311	135	135	56	56	153	153	93	93
R-squared	0.409	0.374	0.525	0.529	0.411	0.42	0.741	0.724	0.544	0.542	0.574	0.552

Notes. Absolute value of t statistics in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. Gross and Net refer to gross and net versions of effort.

Table 2.6: Wage equation - regions 2/4

	Auvergne		**Poit-Charente**		**PACAC**		**Franche-Comté**		**Bourgogne**	
	Gross	Net	Gross	Net	Gross	Net	Gross	Net	Gross	Net
Fathed1ary	0.209	0.164	-0.07	-0.170**	0.009	0.052	0.111	0.063	-0.293	-0.228
	[0.446]	[0.354]	[1.091]	[2.455]	[0.110]	[0.645]	[0.640]	[0.321]	[1.522]	[1.281]
Fathed2ary	0.121	0.097			0.056	0.164*	0.112	0.205	-0.244	-0.068
	[0.255]	[0.199]			[0.654]	[1.848]	[0.620]	[0.851]	[1.219]	[0.365]
Fathed3ary	0.27	0.347	0.096	0.259	-0.003	0.268***	0.353	1.248	-0.255	0.05
	[0.503]	[0.631]	[0.567]	[1.649]	[0.035]	[2.725]	[1.361]	[1.003]	[1.017]	[0.211]
Fin prb ado	0.091	0.049	-0.004	-0.053	-0.019	-0.06	0.001	-0.032	0.120*	0.076
	[0.927]	[0.501]	[0.071]	[0.947]	[0.456]	[1.456]	[0.017]	[0.421]	[1.663]	[1.111]
Gender	0.285***	0.257***	0.018	0.085	0.183***	0.184***	0.240***	0.229***	0.039	0.038
	[2.762]	[2.653]	[0.294]	[1.523]	[4.230]	[4.640]	[3.387]	[3.214]	[0.530]	[0.569]
Age	-0.006	-0.006	0.018**	0.019**	0.008*	0.008*	-0.01	-0.009	0.011	0.012
	[0.356]	[0.351]	[2.420]	[2.525]	[1.692]	[1.710]	[1.125]	[0.731]	[0.933]	[0.997]
Industry	0.479*	0.267*	0.393	0.241	0.428***	0.217**	-0.906***	-0.418**	0.205	0.078
	[1.881]	[1.810]	[1.184]	[1.340]	[2.709]	[2.424]	[3.155]	[2.601]	[0.867]	[0.611]
Constr	0.138	0.071	0.377	0.206	0.12	0.024	-0.927***	-0.386**	-0.015	-0.034
	[0.534]	[0.547]	[1.131]	[1.323]	[0.709]	[0.293]	[3.102]	[2.515]	[0.057]	[0.278]
Services	0.424*	0.248*	0.385	0.239	0.22	0.093	-0.834***	-0.386**	0.073	0.007
	[1.836]	[1.819]	[1.177]	[1.333]	[1.450]	[1.050]	[2.945]	[2.353]	[0.322]	[0.055]
Public	0.381	0.22	0.322	0.213	0.171	0.066	-0.957***	-0.464***	0.007	-0.032
	[1.625]	[1.541]	[0.981]	[1.150]	[1.135]	[0.742]	[3.377]	[2.800]	[0.030]	[0.236]
Edu1ary	-0.028	0.045	0.102	0.08	0.121	0.05	0.111	0.08	0.411*	0.199*
	[0.085]	[0.288]	[0.477]	[0.687]	[0.950]	[0.761]	[0.756]	[0.995]	[1.821]	[1.711]
Lower2ary	0.47	0.285*	0.227	0.143**	0.141	0.069	0.087	0.064	0.209	0.106
	[1.433]	[1.781]	[1.639]	[2.110]	[1.380]	[1.410]	[0.418]	[0.618]	[1.038]	[1.091]
Upper2ary	0.228	0.208	0.197*	0.150**	0.185**	0.114**	0.196**	0.139**	0.227*	0.147*
	[0.850]	[1.233]	[1.886]	[2.259]	[2.442]	[2.503]	[2.054]	[2.296]	[1.769]	[1.935]
Lower3ary	0.573**	0.406**	0.469***	0.299***	0.537***	0.316***	0.593***	0.376***	0.597***	0.367***
	[2.120]	[2.495]	[4.106]	[4.302]	[6.729]	[6.761]	[4.978]	[5.179]	[4.464]	[4.666]
PhD	1.114***	0.494***	0.578*	0.413**	1.056***	0.474***			0.803***	0.323***
	[3.522]	[3.774]	[1.906]	[2.309]	[7.826]	[7.733]			[3.170]	[3.133]
Experience	0.021	0.006	0.038***	-0.005	0.028***	0.004*	0.036**	0.023***	0.033	-0.006
	[0.823]	[1.335]	[2.851]	[4.729]	[3.193]	[0.792]	[2.471]	[0.045]	[1.591]	[1.772]
Experience2	0	0	-0.001***	-0.001**	-0.001***	-0.001***	0	0	-0.001**	-0.001***
	[0.154]	[0.128]	[3.594]	[1.994]	[2.998]	[3.120]	[1.068]	[0.863]	[2.332]	[3.045]
Wrkng hours	0.007	0.016	0.023***	0.023***	0.004*	0.003	-0.001	0	0.010*	0.009*
	[1.510]	[1.063]	[4.976]	[0.706]	[1.775]	[1.837]	[0.386]	[2.637]	[1.840]	[0.535]
Constant	8.487***	9.972***	7.523***	9.386***	8.728***	9.713***	10.327***	10.451***	8.785***	9.788***
	[10.178]	[11.087]	[15.844]	[29.669]	[35.543]	[42.229]	[20.876]	[23.183]	[18.646]	[17.527]
Observ.	88	88	148	148	280	280	113	113	116	116
R-squared	0.499	0.5	0.482	0.452	0.457	0.458	0.403	0.389	0.408	0.426

Notes. Absolute value of t statistics in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. Gross and Net refer to gross and net versions of effort.

Table 2.7: Wage equation - regions 3/4

	Bretagne		**Midi-Pyrenees**		**Centre**		**Aquitaine**		**Haute-Normandie**	
	Gross	Net	Gross	Net	Gross	Net	Gross	Net	Gross	Net
Fathed1ary	-0.027	0.062	0.136	0.155	0.119	0.133	0.201**	0.262**	-0.063	-0.017
	[0.109]	[0.288]	[1.076]	[1.178]	[0.950]	[1.108]	[1.987]	[2.502]	[0.230]	[0.061]
Fathed2ary	0.051	0.197	0.043	0.149	0.135	0.203	0.179*	0.318***	0.004	0.115
	[0.206]	[0.907]	[0.318]	[1.033]	[0.986]	[1.540]	[1.729]	[2.912]	[0.014]	[0.386]
Fathed3ary	0.225	0.493**	-0.012	0.307*	-0.042	0.139	0.331**	0.577***	0.132	0.418
	[0.858]	[2.098]	[0.073]	[1.721]	[0.237]	[0.819]	[2.552]	[4.392]	[0.424]	[1.340]
Fin prb ado	-0.03	-0.054	0.103	0.062	-0.005	-0.035	-0.028	-0.065	-0.033	-0.086
	[0.644]	[1.170]	[1.496]	[0.907]	[0.109]	[0.736]	[0.542]	[1.270]	[0.367]	[0.958]
Gender	0.269***	0.278***	0.228***	0.231***	0.267***	0.243***	0.094*	0.160***	0.131	0.150*
	[5.624]	[5.925]	[2.950]	[3.576]	[5.210]	[5.132]	[1.680]	[3.035]	[1.347]	[1.669]
Age	0.010*	0.011*	0.005	0.004	-0.002	-0.003	0.021***	0.018***	0.023**	0.016
	[1.735]	[1.777]	[0.807]	[0.676]	[0.368]	[0.556]	[2.976]	[2.611]	[2.255]	[1.538]
Industry	-0.159	-0.114	0.087	0.058	0.054	0.031	0.236	0.119	0.19	0.04
	[0.543]	[0.725]	[0.313]	[0.359]	[0.277]	[0.282]	[1.583]	[1.378]	[0.647]	[0.238]
Constr	-0.306	-0.177	0.195	0.102	-0.07	-0.03	0.165	0.059	0.054	-0.026
	[1.007]	[1.183]	[0.651]	[0.658]	[0.323]	[0.293]	[0.902]	[0.641]	[0.156]	[0.150]
Services	-0.346	-0.23	0.153	0.086	0.107	0.066	0.029	-0.001	0.213	0.073
	[1.184]	[1.432]	[0.552]	[0.532]	[0.548]	[0.593]	[0.193]	[0.010]	[0.701]	[0.413]
Public	-0.224	-0.156	-0.004	-0.007	0.026	0.017	-0.008	-0.025	0.047	-0.026
	[0.770]	[0.958]	[0.014]	[0.047]	[0.131]	[0.145]	[0.056]	[0.286]	[0.162]	[0.150]
Edu1ary	0.148	0.116	0.007	-0.046	0.094	0.05	0.039	0.039	-0.573**	-0.295**
	[0.535]	[0.780]	[0.029]	[0.380]	[0.871]	[0.977]	[0.203]	[0.416]	[1.999]	[2.010]
Lower2ary	0.199	0.136	0.309	0.09	0.089	0.052	0.11	0.065	0.074	0.04
	[0.781]	[1.141]	[1.444]	[0.901]	[0.739]	[0.908]	[0.707]	[0.885]	[0.309]	[0.328]
Upper2ary	0.117	0.119	0.23	0.088	0.177**	0.117**	0.238*	0.160*	0.188	0.119
	[0.490]	[0.898]	[1.194]	[0.766]	[2.039]	[2.243]	[1.788]	[1.963]	[1.281]	[1.291]
Lower3ary	0.430*	0.300**	0.694***	0.352***	0.545***	0.322***	0.558***	0.347***	0.473***	0.282***
	[1.809]	[2.370]	[3.537]	[3.150]	[5.688]	[5.835]	[3.999]	[4.200]	[2.964]	[2.983]
PhD	0.779**	0.359***	1.457***	0.664***	1.305***	0.600***	0.337	0.143	0.676**	0.272**
	[2.496]	[2.851]	[6.107]	[6.063]	[5.344]	[5.420]	[1.372]	[1.349]	[2.425]	[2.225]
Experience	0.019**	0.003	0.030*	0.007	0.039***	0.009*	0.007	-0.001	0.009	0.002
	[1.976]	[0.608]	[1.871]	[1.955]	[3.334]	[0.063]	[0.454]	[4.100]	[0.518]	[1.601]
Experience2	-0.000*	0	-0.001	0	-0.001***	-0.001***	0	0	0	0
	[1.667]	[1.331]	[1.622]	[1.345]	[3.034]	[2.959]	[0.642]	[0.203]	[0.521]	[0.528]
Wrkng hrs	0.012***	0.012***	0.007	0.009*	0	0	0.016***	0.016***	0.009	0.009
	[3.803]	[3.752]	[1.592]	[1.073]	[0.015]	[1.683]	[4.003]	[0.223]	[1.632]	[0.255]
Constant	8.856***	9.406***	8.540***	9.688***	9.122***	10.039***	7.966***	8.975***	8.311***	9.340***
	[20.523]	[27.761]	[19.103]	[28.767]	[29.541]	[34.682]	[24.852]	[28.086]	[15.936]	[17.203]
Observ.	179	179	157	157	188	188	186	186	92	92
R-squared	0.474	0.478	0.495	0.495	0.44	0.45	0.477	0.477	0.573	0.568

Notes. Absolute value of t statistics in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. Gross and Net refer to gross and net versions of effort.

Table 2.8: Wage equation - regions 4/4

	Pays Loire		**Champ-Ardenne**		**Alsace**		**Rhone-Alpes**		**Ile de Fr.**	
	Gross	Net	Gross	Net	Gross	Net	Gross	Net	Gross	Net
Fathed1ary	0.019	0.082	-0.11	-0.053	0.230*	0.264**	0.365***	0.438***	0.150***	0.200***
	[0.148]	[0.641]	[0.401]	[0.181]	[1.863]	[2.173]	[3.889]	[4.696]	[3.015]	[4.007]
Fathed2ary	0.048	0.230*	0	-0.002	0.282**	0.368***	0.402***	0.542***	0.169***	0.277***
	[0.370]	[1.734]	[0.000]	[0.006]	[2.255]	[2.731]	[4.031]	[5.411]	[3.258]	[5.092]
Fathed3ary	0.023	0.363**	0.083	0.147	0.077	0.288*	0.390***	0.705***	0.232***	0.473***
	[0.156]	[2.465]	[0.254]	[0.422]	[0.505]	[1.839]	[3.393]	[6.171]	[4.028]	[7.965]
Fin prb ado	-0.018	-0.057	0.036	0.014	-0.031	-0.057	0.007	-0.022	-0.019	-0.053**
	[0.472]	[1.489]	[0.389]	[0.147]	[0.561]	[1.044]	[0.167]	[0.521]	[0.726]	[2.014]
Gender	0.210***	0.244***	0.037	0.118	0.195***	0.266***	0.239***	0.276***	0.131***	0.198***
	[5.024]	[6.486]	[0.436]	[1.383]	[3.531]	[4.906]	[5.472]	[6.762]	[4.953]	[7.647]
Age	0.019***	0.017***	-0.025**	-0.026**	0.013	0.013	0.005	0.004	0.011***	0.009**
	[3.374]	[3.082]	[2.531]	[2.430]	[1.578]	[1.546]	[0.991]	[0.777]	[3.331]	[2.580]
Industry	0.284**	0.150*	0.049	-0.025	0.07	-0.123	0.008	0.087	0.329	0.181
	[2.002]	[1.900]	[0.126]	[0.117]	[1.069]	[0.467]	[0.097]	[0.308]	[1.203]	[1.569]
Constr	0.143	0.059	0.002	-0.055	-0.064	-0.171		0.064	0.25	0.121
	[0.969]	[0.833]	[0.006]	[0.308]	[0.713]	[0.769]		[0.263]	[0.907]	[1.181]
Services	0.207	0.114	-0.018	-0.081		-0.177	0.015	0.086	0.265	0.15
	[1.450]	[1.378]	[0.045]	[0.370]		[0.627]	[0.172]	[0.299]	[0.974]	[1.238]
Public	0.146	0.078	-0.065	-0.106	-0.167**	-0.274	-0.022	0.065	0.112	0.062
	[1.023]	[0.923]	[0.168]	[0.470]	[2.211]	[0.969]	[0.253]	[0.221]	[0.413]	[0.497]
Edu1ary	0.119	0.083*	-0.056	-0.019	0.022	0.02	0.038	0.022	-0.073	-0.037
	[1.131]	[1.690]	[0.258]	[0.174]	[0.129]	[0.267]	[0.290]	[0.355]	[1.034]	[1.061]
Lower2ary	0.024	0.029	0.384**	0.207**	0.045	0.024	0.214**	0.108**	0.134**	0.063**
	[0.237]	[0.595]	[2.151]	[2.304]	[0.278]	[0.298]	[2.055]	[2.177]	[2.168]	[2.101]
Upper2ary	0.213***	0.150***	0.237*	0.155*	0.079	0.052	0.288***	0.183***	0.238***	0.148***
	[2.843]	[3.330]	[1.726]	[1.785]	[0.959]	[1.024]	[4.330]	[4.506]	[4.870]	[5.015]
Lower3ary	0.485***	0.307***	0.679***	0.387***	0.306***	0.185***	0.625***	0.370***	0.545***	0.324***
	[5.689]	[6.145]	[4.461]	[4.194]	[3.125]	[3.145]	[8.322]	[8.397]	[10.989]	[11.093]
PhD	0.655***	0.309***	0.841***	0.368***	0.435*	0.215*	0.840***	0.366***	0.720***	0.336***
	[4.583]	[4.920]	[2.878]	[2.680]	[1.858]	[1.820]	[7.422]	[7.343]	[9.926]	[9.719]
Experience	0.014	-0.008	0.074***	0.023***	0.017	0	0.026***	0.018***	0.029***	0.007**
	[1.493]	[1.611]	[4.184]	[3.781]	[1.395]	[3.458]	[2.742]	[6.219]	[5.372]	[10.052]
Experience2	-0.000***	-0.000***	-0.001***	-0.001*	-0.000*	-0.000*	-0.000**	0	-0.001***	-0.000***
	[3.005]	[2.698]	[2.666]	[1.925]	[1.750]	[1.829]	[2.246]	[1.209]	[4.883]	[2.984]
Wrkng hours	0.014***	0.014***	0.022***	0.036***	0.017***	0.018***	0.018***	0.007	0.015***	0.015***
	[5.180]	[5.099]	[3.687]	[3.855]	[3.413]	[0.028]	[6.121]	[1.560]	[10.140]	[2.291]
Constant	8.098***	9.132***	8.941***	11.256***	8.410***	9.308***	8.101***	9.442***	8.152***	9.642***
	[29.991]	[33.302]	[14.116]	[18.282]	[26.455]	[22.898]	[36.320]	[38.608]	[27.161]	[58.729]
Observ.	285	285	101	101	148	148	375	375	958	958
R-squared	0.469	0.473	0.564	0.536	0.497	0.498	0.47	0.466	0.457	0.448

Notes. Absolute value of t statistics in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. Gross and Net refer to gross and net versions of effort.

2.4.2 Standard Gini versus unfair Gini

In almost all regions, the unfair Gini is smaller than the standard Gini (the Gini reduces from 0.27 to 0.22). This first result means that the distribution of actual income is closer to the distribution of fair income rather than to the egalitarian distribution. In general the unfair Gini increases as we take a net effort and not the gross effort. This is explained by the fact that circumstance impact more on income when the effort is cleaned from the effect of circumstances. Nevertheless, the difference between the standard Gini (0.27) and the unfair Gini with the gross effort ("gross Gini") (0.23) is bigger than the difference between the unfair "gross Gini" (0.23) and the unfair "net Gini" (0.22). It indicates that the indirect effect is relatively small with respect to the direct effect of circumstances.

In addition, when we turn to the analysis per region, the differences are very strong. Firstly, Ile de France is the unique region for which the unfair Gini is bigger than the standard Gini. It means that the income distribution is closer to the equal distribution than to the fair income distribution. The relationship between effort and income in Ile de France appears to be quite weak given these results. The same holds for Rhones-Alpes where the unfair Gini is equal to the standard Gini. In these two regions only, the unfair Gini reduces when accounting for the net effort and not the gross effort. It makes sense to the extent that effort does not explain much income on average, therefore, the less effort we account for, the less inequality we have in these regions.

For the rest of the regions, the unfair Gini is lower than the standard Gini but the difference between both measures is large. In 6 regions the unfair Gini becomes lower than 0.1. This means that actually, when using the theory of equality of opportunity, unfair inequalities are very small, so few redistribution is needed. Instead, in 5 regions, the unfair Gini is smaller than the standard Gini by less than 25%. As a consequence, much more redistribution should be required to reach a situation of equality of opportunity.

A first important conclusion is that, when turning to an ex-post approach, differences in regions regarding the magnitude of inequality of opportunity definitely justify decentralized policies to reduce unfair inequalities. Also, it displays a suprising but interesting result concerning Ile de France: in the biggest region in terms of population and the richest part of this country, inequalities are bigger when using the concept of equal opportunities than when using the standard analysis.

Regarding the relationship between income inequality and inequality of opportunity, there are only three big changes in terms of rankings: Auvergne and Bourgogne that display high level of income inequality perform much better in terms of equality of opportunity. On the contrary, Alsace that has a quite low

level of income inequality displays high level of inequality of opportunity. For the rest, the two first maps of France illustrated in Figure 2.3 show a strong correlation between income inequality and inequality of opportunity. Because the unfair Ginis are lower than the standard Ginis, less redistribution would be required according to the theory of equality of opportunity. To explore further our case, we also show in Figure 2.3 the rankings of regions in terms of gross labour income: the correlation between income level and inequality of opportunity appears to be much weaker among the French regions. In the last subsection we propose to explain diffrences in income between one region and the rest of the regions by decomposing this difference between what is due to effort and what is due to circumstances. In this way, we should explain in further detail the differences in the unfair and fair inequalities across regions of France.

Table 2.9: Inequality of opportunity across the regions - Gini measures

Region	**Gini**	**Unfair Gini (gross)**	**Unfair Gini (net)**	**Wage**
Ile de France	0.299	0.399	0.377	36,558
Champagne-Ardenne	0.271	0.157	0.147	28,779
Picardie	0.238	0.141	0.160	25,705
Haute-Normandie	0.294	0.227	0.253	30,773
Centre	0.212	0.066	0.097	26,360
Basse-Normandie	0.280	0.211	0.236	26,724
Bourgogne	0.222	0.046	0.053	26,824
Nord-Pas-de-Calais	0.248	0.153	0.158	28,352
Lorraine	0.190	0.100	0.086	26,215
Alsace	0.215	0.191	0.158	29,199
Franche-Comté	0.200	0.050	0.035	25,483
Pays de la Loire	0.236	0.082	0.099	26,233
Bretagne	0.218	0.064	0.069	26,856
Poitou-Charentes	0.236	0.130	0.114	25,940
Aquitaine	0.248	0.149	0.138	27,308
Midi-Pyrénées	0.277	0.186	0.173	29,750
Limousin	0.241	0.096	0.147	25,433
Rhône-Alpes	0.288	0.288	0.282	31,942
Auvergne	0.266	0.043	0.083	27,369
Languedoc-Roussillon	0.288	0.192	0.197	27,365
PACA and Corse	0.250	0.180	0.174	28,714
France	0.269	0.228	0.219	29,852

Notes.

2.4.3 Decomposing income inequalities across regions

Finally, we propose an Oaxaca decomposition [50] of income across regions in order to understand better what drives our results on the unfair Gini. To this end, we compare the average income in one region, say region j, with the average income in the rest of the regions, say regions $-j$, and explain these differences by the effect due to "endowments", "coefficients" and the interaction between coefficients and endowments. The decomposition is formalized as follows:

$$\begin{aligned} log(y^j) - log(y^{-j}) &= \Delta X^E \beta^E_{-j} + \Delta X^C \beta^C_{-j} \\ &\quad + \Delta\beta^E X^E_{-j} + \Delta\beta^C X^C_{-j} \\ &\quad + \Delta X^E \Delta\beta^E + \Delta X^C \Delta\beta^C \\ \Delta X^k &= X^k_j - X^k_{-j} \\ \Delta\beta^k &= \beta^k_j - \beta^k_{-j} \end{aligned}$$

With this decomposition, we can explain if differences in the mean income are driven by differences in circumstances or effort. And for each of these two kinds of variables, we decompose the difference between what is due to the distribution of the variables in the region with respect to the other ones and what is due to the impact of the variable on income. This decomposition does not provide a decomposition of the unfair Gini. Nevertheless, it gives a better understanding of the role of circumstances and effort in explaining differences between regions. In this way, it provides complementary conclusions concerning ex-post inequality of opportunity. The results are given in the Tables 2.10 and 2.11.

Our conclusions depend strongly on the measure of effort we adopt. If we focus on the gross measure of effort, we would conclude the following:

In Ile de France whose average income is higher than the average income of the other regions, we find that the role of circumstance plays against this difference in the sense that Ile de France has circumstances that impact negatively on its average income. The distribution of effort and its impact on income is what explained the superior average income of this region. The explanation that conciliates this finding with our previous results on the unfair Gini is the fact that the highest income are located in Ile de France and they are not well explained by the effort variables. This makes the unfair Gini to be big. But if we would exclude the very high incomes not explained by effort, the rest of the population of Ile de France earn more with respect to the rest because they exert more effort and because the reward scheme is different: the effort variables are much more rewarded than in the rest of France.

The same is true for Rhones-Alpes. Given we also find similarities between both regions in terms of the unfair Gini, this analysis allows us to group these

Figure 2.3: Regional Gini, unfair Gini and GDP (the darker the filling, the larger the Gini/GDP)

two regions together. These two regions have large income inequality and large ex-post inequality of opportunity. Their higher level of income with respect to the rest of France are due to the impact of effort variables measured in gross terms, and to a lower extent it is due to better endowments in effort. The effect of circumstances is negative in both regions.

However, if we take the net measure of effort, the conclusions are very different. In both regions, the impact of circumstances become positive and this comes from the coefficients of circumstances. This means that, when we clean the effort from its correlation with circumstances, circumstances turn to explain most of inequalities in income between Ile de France (and Rhones-Alpes to a lower extent) and the rest of France. And this is compatible with our findings on the unfair Gini in both regions. What makes the unfair Gini so high in these regions is the correlation between circumstances and effort and the impact of circumstances on income. These regions exert also more effort but the differences in effort explain less the gap in income than the differences in circumstances when considering the net effort.

The third region that has a higher income than the rest of France is Alsace. In this case, this higher income is due to differences in circumstances when we take the gross and net measure of effort. The distribution of circumstances and the impact of circumstance explain why Alsace has a higher level of income with respect to the other regions when taking the net measure of effort. Moreover, the effect of effort endowments is negative which means that Alsace displays a lower level of net effort than the rest of the regions. We observe exactly the opposite for the fourth region with a higher income than the rest of France: In Haute Normandie, the higher income is explained by effort and not by circumstances.

The other regions have a lower income than the rest of the regions in France. When we take the decomposition made with the gross effort, the effect of circumstances is positive but it turns to be negative when taking the net effort. This indicates that circumstances are negatively correlated with effort in these regions. With the net measure of effort, we conclude that regions that have a lower income than the rest of regions are composed of individuals with "lower" circumstances and exerting a lower degree of effort because the differences in the endowments in circumstances and the differences in the endowments in effort are both slightly negative. Then, regarding the effect of the coefficients, it is more difficult to find a common pattern for all the regions. This shows that the impact of circumstances and effort on income are distinct across regions. Nevertheless, when we consider the decomposition made with the net effort, it seems that, in most of these regions, the impact of circumstances on income explain more the difference in income than the impact of effort on income.

Table 2.10: Oaxaca decomposition of income differences into effort and circumstance factors - region j against all other regions (gross effort)

Region j	Wage region j	Wage France less region j	dlog wage	Total		Endowments		Coefficients		Interaction	
				Circ	Effort	Circ	Effort	Circ	Effort	Circ	Effort
Ile de France	36,558	27,918	0.227	-0.120	0.347	-0.004	0.056	-0.112	0.284	-0.004	0.007
Champagne-Ardenne	28,779	29,878	-0.046	-1.153	1.107	-0.069	0.037	-1.176	1.139	0.093	-0.069
Picardie	25,705	29,988	-0.122	-0.012	-0.110	-0.028	-0.097	-0.009	-0.008	0.025	-0.005
Haute-Normandie	30,773	29,832	0.012	0.446	-0.434	-0.002	0.038	0.437	-0.456	0.010	-0.016
Centre	26,360	30,013	-0.087	0.406	-0.492	0.012	-0.026	0.396	-0.458	-0.002	-0.008
Basse-Normandie	26,724	29,922	-0.132	0.015	-0.147	-0.031	-0.116	0.027	-0.090	0.019	0.060
Bourgogne	26,824	29,937	-0.069	0.176	-0.245	-0.012	0.013	0.182	-0.260	0.006	0.001
Nord-Pas-de-Calais	28,352	29,970	-0.034	0.048	-0.082	0.005	-0.026	0.050	-0.067	-0.007	0.011
Lorraine	26,215	29,987	-0.071	0.253	-0.324	-0.026	-0.065	0.257	-0.249	0.023	-0.010
Alsace	29,199	29,876	0.026	0.377	-0.351	-0.006	0.004	0.383	-0.328	0.000	-0.027
Franche-Comté	25,483	29,971	-0.107	1.275	-1.383	-0.006	-0.017	1.281	-1.342	0.000	-0.023
Pays de la Loire	26,233	30,111	-0.108	0.143	-0.251	-0.006	-0.010	0.148	-0.243	0.001	0.003
Bretagne	26,856	29,983	-0.062	0.562	-0.625	-0.007	0.008	0.562	-0.643	0.007	0.010
Poitou-Charentes	25,940	29,993	-0.118	-0.649	0.531	0.001	-0.038	-0.647	0.563	-0.003	0.006
Aquitaine	27,308	29,968	-0.073	0.200	-0.273	-0.003	-0.009	0.197	-0.263	0.006	-0.001
Midi-Pyrénées	29,750	29,856	-0.009	0.134	-0.143	-0.008	0.038	0.127	-0.159	0.014	-0.022
Limousin	25,433	29,911	-0.133	-0.391	0.258	-0.085	-0.061	-0.358	0.334	0.051	-0.015
Rhône-Alpes	31,942	29,652	0.054	-0.113	0.167	0.008	0.014	-0.120	0.148	0.000	0.005
Auvergne	27,369	29,905	-0.092	-0.260	0.167	0.027	0.003	-0.261	0.136	-0.026	0.028
Languedoc-Roussillon	27,365	29,922	-0.111	0.197	-0.308	-0.009	-0.037	0.192	-0.278	0.015	0.006
PACA and Corse	28,714	29,932	-0.021	0.299	-0.319	-0.004	-0.014	0.300	-0.315	0.002	0.009

Table 2.11: Oaxaca decomposition of income differences into effort and circumstance factors - region j against all other regions (net effort)

Region j	Wage region j	Wage France less region j	dlog wage	Total		Endowments		Coefficients		Interaction	
				Circ	Effort	Circ	Effort	Circ	Effort	Circ	Effort
Ile de France	36,558	27,918	0.227	0.183	0.043	0.011	0.045	0.170	0.002	0.002	-0.004
Champagne-Ardenne	28,779	29,878	-0.046	-0.055	0.009	-0.068	0.047	-0.071	0.028	0.084	-0.066
Picardie	25,705	29,988	-0.122	-0.024	-0.098	-0.040	-0.083	0.003	-0.018	0.013	0.003
Haute-Normandie	30,773	29,832	0.012	-0.090	0.102	-0.007	0.053	-0.094	0.079	0.011	-0.030
Centre	26,360	30,013	-0.087	-0.045	-0.042	-0.005	-0.005	-0.034	-0.030	-0.006	-0.007
Basse-Normandie	26,724	29,922	-0.132	-0.040	-0.091	-0.060	-0.085	-0.009	-0.057	0.029	0.051
Bourgogne	26,824	29,937	-0.069	-0.011	-0.058	-0.013	0.020	-0.003	-0.074	0.004	-0.004
Nord-Pas-de-Calais	28,352	29,970	-0.034	-0.058	0.024	-0.005	-0.017	-0.051	0.035	-0.003	0.007
Lorraine	26,215	29,987	-0.071	-0.022	-0.050	-0.010	-0.085	-0.025	0.030	0.014	0.005
Alsace	29,199	29,876	0.026	0.057	-0.030	0.007	-0.019	0.048	0.007	0.002	-0.018
Franche-Comté	25,483	29,971	-0.107	0.121	-0.229	-0.090	-0.240	0.149	-0.208	0.062	0.219
Pays de la Loire	26,233	30,111	-0.108	-0.095	-0.013	-0.019	0.004	-0.082	-0.014	0.006	-0.003
Bretagne	26,856	29,983	-0.062	-0.079	0.017	-0.003	0.015	-0.083	-0.001	0.006	0.003
Poitou-Charentes	25,940	29,993	-0.118	-0.062	-0.056	-0.033	-0.034	-0.048	-0.033	0.019	0.010
Aquitaine	27,308	29,968	-0.073	-0.096	0.023	0.001	-0.015	-0.102	0.040	0.005	-0.002
Midi-Pyrénées	29,750	29,856	-0.009	-0.019	0.010	-0.006	0.028	-0.025	-0.005	0.012	-0.014
Limousin	25,433	29,911	-0.133	-0.071	-0.062	-0.120	-0.007	-0.031	0.000	0.080	-0.054
Rhône-Alpes	31,942	29,652	0.054	0.029	0.025	0.006	0.017	0.023	0.006	-0.001	0.003
Auvergne	27,369	29,905	-0.092	-0.145	0.052	0.012	0.027	-0.134	0.000	-0.023	0.025
Languedoc-Roussillon	27,365	29,922	-0.111	-0.057	-0.054	-0.007	-0.021	-0.061	-0.027	0.011	-0.005
PACA and Corse	28,714	29,932	-0.021	0.027	-0.048	0.003	-0.020	0.025	-0.038	0.000	0.010

2.5 Conclusions

In this paper, we aimed to analyze inequality of opportunity across the French regions by following an ex-post approach to better consider the effort. To this end, we used a French dataset that includes several indicators of effort and explain income inequalities by differences in responsibility and non-responsibility factors. We define the fair income by using Almas et al. methodology [4] and measure the magnitude of ex-post inequality of opportunity in each region by the distance between the actual income and the fair income. Finally, we use Oaxaca decomposition to further explore the sources of unfair inequalities and the heterogeneity in unfair inequalities across regions.

We find that reward schemes differ across regions. For instance, the effect of education is in some regions three times as much as in other regions. The same is true for the effect of circumstances on income, for which large differences are displayed concerning the impact of gender and father's education. In total the unfair Gini goes from about 0.1 to 0.4 depending on the region. This allows us to conclude that unfair inequalities are not equally distributed across regions and this justifies our regional approach and this may even indicate that policies aiming to reduce inequality of opportunity should be designed at a decentralized level.

Moreover, when using the Oaxaca decomposition to explain the difference between the average income in one region and the average income in the rest of the regions, we obtain that differences are not explained by the same factors. Ile de France and Rhone-Alpes have higher income than the average because their endowments in circumstances and effort are better. On the contrary, Alsace has also a higher income than the rest of France but this is only due to its better endowment in circumstances whereas the higher income in Champagne-Ardenne is explained by the fact that its inhabitants exert more effort. For the regions whose income is lower than the average of the rest of the regions, no common pattern can be identified. This exercice has been complementary to the comparison between the unfair Gini to the extent that it helps to better understand the differences in unfair inequalities across regions.

Finally, the ranking of regions in terms of ex-post inequality of opportunity is found to be similar to the ranking in terms of income inequality which confirm previous results given in the literature, but no similar correlation appears between inequality of opportunity and income level.

Chapter 3

Who Are The Worst-Off When Preferences Matter?

Abstract

Fleurbaey and Maniquet have proposed the criteria of conditional equality and of egalitarian equivalences [33, 34] to assess the equity among individuals in an ordinal framework. They have not been applied as often as the criteria proposed by Roemer or Van de gaer [58, 61] who adopt a cardinal approach. This paper proposes a model that is consistent with these ordinal criteria and enables to compare them with the cardinal criteria. We estimate a utility function that incorporates heterogeneity in groups' preferences from which we approximate individuals' preferences. We obtain ordinal measures of well-being, apply conditional equality and egalitarian equivalence and propose two cardinal measures of well-being that are comparable with the ordinal model to compute Roemer's and Van de gaer's criteria. Finally we compare the characteristics of the worst-off displayed by each criterion. We apply this model to a sample from US data and obtain that only 10% of the worst-off belong to all criteria.

3.1 Introduction

Usually, inequality of opportunity is measured by selecting a particular outcome and decomposing it into two kinds of determinants: the factors beyond the individuals' responsibility, often called circumstances, and the factors for which the individual can be held responsible, the responsibility factors. Equality of opportunity requires erasing unfair inequalities due to circumstances but maintaining fair inequalities due to responsibility factors. As a result, outcomes should be a function of responsibility factors only.

If one takes a cardinal measure for the outcome, interpersonal comparisons are straightforward since everybody is assumed to value equally the outcome. In such a case, we can easily make social orderings based on the concept of equality of opportunity. Two main criteria rely on cardinal measures of outcomes. Firstly, Roemer's criterion requires that individuals who exert the same effort (being the responsibility factor) obtain the same outcome. This corresponds to the ex-post view of equality of opportunity. If this is not the case, an equality of opportunity policy would consist in maximizing the outcome of those who obtain the lowest outcome at each level of effort. A second criterion proposed by Van de gaer's states there is equality of opportunity if the mean outcomes conditional to circumstances are equal. As people who share the same circumstances are called a type, the target of a policy aiming at reducing inequality of opportunity should give the priority to the type with the lowest average outcome. This latter criterion expresses an ex-ante view of equality of opportunity.

In both cases, these criteria assume that individuals have homogeneous preferences over the outcome. If we take the income for instance, it means that two individuals a and b value equally the same amount of income. But it may not be the case. If the individual a values more leisure than the individual b, a lower income for the individual a may give him an equal well-being as a higher income for individual b because the individual a prefers to earn less in order to have more free time. This illustrates the fact that these measures do not account for the heterogeneity in preferences.

Even though it is difficult to identify empirically heterogeneous preferences, it is a noble objective. Indeed, fairness could aim to respect heterogeneous preferences such as to design policies that do not distort preferences. Representing heterogeneous preferences may be done in an ordinal setting in which the way individuals rank situations would reveal their preferences.

However, because it it is impossible to obtain a social ordering that respect fully heterogeneous preferences, it is only possible to rank individuals' situations by making normative choices that enable interpersonal comparisons. This is the proposal made by Fleurbaey and Maniquet [31, 33, 34, 35]. It presents the

advantage of accounting for heterogeneous preferences when defining criteria of fairness and makes explicit the normative choices on which rely fair orderings.

Their work in the field of welfare economics is strongly related to the concept of equality of opportunity [32] because the social rankings they propose are based on equity criteria very similar to the definition of equality of opportunity. They consider that individuals' well being is the relevant outcome on which should be based social rankings and the observed choices made by the individual should be the method to identify their preferences. Secondly, given individual's well-being is a function of preferences and non-responsibility factors, policies should erase inequalities due to non-responsibility factors and be neutral[1] with respect to preferences. In particular, fair policies should rely on two principles that derive from the non-envy principle.

The first principle corresponds to the compensation principle: compensation means we should compensate for factors for which the individual is not responsible. Therefore, people with the same preferences should achieve the same well-being. This is very similar to Roemer's criterion that requires people with the same responsibility factors should end up with an equal outcome.

The second principle corresponds to the neutrality principle: neutrality refers to the neutral treatment of individuals with respect to their preferences. This neutrality principle means that individuals should be treated equally with respect to their preferences[2]. As a result, redistribution mechanisms should be designed in such a way that individuals with equal circumstances will pay/receive the same taxes/transfers. This principle is closer to Van de gaer's criterion that recommends people to have equal opportunities whatever their non-responsibility factors. Because people should have the same opportunities before making their own choice (i.e. deciding on their responsibility factor), this principle encapsulates the idea of neutrality according to which the treatment of the individuals should be independent from their responsibility factors.

The compensation and natural reward principles, probably equally appealing, cannot be both satisfied when individuals have heterogeneous preferences [35, 11]. That is why, Fleurbaey and Maniquet measure unfairness through two criteria, each one giving the priority to one principle and fulfilling only partially the second one. The criterion of conditional equality fulfills the natural reward principle and compensates partially inequalities due to non-responsibility factors. The criterion of egalitarian equivalence gives priority to

[1]Here neutral means that policies should not aim to distort preferences, in such a case, inequalitites between individuals having distinct preferences do not give room to redistribution.

[2]This idea could be questioned if we consider that some goals or preferences should be avoided according to some prevalent values but this discussion falls outside the scope of this paper.

the compensation principle but does not satisfy neutrality with respect to all preferences.

Given the existence of these four criteria, the aim of the paper is twofold:

Firstly, we propose a model in an ordinal set-up that approximates individuals' preferences through the observable choices made by the individuals. This enables us to apply Fleurbaey and Maniquet by respecting their theoretical approach. In fact, previous empirical applications of Fleurbeay and Maniquet's criteria have been adapted in a cardinal framework: Almas [4] and Devooght[27] have used the definition of conditional equality and egalitarian equivalence to measure unfair inequalities for income, therefore, they did not explicitely account for preferences. Then, Decoster and Haan [25] have proposed the first application that follows the key points of the approach, especially the identification of heterogeneous preferences and the use of an ordinal framework [25] but preferences are not individualistic, they are common to groups of people depending on their socio-demographic variables. Moreover, they estimate a discrete choice model to identify groups' preferences but then they translate it into a continuous model when making the social orderings.

In this paper, we propose a extension of Decoster and Haan's model such as to (1) approximate the individualistic component of preferences and (2) apply in a discrete framework the two criteria to be consistent with the original discrete model on which relies our estimates of preferences.

Our second objective is to use the same model to apply Roemer's and Van de gaer's criteria[3]. To this end, we have to solve two issues. Firstly, we need a cardinal measure of well-being. To obtain it, we use two distinct strategies. On the one hand, we erase heterogeneity in preferences such as to obtain a cross-individual comparable measure of well-being. This corresponds to the assumption made to apply conditional equality. On the second hand, we take a money metric for utility that is used for implementing the egalitarian equivalence criterion. In both cases, we rely on explicit assumptions that make comparable the four criteria. Secondly, we have to split between effort and non-responsibility factors instead of the split preference/non-responsibility factors used in the ordinal model. We take the same non-responsibility factors as the one used for conditional equality and egalitarian equivalence. Concerning effort, because here it is not observable direcly, we use Roemer's Identification Axiom: assuming the outcome is a monotonous increasing function of effort, individuals who sit at the same percentile of the outcome's distribution function of their type have exerted the same effort.

[3]To our best knowledge, it exists only a theoretical comparison of the four criteria [32]

This model is finally applied to a sample of US singles from the 2005 CNEF dataset. We identify who are the worst-off according to each criterion such as to know if there is any worst-off regardless of the criterion used. For each criterion, we build a distribution by ranking the individuals (or groups of individuals) from the worst-off to the better-off and the worst-off refer to the individuals who belong to the first decile of the relevent distribution. This comparison across criteria informs us about the impact of the normative choices that enable interpersonal comparisons on the measures of fairness.

The rest of this paper is organized as follows: Section 3.2 presents the model. Section 3.3 details the empirical estimation. In Section 3.4 are presented the data and main results. Finally, Section 3.5 provides a conclusion.

3.2 The Four Criteria

3.2.1 Conditional Equality and Egalitarian Equivalence

A simplified budget set

In a population of N agents, we have a profile of skills s_N equal to the set of wage rates w and a set of utility functions u_N, that depend on consumption and leisure. An economy is $e = (s_N, u_N)$ where every agent maximizes his utility u_i over consumption and leisure $(C_i, L_i) \in X = (0, 1)$.

In brief, well-being is assumed to depend on disposable income, C and leisure L can be represented through an utility function $u(C, L)$. The way people value leisure affect their well-being such that if two individuals have heterogeneous preferences for leisure, a same bundle (C, L) does not mean an equal well-being.

Every individual maximizes his utility subject to a budget constraint $B \subseteq X$. The budget represents the level of consumption accessible to the individual. It varies for each individual because it depends on (1) the individual's wage rate w that is assumed to be constant whatever the individual's labour time, (2) the amount of labour time $l = 1 - L$ and (3) the tax rate t applied to the gross labour earnings.

This budget is actually not linear since t varies along with $w * l$. But when implementing conditional equality and egalitarian equivalence criteria, we aim to preserve as far as possible the neutrality principle. This states that the treatment of individuals should be neutral with respect to preferences and this is incompatible with the actual non linear budget sets. We can compute the redistribution rules that fulfill the criteria in a neutral setting if we worked with nested budget sets, and rank the budget sets in an unequivocal way when defining who the worst-off are.

This is why, before implementing the conditional equality and egalitarian equivalence criteria, a preliminary transformation of the actual budget set into a simplified budget set is necessary. The simplified budget set is composed by the lump sum transfer that, with his observed wage rate and free to choose the bundle (C,L) according to his preferences and this transfer, would make the individual just as well-off as he is in his current situation. We obtain nested budget sets by replacing the actual budget set by a lump-sum transfer in the following way:

Given every individual maximizes his utility subject to a budget constraint $B \subseteq X$, the utility function derived from any subset B is:

$$u_i(B) = \{\max u_i(c,l) | (c,l) \in B\} \tag{3.1}$$

The simplified budget set $B^* \subseteq X$, is determined by the gross income and a lump-sum transfer $\hat{t}$ such that:

$$B^*(w_i, \hat{t}) = \left\{(c,l) \in X | c \leq w \times l + \hat{t}\right\} \tag{3.2}$$

and:

$$u_i(c,l) = u_i(B^*(w_i, \hat{t})) \tag{3.3}$$

Responsibility versus non-responsibility factors

Before defining the equity criteria, we need to define the split between responsibility and non-responsibility factors.

In Fleurbaey and Maniquet setting, the non-responsibility factor is the wage-rate. The wage-rate is likely to represent skills that are mostly the product of genetics, family background, luck but also effort. Despite this choice appears to be quite controversial, we aim to implement the criteria following the spirit of their authors, therefore we take the same hypothesis. The responsibility factors are individuals' preferences and correspond to the preferences for leisure and consumption as preferences affect well-being through consumption and leisure[4]. Once identified the wage-rate, the individual's preferences and the implicit budget set, we can explain in further detail the equity criteria.

Conditional equality

Conditional equality permits to fulfill completely the neutrality principle and achieve partially the compensation for circumstances. Formally defined by Fleurbaey [31], conditional equality can be computed as follows:

[4]In the estimation, we only estimate heterogeneity in preferences for a question of simplicity given it does not change the results.

> "Define a reference value for responsibility characteristics and give priority (according to the leximin criterion) to individuals, who, with their current resources and circumstances and this reference value of responsibility characteristics, would be the worst-off."

In a nutshell, the idea enclosed in this principle is the following: in a fair economy, if individuals had the same preferences, they would end up with the same well-being. Thus, the individuals who would obtain the lowest well being when fixing a reference value for preferences for all are those whose non-responsibility factors impact negatively on their well being. They are the worst-off. Then, a conditional equality rule consists in fully neutralizing the effect of circumstance for the people having the reference value for preferences and partially neutralizing the effect of non-responsibility factors for the rest of individuals because of the restriction according to which this principle always imposes a same treatment for individuals having the same circumstances.

In the framework defined above, implementing conditional equality requires defining a reference value for preferences. This is done by fixing a reference utility function. Then, when we suppose that individuals have the same preferences, according to the fairness principles, they should obtain the same well-being. This turns to have equal simplified budget sets because individuals with equal preferences obtain the same well-being if they have the same budget sets. Therefore, when ranking in ascendent order the simplified budget sets, we obtain the rankings of individuals from the worst-off to the beter-off:

Formally, individuals are ranked according to:

$$\widetilde{u}(B(w_i, \hat{t}_i)) \tag{3.4}$$

where $\hat{t}_i$ is the lump-sum transfer corresponding to the simplified budget set such that:

$$u_i(C_i, L_i) = u_i(B(w_i, \hat{t}_i)) \tag{3.5}$$

and $\widetilde{u}$ is the reference utility function that results from fixing a reference value for preferences.

Instead of fixing a reference value for the form of the utility function, we could also fix a reference value for labour time. It is equivalent with stating that all individuals have preferences such that they decide to work the same amount of working hours. In this case, implementing conditional equality consists in ranking the individuals' well-being when all individuals work the same amount of hours.

Egalitarian Equivalence

The second equity criterion, called egalitarian equivalence, fulfills completely the compensation principle and partially the neutrality principle. As defined

by Fleurbaey [31], it requires the following:

> "Define a reference type of circumstances and give priority (leximin) to individuals whose current level of well-being would be obtained with the least resources if their circumstances were of the reference type"

This method mimics a situation where inequalities would be due to responsibility characteristics only. As a consequence, the redistribution rule that emerges from this principle fully fulfills the compensation principle but it may treat unequally individuals with the same circumstance.

Concerning the application of this criterion, Fleurbaey proposes to use different reference value for the wage rate as the reference value affects people ranking. The ranking depends on the apportion between preferences and income: the higher the reference wage rate, the larger the priority given to people having lower preferences for leisure, that is to say the hard-working people.

If the reference wage rate is equal to zero, we obtain the zero egalitarian equivalence. To this end, we rank $\hat{t_i}$ in ascending order as defined by :

$$u_i(x_i) = u_i(B(0, \hat{t_i})) \tag{3.6}$$

If the reference wage rate is equal to the minimum wage rate, we will obtain the min egalitarian equivalence: It consists in ranking in ascending order the $\hat{t_i}$ as defined by

$$u_i(x_i) = u_i(B(min_{j \in N} s_j, \hat{t_i})) \tag{3.7}$$

Lastly, the wage egalitarian equivalence or equivalent wage is defined in Fleurbaey [31] as:

> "For each individual, compute the counterfactual wage rate (with no transfer) that would make the individual as happy as in his current situation, and give priority (leximin) who are the worst-off in these terms."

Formally, it is equivalent to rank the individuals in ascending order according to $\hat{w_i}$ that satisfies:

$$u_i(x_i) = u_i(B(\hat{w_i}, 0)) \tag{3.8}$$

3.2.2 Roemer's and Van de gaer's criteria

Methodological questions

For the application of conditional equality and egalitarian equivalence, the form of the utility function itself does not matter, we only need to know the

way people rank the bundles (C, L) according to their preferences in order to make orderings. Instead, Roemer's and Van de gaer's criteria use a cardinal measure of well being such as to make interpersonal comparisons. Taking the observed earnings is not an appropriate solution as the first two ones would account for another dimension of well-being and the last two ones would not. A possible solution is to take a measure of well-being. This is consistent with the hypothesis made above about individuals' behavior.

Then, the question is to select an appropriate cardinal measure of well-being. On the one hand, to implement the conditional equality criterion, we need to choose a reference utility function for everybody in order to remove heterogeneity in preferences. It means that the value of $\widetilde{u}(B(w_i, \hat{t_i}))$ that is obtained for conditional equality is (1) comparable across individuals because the same utility function is used for everybody, and is (2) an approximation of the current individual's well-being. In fact, we erased the impact of heterogeneous preferences on well-being, but this limitation is imposed by the Roemer's and Van de gaer's criteria. In consequence, this measure can be used to implement Roemer's and Van de gaer's criteria and will be easily comparable with the results given by the conditional equality criterion.

On the other hand, the egalitarian equivalence consists in obtaining the amount of resources that should be given to an individual such that he obtains his current level of well-being if all individuals had the same circumstance. In consequence, this amount of resources reflect the current well-being of the individuals and it respects individual's preferences. This is a money metric for individual's well-being. For example, if we take the zero equivalance criteria, we obtain the resources that an individuals needs to be as well off as he is in his current situation if he does not work. One limit is that this money metric corresponds to a virtual situation that may not occur ever. Still, this measure enable us to implement Roemer's and Van de gaer's criteria and is directly comparable with the results given by the egalitarian equivalence criterion.

The second problem we face is the fact that Roemer's and Van de gaer's criteria make a distinction between effort and circumstance, instead Fleurbaey and Maniquet distinguish between preferences and non-responsibility factors.

To make the comparison more reliable among the four criteria, it is natural to define the circumstance in the same way as the non-responsibility factors because both terms reflect the factors that are beyond individual's responsibility. Thus, the circumstance is unique and corresponds to the wage rate[5].

[5]More complex specifications could be of main interest but are left to other studies since here the puropose is to offer a consistent comparison among several criteria and not a detailed implementation of one specific criterion.

Regarding effort, from the previous framework, effort is multidimensional, it is made of the preferences for leisure and consumption but also the amount of leisure chosen and the form of the utility function, so it is impossible to find one single continuous value for effort. A way out is to use Roemer's Identification Axiom to determine one index of responsibility: firstly, the population is partitioned into types according to the wage rate. Then, we draw the outcome's distribution function of each type, the outcome being the well-being as defined above. Finally, we assume that the individuals belonging to the same percentile of their outcome's distribution function have exerted the same effort. In this way, we obtain an index of individual's effort. This method is coherent with Roemer's definition of effort and allows us to compute Roemer's and Van de gaer's criteria consistently within their original framework.

Implementing Roemer's and Van de gaer's criteria

There is equality of opportunity according to Roemer when people exerting the same effort obtain the same outcome. Measuring the fairness of any distribution consists in comparing individuals' well-being at each level of effort across types, and improving equality of opportunity requires maximizing the mean outcome of the individuals who have the lowest well-being at each level of effort.

With such a criterion, the worst-off may be defined as the individuals with the lowest well-being at each level of effort to the extent that an increase in their well-being would improve the fairness of the distribution.

To implement it, we firstly divide the population into types. Then we take our cardinalization of well-being and apply Roemer's Identification Axiom to obtain an indirect measure of effort. Lastly, we take the individuals with the lowest well-being for each value of effort and obtain who the worst-off are according to Roemer's criterion.

Equality of opportunity defined by Van de gaer's is characterized by an economy where the average well-being of each type is equal. With such a definition, the worst-off cannot be identified individually. Instead they can be defined as the type with the lowest average well-being.

To identify the worst-off, we use the cardinalization of the well-being detailed above. Then, we measure the average well-being conditional on the type and the worst-off are the individuals who belong to the type with the lowest average well-being.

3.3 A discrete choice model to apply the criteria

3.3.1 Estimation of groups' preferences

The first objective is to identify heterogeneous preferences in an ordinal frawework where individual's well-being depends on consumption and leisure. To this end, we start with Aaberge et al. [1, 2] model.We assume that individual's well-being is representable through a utility function that depends on disposable income C, leisure L, socio-demographic variables X and a random error term ϵ_{ij} that varies independently among individuals and alternatives. It is not observed but affects individuals' choices. The utility function can be written as follows:

$$V_{ij}(C, L) = U(C_{ij}, L_{ij}, X_i) + \epsilon_{ij} \tag{3.9}$$

The two sub-indices may appear redundant. They aim to show that utility differs across individuals i and across alternatives over the working time j[6]. Indeed, the earnings C depend on the wage rate that varies across the individuals i and depend on the amount of working time. Also, the amount of leisure L depend on preferences for leisure, defined individually, and the working hours. Finally, ϵ_{ij} is also assumed to vary across individuals and alternatives.

The individual i maximizes his utility by choosing his amount of working hours ($l = 1 - L$) from a set of the alternatives on working time $j \in J$. Empirically, as we use a dicrete choice model, we restrict the individual's choices such that the individual is free to select his labour time among 12 alternatives:

$$j \in J = (0, 5, 10, 15, .., 55) \tag{3.10}$$

The individual is subjected to a budget constraint that depends on the wage rate received by the individual, the amount of labour time and the taxes he pays on his gross earnings. The wage rate is supposed to be constant whatever the amount of labour time picked by the individual. As the actual taxes are not linear, we use information on income taxes given by the OECD such as to reproduce the actual individual's budget constraint, that is to say, to calculate the net earnings an individual would receive for each of the 12 alternatives.

The budget constraint can be written as:

$$C = w \times l - t(w, l) \tag{3.11}$$

[6] j and l both represent the working hours. Nevertheless we use both to maintain the intuitive meaning of l being the labour time, and j the discrete alternative.

Where $t(w,l)$ is the tax function that transforms gross earnings into net earnings.

Regarding the deteriministic part of the utility function, we use the same specification as Aaberge et al. [1, 2] and Decoster and Haan [25]:

$$U(C,L) = \beta_C \frac{C^{\alpha_C} - 1}{\alpha_C} + \beta_L \frac{L^{\alpha_L} - 1}{\alpha_L} \tag{3.12}$$

The parameters β_C, α_C, β_L and α_L indicate preferences for consumption and leisure. β_C, α_C and α_L are common to all individuals. Heterogeneity among groups of individuals is introduced through β_L that depends on socio-demographic variables:

$$\beta_L = \beta_{L_0} + \beta_{L_1} X_i \tag{3.13}$$

Where X_i are gender, age, education and ethnic group.

The variables that explain the differences in the preference for leisure have been widely used in the literature to explain the determinants for labour supply. They are not really under the individual's control but they are expected to explain differences in preferences for leisure. In fact, Dworkin [28, 29], Rawls [55], Fleurbaey and Maniquet [34] share the view according to which as long as people identify themselves with their preferences, no matter what explains these preferences, we should be neutral with respect to them. The aim of the paper is to implement the criteria according to the view of their respective authors, therefore β_L represents individuals' preferences and we do not want to compensate for them.

To obtain the parameters of the deterministic part of the utility function, the estimation relies on a rationality assumption that states that if the ith individual makes the choice j in particular, it means that V_{ij} is the maximum among the j alternatives. In other words, the probability that the the ith individual makes the choice j is:

$$Prob(y_i = j) = Prob(V_{ij} > V_{ik}) \forall k \neq j \tag{3.14}$$

We replace V_{ij} and V_{ik} by its expression and obtain:

$$Prob(y_i = j) = Prob(\epsilon_{ik} - \epsilon_{ij} < -(U_{ik} - U_{ij})) \forall k \neq j \tag{3.15}$$

The resulting multinomial model is treatable if we assume that ϵ_{ij} is $i.i.d$ random value with type I extreme value distribution, then differences in epsilon follow a standard logistic distribution [16]. In this case, we estimate the parameters of the utility function by maximum likelihood as a conditional logit model where:

$$Prob(y_i = j) = \frac{\exp U(C_{ij}, L_{ij})}{\sum_{k=0}^{n} \exp U(C_{ik}, L_{ik})} \forall k \neq j \tag{3.16}$$

We obtain as many utility functions as numbers of groups having the same socio-demographic characteristics. Formally, we have:

$$V_{ij}(C,L) = \beta_C \frac{C^{\alpha_C} - 1}{\alpha_C} + (\beta_{L_0} + (\beta_{L_1} X_i)) \frac{L^{\alpha_L} - 1}{\alpha_L} + \epsilon_{ij} \tag{3.17}$$

Individuals who have the same leisure and the same consumption may obtain a distinct well-being if they have a disctinct β_L. That is to say that the form of the utility function varies accross individuals who have a different age or gender or education or ethnic group.

In the end, with this estimation proposed by Aaberge et al. [1, 2] and Decoster and Haan [25], we estimated groups' preferences for leisure. However, Fleurbaey and Maniquet recommend to identify individuals' preferences as the criteria they propose allow us to establish individuals' rankings. Instead, identifying groups' preferences is not really satisfactory to establish individuals' rankings. This is why, we propose a extension of this model.

3.3.2 Approximation of individual's preferences

β_L does not include information specific to the individuals. On the contrary, ϵ_{ij} is specific to the individual and to the alternative and it impacts on the well-being V_i. As a result, it is reasonable to assume that the error term, ϵ_{ij}, may capture the individualistic component of preferences not explained by β_L.

After estimating the parameters of $U(C,L)$, we can compute the utility an individual would obtain for each alternative $j \in J$. It may occur that the individual maximises his utility for the amount of labour time he actually chose. In this case, we assume the individual has the same preferences as the group to which he belongs.

But, we can also obtain that the individual maximises his utility for an amount of labour time he did not pick. Formally, we can observe $U(C_{ij}, L_{ij}) < U(C_{ik}, L_{ik})$ where j is the amount of labour time the individual has actually chosen and k is one or some other possible alternatives. However following the assumption of our model according to which the individual is rational, we should have $V(C_{ij}, L_{ij}) > V(C_{ik}, L_{ik})$ for all the other possible alternatives k.

Assuming there is no problem of specification of our model and using the expression of V_{ij} we deduce that ϵ_{ij} may explain why, at the same time, the individual does not maximise his utility U_{ij} for his actual labour time j, whereas he actually maximizes his utility V_{ij} when picking a labour time j. In such a case, ϵ_{ij} captures the individualistic component of preferences not explained by β_L. Still, ϵ_{ij} is not observable after the estimation. In this

sense, we cannot capture the individual's preferences. Nevertheless, we can use the assumptions of the model to get a proxy for differences in epsilon as follows:

(1) if the individual has picked the alternative j it means that the utility he obtains with this option is superior to any other alternative. Formally, it means that:

$$\epsilon_{ik} - \epsilon_{ij} < -(U_{ik} - U_{ij}) \forall k \neq j \tag{3.18}$$

Omitting indexes, we can say that because individuals are rational, we have to satisfy:

$$\Delta\epsilon < -\Delta U \tag{3.19}$$

(2) The hypothesis of the estimation according to which differences in epsilon follow a standard logistic distribution means that the density function is:

$$f(\Delta\epsilon) = \frac{\exp \Delta\epsilon}{1 + \exp \Delta\epsilon^2} \tag{3.20}$$

Using (1) and (2) permits us to rescale the density function of $\Delta\epsilon$ so that:

$$f(\Delta\epsilon|\Delta\epsilon < \Delta U) = \frac{f(\Delta\epsilon)}{Prob(\Delta\epsilon < \Delta U)} = \frac{f(\Delta\epsilon)}{P} \qquad \forall\epsilon < -\Delta U \tag{3.21}$$

$$f(\Delta\epsilon|\Delta\epsilon < \Delta U) = 0 \qquad \forall\epsilon \geq -\Delta U \tag{3.22}$$

where P is:

$$P = 1 - \frac{\exp \Delta U}{1 + \exp \Delta U} \tag{3.23}$$

This rescaling matters for our purpose because it allows us to better identify the form of the indifference set for every individual and this is the starting point for the application of the conditional equality and egalitarian equivalence criteria.

With our estimation we can compute the utility received by an individual for the bundle (C,L) he actually chose. With this information, we can compute the level of consumption an individual would need to obtain the same utility if he choose another labour time. That is to say, we can compute the 12 points of the indifference set (to the extent we the individual has 12 alternatives of labour time).

The points of an indifference set are such that $V_{ik} = V_{ij}\ \forall k \neq j$. By replacing V_{ik} and V_{ij} by their expressions and omitting the index i, each point of the indifference set is as follows:

$$IC_k = [C_j^{\alpha_c} + \frac{\alpha_c}{\beta_c}(\Delta\epsilon + \frac{\beta_l}{\alpha_l}(L_j^{\alpha_l} - L_k^{\alpha_l}))]^{\frac{1}{\alpha_c}} \tag{3.24}$$

We observe we need a value for $\Delta\epsilon$. This is why the rescaling matters. Given the unconditional density function of $\Delta\epsilon$, we generate one million drawings of a random variable that follows a standard logistic distribution so as to obtain one million possible values for $\Delta\epsilon$. Then, for each alternative and each individual, we measure ΔU. From each of these values, we compute the expected value of our generated random variable given in Equation 3.21.

To summarize, what we have done is simply to use the assumptions about the distribution of $\Delta\epsilon$ and the rationality hypothesis to define the conditional distribution $\Delta\epsilon$. This allows us to better approximate the indifference set of every individual as the form of the indifference set is specific to the individual when $E(\Delta\epsilon)$ is different from zero. Still, two individuals who (1) share the same socio-demographic characteristics and (2) maximize U for their chosen labour time have the same indifference set.

3.3.3 Conditional Equality

First of all, we use the actual budget set to estimate the deterministic part of the utility function. Then, we need to determine the simplified budget set to implement both criteria. Given the definition given above and our discrete choice model, we take the observed individual's wage rate and the points of the indifference set to compute $\hat{t}$ such that:

$$\hat{t} = IC_k - w_i \times k \tag{3.25}$$

The minimum value of $\hat{t}$ among the twelve k is $\bar{t}$. It is the minimal lump sum transfer that makes individual as well off as he is in his current situation with his observed wage rate and free to choose his labour time. It allows us to compute the simplified budget set BC_k defined as:

$$BC_k = w_i \times k + \bar{t} \tag{3.26}$$

With the simplified budget set, we can turn to the application of the conditional equality criterion: we fix a reference value for β_L and ϵ, and obtain for each individual the utility he would obtain under each alternative of labour time and given the simplified budget constraint.

The maximum value for utility we obtain for each individual is labelled $\bar{U}_i$ and these values are now comparable across individuals since all individuals share the same utility function form. Ranking $\bar{U}_i$ gives us the ranking of individuals according to the conditional equality criterion. The worst-off are the persons with the 10% lowest value for $\bar{U}_i$.

As the results may change depending on the reference value for β_L and ϵ, we vary the reference values. We use four different technics for fixing reference

values for the utility function. When selecting a utility function form common to all individuals, well-being varies across alternatives only, this can be done directly through β_l, thus, we assume epsilon is equal to zero for every individual and every alternative. We fix two distinct values for β_L in order to know by how much the characteristics of the worst-off vary when we take a low or a high reference value for the preference for leisure. Secondly, we fix l=20 and alternatively l=40 and compute the utility of every individual if he opted for this labour time, given his observed wage rate and his own preferences.

3.3.4 Egalitarian Equivalence

Departing from the indifference set IC_k, egalitarian equivalence requires fixing a reference value for the wage rate and compute the minimum lump-sum transfer that gives to the individual his actual level well-being with this reference value for the wage rate. To do this, we compute $\forall k$:

$$EE_k = IC_k - \tilde{w} \times k \tag{3.27}$$

where $\tilde{w}$ is the reference wage rate. It is equal to 0 in the case of the zero egalitarian equivalence criterion and $\tilde{w}$ is equal to the minimum wage rate (fixed to 5 dollars per hour) when computing the min egalitarian equivalence criterion.

We call $\bar{EE}$ the minimum among the twelve values EE_k for every individual and rank $\bar{EE}$ in ascending order. The worst-off are the persons with the 10% lowest value for $\bar{EE}$.

For the computation of the wage egalitarian equivalence, we find the wage rate the individual should receive to achieve the lowest point of his actual indifference set if there were no tax.

For each point of the indifference set IC_k, except for $k = 0$, we generate:

$$WEE_k = \frac{IC_k}{k} \tag{3.28}$$

We take the minimum WEE_k among the eleven values for k, it is labelled $\bar{WEE}$, and the worst-off are the individuals with the 10% lowest value for $\bar{WEE}$.

3.3.5 Roemer's and Van de gaer's criteria

We first have to use a cardinalization of the utility function. As explained above, we propose two technics to obtain a cardinal measure of the actual well-being.

The first cardinal measure of the well-being corresponds to the point of the indifference set when $k = 0$. It is the consumption level that would give to the individual his actual level of well-being if he did not work, called CI_0. It is a money metric of the actual well-being, it is directly comparable with the egalitarian equivalence and it is also related to a special case of conditional equality when individual have strong preference for leisure so as to choose not working. The second method consists in imposing to all individuals the reference value of β_L and ϵ we have used for the conditional equality criterion. These two cardinalizations will make the cardinal criteria more comparable alternatively with the egalitarian equivalence and with the conditional equality criteria.

Regarding the split effort/circumstance: types are defined according to the wage rate. Precisely, we build ten types according to the value of each decile of the distribution function. Effort is measured indirectly by using Roemer's Identification Axiom: people who sit at the same percentile of the utility distribution function of his type have exerted the same effort. Given the value of each decile, we build ten levels of effort. The worst-off are the individuals with the 10% lowest utility at each level of effort. Given our decomposition betwen types and effort, we obtain the same number of worst-off for Roemer's criterion and for the other three criteria.

To implement Van de gaer's criterion, we simply compute the mean utility conditional to the type. The worst-off is the type with the lowest average well-being.

3.4 Data and Results

3.4.1 Data

The empirical analysis is based on US data from the Cross National Equivalent File (CNEF) for 2005 that provides information on incomes in 2004. It includes detailed information on socio-demographic variables. Given we do not have at our disposal a microsimulation model describing the exact model of tax and transfers the individuals face, we focus on singles without children because the OECD report on income taxes gives detailed information on taxes for this population. As a result, our measures are restricted to this specific sample and the results cannot be generalized to the whole population. In addition, we restrict our sample to people who work at least twenty hours a week[7]. We also restrict our sample to individuals aged between 25 and 65 years old who are not self-employed, retired or fully engaged in education.

[7]As we have poor information on transfers given to people who do not work, we are not able to well explain the decision of not working. To avoid mis-specification, we exclude this population that represent 8% of the sample.

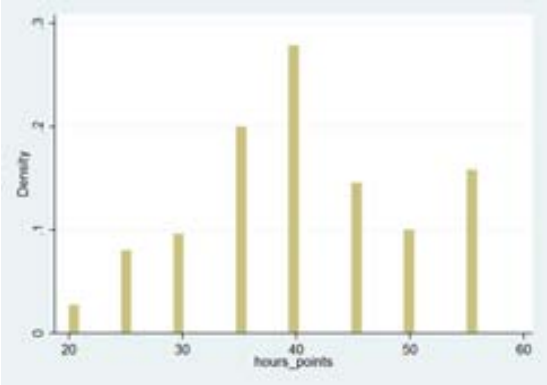
Table 3.1: Descriptive statistics

Characteristics	Frequence	mean labour time	mean monthly net income
Education: less than high school	9.73%	40.26	1680.53
Education: high school	37.75%	39.73	2080.68
Education: more than high school	52.52%	40.78	2738.07
White	51.93%	40.31	2581.76
Black	42.71%	40.41	2185.45
Other	5.37%	40.47	2120.07
Women	43.34%	38.70	2307.40
Men	56.66%	41.62	2449.29
Sample	100%	40.36	2387.73

This gives us a sample of 597 individuals.

Even though we restrict our sample to individuals who work at least 20 hours, we estimate our model by assuming that individuals can choose freely among twelve alternatives on working hours[8]. The wage rate is assumed to be constant whatever the number of working hours, this is in line with Decoster and Haan model [25]. Distribution of working hours is given in Figure 3.1.

Figure 3.1: Working hours



To build the budget constraint, we use the 2004 OECD report for the US to derive the net income for each possible working hours. Precisely, we first calculate the gross wage rate by dividing the individual gross labour earnings by the annual working hours. Then, we make discrete the labour time such as to vary between 0 and 55. For each discrete alternative, we compute the corresponding gross total earnings and we use the report on tax

[8]This does not affect in a large extent our estimates since individuals are found to maximize their utility when working between 30 et 55 hours a week.

income to simulate the corresponding net earnings. For every individual, we fix a maximal amount of time endowment T=80, and leisure is T minus the amount of weekly working hours. Then, in the empirical estimation, leisure is normalized to one and net income is expressed in units of 10000 dollars per year.

We do not include capital income in the estimation. The gross amount of capital income the individual receives is available but there is no desaggregation by type of capital income and no data on net capital income. As taxes depend on the sources of this income, net capital income cannot be properly computed[9].

Figure 3.1 indicates the distribution of the sample in terms of education, gender and ethnicity, these characteristics are used to describe who the worst-off are according to each criterion. We also indicate the mean labour time and mean net income.

3.4.2 Estimation Results

Table 3.2 presents the estimates of the parameters of the utility function. We find individuals have positive preferences for consumption (α_c is positive and significantly different from zero) and leisure because β_l is found to be positive for every individual. On the other hand, despite the parameters α_c and α_l are not both inferior to one (but significantly different from zero), the utility function is still concave because the value for labour time is never bigger than $11/16$.

Regarding the components of β_l, we find that being a women, and having a low education level have a significant positive impact on preference for leisure, which is also found in Decoster and Haan study for Germany [25].

There is no significant heterogeneity among individuals explained by the other observable characteristics. This may be explained econometrically by the small sample size (597 individuals). Moreover, this may be due to the fact we have restricted our sample to individuals without any children. Within this particular sample, most individuals tend to work full-time, so it is likely that this sample lacks of heterogeneity in preferences. In this context, it appears even more appropriate to correct groups' preferences with our proxy for the differences in epsilon.

In addition, we propose to compute by how much our model fits well with our data in the following sense: our model will be satisfactory if our estimation results show that the individual obtains a maximum value for U_i for his chosen labour time. It would mean that our model explains well

[9]Capital income would affect the level of the budget constraint and probably the decision of working or not. Therefore, not taking into account capital income should affect the slope of the utility function for low level of working hours. As most individuals of the sample work full time, our estimates should not be affected by the fact we exclude capital income.

Table 3.2: Parameters of the utility function

	Coefficient	z
Preference for Consumption		
β_c	16.21	4.61
α_c	0.017	0.79
Preference for leisure		
β_l :		
Age	-32.22	-0.90
Age squared	37.59	0.87
Women	3.50	3.39
Education (ref: more than high school degree)		
Less than high school	1.85	1.09
High school	2.02	1.92
Ethnicity (ref:white)		
Black	-0.08	-0.09
Other	0.93	0.45
Intercept	41.84	2.62
α_l	1.33	6.45

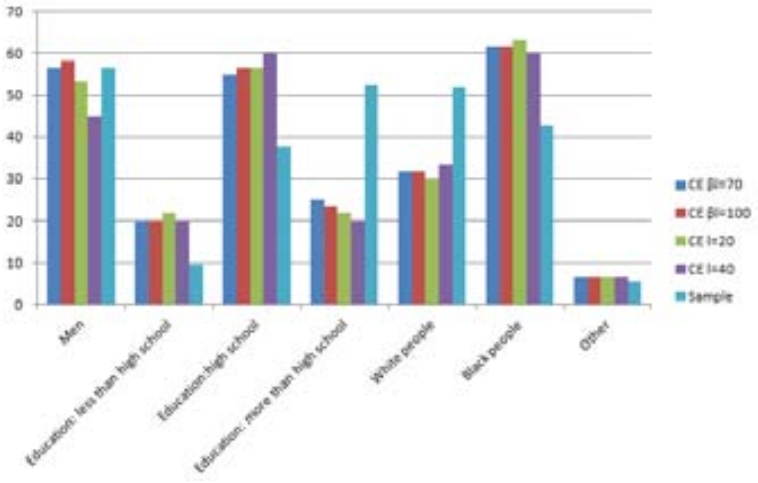
individuals' preferences with the socio-demographic variables. To this end, we compute the value for U_{ik} for each labour time and check if $U_{ij} > U_{ik}$, j being the amount of working hours the individual chose and k the other possibilities. Given we identify groups' preferences and have a discrete choice model, it is unlikely to find a 100% fit. Therefore, we also extend our comparison: if the individual does not maximize U_{ik} for the labour time j, we check if the maximum value for U_{ik} corresponds to a "close" alternative. For instance, if the individual works 40 hours a week, in case U_{ik} is not maximum for $k = 40$, we check if it is maximum for $k = 35$ or $k = 45$.

We find that our estimation results display that the individual obtained a maximum value for his chosen labour time in 1/3 of the sample. More satisfactory is the fact that in 60% of the sample, U_{ik} is maximum for the chosen labour time or the "close" alternative.

3.4.3 Who are the worst-off?

Firstly, we fix two different reference values for β_l. The first one is 70, the second one is 300. The mean of β_l among individuals in the sample is around 37, but we decide to take a larger value because ϵ is assumed to be equal to 0. Indeed, with our estimation, we underestimate the quantity of people who maximize their utility for a labour time inferior to 40 hours a week. This means that our value of β_l is probably underestimated. With the computation of the expected value of the differences in ϵ , we can correct this. In order to

Figure 3.2: Conditional equality



take it into account for the computation of the conditional equality criterion, we take a value of β_l that is bigger than the estimated value.

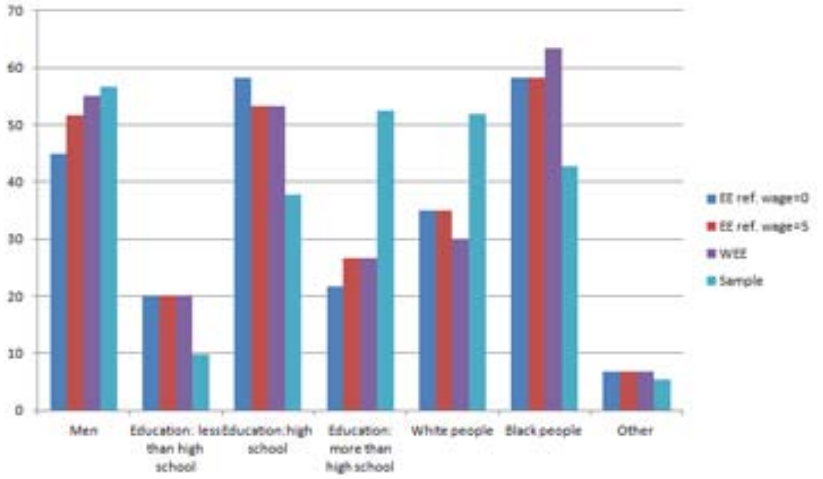
Intuitiveley, when increasing the value of β_l the worst-off should turn to be individuals with higher wage-rate. That is concordant with our results. As we increase the reference value for β_l the results change significatively. If we assume strong preferences for leisure, the worst-off turn to be men, individuals with high education and the white people. Instead, when we took a reference value of 70, the worst-off tend to be people with a rather low educational level (high school or less), the women and the black people.

When we fix an amount of labour time equal to 20 or 40, the results are quite similar with the conditional equality criterion when β_l is equal to 70. This is explained by the fact that determining a positive amount of labour time means the individuals do not have very high preference for leisure. If this were the case, they would not choose working part time or full-time.

Regarding the egalitarian equivalence criteria, results change slightly depending on the reference value for the wage rate. When the reference value is equal to zero, the worst-off tend to be people with high distate for working, so that they need low transfer to be as happy as they would be if they did not work. Here, in comparison with the composition of the sample, the worst-off are composed by more women and people with a degree equal to high school who are those with high preference for leisure.

On the other side, results given by the wage egalitarian equivelence criterion

Figure 3.3: Egalitarian equivalence



and by the egalitarian equivalence when the reference value for the wage rate is equal to 5 dollars per hour are quite similar. The proportion of men increases slightly as well as the proportion of white people. We could expect bigger differences but given our restricted sample, we fail to obtain significant differences when we vary the reference value.

Regarding Roemer's criterion, the result vary significantly depending on the cardinal measure of well-being we select. When we fix $\beta_l = 70$ and $\epsilon = 0$, we make the results more comparable with the conditional equality criterion. However, the worst-off according to Roemer's criterion are not comparable with the ones identified by the conditional equality criterion. In comparison with conditional equality, there are more men, individuals with high educational level and more white people among the worst-off using Roemer's criterion. The reason why we obtain more people with high net observed income could be found in the definition of the worst-off. Given they are the people with the lowest outcome at each level of effort, we have automatically some worst-off among people exerting a high level of effort. In consequence, within the group at the top level of effort, the worst-off have a relatively low outcome but still they have a higher outcome than the individuals exerting a low level of effort.

The worst-off according to the min egalitarian equivalence criterion are also different from the worst-off according to Roemer's criterion when the outcome is the min egalitarian equivalence. We can give the same explanation as the one provided above. This shows that not only the normative choice induced by the cardinalization modifies the identification of the worst-off but also the split between preference/circumstance versus effort/circumstance lead to

distinct results.

Figure 3.4: Roemer's criterion

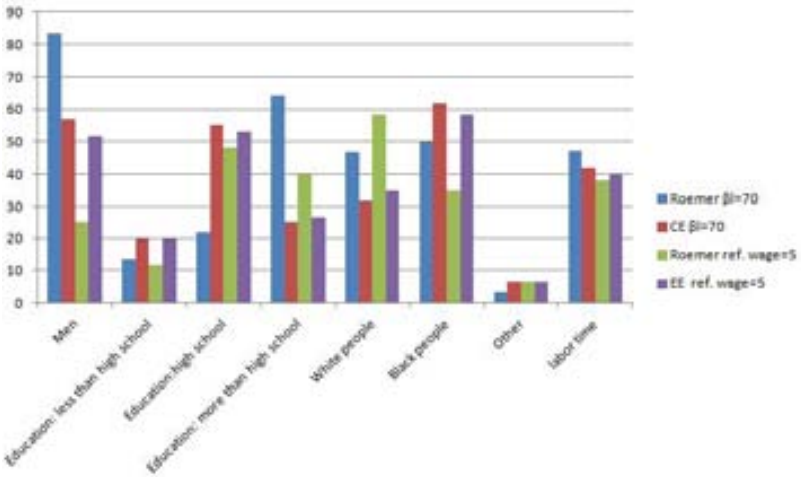
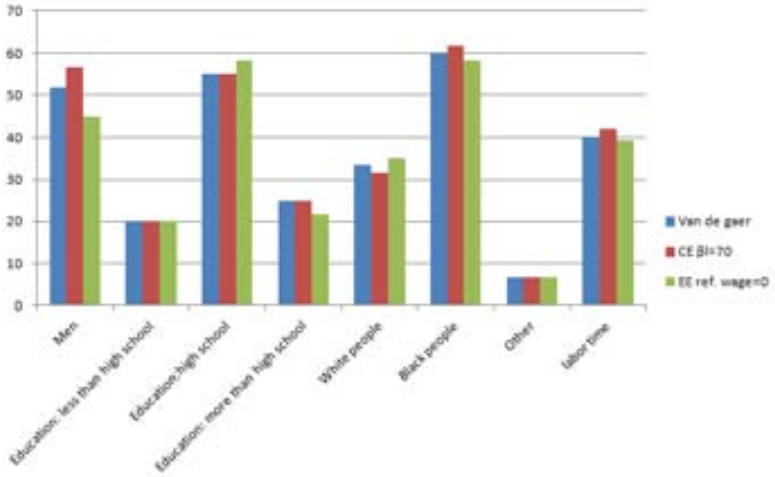


Figure 3.5: Van de gaer's criterion



When turning to Van de gaer's criterion, we find that the worst-off according to this criterion are the same regardless the cardinalization. In any case, they are the individuals who belong to the type with the lowest wage-rate. The composition of the worst-off are depicted in the Figure 3.5 and we observe that the composition of the worst-off coincides in a larger extent with the

one depicted by the conditional equality, reflecting the theoretical similarity between both criteria. Also, we obtain that if we took the mean observed net income, the individuals with the lowest wage rate would still be the worst-off.

3.4.4 Differences and Similarities among the criteria

Figure 3.6: Characteristics of the worst-off across criteria

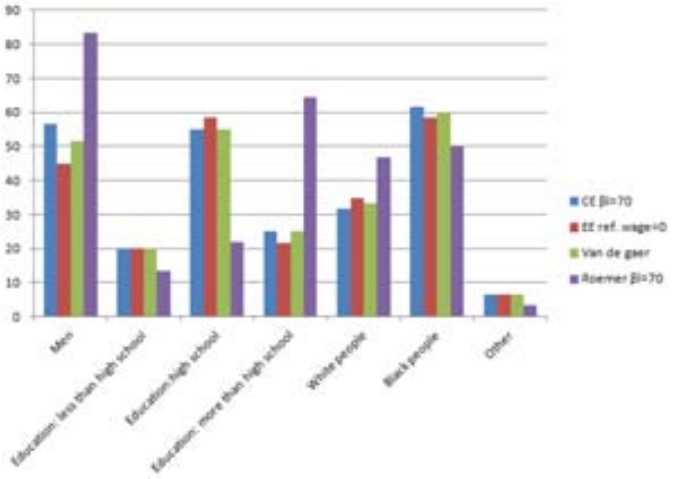


Figure 3.7: The worst-off according to each criterion

The final question we raise here is the following: Is there any worst-off common

to all the criteria? In other words, can we raise a unique conclusion about the persons who suffer more unfairness regardless the criteria of inequality of opportunity and the cardinalization we use?

To this end, we take the four criteria and for each criterion, we take one particular version of it. We compare the zero egalitarian equivalence with the conditional equality when β_l is equal to 70, with Van de gaer's criterion and with Roemer's criterion when the cardinal outcome is obtained by fixing a common preference for leisure equal to 70. In fact, the results do not vary significantly when we change the version of each criterion[10]. Then, we identify the worst-off according to each criterion and we compare how many individuals are common to the four criteria and/or to some of them.

To summarize the results, Figure 3.6 shows the characteristics of the worst-off across criteria and the Figure 3.7 displays the percentage of individuals who belong to multiple criteria simultaneously. We obtain that only 10% of the individuals are worst-off according to all the criteria.

This exercise confirms previous results on the difference between Van de gaer's and Roemers's criteria [46]. Here, the two criteria differ substantially even when we use the same cardinalization. This means that the distribution functions of well-being by type cross so that the individuals with the lowest wage rate are not always the worst-off when we use Roemer's criterion.

Regarding the difference between the two ordinal criteria, most of the worst-off are common to the conditional equality and egalitarian equivalence criteria. In fact, both criteria qualify unfair inequalities as due to the wage rate. This is why we obtain that the worst-off tend to be the individuals with low wage rate and this explains also why Van de gaer and the ordinal criteria coincide in a large extent. Indeed, the two ordinal criteria will exhibit more differences regarding the redistribution policy they recommend because each of them focuses on distinct principles (compensation versus neutrality). The difference is less obvious when we identify the worst-off.

Nevertheless, this exercise is instructive in order to evidence the impact of the normative choice we make when enabling interpersonal comparisons. Firstly, the fact that only 10% of the individuals are worst-off regardless the criteria show that the hypothesis that allows us to make interpersonal comparisons have a large impact on the identification of the proper target of any redistribution policy aiming at reducing unfairness. Secondly, the fact that the worst-off according to Roemer's criterion are so distinct from the ones identified by the other criteria show how much the definition of equality of opportunity matters. In fact, stating that equality of opportunity consists in

[10] robustness checks are available upon request.

erasing inequalities due to factors for which the individual is not responsible is not a precise enough definition. This confirms the difference between the ex-post and ex-ante approaches of equality of opportunity.

3.5 Conclusion

This paper proposes an empirical application of distinct fairness criteria. Our contribution is twofold: Firstly, we define a model that allows us to better approximate individuals' preferences such as to apply the criteria of conditional equality and egalitarian equivalence. Secondly, we propose to compare these criteria with the criterion proposed by Roemer and the one proposed by Van de gaer. In this way, we propose two methods for translating the ordinal model into a cardinal model. Each method derives from the normative choices raised by the criteria of conditional equality and egalitarian equivalence in order to make the comparisons as reliable as possible.

We apply our model to a sample of singles from the CNEF dataset and identify who are the worst-off according to each criterion. We find that our model makes possible refining groups' preferences to better approximate individuals' preferences. Then, we find that only 10% of the individuals are worst-off regardless of the criterion used.

We conclude (1) our model for identifying individuals' preferences may be improved but still offers a valuable method for better applying the ordinal criteria of equality of opportunity and (2) the discrepancy between the criteria in terms of the worst-off we identify shows how important are the normative choices on which rely interpersonal comparisons.

This first application leads to results of limited scope as we restrict our analysis to singles without children. A micro-simulation model would allow to enlarge our conclusions. Nevertheless, our results provide a first insight on the divergence between the criteria of fairness from an empirical perspective.

General Conclusion

This thesis studies (in)equality of opportunity from an empirical perspective. This field has considerably expanded during the last two decades since the publication of seminal works (Arneson, [6], Cohen [22], Dworkin [28, 29], Roemer [59]) that defend a responsibility-sensitive egalitarianism. From this work, it emerges an appealing concept of equal opportunities that combines a real concern for justice with the recognition of acceptable inequalities on one hand, and a diversity of measurement methods on the other hand (Almas et al. [4], Bjorklund et al. [10], Bourguignon et al. [14], Cogneau et al. [21], Devooght [27], Lefranc et al. [44], Peragine [52], Pistolesi [53], Van de gaer [61]). Moreover, the increasing availability of detailed information on socio-economic background of individuals made possible the application of this concept to assess the justice in the distribution of ressources. Thus, in this thesis we propose to use different approaches, to confront them with actual datasets to better understand to which extent the empirical conclusions rely on theoretical questions.

This thesis presents different perspectives on inequality of opportunity, by following three different methodologies (ex-ante, ex-post and an ordinal approach) and by analyzing three different datasets (France, Spain and the United States). Each chapter brings an original contribution on the measurement of unfairness that we detail below. In a nutshell, the first chapter proposes for Spain a finer understanding of the role of circumstances on individuals' outcomes by relying on a broader set of circumstances; the second chapter studies how regions can differ in the reward of effort and in the way circumstances shape inequality; the last chapter proposes a refinement to the methodology used to apply the ordinal approach of Fleurbaey and Maniquet [31, 34] and compares the results with ex-ante and ex-post approaches.

In the first chapter, we adopt the ex-ante view of equality of opportunity to measure the magnitude of inequality of opportunity in the acquisition of income in Spain. The ex-ante approach focuses on the impact of circumstances on outcomes prospects. This gives room to many measurement methods, which complicates the comparison of empirical studies. The aim of this chapter

is to measure to which extent the magnitude of inequality of opportunity is sensitive to changes in the measurement methods. We focus on three questions: we study by how much the magnitude of inequality of opportunity varies when we include one or several circumstances, when we account or not for the direct and indirect effect of circumstances on outcome, when we change the responsibility set.

To this end, we take the 2005 EU-SILC dataset for Spain and we propose a baseline specification where earnings depend on one circumstance and several effort variables. We equalize the circumstances among individuals and compute the reduction of inequality we would obtain. Then, we enlarge the circumstance set and compare with the baseline specification. We also use Bourguignon et al.[13, 14] to measure the direct and indirect effect of circumstances on earnings and finally we propose different responsibility cuts and compare the magnitude of inequality of opportunity according to these views of individual's responsibility.

We found that not only family background shape inequalities among individuals, but also gender, country of birth and genetics and that the magnitude of inequality of opportunity is sensitive to the inclusion of these variables. In this sense, enlarging the set of circumstance improves our understanding of the origin of unfair inequalities. We also distinguished between the direct effect of circumstances on outcomes and the indirect effect through effort and found that the indirect effect should be taken into account as it increases significantly the magnitude of inequality of opportunity. On the contrary, the change in the responsibility cut does not appear to be crucial because only the variables that are obviously out of the scope of individual's responsibility cause large changes in the magnitude of inequality of opportunity. Finally, when ranking the circumstances according to their impact on inequality, we found that family background remains the most impacting circumstance.

The second chapter also examines the determinants of inequality of opportunity, but we adopt here the ex-post approach, that is to say, we study whether individuals who exert the same effort obtain the same outcome. The application concerns the regions of France and the objective is to show whether (1) regions reward equally effort, (2) inequality of opportunity is equally distributed among regions and (3) it is correlated with income inequality. One novelty of the approach consists in applying for France a direct mesure of effort thanks to the 2005 EU-SILC dataset that include indicators of effort. And, to our best knowledge, we propose the first measure of inequality of opportunity for the regions of France.

To perform such an analysis, we first estimate an earning equation where income is a function of effort and circumstance variables. We include several

circumstance variables and account for the indirect effect of circumstances, as suggested in the first chapter. Effort is composed by education, sector of activity, experience and working hours. This is a minimal version of effort as residuals of the equation are included into the circumstance set. We opt for this in order to be consistent with our objective of measuring directly effort. Then, we follow Almas et al. [4] methodology to define a fair income that fulfills ex-post equality of opportunity requirements: this is an income that depends on effort variables only. In this way, unfairness is measured by an unfair Gini based on the distance between the actual income and the fair income. Finally, we use Oaxaca decomposition to further explore the sources of unfair inequalities across regions.

Our findings reveal that the impact of the non-responsibility factors on income differ across regions as well as the reward schemes (that is to say the effect of effort variables). For instance, the effect of education is in some regions three times larger than in other regions and the same is true for the effect of circumstances on income for which large differences are displayed concerning the impact of gender and father's education. Mainly, unfair inequalities are found not to be equally distributed across regions and this justifies our regional approach. This may even indicate that policies aiming to reduce inequality of opportunity should be designed at a decentralized level. Moreover, when using the Oaxaca decomposition to explain the difference between the average income in one region and the average income in the rest of the regions, we obtain that differences are not explained by the same factors. Some regions have a higher average income because of its endowments in circumstance, other for its endowments in effort variables. Finally, the ranking of regions in terms of ex-post inequality of opportunity is found to be close to the ranking in terms of income inequality which confirm previous results given in the literature.

Through these two chapters, we found that accounting for the different channels through which inequality of opportunity is generated permits to get a better understanding of the origin of unfair inequalities. In this way, the access to a large set of opportunity determinants is key to better measure inequality of opportunity. Moreover, these chapters tend to show that parents' education remains the main determinant of inequality of opportunity for income. This result indicates that the transmission of opportunities goes through the family first. As a result, it might be concluded that inequality of opportunity will not be tackled through any educational or labour market policy without improving first our understanding of the channels by which families transmit unequal opportunities. Another finding is that responsibility factors explain well inequalities. The effort displayed by individuals determine well their income, therefore the theory of equality of opportunity could justify a limited and precise scope for redistribution policies.

In the first two chapters, taking earnings as the measure of well-being allows us to make straightforward interpersonal comparisons since everybody is assumed to value equally the outcome. However, it is likely that individuals have heterogeneous preferences, and this makes more challenging the comparison of individuals' outcomes. Fleurbaey and Maniquet [31, 33, 34, 35] have defined a set-up that requires first identifying heterogeneous individual's preferences to obtain ordinal measure of well-being, and then permits to rank individuals' allocation according to principles of fairness. These principles are very close to the concept of equality of opportunity but applications are still scarce [4, 25, 27]. This explains our interest for developping, in the third chapter, an empirical strategy that allows us to apply the criteria developed by Fleurbaey and Maniquet.

The third chapter proposes a refinement of Decoster and Haan econometric model [25]. Their model makes possible the identification of groups preferences according to socio-demographic characteristics, but does not allow us to indentify individual's preferences. The refinement consists in using information on individuals' choice in terms of earnings and leisure to better approximate the individualistic component of preferences. Then, we apply the criteria of egalitarian equivalence and conditional equality in this new set-up to rank individuals from the worst-off to the better-off. In addition, we propose cardinalizations of the ordinal measures of well-being to apply Roemer's and Van de gaer's criteria that correspond to the ex-post and ex-ante approaches and compare the worst-off according to the four criteria.

Firstly, we find that our empirical strategy offers a valuable method for applying the ordinal criteria. Secondly, we find that only 10% of the worst-off are common to all criteria. This discrepancy between the criteria in terms of the worst-off shows how important are the normative choices on which rely interpersonal comparisons. This application leads to results of limited scope, since we restrict our sample to US singles from the CNEF dataset, but still it provides a first empirical result on the divergence between ordinal and cardinal criteria of fairness.

These three chapters offer a wide view on the current state of the research on inequality of opportunity. By striving to give answers to some key questions, many new interrogations arose during our work. We share here three important ones, which could call further research work.

First, responsibility factors, measured by effort or preferences, always depend on the non-responsibility factors set. This problem arose in the three chapters and required sensibility analyses and discussions. We believe that the dependence of responsibility factors on non-responsibility factors should be

more central in the development of any model or methodology investigating inequality of opportunity. Though this was out of the scope of our thesis projects, the importance of this question is explicit in each of our chapters. This thesis does not provide an answer, instead it aims to motivate a debate on this question.

Second, whatever the methodology, whatever the approach, strong normative choices underly any work in the field of inequality of opportunity. The most obvious example is the cut between circumstances and effort. Many variables are on the border, and considering them as effort or circumstances often changes dramatically the results. In addition, residuals often have the largest explanatory power. Considering them as circumstances is not anodine at all. Any work in this field is normative. We tried to make the normative assumptions explicit along the three chapters, in order to defend the view according to which any empirical analysis of equality of opportunity should be very explicit on the assumptions used to carry its results.

Finally, econometric models encompassing the distinct approaches would be very promising such as the one developed by Peragine [52, 19] since they build links in a field where empirical studies are hardly comparable. Even though we proposed in the last chapter a first way out for comparing the ordinal criteria with the cardinal criteria, an unifying model could be of strong appeal to better grasp the differences and similarities between all approaches.

Bibliography

[1] Aaberge R., Colombino U., Strom S.: *Do more equal slices shrink the cake? An empirical investigation of tax-transfer reform proposals in Italy*, Journal of Population Economics, 17(4), 767-785 (2004)

[2] Aaberge R., Colombino U., Strom S. : *Joint Labour Supply of Married Couples: Efficiency and Distribution Effects of Tax and Labour Market Reforms*, Labor and Demography, 0501004 EconWPA (2005)

[3] Aaberge R., Mogstad M., Peragine V.: *Measuring Long-Term Inequality of Opportunity*, ECINEQ Working Paper, 158 (2010)

[4] Almas I., Cappelen A.W., Lind J.T., Sorensen, E.O., Tungodden B.: *Measuring unfair (in)equality*, Journal of Public Economics, 95(7-8), 488-499 (2011)

[5] Angrist J.D., Krueger A.B.: *Does compulsory school attendance affect schooling and earnings?*, Quarterly Journal of Economics, 106, 979-1014 (1991).

[6] Arneson R.J.: *Equality of opportunity for welfare*, Philosophical Studies, 56, 77-93 (1989)

[7] Becker G.S.: *Human Capital: A Theoretical and Empirical Analysis, with special Reference to Education*, New York: Columbia University Press (1967)

[8] Behrman J.R and Rosenzweig M.R.: *"Ability biases in schooling returns and twins: a test and new estimates*, Economics of Education Review, (1999)

[9] Betts J.R., Roemer J.E.: *Equalizing Opportunity for Racial and Socioeconomic Groups in the United States Through Educational Finance Reform*, in Ludger Woessmann and Paul E. Peterson (Eds.) Schools and the Equal Opportunity Problem, Cambridge, MA: MIT Press (2007)

[10] Bjorklund A., Jantti M., Roemer J.E.: *Equality of Opportunity and the Distribution of Long-Run Income in Sweden*, Institute for the Study of Labor (IZA), Discussion Papers 5466 (2011)

[11] Bossert W., Fleurbaey M.: *Redistribution and compensation*, Social Choice and Welfare, 13(3), 343-355 (1996)

[12] Bound J., Jaeger D. A., Baker R.M.: *Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak*, Journal of the American Statistical Association, 90(1), 443-450 (1995)

[13] Bourguignon F., Ferreira G.H., Menendez M.: *Inequality of Outcomes and Inequality of Opportunities in Brazil*, DELTA (Ecole normale superieure) Working Papers 24 (2003)

[14] Bourguignon F., Ferreira G.H., Menendez M.: *Inequality Of Opportunity In Brazil*, Review of Income and Wealth, 53(4), 585-618 (2007)

[15] Calsamiglia C.: *Decentralizing Equality of Opportunity*, International Economic Review, 50(1) (2009)

[16] Cameron A.C., Trivedi P.K.: *Microeconometrics: Methods and Applications*, Cambridge University Press, New York (2005)

[17] Card D.: *The Causal Effect of Education on Earnings*, Handbook of Labor Economics, Elsevier Science, 3 (1999)

[18] Card D.: *Estimating the return to schooling: Progress on some persistent econometric problems*, Econometrica, 69, 1127-1160 (2001)

[19] Checchi D., Peragine, V., Serlenga L.: *Fair and Unfair Income Inequalities in Europe*, Institute for the Study of Labor (IZA), 5025 (2010)

[20] Checchi D., Peragine V.: *Regional Disparities and Inequality of Opportunity: The Case of Italy*, Institute for the Study of Labor (IZA), Discussion Papers 1874 (2005)

[21] Cogneau D., Mesple-Somps S. : *Inequality of Opportunity for Income in Five Countries of Africa*, DIAL WP 04 (2008)

[22] Cohen GA : *On the currency of egalitarian justice*, Ethics 99, 906-944 (1989)

[23] Combes P-P., Lafourcade M., Thisse J-F., Toutain J-C : *The rise and fall of spatial inequalities in France: A long-run perspective*, Explorations in Economic History, Elsevier, 19(2), 243-271 (2011)

[24] Dearden L.: *The effects of families and ability on men's education and earnings in Britain*, Labour Economics, Elsevier, 6(4), 551-567 (1999)

[25] Decoster A., Haan P.: *Empirical Welfare Analysis in Random Utility Models of Labour Supply* IZA Discussion Papers, 5301 (2010)

[26] Devaux M., Jusot F., Tubeuf S., Trannoy A. : *Inequality of opportunities in health in France: a first pass*, Health Economics, 19(8), 921-938 (2010)

[27] Devooght K.: *To Each the Same and to Each His Own. A Proposal to Measure Responsibility-Sensitive Income Inequality*, Economica, London School of Economics and Political Science, 75(298), 280-295 (2005)

[28] Dworkin R.: *What is Equality: Part 1: Equality of Welfare*, Philosophy and Public Affairs, 10(3), 186-245 (1981)

[29] Dworkin R.: *What is Equality: Part 2: Equality of Resources*, Philosophy and Public Affairs, 10(3), 283-345 (1981)

[30] Ferreira G.H., Gignoux J.: *The measurement of inequality of opportunity : theory and an application to Latin America*, The World Bank, Policy Research WP4659 (2008)

[31] Fleurbaey M.: *Fairness, Responsibility, and Welfare*, Oxford: Oxford University Press, (2008)

[32] Fleurbaey M.: *Four approaches to equal opportunities*, in Fairness and redistributive Justice, Oxford University Press, (2011)

[33] Fleurbaey M., Maniquet F.: *Fair Income Tax*, Institute for Advanced Study, School of Social Science, Economics Working Papers 0021 (2002)

[34] Fleurbaey M., Maniquet F.: *Fair social orderings when agents have unequal production skills*, Social Choice and Welfare, 24(1), 93-127 (2005)

[35] Fleurbaey M., Maniquet F.: *Compensation and responsibility*, Working Papers, HAL (2006)

[36] Fleurbaey M., Schokkaert E., Decancq, K.: *What good is happiness?*, CORE Discussion Papers 2009, Universite catholique de Louvain, Center for Operations Research and Econometrics (CORE) (2009)

[37] Forsé M., Parodi M.: *La perception des Inégalités en France depuis dix ans*, Revue de l'OFCE 118 (2011)

[38] Godechot O.: *Finance and the rise in inequalities in France*, Paris School of Economics WP 13 (2011)

[39] Hild M.: *Roemer on Equality of Opportunity*, Social Science Working Paper 1128, California Institue of Technology. (2001)

[40] Jones A., Rice N., Rosa Dias P.: *Quality of Schooling and Inequality of Opportunity in Health*, Empirical Economics, forthcoming (2011)

[41] Jusot F., Tubeuf S., Trannoy A. : *Effort or Circumstances: Does the Correlation Matter for Inequality of Opportunity in Health*, IRDES Document de travail 33 (2010)

[42] Lefranc, A., Trannoy, A.: *Intergenerational earnings mobility in France: Is France more mobile than the US?*, IDEP Working Papers 0401 (2004)

[43] Lefranc, A., Pistolesi, N., Trannoy, A.: *Inequality of Opportunities versus Inequality of Outcomes: Are Western Societies all alike?*, Review of Income and Wealth, 54(4), 513-546 (2008)

[44] Lefranc A., Pistolesi N., Trannoy A.: *Equality of Opportunity and luck: Definitions and testable conditions, with an application to Income in France*, Journal of Public Economics, 93, 1189-1207 (2009)

[45] Martin P.: *The geography of inequalities in Europe*, Swedish Economic Policy Review, 12, 83-108 (2005)

[46] Martinez M., Schokkaert E., Van de Gaer D. : *Measuring Intergenerational Mobility and Equality of Opportunity*, Public Economics Working Paper Series (1998)

[47] Moreno-Ternero J.D.: *On Roemer's Equality of Opportunity*, Working Paper, Yale University (2004)

[48] Niaz A., Yalonetzky G.: *Inequality of Educational Opportunity in India: Changes over Time and across States*, IZA Discussion Papers 5146 (2010)

[49] Nilsson W.: *Opportunities, Preferences and Incomes*, Umea Economic Studies 649 (2005)

[50] Oaxaca, R.: *Male-Female Wage Differentials in Urban Labor Markets*, International Economic Review, Department of Economics, University of Pennsylvania and Osaka University Institute of Social and Economic Research Association, 14(3), 693-709 (1973)

[51] Ooghe, E, Schokkaert, E., Van de gaer, D.: *Equality of Opportunity versus Equality of Opportunity Sets?*, Social Choice and Welfare, 28(2), 209-230 (2007)

[52] Peragine V.: *Measuring and implementing equality of opportunity for income*, Social Choice and Welfare, 22(1), 187-210 (2004)

[53] Pistolesi N.: *Inequality of opportunity in the land of opportunities, 1968-2001?*, Journal of Economic Inequality, 7(4), 411-433 (2009)

[54] Prieto-Rodriguez J., Rodriguez-Gutierez C.: *Participation of Married Women in the Labour Market and the "Added Worker Effect" in Europe*, WP Series 2000-12 IRISS (2000)

[55] Rawls J.: *Justice as Fairness*, Contemporary Political Philosophy, Blackwell Philosophical Anthologies, Goodin and Pettit eds. (1997)

[56] Rodriguez J.G.: *Partial and complete equality-of-opportunity orderings*, ECINEQ, Society for the Study of Economic Inequality, WP 70 (2007)

[57] Roemer J.E.: *A Pragmatic Theory of Responsibility for the Egalitarian Planner*, Philosophy and Public Affairs, 22(2), 146-166 (1993)

[58] Roemer J.E.: *Equality of Opportunity*, Harvard: Harvard University Press (1998).

[59] Roemer J.E.: *To What Extent Do Fiscal Regimes Equalize Opportunities for Income Acquisition among citizens?*, Journal of Public Economics, 87, 539-565 (2003)

[60] Sen R.: *Inequality Reexamined*, Oxford: Oxford University Press (1992)

[61] Van de gaer D.: *Investment in Human Capital and Equality of Opportunity*, doctoral dissertation, KULeuven, Leuven (1993)

[62] Woessman L.: *How equal are Educational Opportunities? Family Background and Student Achievement in Europe and the United States*, IZA Discussion Paper No. 1284 (2004)