NBER WORKING PAPER SERIES

THE CAUSES AND EFFECTS OF INTERNATIONAL MIGRATIONS: EVIDENCE FROM OECD COUNTRIES 1980-2005

Francesc Ortega
Giovanni Peri

Working Paper 14833
http://www.nber.org/papers/w14833

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2009

We are thankful to Greg Wright and Tommaso Colussi for excellent research assistance. Peri gratefully acknowledges generous funding from the John D. and Catherine T. MacArthur Foundation. This paper was commissioned as background research study for the United Nation Human Development Report, 2009. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

The Causes and Effects of International Migrations: Evidence from OECD Countries 1980-2005
Francesc Ortega and Giovanni Peri
NBER Working Paper No. 14833
April 2009
JEL No. E25,F22,J61

ABSTRACT

This paper contains three important contributions to the literature on international migrations. First, it compiles a new dataset on migration flows (and stocks) and on immigration laws for 14 OECD destination countries and 74 sending countries for each year over the period 1980-2005. Second, it extends the empirical model of migration choice across multiple destinations, developed by Grogger and Hanson (2008), by allowing for unobserved individual heterogeneity between migrants and non-migrants. We use the model to derive a pseudo-gravity empirical specification of the economic and legal determinants of international migration. Our estimates clearly show that bilateral migration flows are increasing in the income per capita gap between origin and destination. We also find that bilateral flows decrease when destination countries adopt stricter immigration laws. Third, we estimate the impact of immigration flows on employment, investment and productivity in the receiving OECD countries using as instruments the "push" factors in the gravity equation. Specifically, we use the characteristics of the sending countries that affect migration and their changes over time, interacted with bilateral migration costs. We find that immigration increases employment, with no evidence of crowding-out of natives, and that investment responds rapidly and vigorously. The inflow of immigrants does not seem to reduce capital intensity nor total factor productivity in the short-run or in the long run. These results imply that immigration increases the total GDP of the receiving country in the short-run one-for-one, without affecting average wages and average income per person.

Francesc Ortega
Department of Economics and Business
University Pompeu Fabra
Ramon Trias Fargas 25-27
Barcelona, 08005, Spain
francesc.ortega@upf.edu

Giovanni Peri
Department of Economics
University of California, Davis
One Shields Avenue
Davis, CA 95616
and NBER
gperi@ucdavis.edu

1 Introduction

The present paper advances the literature on the economic determinants and effects of international migrations. We make three main contributions. First, we gather and organize annual data on bilateral immigration flows from 74 countries of origin into 14 OECD countries from 1980 to 2005 and on immigration laws in those OECD countries in order to analyze the economic and legal determinants of migration flows. We first update the data used in Mayda (forthcoming) from the OECD international migration statistics. These data were discontinued in 1994. For the period 1995-2005 it has been substituted with a new database on immigration flows and stocks in OECD countries.[1] We merge these two datasets on flows covering the period 1980-2005 with data on the stock of immigrants residing in the 14 OECD destination countries from the same 74 countries for the period 1990-2000. This also allows us to impute the "net" migration flows to the OECD countries–that is, immigration net of re-migration out of the country. For the same 14 OECD countries we also collect, organize, and classify information on immigration laws, distinguishing between laws regulating *entry*, *stay*, *asylum*, and a few specific multilateral treaties with implications for international labor mobility. The richness of our data allows us to control for a very large set of fixed effects when analyzing the determinants of bilateral flows. Furthermore, it allows us to identify the effects of economic variables and immigration laws using variation by destination country over time only.

The second contribution is that we use an empirical "generalized gravity equation", derived from a model in which potential migrants maximize utility by choosing where to migrate. We use such a model to estimate the effects of variation in geographic, economic and policy variables in the destination countries on immigration flows. Our empirical model adapts and generalizes the one proposed in Grogger and Hanson (2007, 2008). In contrast to them, however, we do not focus (as they do, following Borjas, 1987) on the selection of immigrants according to skills but rather on the total size (scale) of bilateral migration flows. On the other hand, we allow for a more general empirical specification that is consistent with several different discrete choice models (simple logit as well as nested logit) and requires only data on bilateral stocks (or flows) of migrants in order to be implemented. Importantly, we allow for unobserved individual heterogeneity between migrants and non-migrants. Also, since we have data on bilateral flows over time we can control for unobserved, time-varying, sending-country characteristics and focus mainly on income per person, employment, and immigration policies in the *destination countries* as determinants of migrations.

Third, and most importantly, we can identify the *aggregate effects* of these immigrant flows on the economy of the receiving country, specifically on total employment, total hours worked, physical capital accumulation and total factor productivity. While the recent literature on the impact of immigrants on labor markets (Borjas and Katz 2007, Ottaviano and Peri 2008) acknowledges that the country is the appropriate unit with which

[1] Publicly available at http://stats.oecd.org/wbos/Index.aspx?datasetcode=MIG.

to analyze such effects (due to the high degree of mobility of workers and capital within a country) there are extremely few cross-country (or panel) studies of those effects. The reason is that in order to do this one needs to overcome two problems. First, we need to gather consistent, yearly data on hours worked, employment, capital stock for each of the 14 OECD countries of destination, over the period 1980-2005. Second, we need to isolate the impact of immigration on those variables when we know that productivity, investment and employment growth are also determinants of immigration flows (through their effects on income and wages). We address the first issue by employing data from different OECD datasets, while to solve the second issue we use our bilateral migration equation estimated below. Restricting the explanatory variables of the bilateral migration flows to factors specific to the country of origin and to bilateral costs only, we obtain a predicted flow of migrants to OECD countries that can be used as an instrument, since it isolates the push-driven flows. Those flows vary across country of destination due to the different bilateral costs (due to geography and networks) of migrating from one country to another, which are independent of any destination country variable. For instance, a boom in emigrants from Poland due to the opening of its border is more likely to generate large migration to Germany than to Canada (for geographical and historical reasons), while a boom of emigrants from the Philippines is more likely to generate large immigration to Japan (proximity) and the US (previous networks) than to France. Using such push-driven flows we track their effects on the employment, capital and productivity of the receiving countries.

The paper has three main findings. First, confirming previous literature (e.g. Mayda, forthcoming), our regressions consistently show that differences in the level of income per person between the destination and origin country have a positive and significant effect on bilateral migration flows. An increase in the gap by 1000 PPP$ (in 2000 prices) increases bilateral migration flows by about 10% of their initial value. Also, we find that stricter entry laws significantly discourage immigration. Each reform which introduced tighter *rules of entry* for immigrants decreased immigration flows by about 6% on average. Second, we find that time-varying push factors specific to countries of origin and interacted with bilateral fixed costs of migration, predict a significant share (between 30 and 40%) of the variation in migration to the OECD receiving countries. Such variation of immigration flows for a receiving country over time can legitimately be consider as "exogenous" to the economic and demographic conditions of the receiving country. Third, consistent with an increase in the labor supply in the neoclassical growth model with endogenous capital adjustment, we find that the “exogenous” inflow of immigrants increases *one for one* employment, hours worked and capital stocks in the receiving country, implying no crowding-out of natives and a speedy and full adjustment of capital. Hence, even in the short run (one year), the capital-labor ratio at the national level fully recovers from an immigration shock. We note that in most instances, immigration flows are only a fraction of a percentage point of the labor force of the receiving country. Moreover, the largest part of these flows is easily predictable, implying that full capital adjustment is a very reasonable finding even in the short run. Also, immigration does not seem to have any significant effect on total factor productivity. These effects, taken together, imply no significant effect of immigration on average wages and on the return to capital in the receiving countries. Instead, immigration shocks lead to an increase in total employment and a proportional response of GDP.

The rest of the paper is organized as follows: section 2 reviews the existing literature on the determinants and effects of international migrations and puts the contribution of this paper into perspective. Section 3 describes and presents the data, especially those on migration flows and immigration laws. Section 4 justifies the empirical model used to analyze the determinants of bilateral migrations and estimates the effect of income differences (between sending and receiving country) and immigration laws (in destination countries) on bilateral flows. Section 5 presents the estimates of the effect of immigration on employment, physical capital accumulation and productivity of the receiving country. Using an instrumental variable approach which isolates only the push-driven part of immigrant flows, exogenous to the economic conditions of destination countries, we can provide a causal interpretation of the estimated effect. Section 6 discusses the main implications of our findings and provides some concluding remarks.

2 Literature Review

This paper contributes to two strands of the literature on international migration that, so far, have developed separately. One analyzes the determinants of international migrations (mostly by international economists) and the other analyzes the impact of immigration on the receiving countries (mostly by labor economists and limited to labor market effects). On the first front we improve on the existing literature regarding the determinants of bilateral migrations by applying a simple model of optimal choice similar to Grogger and Hanson (2008) as the basis of our estimating equation. A large part of the literature on migration flows had previously either estimated a gravity or "pseudo-gravity" equation between many origins and one destination (e.g. Clark et al 2008, Karemera et al 2000, Pedersen et al 2004) with no foundation in the individual choices of migrants. Other papers have derived predictions on the selection of migrants from a Roy model and estimated some of its implications (Borjas 1987, Dahl 2002). Recently, Grogger and Hanson (2008) have analyzed the scale, selection and sorting across destinations of migrants with different education levels using a model based on optimal discrete choice. Their contribution is part-way between the theory of optimal choice and an empirical, pseudo-gravity equation. In particular, their specification for the "scale" of migration uses as the dependent variable the difference between the logs of the odds of migrating to a specific country and the odds of not migrating at all.

Gravity regressions have become very popular in analyzing trade flows (Anderson and Van Wincoop 2003, Chaney 2008 and Helpman, Melitz and Rubinstein, forthcoming) primarily because they can be derived from an equilibrium model with optimizing firms. Building on Grogger and Hanson (2008), we employ an extension of their model that allows for unobserved individual heterogeneity between migrants and non-migrants in order to derive an empirical specification that is fully consistent with a generalized gravity model. Unlike them we do not distinguish between education groups. The model delivers an equation in which the log of bilateral migration (stocks or flows) is a function of sending and receiving country effects, expected income differentials and migration costs. Moreover, this pseudo-gravity equation can be seen as the result of a simple multinomial logit model in which the migrant makes a comparison between migrating to any other country or staying at home, assuming bilateral and destination-specific migration costs. The empirical specification can also be derived from a more general nested logit model in which migrants first decide whether to migrate and then decide among the potential destinations. Importantly, the nested logit model allows for unobserved individual heterogeneity between migrants and non-migrants or, equivalently, for idiosyncratic shocks that may be correlated across destinations.

We test the predictions of the model with aggregate panel data on stocks and flows of migrants. Our empirical specification allows us to focus on the determinants of migration in the destination countries (while fully controlling for any factor depending on country of origin and year). Another contribution of this paper (with the exception of Mayda, forthcoming) is the careful analysis of the effects of *immigration laws* on *immigration flows*.[2] In this respect, we present new data on several hundred immigration reforms in the 14 OECD countries analyzed. Following some mechanical rules and by reading carefully the content of these laws we classify them based on whether they tighten the requirements to enter or stay in the country, separating laws that concern asylum seekers from laws dealing with other types of immigrants. The effects of these laws on subsequent immigration flows turn out to be quite significant, especially in the case of entry laws, and precisely estimated. Our dataset on immigration laws over the 1980-2005 period, documented in the "Immigration Reform Appendix", may become an important point of reference toward building a systematic classification of immigration laws across OECD countries. In particular, we hope our data stimulates the literature on the determinants of immigration policy that so far has remained mainly theoretical (Benhabib 1996, Ortega 2005) for lack of data measuring the “tightness" of immigration policies.[3]

The second part of this paper analyzes the impact of migration on the employment, investment and productivity of the receiving country using a panel of 14 countries over time. Most of the existing papers tracking the impact of immigration focus only on labor market implications and on one or only a few receiving countries (e.g. Aydemir and Borjas 2007, Borjas 2003, Ottaviano and Peri 2008, Manacorda et al. 2006). Angrist and Kugler (2005) use a panel of European countries and analyze the labor market effects of immigration. Related to this

[2] See also Bertocchi and Strozzi (2008) for a historical analysis of the effects of institutions on migration flows for a reduced number of countries.

[3] A notable exception is Bertocchi and Strozzi (2010) that looks at the economic and demographic determinants of citizenship laws.

paper, Peri (2008) and Ortega (2008) analyze the effects of immigration on employment, capital accumulation and productivity, respectively, across US states and Spanish regions. The literature on the aggregate effects of migration using cross-country panel analysis is extremely scant. In particular, there are no estimates, so far, of the effect of immigration on total employment, capital accumulation or productivity based on country level data. Two major reasons that such analysis has not been performed are that consistent data on migration across countries and over time are hard to find and, since immigration is endogenous to income levels and to their changes, the lack of plausible instruments has limited the ability to draw any inference on the effect of immigration on national income. This paper addresses both issues, providing estimates of the effects of immigration on aggregate employment, the capital stock, productivity and, consequently, income per capita at the country level. Hence, though the paper builds on a rigorous model which can explain migration flows, the main contribution is to estimate the aggregate impact of these flows on the receiving economies.

3 Data

This section describes the data that are novel to this paper, namely those on yearly migration flows into 14 OECD countries over the period 1980-2005 and those on immigration laws and reforms in the same countries over the same period.

3.1 Migration Flows

The data on yearly migration flows come from the International Migration Dataset (IMD) provided by the OECD. Data for the period 1980-1995 relative to 14 OECD destination countries and for close to 80 countries of origin were collected and organized by Mayda (forthcoming)[4]. We merged these data with the new data relative to the period 1995-2005 for 25 OECD receiving countries and more than one hundred sending countries, available at OECD (2007). In order to obtain a balanced and consistent panel we select 14 OECD destination countries[5] and 74 countries of origin (listed in table A1 of the Appendix). The data on migration flows collected in the IMD are based on national statistics, gathered and homogenized by the OECD statistical office[6]. The national data are based on population registers or residence permits. In both cases these are considered to be accurate measures of the entry of legal foreign nationals. We consider the data relative to the total inflow of foreign persons, independently of the reason (immigration, temporary or asylum). While the OECD makes an effort (especially since 1995) to maintain a consistent definition of immigrants across countries, there are some

[4] We refer to Mayda (forthcoming) for specific descriptions of the data relative to the 1980-1995 period. The source (OECD International Migration Data) and the definitions, however, are the same as those provided by the OECD for the statistics relative to the 1995-2005 period. Hence, we simply merged the two series.

[5] Australia, Belgium, Canada, Denmark, France, Germany, Japan, Luxembourg, Netherlands, Norway, Sweden, Switzerland, UK and USA.

[6] More details on the immigration data and their construction is provided in Appendix A.

differences between destination country definitions. An important one is that some countries define immigrants on the basis of the place of birth, and others on the basis of nationality. While this inconsistency can make a pure cross-country comparison inaccurate, our analysis focuses on changes within destination countries over time. Therefore it should be exempt from large mis-measurement due to the classification problem. The total inflow of foreign persons each year for each country of destination, as measured by these OECD sources, constitutes what we call total (gross) immigration. We also construct a measure of total *net* immigration for each receiving country. In this measure we try to correct for the outflow of foreign persons, due to re-migration or return migration. [7] Those flows, however, are harder to measure as people are not required to communicate to the registry of population their intention to leave the country. Hence we infer the net immigration flows using the gross immigration data and the data on immigrant stocks (by country of origin) from Docquier (2007) for 29 OECD countries in years around 1990 and around 2000. Therefore, for each of our 14 countries of destination we know the yearly inflow and the stock circa years 1990 and 2000. For each receiving country we impute a yearly out-migration rate of the stock of immigrants that, using the stock in 1990 and the measured yearly flows between 1990 and 2000, would produce the measured stock in 2000[8]. We apply this constant, destination-specific, re-migration rate to all years and obtain the stock of immigrants each year (between 1980 and 2005) and the net immigration rates each year. Panel A1 in the Appendix reports the gross and net immigration rates (i.e. immigration flows as a percentage of the population at the beginning of the year) for our 14 destination countries over the 25 years considered. For most countries gross and net immigration rates are similar and move together over time. We note that our net immigration rates are probably much less precise than our measures of gross immigration. Recall that we assumed constant re-migration rates for all years, while gross immigration flows and re-migration rates are likely to be correlated[9]. Second, any difference between stocks and flows could also be due to undocumented immigration, their somewhat different classification systems, or other discrepancies, rather than to re-migration only. Third, for some countries the implied re-migration rate is extremely high and not very plausible[10]. Hence, while we will use the net immigration flows to check some regression results (see Table 3 and 5) the preferred specifications which analyze the impact of immigration on the receiving economy will be based on gross inflows of immigrants.

A preliminary look at Panel 1 reveals two facts. First, immigration rates have displayed an increasing trend in many countries but for some countries, such as the US and Germany, they peaked in the middle of the period (corresponding to the regularization of the late 1980s for the US and to immigration from the East in the early 1990s in Germany). Therefore it is hard to establish a common trend of immigration flows over time. Second,

[7] This phenomenon can be significant–depending on the country, we estimate that every year between 0.5 and 10% of the existing stock of migrants will migrate out.

[8] This procedure is like finding the unknown "depreciation rate" when we have a measure of a stock variable in 1990 and 2000 and a measure of yearly flows between them.

[9] Coen-Pirani (2008) analyzes migration flows across US states. He finds that gross inflow and outflow rates are strongly, positively correlated.

[10] Appendix A reports the calibrated re-migration rates for each country of destination.

there is a lot of idiosyncratic fluctuation in immigration rates across countries. Hence, in principle, the variation within country over time is large enough (and independent across countries) to allow us to identify the effects of immigration on employment, capital accumulation and TFP. Table A2 in the Appendix reports the summary statistics and the data sources for the other economic and demographic variables in the empirical analysis. Note that the average GDP per person was more than double in the receiving countries relative to the countries of origin in each year; furthermore, the employment rate was also consistently higher and income inequality (Gini coefficient) consistently lower in the countries of destination. Countries of destination also typically had a lower share of young persons in their population, reflecting the fact that most international migration is by young workers from countries where they are abundant to countries where young workers are scarce.[11]

3.2 Immigration Laws

An important contribution of this paper is the updating of a database on immigration laws for the 14 OECD countries in our sample and the codification of a method to identify an immigration reform as increasing (+1) or decreasing (-1) the tightness of immigration laws. The starting point for the database is the laws collected by Mayda and Patel (2004) and the Fondazione Rodolfo DeBenedetti (FRDB) Social Reforms database (2007). Mayda and Patel (2004) documented the main characteristics of the migration policies of several OECD countries (between 1980 and 2000) and the year of changes in their legislations. The FRDB Social Reforms Database collects information about social reforms in the EU15 Countries (except Luxembourg) over the period 1987-2005.

We merged and updated these two datasets obtaining the complete set of immigration reforms in the period 1980-2005 relative to all the 14 OECD countries considered, for a total of more than 240 laws. The list of immigration laws by country and year and a brief description of what each of them accomplished can be found in the "Immigration Reform Appendix" to the paper[12]. We then constructed three separate indices of "tightness" for every reform mentioned in the database. The first index includes only those measures tightening or loosening the "entry" of non-asylum immigrants. The second is a more comprehensive index that includes measures tightening or relaxing provisions concerning the entry and/or the stay of non-asylum immigrants. The third is an index that includes changes in immigration policy concerning the entry and/or the stay of asylum seekers only. In general, we consider as "loosening" entry laws (implying a change in the tightness variable of -1) those reforms that (i) lower requirements, fees or documents for entry and to obtain residence or work permits or (ii) introduce the possibility or increase the number of temporary permits. We consider as a loosening in stay laws those legal changes that (iii) reduce the number of years to obtain a permanent residence permit and those that (iv) foster the social integration of immigrants. On the other hand, a reform

[11] The other variables used in the bilateral regressions are Log Distance, Border, Common Language and Colony dummies and are taken from Glick and Rose (2001).

[12] Available at the website: http://www.econ.ucdavis.edu/faculty/gperi/Papers/immigration_reform_appendix.pdf

is considered as tightening entry laws (+1 in the variable capturing tightness of entry) if (i) it introduces or decreases quotas for entry, and (ii) increases requirements, fees or documents for entry and to obtain residence or work permits. It is considered as tightening the stay-laws if (iii) it raises the number of years to obtain a permanent residence permit/citizenship or (iv) it introduces residence constraints. We also apply the same definitions for the tightening of entry and stay to asylum seekers in order to produce tightness variables for this group. In spite of these rules there are several reforms that do not explicitly fit any of the categories above. In those cases we classified them as "loosening" or "tightening", or no change, by scrutinizing the content of each regulation. [13]

Panel A2 in the Appendix plots the variables for immigration policy tightening with respect to entry for immigrants (solid lines) and asylum seekers (dashed lines) for each of the 14 countries of destination. The initial value of each variable in each country is 0. Hence the variables only capture the variation in laws over time within a country. In the regressions which include the bilateral migration flows we always include a country of destination effect which captures initial cross-country differences in tightness of immigration laws. A preliminary inspection of the variables reveals that countries such as Australia, Germany, Luxembourg, Sweden and Canada significantly loosened their entry laws beginning around 1990, (with less of a change for their asylum laws). Denmark and Japan tightened their entry laws. The US loosened its immigration policy regarding entry during the eighties and nineties and tightened policy beginning around 2000. The remaining countries did not change the tightness of their immigration policies regarding entry very much. As it is hard to detect any clear correlation between the change in laws over time and the change in immigration flows, we move to more formal regression analyses of the determinants of bilateral migration flows, basing the estimating equation on a simple theory of the discrete choices of migrants.

4 Determinants of Immigration

This section presents a model of migration choice across multiple locations and derives an estimating equation from the model. Our estimating equation is consistent both with a simple logit model (McFadden, 1974) as well as with a nested logit model (McFadden, 1978). Our migration model extends Grogger and Hanson (2007, 2008) by allowing for unobserved individual heterogeneity between migrants and non-migrants. Potentially, this is an important omission. It is plausible that migrants systematically differ from non-migrants along important dimensions that are hard to measure, such as ability, risk aversion, or the psychological costs of living far from home. An additional attractive feature of our empirical specification is that it is reminiscent of a generalized gravity equation in which the logarithm of bilateral migration flows is a function of origin and destination

[13] Three research assistants read the laws and provided us with a brief summary of each law. These summaries were read by the two authors and discussed until converging on the sign of the policy change.

country fixed effects and bilateral migration costs.

4.1 Migration model

Following Grogger and Hanson (2007, 2008), we study the problem of a potential migrant that makes a utility-maximizing migration decision among multiple destinations. Agent i, in country of origin $o \in O$, decides whether to stay in o or to migrate to any of $d \in D = \{1, ..., D\}$ potential destination countries.

The utility from a given destination d depends on the potential migrant's expected permanent value of labor income in that country and on the costs associated with migrating to d. Specifically, individual i's utility (net of costs) associated with migrating from country of origin o to country d is given by:

$$U_{odi} = \delta_{od} - v_{odi} = f(\overline{W}_d) - g(C_{od}) - v_{odi}, \tag{1}$$

where δ_{od} is a country-pair-specific term shared by all individuals migrating from the same origin to the same destination, and v_{iod} is individual-specific. In particular, the term $\overline{W}_d$ is the permanent expected earnings of individual i in country d and C_{od} is the cost of migration, which may include destination-specific terms and bilateral costs that vary by country pair.

We assume separability between costs and benefits of migration. We also assume that the average expected labor income in the country of destination $\overline{W}_d$ can be decomposed into the product of the probability of employment in that country (p_d) times the average wage when employed (W_d). We explicitly allow migration costs to depend on specific destination country factors θ_d (such as immigration laws), and on specific bilateral country factors X_{od} (such as geographical or cultural distance). We normalize the average expected utility from not migrating (remaining in o) $f_1(p_o W_o)$ to zero. Obviously, migration costs are zero for individuals that choose to stay in the country of origin.

We also assume that f and g are increasing functions. If these functions are approximately linear, we can interpret them as monetary costs that reduce expected income. If f and g are better approximated by logarithmic functions then migration costs can be viewed as time costs, which can be subtracted from log real wages. Grogger and Hanson (2008) argue that their estimation results are inconsistent with utility maximization under logarithmic f and g, implying that the logarithmic model is mis-specified and produces omitted variable bias[14]. To keep our estimates comparable to theirs we proceed by assuming that functions f and g are approximately linear. Hence, we can write (1) as:

$$U_{odi} = f_1(p_d W_d) - g_1 \theta_d - g_2 \beta X_{od} - \nu_{odi}, \tag{2}$$

[14] Our empirical specification is much richer, in terms of fixed effects, than the one used by Grogger and Hanson (2008). Hence, we do not expect such a large bias from the log utility model. This is confirmed by the fact that our linear and logarithmic estimates (see Table 1) are not too different.

where f_1 and g_i are positive constants.

The idiosyncratic term ν_{odi} captures any other individual, unobservable characteristics that are important to migration decisions. There is substantial evidence suggesting that migrants and non-migrants are systematically different in important dimensions. For example, it is plausible to expect migrants to have higher ability, lower risk aversion, or lower psychological costs from being in a foreign country than non-migrants from the same country of origin. A convenient way to capture these differences is by adapting the nested logit discrete-choice model first proposed in McFadden (1978) to our problem. Specifically, we follow the rendition by Cardell (1991), which frames the nested logit model in the language of the random coefficients model.[15] Let

$$\nu_{odi} = (1-\sigma)\varepsilon_{iod}, \text{for } d = o \tag{3}$$

$$\nu_{odi} = \zeta_i + (1-\sigma)\varepsilon_{iod}, \text{for } d \in D, \tag{4}$$

where ε_{iod} is *iid* following a (Weibul) extreme value distribution, and ζ_i is an individual-specific term that affects migrants only, and its distribution depends on $\sigma \in [0, 1)$. As shown by Cardell (1991), ν_{odi} has an extreme value distribution as well. Two points are worth noting. First, we note that term ζ_i is individual-specific but constant across all possible destinations. Thus, it can be interpreted as differences in preferences for migration. Second, this model nests the standard logit model used in Grogger and Hanson (2007, 2008) when we set $\sigma = 0$.[16]

Utility maximization under our distributional assumptions delivers a neat way to identify the utility (net of costs) associated with migration decisions from data on the proportion of individuals that migrate to each destination, or choose to stay in the country of origin. Namely,

$$\ln s_{od} - \ln s_{oo} - \sigma \ln s_{dD} = f_1 \overline{W}_d - g_1\theta_d - g_2\beta X_{od}, \tag{5}$$

where $s_{od} = n_{od}/(n_{oo} + \sum_{d=1}^{D} n_{od})$ is the share of people born in o who migrate to d (n_{od}) in the total population born in o, s_{oo} is the share of those who stay in o (n_{oo}) among those born in o, and $s_{dD} = n_{od}/\sum_{d=1}^{D} n_{od}$ is the proportion of people born in o migrating to destination d over the total number of people born in o who migrate ($\sum_{d=1}^{D} n_{od}$).[17]

Keeping in mind our normalization, assigning a utility of zero to staying in the home country, we note that coefficient f_1 measures the effect of an increase in the expected earnings *gap* between the origin-destination

[15] See also Berry (1994).
[16] In this case, the distribution of ζ_i collapses and $\nu_{odi} = \varepsilon_{iod}$.
[17] If we did not normalize the utility from staying in the origin to zero we would have

$$\ln s_{od} - \ln s_{oo} - \sigma \ln s_{dD} = f_1(\overline{W}_d - \overline{W}_o) - g_1\theta_d - g_2\beta X_{od}. \tag{6}$$

pair on the left-hand side variable. We also point out that the standard logit model leads to a very similar expression: simply substitute $\sigma = 0$ in equation (5). Intuitively, the term σ corrects for the fact that there is some information in the total share of migrants that helps identify the average value of the difference in utilities (due to costs or expected benefits) between migrants (to somewhere) and non-migrants. After this correction, the difference in log odds equals the difference between the average utility net of cost associated to destination d and the utility from staying in o, which we normalized to zero.

Substituting the definition of the shares and solving for $\ln n_{od}$ the logarithm of migrants from o to d, equation (5) can be rearranged into

$$\ln n_{od} \quad = \quad \frac{1}{1-\sigma}\left(f_1 \overline{W}_d - g_1\theta_d - g_2\beta X_{od}\right) + \frac{1}{1-\sigma}\ln n_{oo} - \frac{\sigma}{1-\sigma}\ln\sum_{d=1}^{D} n_{od} \tag{7}$$

.

Noting that the last two terms on the right-hand side are constant across all destinations d, we can write

$$\ln n_{od} = D_o + \phi_w \overline{W}_d - \gamma_1\theta_d - \gamma_2\beta X_{od}, \tag{8}$$

where D_o is a constant that collects all terms that do not vary by destination d, $\phi_w = \frac{f_1}{1-\sigma}, \gamma_1 = \frac{g_1}{1-\sigma}$ and $\gamma_2 = \frac{g_2}{1-\sigma}$. Equation (8) is the basis of our estimating equation, which obviously encompasses both the logit and the nested logit models. In the former case, fixed effect D_o captures the size of the group of stayers (n_{oo}). In the case of the nested logit, the fixed effect also includes the size of the group of migrants ($\sum_{d=1}^{D} n_{od}$), which provides a correction for the average unobserved heterogeneity between migrants and non-migrants. At any rate, term D_o allows for identification of coefficient ϕ_w, which measures the effect of an increase in the gap between the expected earnings in the home country and in destination d.

Assume that we observe, with some measurement error, the share of people born in country o and residing in destination country d for a set of countries of origin O, destinations D, and for different years t. The log of the migration flow from o to destination d is given by

$$\ln n_{odt} = D_{ot} + D_d + \phi_w \overline{W}_{dt} + \phi_1 Y_{dt} + \phi_2\beta X_{od} + e_{odt}. \tag{9}$$

Term e_{odt} in (9) is the zero-mean measurement error. Coefficient ϕ_w equals $f_1/(1-\sigma)$. Term D_{ot} is a set of country-of-origin by time effects and D_d are destination-country dummies. Note that we are allowing for time-invariant, destination-specific migration costs (through dummies) as well as time-varying ones (Y_{dt}), which will proxy for changes in the tightness of immigration laws or in variables that may affect these laws (population, income inequality and the share of young people in the destination country).

As emphasized above, the set of dummies D_{ot} absorbs any effect specific to the country of origin by year. Justified by our theoretical model, this term serves the purpose of controlling for, among other factors, specific features common to all migrants, for the average migration opportunities/costs in each country of origin in each year. Potential migrants in country o and year t compare average expected utility across destinations and choose the one that maximizes their expected utility. However, besides the average wage there are many other features of the country of origin affecting the cost and opportunity of migrating over time (such as the sudden fall of the Iron curtain in Europe, the loosening of emigration controls in China, and so on) and that specification accounts for them.

Finally, let us note that the theoretically grounded empirical specification (9) can be interpreted as determining a relationship between *stocks* of migrants from each country o to each country d in each year t, or the analogous *flows*. Given our interest in the economic effects of immigration flows in the second part of the paper, we shall focus on explaining immigration flows, and estimate the model using stocks as a robustness check. Having data both on flows and stocks is a strength of our analysis. Data availability constrained previous studies to the analysis of data on stocks only (e.g. Grogger and Hanson, 2008).

4.2 Economic and Geographic determinants of bilateral migration stocks

The basic empirical specification that we estimate on the data and its variations are all consistent with (9). In particular, Table 1 shows the coefficients for several different variations of the following basic specification:

$$\begin{aligned} \ln(Migrant\ Stock)_{odt} &= \phi_w \overline{W}_{dt-1} + D_d + D_{ot} + \phi_d \ln(Distance)_{od} + \phi_b (Land\ Border)_{od} + \\ &\quad + \phi_c (Colonial)_{od} + \phi_l (Language)_{od} + e_{odt} \end{aligned} \quad (10)$$

Specification (10) captures variables specific to the country-of-origin by year with the set of dummies D_{ot}. The fixed migration costs specific to country of destination d are absorbed by the dummies D_d and we explicitly control for distance, colonial ties, common land border and common language as variables affecting the pair-specific bilateral migration costs X_{od}. The term $\overline{W}_{dt}$ captures explicitly the effect of the linear difference in income between destination and origin country, measured as PPP gross domestic product per person in USD, 2000. The theory implies a positive and significant coefficient ϕ_w. At the same time, if we assume that costs of migration increase with distance, a negative value for ϕ_d is expected, while if sharing a border, having colonial-era connections and speaking a common language decrease the costs of migration, ϕ_b, ϕ_c and ϕ_l should be positive. The measures of $(Migrant\ Stock)_{odt}$ used in Table 1 are obtained from the bilateral stocks of immigrants circa year 1990 (from Docquier 2007 data) updated backward and forward using the bilateral, yearly migration flows data (described in section 3.1). In doing so we allow for receiving-country-specific re-migration rates calibrated so that the stock of immigrants for each country of destination match the stock measured around year 2000, also from the Docquier (2008) data. Specification (1) in Table 1 reports the estimates of the coefficients for the basic regression (10). In all regressions, unless otherwise specified, we lag the explanatory variables one period, allowing them to affect the stock of immigrants in the following year. Our method of estimation is least squares, always including the destination countries and the country-of-origin by year fixed effects. We add one to each observation relative to stock and flows of immigrants so that when taking logs we do not discard the 0 observations[18]. Finally we weight observations by the population of the destination country to correct for heteroskedasticity of the measurement errors and we cluster the standard errors by country of destination to account for the "within-destination country" correlation of the errors.

The estimated coefficients on the income differences (first row of Table 1) are always significant (most of the time at the 5% confidence level) and positive. The magnitude of the coefficient in the basic specification (1) implies that the increase in the average income differences between destination and origin countries experienced over the period 1980-2000 (equal to +7,000 US $ in PPP, calculated from Table 1A) would generate an increase of 42% (=0.06*7, since the income per capita is measured in thousands) in the stock of migrants to the destination countries. This is equal to two thirds of the observed increase in the stock of immigrants from those 74 countries in the 14 OECD countries, which grew by 60%. Hence, both statistically and economically the absolute real income differences between sending and receiving countries, and their changes over the considered period, can explain a very large fraction of the growth in the stock of immigrants.

As for the effect of geographic variables on migration costs, the variable "colonial relations" and the natural logarithm of distance have very significant effects with the expected signs. Having had colonial connections more than doubles the average stock of immigrants from origin to destination, and that stock decreases by 80% any time the bilateral distance increases by 50%. On the other hand, sharing a land border and speaking a common language do not significant affect bilateral migration flows. This is hardly surprising as most of the large migratory flows to the OECD (except for Mexico-US) take place between countries that do not share a land border or a common language. These two results are also found by Mayda (forthcoming) who does not find any significant effects for common border and common language dummies. Specification (2) checks whether including the logarithm of the destination country wage $\ln(\overline{W}_{dt})$ instead of its level results in similar effects.[19] The sign and significance of the income difference variable is as in specification (1), though the magnitude of the coefficient is smaller. In fact, a change by 1 (100%) in the log difference would only produce an increase of 29% in the stock of immigrants. Notice, also, that in terms of log-difference (percentage difference) the gap between origin and destination countries has barely changed between 1980 and 2000. This may imply that the logarithmic specification is not the optimal approach; still, we are reassured that the sign and significance of

[18] Except for Specification (6) of Table 1 where we explicitly omit zeros.

[19] Recall that W_{ot} or its log are absorbed into the country of origin by year fixed effects.

the income effect does not depend on the specific functional form chosen.

Specification (3) decomposes the effect of the expected (logarithmic) income difference (between destination and origin) into the effect of differences in (the logarithm of) GDP per worker and differences in (the logarithm of) the employment rate (probability of employment)[20]. Both variables turn out to be significant, confirming that the expected destination-country income, on which potential migrants base their decisions, depends on potential wages and on the probability of being employed.

Specification (4) adds three destination-country variables that can plausibly affect the willingness of the country to accept immigrants and hence its immigration policies (and immigration costs). The first is total population, the second is a measure of income distribution (Gini Coefficient) and the third is the share of young (aged 15 to 24) individuals in the population. A country whose population is growing may find it easier to absorb new immigrants with little consequence for its citizens. Similarly, in periods when the income distribution is more equal, the opposition to immigration may be milder. There is weak evidence of a positive effect of population on immigration flows and of a negative effect of inequality: the point estimates have the expected sign but the coefficients are not significant at standard levels of confidence. Also, the share of young workers does not seem to be significant at all, possibly because young workers may fear the competition from immigrants (who are typically younger than the average native) or, alternatively, they may be more flexible and mobile in adjusting their occupation in response to immigrants, and hence suffer less from the competition.

In specification (5) we consider whether including longer lags of the income variable changes its impact on immigration. As it may take more than one year before income differences put in motion a migration response, including a longer lag may strengthen the effect. The coefficient on log income, lagged two years, is only marginally different from that of the one year lag. If one includes both lags (not reported) or two lags and the contemporaneous value (also not reported) only the two-year lagged income difference is significant (with a coefficient of 0.06). This implies that it takes at least one year and possibly up to two years for income differentials to stimulate migrations.

Specification (6) drops all the 0 observations. Note that we are using stocks as the dependent variable and there are not many zeros (only 10% of the observations), and therefore the estimates do not change much. Finally, we show in specifications (7) and (8) the results omitting the UK, whose immigration flows before 1990 look suspiciously small (see Panel 1A), and the US, whose large undocumented immigration from Mexico is not included in our data. Neither omission affects the results. We also run other checks changing the weighting of the observations and the clustering of the residuals or using only the observations after 1990. All estimates of the income and geography variables are quite stable and similar to those in the basic specification. A particularly interesting robustness check (that will be systematically incorporated in Table 2) is the introduction of a full

[20] We decompose the effects of GDP per worker and employment rates in the logarithmic specification because the logarithm of GDP per person is the sum of those two logarithmic components.

set of origin-destination pair dummies. Such a specification adds 1022 fixed effects and removes the geographic controls (absorbed in the dummies). The estimated effect of wage differentials on migration flows is equal to 0.054 with a standard error of 0.02 . Hence, still significant and very similar to the estimate obtained in the basic specification of Table 1.

4.3 Effect of Immigration laws on bilateral migration flows

In evaluating the effects of immigration reforms, it is easier to look at the effect on subsequent immigration flows. After all, the immigrant stocks are the long-run accumulation of yearly flows, so the determinants of the first should also determine the second. Hence we simply adopt the specification in (9) and use as the dependent variable the logarithm of the flow of immigrants from country o to country d in year t, adding immigration laws as an explanatory variable. Column (1) of Table 2, Panel A reports the relevant estimates for the following specification:

$$\begin{aligned}\ln(Migrant\ Flow)_{odt} &= \phi_w \overline{W}_{dt-1} + \phi_R (Tightness)_{dt-1} + D_{ot} + \\ &+ \phi_d \ln(Distance)_{od} + \phi_b (Land\ Border)_{od} + \phi_c (Colonial)_{od} + \phi_l (Language)_{od} + \epsilon_{odt} \end{aligned} \tag{11}$$

Our data on $(Migrant\ Flow)_{odt}$ are from the OECD International Migration Database, from 74 countries of origin into 14 OECD countries. The variable "Immigration policy tightness" is the measure of tightness of immigration (and asylum) laws described in section 3.2[21]. The other columns of Table 2 Panel A perform variations and robustness checks on this basic specification. In Panel B of Table 2 we estimate a similar specification but now include a full set of (73x14) country-pair fixed effects, D_{od}, rather than the four bilateral variables (*Distance*, *Land Border*, *Colonial*, *Language*) in order to capture any specific time-invariant bilateral costs of migration.

Moving from left to right in Table 2 we modify our basic specification (1) by including income on logarithm, rather than in levels, (specification 2), then using a broader measure of tightness (specification 3), or longer lags of the explanatory variables (specification 4). Specification (5) includes extra destination country controls, (6) omits observations with 0 flows and (7) omits the UK data, whose immigration flows recorded before 1990 appear suspiciously small. In all these specifications we include four variables that capture aspects of the immigration laws. The first variable is our constructed measure of "Tightness of entry laws", the second is our measure of "Tightness of asylum laws". Both are described in section 3.2 and their values for each country and year are shown in Panel 2A. We also include dummies for the two most important multilateral treaties affecting several

[21] Notice that all the explanatory variables (that vary over time) are included with one lag.

of the considered countries[22].

The "Maastricht" treaty was ratified by most EU countries in 1992. Among other things, it introduced free labor mobility for workers of the member states and it led to the introduction of the Euro, which may have reduced migration costs within the European Union. The corresponding dummy takes a value of one for those countries and years in which the agreement is in place and 0 otherwise. The "Schengen" agreement, adopted in different years by 22 European countries, regulates and coordinates immigration and border policies among the signatory countries. While it eases intra-EU movement for citizens of the signatory countries, the agreement also implies more restrictive border controls to enter the "Schengen" area. The corresponding dummy takes a value of one for countries and years in which the agreement is in place. Three main results emerge from Table 2. First, income differences between origin and destination country (whether in logs or in levels) have a positive and significant effect on immigration flows to OECD countries in almost every specification. Second, the "Tightness of entry" has a significant negative effect on immigration flows in most specifications. Each reform that introduced less restrictive measures increased, on average, immigration flows by 5 to 9%. For instance, this implies that a country like Canada, whose immigration policy loosened by 6 points between 1985 and 2005 (see Panel 2A), should exhibit an increase in immigration rates of 25 to 54%. The yearly immigration rates, in Canada, went from 0.5% of population in the early eighties to 0.7-0.8% in the early 2000's. That is, the entire increase in immigration flows can be attributed to the change in the laws. Third, among the other laws the most significant effect is associated with the Maastricht treaty which increased, on average, the immigration of signatories between 50 and 60%. Tightness of asylum laws had a negative (but rarely significant) impact on immigration and Schengen had no effect at all. Interestingly, column (3) in both Panel A and B reveals that combining immigration entry- and stay- laws decreases the precision of the estimated coefficient, suggesting that mainly entry laws had an effect on the actual inflow of immigrants. At the same time the effect of entry laws is less significant when we include population, income distribution and the share of young among the receiving country variables (specification 5, both in Panel A and B). This may imply that some of those variables affect immigration laws, and indirectly immigration, so that including them reduces the effect of the laws. Finally, omitting the cells with 0 immigration flows (specification 6) reduces drastically the effect of wage differentials, while the effect of entry laws is still significant. Since almost 70% of the cells are zeros, because we are looking at bilateral flows (rather than stocks), it is remarkable that the immigration laws variable maintains its sign and significance. Omitting the UK (column 7) does not change the results much. The estimated effects on the geographic variables (not reported in Table 2 and available only for Panel A) are qualitatively and quantitatively close to the estimates reported in Table 1. In particular, sharing a land border (point estimate -1.6 and standard error 1.3) and sharing a common language (point estimate 0.4, standard error 0.5) have no significant impact

[22] We have run a few other specifications such as a Tobit regression with censoring at 0, to account for the clustering of observations at 0, and obtained a coefficient of 0.25 on $\overline{W}_{dt-1}$ and of -0.14 on $Tightness$ confirming the results in Table 2.

on migration flows, while having had colonial ties (point estimate 3.88 and standard error 0.46) and the log of distance (point estimate -2.2 standard error 0.46) are both very significant in their impact on migration flows[23].

Let us emphasize that the estimates in Table 2 Panel B include 1022 country-pair fixed effects and 1825 country-of-origin by year fixed effects. Hence any variation is identified by the change over time in a specific bilateral migratory flow, after controlling for any country-of-origin by year specific factor. We are not aware of any previous analysis that could run such a demanding specification on bilateral migration panel data. All in all, our analysis finds statistically and quantitatively significant effects of income differentials on bilateral immigration stocks and flows. These effects are very robust to sample choice, specification and inclusion of controls. We also find strong evidence that the receiving country laws, particularly those relative to the entry of immigrants, significantly affected the size of yearly inflows. The inclusion of income differences in levels or in logs does not produce very different effects.

5 Impact of Immigration on OECD countries

5.1 A Production Function Framework

In order to evaluate the impact of immigration on the receiving economy's income, average wages, and return to capital, we use an aggregate production function framework, akin to the one used in growth accounting (see for instance Chapter 10 of Barro and Sala-i-Martin 2004). Suppose that total GDP in each destination country and year, Y_{dt}, is produced using a labor input represented by total hours worked, L_{dt} (that can be decomposed into $Employment_{dt}$ times $Hours\ per\ worker_{dt}$), services of physical capital represented by K_{dt} and total factor productivity A_{dt}. According to the popular Cobb-Douglas production function:

$$Y_{dt} = A_{dt} K_{dt}^{\alpha} L_{dt}^{1-\alpha} \tag{12}$$

where α is the capital income share and can be approximated for the destination countries in our sample by 0.33[24]. In such a framework if we intend to analyze how immigration flows affects income or wages (marginal productivity of labor), we need to identify first how immigrations affects the supply of each input and of total factor productivity. Then we can combine the effects of immigration using the implications of the model. Specifically, the percentage changes in total real GDP, Y_{dt}, real GDP per hour, y_{dt}, and the average real wage, w_{dt}, are given, respectively, by:

$$\frac{\Delta Y_{dt}}{Y_{dt}} = \frac{\Delta A_{dt}}{A_{dt}} + \alpha \frac{\Delta K_{dt}}{K_{dt}} + (1-\alpha) \frac{\Delta L_{dt}}{L_{dt}} \tag{13}$$

[23] The reported point estimates and standard errors are from the basic specification of column 1, Panel A, Table 2.
[24] See Jones (2008) page 24 and Gollin (2002) to justify this assumption.

$$\frac{\Delta y_{dt}}{y_{dt}} = \frac{\Delta w_{dt}}{w_{dt}} = \frac{\Delta A_{dt}}{A_{dt}} + \alpha(\frac{\Delta K_{dt}}{K_{dt}} - \frac{\Delta L_{dt}}{L_{dt}}) \tag{14}$$

If we can identify the percentage changes in A_{dt}, K_{dt}, and L_{dt} in response to exogenous immigration flows to the country we will be able to evaluate the impact of immigration on total income, labor productivity and average wages.

Clearly, immigration flows directly affect labor input L_{dt} by adding potential workers. However, the increase in employment may be less than one-for-one if immigrants displace native workers (out of the country or out of the labor market). In addition, there may also be composition effects if immigrants' employment rates or hours worked are lower than those of natives.

Regarding the capital input, standard models with endogenous capital accumulation imply that immigration-induced increases in the labor force will generate investment opportunities and greater capital accumulation, up to the point that the marginal product of capital returns to its pre-shock value. However, the short-run response of the capital stock to an international immigration flow can be less than complete and it has yet to be quantified empirically.

Concerning TFP, on the one hand immigrants may promote specialization/complementarities (Ottaviano and Peri 2008) which increase the set of productive skills (Peri and Sparber, forthcoming) and increase competition in the labor markets, generating efficiency gains that increase TFP. Or there can be positive scale effects on productivity if immigrants bring new ideas or reinforce agglomeration economies (of the kind measured by Ciccone and Hall, 1996). On the other hand, it is also possible that immigration induces adoption of less "productive", unskilled-intensive technologies (as in Lewis 2005) that lead to reductions in measured TFP. Ultimately, it is an empirical question whether an immigration shock increases, decreases or does not affect TFP.

We denote by $\frac{\Delta F_{dt}}{Pop_{dt}}$ the *immigration rate*, namely the change in the foreign-born population F_{dt} (immigration flows to country d in year t) relative to the total population of country d at the beginning of year t (Pop_{dt}). We then estimate the following set of regressions:

$$\frac{\Delta X_{dt}}{X_{dt}} = D_t + \gamma_x \frac{\Delta F_{dt}}{Pop_{dt}} + e_{st} \tag{15}$$

Where X will be alternatively total hours worked (L_{dt}),[25], services of physical capital (K_{dt}) and total factor productivity (A_{dt}). As a check we also analyze directly the effect of $\frac{\Delta F_{dt}}{Pop_{dt}}$ on aggregate GDP, GDP per hour, and capital per worker. The term D_t captures year fixed effects that absorb common movements in productivity and inputs across countries in each year. In order to assert that the estimated coefficients $\widehat{\gamma_x}$ identify the causal

[25] Also decomposed between employment $Employment_{dt}$ and $Hours\ per\ worker_{dt}$.

effect of immigration on domestic variables we will instrument total immigration flows to a country with the sum of bilateral flows to that country predicted using our empirical model in (11), but excluding variables relative to the destination country[26]. Essentially we predict those flows using only the components that vary by country of origin and time, and the fixed bilateral migration costs.

5.2 Measurement of Employment, Capital Intensity and Productivity

The data on income and factors of production are mostly from OECD datasets. Specifically, GDP data is from the OECD Productivity dataset, and employment and hours worked are from the OECD-STAN dataset. The data cover the whole period 1980-2005 for the 14 countries in our sample.[27]

The capital services data are also from the OECD Productivity dataset, but we make use of the data on aggregate investment in the Penn World Tables (version 6.2) to extend its coverage. Let us provide a bit more detail on the capital data that we use. The conceptually preferred measure of capital for our purposes is the services of the capital stock that contribute to current production. Capital services are computed as follows. For each type of capital (six or seven, depending on the country), we accumulate past investments making two adjustments. First, we take into account that older units of capital provide fewer services than newer ones (efficiency weighting). Secondly, we take into account the productive life of each type of capital (retirement pattern). Finally, we aggregate across all types of capital using the relative productivity of each type to obtain the stock of *productive* capital. The capital services data reported by the OECD is the rate of change of the stock of productive capital and it is interpreted as the flow of capital services that went into production during that period.

The original data on capital services is available annually from 1985 onward and only covers 12 out of the 14 countries in our main sample.[28] In order to expand the data to cover the whole country-year panel we use data on gross fixed capital formation. Specifically, we proceed in three steps. First, we use the long series on real investment provided by the PWT to compute the stock of capital for the 14 countries in our sample between 1980 and 2005. More specifically, we initialize the capital stock in 1970 following the procedure based on the perpetual inventory method used in Young (1995). Next, we iteratively build the entire series of capital values for the period 1980-2005. The main difference between this capital stock and the stock of productive capital derived from capital services data is that here we are imposing the same growth rate across all types of capital.

Second, we build a predictor for productive capital using the data on capital stocks that we just created. In particular, we estimate a regression model where the dependent variable is the change in the log of productive capital and the main explanatory variable is the change in the log of the capital stock. We estimate this

[26] Essentially we omit the term $\left(\overline{W}_{dt-1} - \overline{W}_{ot-1}\right)$ and the term from the basic specification.
[27] The data on Hours for Luxembourg start in 1983. We use employment growth to fill in the missing values.
[28] Norway and Luxembourg are missing.

relationship for the sample period for which we have data on both variables, namely, 1985-2005. The slope coefficient is 1.31, estimated very precisely. A coefficient larger than one makes sense. In good times, firms may increase the rate of replacement of old capital goods for new ones. This automatically leads to the provision of greater capital services, even keeping constant the total capital stock. This is because of the age-flow profile of capital goods used in the calculation of capital services: a new truck is assumed to produce more services than an old one. Finally, we use our predictor to extend the data on capital services to cover the whole sample. For the twelve countries for which we have data on capital services (that is, the growth rate of productive capital), we use our predictor to extend the data back to 1980. For the two countries for which we lack data on capital services we use the prediction rule for the entire period, 1980-2005.

Equipped with a full panel for real GDP and labor and capital inputs, we compute total factor productivity as a Solow residual, imposing a labor share of 0.66 and using total hours worked and capital services as the inputs into production.[29]

Let us now have a descriptive look at our panel data for income, labor, capital services, and TFP. Table A3 reports annualized growth rates of these variables for three sub-periods: the 1980s, the 1990s, and 2000-2005. Three features stand out. First, there is a noticeable slowdown in economic growth between 1980 and 2005 for our sample of OECD countries. In the three sub-periods real GDP grew annually by 2.72%, 2.62%, and 1.98%, respectively. The slowdown is also noticeable in terms of lower employment growth (from 0.68% to 0.34%), lower capital growth (from 3.43% to 3.11%), and lower TFP growth (from 1.14% to 0.73%). Note also the large cross-sectional dispersion.

Secondly, average employment growth was substantially higher than average growth in total hours worked between 1980 and 2005. That is, hours per worker on average fell during the period. Finally, capital intensity on average increased substantially over the period. The average annual growth in capital services (in real terms) was roughly three times as large as the annual growth rate in employment.

5.3 The Effects of Immigration: OLS

Table 3 presents the estimates, using least squares methods, of the coefficients γ_x from equation 15. The dependent variables are, in order, inputs to production (first to fourth row), total factor productivity (fifth row), total GDP (sixth row), capital per worker, and output per hour worked (rows seven and eight). Notice that not all the estimated coefficients are independent of each other due to the relationship between inputs and output provided by the production function. Hence, for instance, in the basic specifications in which no other control variables are included and the selected observations are common between regressions, by

[29] The OECD Productivity dataset features an analogous measure of TFP for some countries covering part of our period of interest. Our own measure is very strongly correlated with theirs. We run a regression of growth rates of the two measures amd find that the estimated coefficient is 0.92 and the standard error is 0.018.

virtue of (14) the estimated coefficient on $\Delta y/y$ in the last row of the table should be equal to the difference between the coefficient on $\Delta Y/Y$ and the coefficient on $\Delta L/L$[30]. Since we regress the percentage change of the dependent variable on the inflow of immigrants as a percentage of the initial population, the interpretation of the coefficients (as elasticities) is straightforward. Different columns of Table 1 correspond to different samples and specifications. Specification (1) is the basic one and it estimates 15 on 25 yearly changes (1980-2005) for 14 OECD countries. The method of estimation is OLS with year fixed effects (since the variables are already in changes we do not include country-level effects[31]). The standard errors in parentheses are heteroskedasticity robust and clustered by country. Specification (2) omits the US, which is one of the most studied cases, to show that the rest of the sample does not behave too differently from the US. Column 3 includes only the continental European countries, excluding the Anglo-Saxon group (US, UK, Canada and Australia) often considered as more "immigration friendly". Specification (4) includes only the more recent years (1990-2005) , for which the most accurate migration data from the OECD are available and specification (5) includes in each regression the lagged level of the dependent variable to control for potential "convergence" behavior of each variable to a balanced growth path or a steady state. Finally, specification (6) uses as explanatory variable the immigration flows net of imputed re-migration of the stock of immigrants. While there is significant potential for endogeneity in these OLS specifications, let us comment on some robust and clear correlations that emerge from Table 3. First, the coefficient on total labor inputs $\Delta L/L$ and on total capital $\Delta K/K$ are in most cases similar to each other and close to one. Except for specification (6) we can never reject that the effect on total labor input is equal to one and in specifications 1 to 4 we cannot reject that the effect on total capital services is equal to one. This implies that the correlations do not show any evidence of crowding-out of native jobs: one newly arrived immigrant worker increases employment by exactly one. Also, the estimates imply that the increase in labor inputs occurs because of an increase in employment (one-to-one) and no changes in average hours worked per person. The estimates on the capital stock imply that investment adjusts to the larger potential worker pool (at constant wages) and capital increases within one year, effectively leaving unchanged the capital-labor intensity in production. Row seven shows that capital labor ratios are not significantly affected by immigration in all six specifications. Finally, the estimates in row 5 imply that there is no significant effect of immigration on TFP, $\Delta A/A$. These effects, combined together, imply that the inflows of immigrants are associated with larger employment, larger total GDP, and unchanged wages, capital intensity and GDP per hour. These correlations also hold when we consider European countries only (specification 3), when we restrict ourselves to the more recent period 1990-2005 (in specification 4) or when we include lagged levels of the dependent variable (specification 5). The results obtained using the net immigration flows, on the other hand (specification 6),

[30] The reader can easily check that these relations hold.

[31] We have also run the panel regression with country fixed effects, obtaining similar qualitative estimates, with larger point estimates and standard errors, however.

show much larger coefficients and standard errors on labor inputs and capital inputs (with similar effects on productivity). This suggests that the imputed re-migration flows are probably a rather noisy measure of actual outflows of immigrants and by subtracting these imprecisely estimated outflows we are reducing the value of flows and increasing the noise to signal ratio. Still, even this specification does not show any evidence of a change in the capital-labor ratio or GDP per person associated with immigration. What seems implausible in specification 6, however, is the very large (more than 1 to 1) response of labor inputs to immigrants, which may indicate measurement error or endogeneity problems. For this reason we prefer the gross flows, which are directly measured in the data, and which we use in the instrumental variable analysis below. Combining the estimated γ_x with the formula **??** would imply that immigration has no significant correlation with average wages (or returns to capital) and that immigration increases employment and GDP in the receiving economy one for one, *even in the short run.*

5.4 Immigration Effects: Instruments and 2SLS approach

The most significant limitation of the estimates presented in Table 3 is that immigration flows are endogenous. In fact, we have shown in section 4 that immigration flows respond vigorously to changes in wage differences between origin and destination. Employment, capital and TFP are the determinants of those wages, hence we cannot consider immigration as exogenous to them. The framework of section 4, however, provides an analysis of the determinants of the international migration flows and lends us a solution to the problem of endogeneity. In particular, consider the bilateral regression model used in Table 2, Panel B:

$$\ln(Migrant\ Flow)_{odt} = \phi_w \overline{W}_{dt-1} + \phi_R (Tightness)_{dt-1} + D_{ot} + D_{od} + e_{odt} \tag{16}$$

The terms D_{ot} capture any economic, demographic and cost determinant of migration out of country o which varies over time t. That set of dummies captures all the so called "push-factors" of immigration that do not depend on specific destination countries but only on conditions in the countries of origin. The terms D_{od}, on the other hand, capture the fixed bilateral costs of migrating from o to d. They mostly reflect geographic factors and the existence of historical networks which provide information and ease the adjustment of immigrants to the destination country. Therefore, only the terms $\phi_w \overline{W}_{dt-1}$ and $\phi_R (Tightness)_{dt-1}$ are specific to the country of destination and in particular to its economic conditions. The wage differential is the primary included economic determinant of immigration, while the tightness of immigration laws can be considered as a determinant of the cost of immigration which is still related to current economic conditions, although to a lesser degree. Hence we can use (16), removing $\phi_w \overline{W}_{dt-1}$ and $\phi_R (Tightness)_{dt-1}$, to predict the log of annual bilateral flows from all countries of origin to their destinations. The remaining factors in the regression, D_{ot} and D_{od} are, by construction, independent of time-varying economic (and legal) factors in the country of destination . Using these predicted values we calculate the imputed immigration rate for each of the 14 destination countries in each year (adding the predicted immigration rates from each country of origin).[32] These imputed immigration rates are what we use as instruments for the actual immigration rates. To the extent that immigration laws (lagged one period) may also be considered as exogenous to the current economic condition of a country, we can also construct predicted immigration flows by including the estimated term $\widehat{\phi}_R(Tightness)_{dt-1}$ in predicting the bilateral flows in regression 16. Table 4 shows the statistics for the first stage regressions using the predicted immigration flows from 16 without wage differentials or immigration laws (first row of Table 4), and those relative to predicted flows omitting only wage differentials (second row of Table 4). We test the significance of the instrument on the whole sample (specification 1) or omitting the US (specification 2), using only European countries of destination (specification 3) or only on the more recent period (specification 4). In each case the coefficient on the instrument is positive and very significant, and the partial R-square of the instrument is between 0.32 and 0.42. Each regression includes time fixed effects. The F-statistic of significance of the instrument is usually above 300. Thus, the instrument is quite powerful and captures only the variation in immigration rates due to the interactions between country-of-origin specific factors and bilateral migration costs (due to geography and historical bilateral networks). For instance, the large increase in Polish emigrants in the period 1990-1995 due to the end of the communist regime produced a large Poland-specific term ($\widehat{D}_{ot}$) for those years in the migration equation. The fact that Poland has smaller bilateral costs of migration to Germany and the UK than to (say) Japan (which is captured by the higher estimated $\widehat{D}_{od}$ for Germany and the UK) implies that the predicted migration rates from Poland to Germany and the UK, using our model, are larger then the predicted migration rates to Japan, and particularly so during the years of large Polish migration. Recall that while they are additive in equation 16, the terms D_{ot} and D_{od} predict the logarithms of immigrant flows. Hence, when we calculate their levels (divided by population to obtain immigration rates) the two effects are multiplicative, so for a given sending country shock, D_{ot}, the effect would be magnified by a large D_{od}. The constructed immigration rate represents the exogenous (push-driven) variation in the immigration rates of the receiving country and will be used as an instrument.

Table 5 shows the 2SLS estimates of the effect of immigration on inputs, productivity and per capita income. The specifications and the dependent variables are as in Table 3. Again, the estimates obtained using net immigration flows (specification 5) seem too large, but all the other specifications (using gross flows) are consistent with the results obtained using OLS in Table 3. In particular, the effect of immigration on total labor supply $\Delta L/L$ is always very close to one (between 0.96 and 1.02) and precisely estimated (standard error around

[32] One further source of error in proxying the actual immigration rates with those predicted from the regression is that in the bilateral regression we only have 74 countries of origin (the most important ones) and add the predicted flows from those. The immigration rates, instead, measure the total immigration flows from those countries plus any other country in the world.

0.09). Similarly, the coefficient on the capital adjustment ($\Delta K/K$) is always larger than one (and in most cases not significantly different from it) suggesting full adjustment of the capital stock within one year, so that the change in the capital labor ratio ($\Delta k/k$) is always equal to 0. Similarly, there seems to be no significant effect of immigrants on productivity changes ($\Delta A/A$).The estimates of these effects are robust to the choice of countries in the sample (specification 2 omits the US, and specification 3 omits Europe) and to the choice of the period (specification 4 considers only 1990-2005). All in all, the results of Table 5 confirm the correlations obtained with the OLS estimates of Table 3. Immigrant flows caused (and predicted) by country-of-origin and geographic factors increase the employment and labor supply in the receiving country one-to-one. Such an increase in the pool of workers induces investments and capital accumulation that, even within one year, adjusts the capital-labor ratio (and therefore the wages and return to capital) to the pre-immigration levels. The economy expands and there is no significant effect on the total productivity of factors but only to the overall size of GDP, which grows in percentage roughly by the same amount as the immigration rates. Hence, for instance, the average yearly inflow of immigrants in the US, recorded between 1995 and 2005 at around 0.3-0.4% of the population, increased US GDP by around 0.3-0.4% each year, with no appreciable effect on average wages and income per person.

The reader may find it puzzling that the capital stock adjusts fast enough to eliminate any effect of immigration on wages, even within one year. Let us emphasize that immigration flows, even those that are push-driven, have been quite predictable and, as a percentage of the population, never too large (mostly around 0.5% of the population). Therefore, with yearly investments on the order of 20-30% of GDP there is ample room to adjust investment by a relatively modest amount in order to accommodate new immigrant workers. Moreover international capital movements may also follow migration and help the adjustment. As a further check that our short-run estimates are not driven by some short-frequency noise in the data we have re-calculated the responses of employment, capital, TFP and income to immigration over 5-year changes (rather than yearly changes). Table 6 reports the estimated coefficients from four different specifications. Notice, importantly, that the coefficients on labor adjustment ($\Delta L/L$) and capital adjustment ($\Delta K/K$) are still close to one and not significantly different from one another (the capital response still seems to be a bit larger than one). The effects on productivity ($\Delta A/A$), on the capital-labor ratio and output per hour worked, are all insignificant. The adjustment within one year seems fairly similar to the adjustment over 5 years and compatible with the adjustment in the neoclassical model with endogenous capital: more workers encourage investment and do not affect productivity so that capital per worker and wages remain stable while the size of the workforce and of the economy grows.

6 Discussion and Conclusions

The impacts of immigration on Western-country economies and labor markets have frequently been analyzed by considering a single receiving country combined with individual or regional data. Similarly, the determinants of international migrations have mostly been analyzed using only a single receiving country. These studies are quite useful, however they have brought to light some issues that are difficult to address in the context of one receiving country, or by focusing exclusively on labor-market effects. For instance, the degree and the speed of adjustment of capital to immigration is a key determinant of the short-run effect of immigration on wages (see Borjas and Katz 2007, Ottaviano and Peri 2008). However, if capital is mobile within a country we cannot estimate its response to immigrants with data from one country only (unless we have a very long time series). Furthermore, the literature recognizes that we would need some "purely push-driven" migration flows to identify the causal effect of immigrants on economic outcomes in the destination country (e.g. Card 2001). Those shocks, however, are hard to identify in the context of one receiving country only. This paper suggests a couple of new approaches to address these issues and provides a new framework to estimate the determinants of migration flows, to isolate the push-driven determinants, and to use them to identify the causal effects of immigration at the country-level. We also organize an extensive dataset of migration flows and immigration laws for OECD countries (1980-2005).

We make three main contributions. First, following Grogger and Hanson (2008) we use a bilateral migration regression model that can be derived from a simple or nested logit model of the migration choices of potential migrants. Migrants decide where to reside based on utility comparison between locations. Such a model can explain the logarithm of the stock (and flow) of migrants from country o (origin) to country d (destination) as a function of the wage differential between d and o, of bilateral migration costs and country-of-origin specific effects. Therefore, conveniently, we are microfounding a pseudo-gravity equation for international migrations. We estimate that an increase in the wage differential between origin and destination of 1000 US \$ (in 2000 PPP prices) increases the flow of migrants by 10-11% of their initial value. We also show that the immigration reforms that made entry laws more restrictive were effective in reducing migration flows by 6%, on average, for each reform.

Second, we use our model to separate between push factors, bilateral costs and pull factors, and construct a prediction of migration flows that is "exogenous" to the economic conditions in the country of destination (pull factors). Finally, using the predicted flows as an instrument we estimate the effect of immigration on employment, capital accumulation, and total factor productivity. We find that, already within one year, employment responds to new immigrants one for one, and capital adjusts in order to maintain the capital labor ratio. We do not find any significant effect of immigrants on total factor productivity. These results, taken together, imply that immigration has no negative impact on average wages, or on income per worker in the short run (one year) or in the long run (five years). The inflow of immigrants only increases the overall size of the economy without altering the distribution of income between workers and capital owners. This is due to the fact that capital owners respond efficiently to a larger labor pool by investing more. We hope that this paper will stimulate the analysis of the effects of international migrations, encouraging improvements and extensions in the collection and organization of data on migration flows and immigration laws.

References

Anderson, James E., and Eric van Wincoop. (2003) "Gravity with Gravitas: A Solution to the Border Puzzle." *American Economic Review*, 93(1): 170–92.

Angrist Joshua and Adriana Kugler (2005) "Protective or Counter-Productive? Labor Market Institutions and the Effect of Immigration on EU Natives," *Economic Journal,* 113(488): 302-331.

Aydemir and Borjas (2007) "A Comparative Analysis of the Labor Market Impact of International Migration: Canada, Mexico, and the United States" National Bureau of Economic Research Working Paper #12327, June 2006.

Barro, Robert and Xavier Sala-i-Martin (2004) "Economic Growth" MIT Press, Cambridge Massachusetts

Benhabib, Jesse, (1996) "On the political economy of immigration" European Economic Review.

Berry, Steven T. (1994) "Estimating Discrete Choice models of Product Differentiation" *RAND Journal of Economics*, 25 (2).

Bertocchi Graziella and Chiara Strozzi. "International Migration and the Role of Institutions", 2008, Public Choice, Vol. 137, No. 1, pp. 81-102.

Bertocchi Graziella and Chiara Strozzi. "The Evolution of Citizenship: Economic and Institutional Determinants", Journal of Law and Economics, forthcoming, 2010.

Borjas, George (1987) "Self Selection and the Earnings of Immigrants" *American Economic Review* 77(4), 531-553.

Borjas, George. (2003). The Labor Demand Curve is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market. *Quarterly Journal of Economics* 118, no. 4: 1335-1374.

Borjas, George, and Lawrence Katz. (2007). The Evolution of the Mexican-Born Workforce in the United States, in Borjas, George, editor "*Mexican Immigration to the United States*" National Bureau of Economic Research Conference Report, Cambridge Ma.

Card, David. (2001). Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration. *Journal of Labor Economics* 19, no 1: 22-64.

Cardell, N.S (1991). "Variance Components Structures for the Extreme Value and Logistic Distributions." Mimeo, Washington State University.

Chaney, Thomas (2008) "Distorted Gravity: The Intensive and Extensive Margin of International Trade" *American Economic Review* 2008, 98:4, 1707–172.

Ciccone, Antonio and Robert Hall (1996) "Productivity and the Density of Economic Activity" *American Economic Review*, 86 (1), 54-70.

Clark X., Hatton T.J. and J. Williamson (2007) "Explaining U.S. Immigration 1971-1998" *Review of Economics and Statistics* 89(2), 359-373.

Coen-Pirani, Daniele (2008) "Understanding Gross Workers Flows Across United States" Manuscript, Canegie Mellon University.

Docquier, Frederic (2007) "International Migration Datasets" available at http://www.ires.ucl.ac.be/CSSSP/home_pa_pers/docquier/oxlight.htm.

Dahl Gordon (2002) "Mobility and the Return to Education: Testing a Roy Model with Multiple Markets" *Econometrica* 10(6): 2367-2420.

Deininger Klaus and Lyn Squire (1996) "A New data set Measuring Income Inequality" The World Bank Economic Review, Vol. 10, N. 3 pp. 565-591.

Fondazione Rodolfo DeBenedetti (2007) "Social reform database" http://www.frdb.org/documentazione/scheda.php

Glick, Reuven and Andrew Rose (2002) 'Does a Currency Union Affect Trade? The Time-Series Evidence" *European Economic Review* 46(6) 1125-1151.

Grogger Jeffrey and Gordon Hanson (2007) "Income Maximization and the Sorting of Emigrants across Destinations" Manuscript UC San Diego, 2007

Grogger Jeffrey and Gordon Hanson (2008) "Income Maximization and the Sorting of Emigrants across Destinations" National Bureau of Economic Research Working Paper #13821, Cambridge, Ma.

Gollin Douglas (2002) "Getting Income Shares Right" *Journal of Political Economy,* vol. 100. pp. 458-474.

Helpman, Elhanan, Marc J. Melitz, and Yona Rubinstein (Forthcoming) "Estimating Trade Flows: Trading Partners and Trading Volumes." *Quarterly Journal of Economics.*

Jones, Charles I.(2008) "Macroeconomics" W.W. Norton, New York.

Karemera D., Oguledo V.I. and Davis B. (2000) " A Gravity Model Analysis of International Migration to North America" *Applied Economics* (32) 1745-1755.

Lewis, Ethan. (2005) "Immigration, Skill Mix, and the Choice of Technique" Federal Reserve Bank of Philadelphia Working Paper no. 05-08 Philadelphia, PA.

Manacorda Marco, Alan Manning and John Wadsworth. (2006). The Impact of Immigration on the Structure of Male Wages: Theory and Evidence from Britain IZA Discussion paper 2352. Bonn, Germany.

Mayda, Anna Maria (Forthcoming): "International Migrations: A Panel Data Analysis of the Determinants of Bilateral Flows" *Journal of Population Economics.*

Mayda Anna Maria and Krishna Patel (2004) "OECD Countries Migration Policy Changes" available at http://www9.georgetown.edu/faculty/amm223/papers.htm

McFadden, Daniel (1974) "Conditional Logit Analysis of Qualitative Choice Behavior" in P. Zarembka Eds. *Frontiers in Econometrics*, New York Acedemic press.

McFadden, Daniel (1978) "Modelling The Choice of Residential Location" in A.Karlgvist et al. Eds *Spatial Interaction Theories and Models*; Amsterdam North Holland.

Ortega, Francesc (2005) "Immigration Quotas and Skill Upgrading" *Journal of Public Economics*, 89 1841-1863.

Ortega, Francesc (2008) "The Short-Run Effects of a Large Immigration Wave: Spain 1998-2008". Manuscript, Universitat Pompeu Fabra.

Ottaviano Gianmarco I.P. and Giovanni Peri (2008) "Immigration and National Wages: Clarifying the Theory and the Empirics" National Bureau of Economic Research Working Paper #14188, July 2008.

Pedersen, P. J., Pytlikova, M., and Smith, N. (2006) Migration into OECD countries 1990- 2000. In Parson and Smeeding, editors, Immigration and the Transformation of Europe. Cambridge University Press, Cambridge.

Peri, Giovanni (2008) "Immigration Accounting: U.S. States 1960-2006" CReAM Working paper # 04/08.

Peri Giovanni and Chad Sparber (forthcoming) ""Task Specialization, Immigration and Wages" *American Economic Journal: Empirical Economics.*

Young, Alwyn, (1995) "The Tyranny of Numbers: Confronting the Statistical Realities of the East Asian Growth Experience," *Quarterly Journal of Economics*, 110: 641-680.

A Data on Migration Flows and Stocks

The International Migration Data (2007) published by the OECD originate from contributions of the national correspondents (National Statistical Agencies) organized in a network called "The Continuous Reporting System on Migration" (SOPEMI). Since the criteria for classifying immigrants and for registering the population may vary significantly across countries the data are not necessarily homogeneous. Also, for each receiving country the IMD (2007) records only the immigrants from the 15 countries of origin with the largest number of immigrants. The OECD statistical annex to the data emphasizes the difficulty of measuring the undocumented/illegal immigrants with this method. Only through censuses or after a regularization program are some of the undocumented immigrants measured.

The total inflows and outflows of the foreign population are derived from population registers and residence and work permits. Due to the fact that removal from the registers due to departure is much less common than the inclusion due to arrival these data are much better at measuring inflows than outflows of immigrants. The countries of origin that we are able to record consistently and that therefore constitute our universe in the bilateral regression analyses are listed in Table A2 of the Appendix.

In the construction of the net immigration flows and immigration stocks for each year and each origin-destination pair we compute the estimated rate of re-migration of the foreign population in each country. These re-migration rates, which measure the percentage of the existing stock of immigrants in a country that leave the country, are calculated to match the stocks of immigrants (in 1990 and 2000) with the flows between those years in each country. The imputed re-migration rates, specific to the country of destination, were: 0.005 for Australia, 0.09 for Belgium, 0.035 for Canada, 0.06 for Denmark, 0.015 for France, 0.05 for Germany, 0.05 for Japan, 0.08 for Luxembourg, 0.02 for Netherlands, 0.12 for Norway, 0.04 for Sweden, 0.04 for Switzerland, 0.02 for the UK and 0.005 for the US.

Tables and Figures: Main Text

Table 1
Destination country determinants of bilateral migration stocks:
1980-2005 in 14 OECD countries

Specification:	*(1) Basic: Income in levels*	*(2) Income in logarithms*	*(3) Decomposition log(wage)-log(employment rate)*	*(4) Including country of destination controls*	*(5) Income is lagged 2 periods*	*(6) Omitting zero migration cells*	*(7) Omitting UK*	*(8) Omitting US*
Income per capita, destination	0.06** (0.01)	0.29** (0.10)		0.04* (0.02)	0.06** (0.01)	0.076** (0.01)	0.06** (0.01)	0.06** (0.01)
Income per worker, destination			0.17** (0.08)					
Employment/population, destination			2.53** (0.93)					
Ln(population), destination				2.39 (1.66)				
Gini, destination				-0.01 (0.02)				
(Percentage of population between 15 and 24), destination				0.002 (0.02)				
Land Border	-1.29 (0.73)	-1.29 (0.73)	-1.29 (0.76)	-1.39 (0.73)	-1.33 (0.73)	-0.12 (0.52)	-1.66* (0.75)	-1.29 (0.73)
Same Language	0.08 (0.40)	0.08 (0.41)	0.08 (0.40)	0.16 (0.42)	0.08 (0.42)	0.09 (0.20)	0.13 (0.40)	0.08 (0.41)
Colonial Ties	2.66** (0.42)	2.65* (0.42)	2.65* (0.42)	2.66** (0.42)	2.63** (0.33)	2.66** (0.33)	2.04** (0.78)	2.65** (0.42)
Log(distance)	-2.02 (0.32)	-2.02** (0.32)	-2.04** (0.32)	-2.02** (0.32)	-2.03** (0.32)	-1.43** (0.14)	-2.14** (0.31)	-2.02** (0.32)
Observations	21,805	21,805	21,805	21,148	19,776	19,343	19,091	21,805

Note: The dependent variable in each regression is ln(Immigrant $Stock_{op}$+1)$_t$. Each regression includes 14 destination country effects and 1825(=73X25) year by country of origin effects. The explanatory variables are all entered with one lag, unless otherwise specified. Income per person is measured in PPP US $ at 2000 prices. The observations are weighted by the destination country population to account for heteroskedasticity of the errors across destination countries and the standard errors are clustered by destination countries to account for correlation within those. **,* imply significance at the 5, 10% level.

Table 2
Bilateral migration flows, income per capita and immigration laws
1980-2005 in 14 OECD destination countries

Specification:	(1) Basic	(2) Log income	(3) Combining entry and stay laws	(4) Laws lagged 2 periods	(5) Adding other destination country controls	(6) Omitting 0 migration cells	(7) Omitting the UK
Panel A: Including bilateral geographic characteristics to control for migration costs							
Income per capita, Destination	0.11** (0.03)	0.75** (0.27)	0.12** (0.03)	0.10** (0.04)	0.11** (0.04)	0.03* (0.01)	0.12** (0.02)
Tightness of immigration entry laws	-0.07** (0.02)	-0.09** (0.02)	-0.04* (0.02)	-0.07** (0.02)	-0.06 (0.04)	-0.06** (0.02)	-0.065* (0.025)
Tightness of asylum laws	-0.19 (0.11)	-0.19 (0.11)	-0.21* (0.10)	-0.25 (0.18)	-0.12 (0.07)	0.12 (0.07)	-0.01 (0.207)
Maastricht	0.58** (0.14)	0.50** (0.18)	0.63** (0.14)	0.51** (0.15)	0.70** (0.12)	0.31 (0.28)	0.79** (0.12)
Schengen	0.38 (0.34)	0.27 (0.34)	0.37 (0.34)	0.37 (0.37)	0.37 (0.38)	-0.66* (0.32)	0.02 (0.25)
Panel B: Including the full set of origin-destination country pair dummies to control for pair-specific migration costs							
Income per capita Destination	0.12** (0.03)	0.77** (0.27)	0.13** (0.03)	0.11** (0.04)	0.11** (0.04)	0.02 (0.01)	0.13** (0.03)
Tightness of immigration entry laws	-0.055* (0.03)	-0.08** (0.03)	-0.03 (0.03)	-0.06** (0.03)	-0.04 (0.03)	-0.044* (0.024)	-0.05 (0.04)
Tightness of asylum laws	-0.20* (0.10)	-0.20 (0.12)	-0.23** (0.10)	-0.26 (0.18)	-0.15* (0.08)	0.13 (0.07)	-0.04 (0.09)
Maastricht	0.68** (0.14)	0.59** (0.18)	0.72** (0.14)	0.60** (0.16)	0.78** (0.13)	0.37 (0.28)	0.81** (0.11)
Schengen	0.25 (0.35)	0.12 (0.35)	0.25 (0.36)	0.24 (0.38)	0.27 (0.40)	-0.45* (0.23)	0.01 (0.26)
Observations	21,805	21,805	21,805	21,148	19,776	5,371	19,332

Note: Dependent variable ln(Immigrant $Flow_{op}$+1)$_t$. All regressions include 14 receiving country fixed effects and 1825(=73X25) year-by-sending-country fixed effects. Regressions of Panel A include border, common language, colonial ties dummies and log distance as bilateral controls while regressions of Panel B include a full set of 1022 (=73X14) country-pair fixed effects. Observations are weighted by the population of the receiving country. Robust standard errors clustered by country of destination are reported in parentheses. **,* imply significance at the 5, 10% level.

Table 3
Impact of yearly gross immigrant flows on production factors, productivity and factors per worker: yearly changes, OLS Estimates

	(1) basic OLS	*(2) Omitting US*	*(3) Europe only*	*(4) 1990-2005*	*(5) including lagged levels*	*(6) Using Net Immigration*
ΔL/L	0.95**	1.03**	1.06**	0.95**	1.09**	3.25**
	(0.12)	(0.08)	(0.08)	(0.17)	(0.29)	(0.32)
ΔEmployment/ Employment	0.99**	1.05**	1.09**	1.06**	1.30**	2.87**
	(0.12)	(0.08)	(0.14)	(0.23)	(0.28)	(0.35)
ΔHours per worker /Hours per worker	-0.03	-0.01	-0.03	-0.11	-0.03	0.36
	(0.08)	(0.08)	(0.07)	(0.13)	(0.07)	(0.24)
ΔK/K	0.75**	0.84**	1.12**	1.04**	0.51**	1.77**
	(0.27)	(0.24)	(0.20)	(0.30)	(0.19)	(0.39)
ΔA/A	-0.06	-0.07	-0.06	-0.23	-0.01	-0.27
	(0.31)	(0.32)	(0.32)	(0.25)	(0.32)	(0.78)
ΔY/Y	0.82**	0.87**	1.00**	0.73**	0.82*	2.48**
	(0.41)	(0.39)	(0.40)	(0.41)	(0.45)	(0.81)
Δk/k	-0.24	-0.20	0.02	-0.02	-0.08	-0.11
	(0.20)	(0.21)	(0.16)	(0.18)	(0.38)	(0.54)
Δy/y	-0.13	-0.15	-0.05	-0.21	-0.06	-0.51
	(0.37)	(0.38)	(0.35)	(0.29)	(0.36)	(0.88)
Observations	350	325	225	210	336	350

Note: Each cell shows the coefficient from a different regression with the dependent variable described in the first cell of the row and the explanatory variable equal to the total flow of immigrants as a share of the initial population of the receiving country. The method of estimation is Least Squares. Each regression includes year fixed effects. The standard errors in parentheses are heteroskedasticity robust and clustered by country. Specification (1) uses all the country-year observations, specification (2) omits the US data, specification (3) includes only the nine continental European countries in our sample; specification (4) uses only observations from the 1990-2005 period, specification (5) includes the lagged value of the level of the dependent variable as explanatory variable and specification (6) uses the "net" immigration flows, namely those obtained after subtracting the estimated re-migration from the existing stock of immigrants. **,* imply significance at the 5, 10% level.

Table 4
First stage of the 2SLS: dependent variable is the gross immigration flow

Instrument:		***(1)*** ***Basic***	***(2)*** ***Omitting US***	***(3)*** ***Europe Only***	***(4)*** ***1990-2005***
Predicted flows from gravity push	***Coefficient***	0.67** (0.03)	0.67** (0.03)	0.67** (0.04)	0.62** (0.03)
	F-test (p-value)	495.1** (0.000)	496** (0.000)	319.2** (0.000)	539.14 (0.000)
	Partial R-Square	0.43	0.44	0.42	0.41
Predicted flows from gravity push and immigration laws	***Coefficient***	0.24** (0.01)	0.24** (0.01)	0.24** (0.01)	0.23** (0.01)
	F-test	301.2 (0.000)	301.5 (0.000)	276.2 (0.000)	301.2 (0.00)
	Partial R-Square	0.32	0.32	0.35	0.43
Observations		350	325	225	210

Note: The dependent variable in all first stage regressions is the immigration rate measured as gross immigration flows relative to initial population. The explanatory variable (instrument) in the first row is the predicted immigration rate using the empirical equation (14) without wage differential or immigration tightness measures. In row 2 we use the predicted immigration flows from (14) omitting wage differentials only. All regressions include year fixed effects. Specification (1) includes all countries and all years 1980-2005, specification (2) omits the US, specification (3) includes only the nine Continental European Countries (not the UK) and specification (4) considers only observations relative to the period 1990-2005. In parentheses below the coefficient estimates we report heteroskedasticity robust standard errors clustered by country. Below the F-statistics we report the probability of rejecting the inclusion of the instruments in the first stage. **,* imply significance at the 5, 10% level.

Table 5
Impact of yearly gross immigrant flows on production factors and productivity:
2SLS estimates, instruments: gravity push factors only

	(1) *basic 2SLS*	*(2)* *Omitting US*	*(3)* *Europe only*	*(4)* *1990-2005*	*(5)* *Using Net Immigration*
ΔL/L	1.02** (0.12)	0.99 (0.08)	1.00 (0.08)	0.96** (0.08)	4.00** (0.29)
ΔEmployment/ Employment	1.22** (0.09)	1.21** (0.10)	1.22** (0.13)	1.22** (0.12)	4.81** (0.37)
ΔHours per worker /Hours per worker	-0.20* (0.10)	-0.20 (0.11)	-0.25** (0.08)	-0.26** (0.10)	-0.80 (0.45)
ΔK/K	1.36** (0.17)	1.37** (0.18)	1.49** (0.20)	1.38** (0.19)	5.37** (0.52)
ΔA/A	-0.13 (0.17)	-0.11 (0.16)	-0.06 (0.14)	-0.37 (0.14)	-0.51 (0.67)
ΔY/Y	0.99** (0.50)	0.99** (0.17)	1.09** (0.20)	0.94** (0.17)	3.91** (0.62)
Δk/k	0.14 (0.13)	0.16 (0.14)	0.23 (0.14)	0.15 (0.14)	0.56 (0.51)
Δy/y	-0.02 (0.20)	0.03 (0.19)	0.08 (0.15)	-0.04 (0.17)	-0.10 (0.80)
Observations	350	325	225	210	350

Note: Each cell shows the coefficient from a different regression with the dependent variable described in the first cell of the row and the explanatory variable equal to the total flow of immigrants as a share of the initial population of the receiving country. The method of estimation is 2SLS using the predicted flow of immigrants from the gravity push factors as instruments. Each regression uses yearly differences by country and includes year fixed effects. The standard errors in parentheses are heteroskedasticity robust and clustered by country. Specification (1) uses all the country-year observations, specification (2) omits the US data, specification (3) includes only the nine continental European countries in our sample; specification (4) uses only observations from the 1990-2005 period, specification (5) uses the “net” immigration flows, namely those obtained after subtracting the estimated re-migration from the existing stock of immigrants. **,* imply significance at the 5, 10% level.

Table 6
5-year differences
2SLS estimates, instruments: gravity push factors only

	(1) ***basic 2SLS***	***(2)*** ***Omitting US***	***(3)*** ***Europe only***	***(4)*** ***1990-2005***
ΔL/L	0.99** (0.09)	0.97** (0.08)	0.97** (0.10)	0.97** (0.08)
ΔEmployment/ Employment	1.18** (0.09)	1.16** (0.09)	1.21** (0.10)	1.18** (0.10)
ΔHours per worker /Hours per worker	-0.19 (0.08)	-0.19 (0.09)	-0.23** (0.06)	-0.22** (0.10)
ΔK/K	1.24** (0.13)	1.25** (0.17)	1.33** (0.18)	1.22** (0.17)
ΔA/A	-0.09 (0.13)	-0.08 (0.12)	-0.03 (0.11)	-0.08 (0.12)
ΔY/Y	0.96** (0.14)	0.97** (0.16)	1.05** (0.18)	0.97** (0.15)
Δk/k	0.06 (0.13)	0.08 (0.13)	0.16 (0.13)	0.04 (0.14)
Δy/y	-0.02 (0.24)	0.01 (0.14)	0.09 (0.12)	-0.04 (0.14)
Observations	70	65	45	56

Note: Each cell shows the coefficient from a different regression with the dependent variable described in the first cell of the row and the explanatory variable equal to the total flow of immigrants as a share of the initial population of the receiving country. The method of estimation is 2SLS using the predicted flow of immigrants from the gravity push factors as instruments. Each regression uses changes over five years between 1980 and 2005 and includes period fixed effects. The standard errors in parentheses are heteroskedasticity robust and clustered by country. **,* imply significance at the 5, 10% level.

Tables and Figures: Appendix

Table A1:
List of the countries of origin of migrants for the bilateral migration data

Countries of Origin		
Algeria	Ghana	Nigeria
Australia	Greece	Norway
Austria	Guatemala	Pakistan
Bangladesh	Guyana	Peru
Belgium	Haiti	Philippines
Bosnia-Herzegovina	Honduras	Poland
Brazil	Hong Kong	Portugal
Bulgaria	Hungary	Romania
Cambodia	Iceland	Russian Federation
Canada	India	Slovenia
Chile	Iran	Somalia
China	Iraq	South Africa
Colombia	Ireland	South Korea
Croatia	Italy	Spain
Cuba	Jamaica	Sri Lanka
Cyprus	Japan	Suriname
Denmark	Kenya	Sweden
Dominican Republic	Laos	Thailand
Ecuador	Lebanon	Tunisia
El Salvador	Malaysia	Turkey
Ethiopia	Mexico	UK
Fiji	Morocco	USA
Finland	Netherlands	Vietnam
France	New Zealand	Zaire
Germany	Nicaragua	

Table A2:
Average values of the variables included in the regressions for bilateral flows;
Separated between countries of origin and countries of destination, in 1980, 1990, 2000, 2004.

variable	1980	1990	2000	2004
GDP per person Origin	7,944	9,442	11,198	12,018
GDP per person Destination	17,979	21,916	28,565	29,022
Employment rate Origin	42%	44%	46%	47%
Employment rate Destination	47%	49%	50%	49%
Gini Origin	0.38	0.39	0.40	0.40
Gini Destination	0.31	0.33	0.33	0.33
Share of population between 14 and 24 years, Origin	9.2%	8.6%	8.82%	8.81%
Share of population between 14 and 24 years, Destination	7.1%	6.1%	5.25%	5.99%
Observations Origin	77	77	77	77
Observations Destination	14	14	14	14

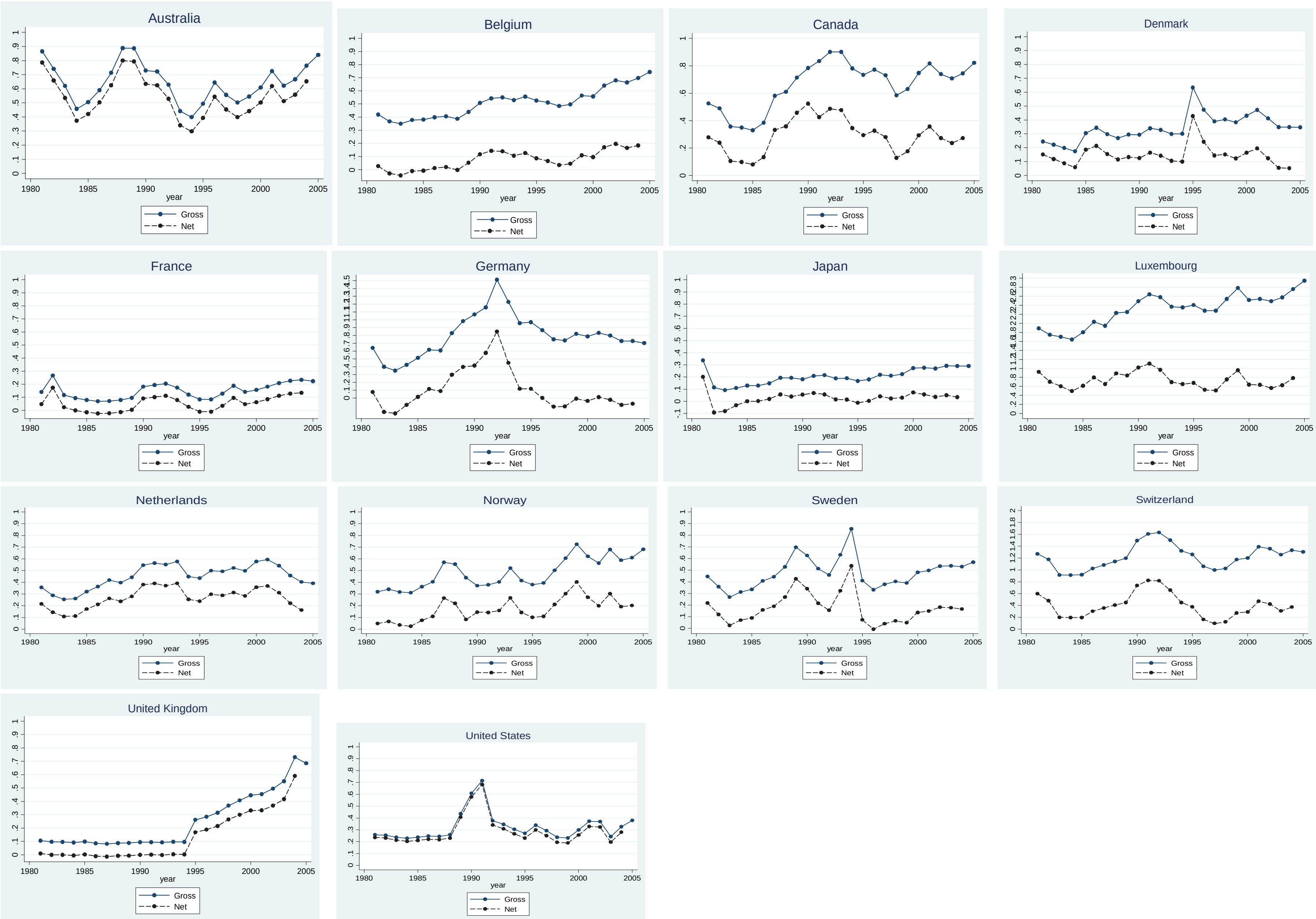
Note:
Per Capita GDP, Employment Rate and Population (available 1980-2004): Penn World Table 6.2. PPP-converted, chain-weighted GDP per worker.
Share of population 15 to 24 (available 1980-2004): United Nations population statistics.
Inequality (Gini Coefficient), (available 1980-2004): Deininger and Squire (1996), only "high quality" observations. The data are available to differing degrees across countries, but we linearly interpolated and extrapolated values for all years as long as there was at least one value given for the country in some year. Sometimes there was only a single value for a country, so we assumed this value for all years. Also, sometimes there were not enough values to generate a reasonable extrapolation, so we assigned the values nearest (in time) to the missing values.

Table A3
Annualized growth rates of inputs, productivity and output

Variable	Observations	Mean	Std. Dev.	Min	Max
		1980-1990			
Real GDP	14	2.72	0.79	1.99	4.83
Total Hours worked	14	0.68	0.84	-0.50	1.78
Employment	14	1.07	0.63	0.18	1.88
Capital Services	14	3.43	1.04	2.25	5.59
Total Factor Productivity	14	1.14	0.71	-0.25	2.50
		1990-2000			
Real GDP	14	2.62	1.02	1.06	4.91
Total Hours worked	14	0.58	0.96	-0.92	2.77
Employment	14	0.92	0.97	-0.67	3.40
Capital Services	14	3.75	1.05	2.06	6.12
Total Factor Productivity	14	1.01	0.65	-0.56	2.38
		2000-2005			
Real GDP	14	1.98	0.86	0.60	3.65
Total Hours worked	14	0.34	0.84	-0.93	1.96
Employment	14	0.73	0.97	-0.41	3.08
Capital Services	14	3.11	1.26	1.37	6.02
Total Factor Productivity	14	0.73	0.54	0.17	1.79

Note: The data were constructed by the authors using the OECD-STAN dataset and the PWT 6.2 data as sources. The exact definition of the variables and of the procedure to construct them is in the main text, section 5.2.

Panel A1: Immigration flows relative to population, Gross and Net for 14 OECD Countries



Panel A2: Tightness of immigration reforms over time. 14 OECD countries 1980-2005