The Effect of Trade and Migration on Income

Francesc Ortega (Queens College CUNY)

Giovanni Peri (University of California, Davis and NBER)*

This Draft: July 2012

Abstract

This paper explores the relationship between openness to trade and to immigration on income per person. To address endogeneity concerns we extend the instrumental-variables strategy first used by Frankel and Romer (1999). We show that distance (geographical and cultural) can be used to build a strong predictor of openness to immigration and to trade. Our instrumental-variables estimates establish a robust, positive effect of openness to immigration on long-run income per capita, using demanding econometric specifications that account for trade openness, the role of institutions, and early development. In contrast the positive effect of trade openness on income is not robust to controlling for the direct effects of geography, providing support for the critique by Rodriguez and Rodrik (2001). We also show that the main effect of migration operates through total factor productivity, consistent with a theory where immigration increases the variety of skills available for production. We provide further evidence in support of this mechanism by showing that the degree of diversity (by origin country) in migration flows has an additional positive effect on income. Finally, we also find that immigration increases (ethnic and linguistic) fractionalization, which are associated to negative effects on income per capita. However, the direct gains from greater skill diversity appear to be larger than the costs arising from increased fractionalization. We do not find evidence of increased income inequality due to openness to immigration or trade.

Key Words: International Migration, Trade, Income per person, Productivity, Geography, Institutions, Diversity.

JEL Codes: F22, E25, J61.

*Francesc Ortega, fortega@qc.cuny.edu. Address: Economics department, 300A Powdermaker Hall, 65-30 Kissena Blvd, Flushing, New York 11367. Giovanni Peri, gperi@ucdavis.edu. Address: University of California, Davis, One Shields Avenue, Davis CA 95616. The authors thank Enrico Moretti, Petra Moeser, Ryuichi Tanaka, Andrei Levchenko for helpful discussions and seminar participants at GRIPS (Tokyo), Collegio Carlo Alberto, UC Berkeley and All UC History Conference for their useful comments.

1 Introduction

Historically, exchange and economic interactions with other countries have been a powerful engine for economic development, technological change, and social progress, particularly for small countries (Alesina, Spolaore and Wacziarg (2000, 2005), Frankel and Romer (1999)). Empirical tests of the effects of economic openness on long-run income have so far focused almost exclusively on *openness to trade*, measured by using observed policies (as in Sachs and Warner 1995, or Lucas 2010), or trade as a share of GDP (as in Frankel and Romer 1999, Rodrik 2000, or Alcala and Ciccone 2004, among many others). Besides the measure issues, these studies have faced important identification challenges due to omitted variables and endogeneity concerns. Specifically, all measures of trade openness are likely to be a function of of a country's level of development, among other factors. It is perfectly plausible that as a country's population gets richer, its consumption pattern shifts toward an increasing (or decreasing) share of foreign-produced goods. In addition factors that stimulate trade openness are likely to affect simultaneously other measures of openness as well as the country's macroeconomic performance.

Our main goal in this paper is to analyze jointly the effects of openness to international trade and migration on income per capita, while taking into account other factors that have been shown to play an important role (namely, institutional quality and early development). Our emphasis is on the role played by *openness to international migration* yet we recognize that it is important to study these two dimensions of openness jointly. The main reason is that international flows of goods and people depend, to a large extent, on the same set of factors.[1] Thus it is possible that the existing estimates of the effects of trade on income are simply a reflection of the effects of immigration or, more generally, those estimates may suffer from ommitted-variable bias.

Between the 1870s and World War I, the so-called First Globalization era, international migration (along with international trade and capital flows) was very high (relative to population). Some economic historians have argued that migration was an important vehicle for economic convergence in terms of factor prices and income levels during that period (Taylor and Williamson 1997).[2] The analysis of the role of immigration in accounting for cross-country differences in income has been neglected because of the relatively low levels of immigration (and trade) in many countries for several decades. However, since the 1980s a Second era of Globalization has begun, rekindling interest in this question. Immigrants contribute to their host countries in a variety of ways: besides raw labor, they bring new ideas and skills, increasing the diversity of productive inputs and becoming a potentially important vehicle for the international diffusion of knowledge. Putterman and Weil (2010) have argued that migration played an important role in the early economic development of many countries and its

[1] Figure 1 reports the partial correlation between trade as a share of GDP and the foreign-born share across the 146 countries included in the Frankel and Romer (1999) sample. Each variable is a residual, after we control for the country size (as measured by the logarithm of population and the logarithm of the area), which may affect openness and income. The Figure illustrates a clear positive and significant correlation between openness to trade and migration, even after controlling for country size.

[2] Taylor (1997a) studies the economic effects of migration on the economic development of Argentina. Taylor (1997b) focuses on the Asia-Pacific region. Immigration seems to have played a much larger role in the former case.

effects have been extremely persistent.[3] In the light of the strong co-movements among the different dimensions of globalization, it seems clear that our understanding of the role of economic openness on development is likely to benefit importantly from the joint analysis of openness to trade as well as to migration.[4]

The empirical challenges entailed by a causal analysis of the effects of migration on income are analogous to those faced by the literature on the effects of trade. It is highly plausible that migrants are not allocated randomly across countries. Rather they choose, to some extent, their destination on the basis of the economic opportunities offered by each country. As a result, one would expect that immigration (both in levels and as a share of population) will be positively correlated with income per capita and other measures of a favorable economic context, such as low unemployment or a healthy institutional environment. To deal with these issues we adapt the identification strategy initially proposed by Frankel and Romer (1999). These authors were interested in estimating the causal effect of trade on income per capita. They proposed an instrumental-variables approach based on the role of geographic distance in accounting for bilateral trade flows, where the main idea is that, *other things equal*, countries that are at a shorter distance from each other face a lower trade cost. This strategy, however, has been criticized by Rodriguez and Rodrik (2001) on the basis of a potential correlation between a country's "relative" geography (e.g. its remoteness relative to other countries) and its "absolute" geography, that is, its location. For historical reasons, countries are larger in some regions of the world than in others. For instance, European countries tend to be much smaller (as well as richer), on average, than countries in the New World. Geography thus not only affects distance (and trade costs) but may also have a *direct* effect on income through climate, soil quality, agricultural productivity, or its disease environment. In fact, several authors have emphasized the role of geography on economic development as mediated by institutional quality. In a series of influential papers, Hall and Jones (1999), Acemoglu et al (2001), and many followers, have explored the interaction between geography, a country's colonization history, and economic development. The essence of the critique by Rodriguez and Rodrik (2001) is that Frankel and Romer (1999) did not suitably control for the direct role of a country's geographic location on economic development. Other authors that have employed this strategy have emphasized the importance of using better data (Noguer and Siscart 2005), controlling for institutional quality, and employing more sophisticated measures of trade openness (Alcala and Ciccone 2004).

On the basis of a large collection of recent empirical studies that analyze the determinants of bilateral migration flows (starting with Karemera et al 2000, and Pedersen et al 2004), we note that bilateral distance (both in a geographic and cultural sense) is also a crucial determinant of migration flows between pairs of countries. Hence, the strategy proposed by Frankel and Romer (1999) can be readily extended to produce an exogenous predictor for openness to international immigration.

This paper offers several contributions. First, we present a simple theoretical framework that extends

[3]We discuss the work by Putterman and Weil (2010) in depth later on.

[4]The inclusion of international capital flows in the analysis is definitely interesting, but is left for future research.

Alesina et al (2000). In this model international trade and migration increase the variety of productive inputs in the economy, which increases total factor productivity and income per person. The model is used to derive our econometric specification and illustrates the need to control for country size. Second, we use new, more comprehensive data on bilateral trade and migration flows (for year 2000), which allow for a more robust implementation of the Frankel and Romer (1999) methodology (henceforth, FR). Furthermore our dataset contains a very extensive set of covariates from a variety of sources, which allows us to estimate much more demanding specifications than previously done in the literature. Third, we extend and evaluate the ability of the FR strategy as a way to uncover the causal effects of immigration (as well as trade) on income exploiting cross-country variation. Fourth, we analyze a large set of outcomes. Besides income per capita, we offer estimates of the effects of trade and immigration on physical capital intensity, human capital, and total factor productivity (following the decomposition proposed by Hall and Jones 1999), income inequality (as measured by the Gini coefficient), ethno-linguistic fractionalization, and patenting activity. Finally, we also provide results for the effects on income of the degrees of diversity (by country of origin) in trade and migration flows, and the education level of immigrants.[5]

Our analysis delivers the following main findings. First, the gravity-based predictor for the share of immigrants in the population performs remarkably well. In contrast to the predictor for the trade share in GDP, it is very robust to alternative specifications and subsamples. Thus the Frankel and Romer (1999) method provides a useful route to try to estimate the causal effects of immigration. Second, our instrumental-variables estimates (based on a cross-country regression of 147 countries) imply that immigration as a share in a country's total population (as well as a share of its human capital) has a large and significant effect on long-run income per capita. A 10 percentage-point difference in the share of foreign born, which is close to the standard deviation across countries, is associated with differences in income per person by a factor between 2.3 and 2.7 (130-170%). We show that this effect is robust to the inclusion of a large series of geographical, climatic, institutional and historical controls. In particular it is not driven by current institutional quality, history or earlier migration, most of which play a direct role in accounting for current cross-country differences in income. In comparison, our estimates suggest a much smaller role for trade openness than found by previous studies. To a large extent this reflects the fact that the trade to GDP ratio is an increasingly poor measure of the degree of trade openness. Third, the immigrant share in the population appears to increase total factor productivity, while leaving unchanged the capital intensity, and possibly incentivizing human capital accumulation (as in Hunt 2012). This pattern is consistent with the predictions of our theoretical model, where immigration increases the diversity of productive skills, leading to higher TFP and income per capita. Fourth, we also show that more diverse immigration flows (by country of origin) and higher human capital in the receiving country are associated to

[5] A paper by Alesina, Harnoss and Rapoport (2012) available as preliminary manuscript and developed independently from this paper also analyzes in detail the connection between diversity of immigrants and income per person across countries.

more beneficial effects of immigration on income per capita. Finally, we find that immigration has no effect on income inequality (Gini coefficient). However, it does appear to increase (linguistic) fractionalization, which appears to have a negative effect on income but its magnitude is small compared to the gains that arise from greater diversity in terms of skills. We also find partial evidence of a positive effect of immigration on innovation (patenting activity), which may be due to the direct contribution of immigrants (as in Gauthier-Loiselle and Hunt 2010) or the result of a general scale effect (as in di Giovanni et al 2012).

The rest of the paper is organized as follows. In section 2 we describe the contributions of our paper in the context of the existing literature. In section 3 we present the model. Section 4 discusses our estimation and identification strategy. Section 5 presents the data and introduces the empirical model. In section 6 we reproduce the analysis of the effect of trade openness on income per person, while in section 7 we introduce our main specifications that includes openness to immigration. Section 8 explores in depth the roles of institutions and early development. Section 9 examines the mediating effect of productivity and diversity by country of origin. Section 10 collects some additional results and Section 11 concludes.

2 Literature Review

There is a vast theoretical literature linking several aspects of openness (or globalization) to income levels and growth.[6] Some authors emphasize the role of openness to trade in promoting innovation, technological diffusion and catch-up (Grossman and Helpman (1991), Rivera-Batiz and Romer (1994), Eaton and Kortum (1996), or Lucas (2010), to name a few). Others have focused on the effect of market size via trade on innovation and growth: Acemoglu (2003) has argued that the size of the market can affect the speed (as well as the direction) of technological adoption; Matsuyama (1992) and Galor and Mountford (2008) have argued that market size may encourage specialization and learning by doing. Finally, Weil (2005) has focused on the efficiency gains experienced by firms subject to international competition.

More closely related to this paper are the empirical studies that attempt to estimate the effects of trade on income per capita. We have already discussed the important contributions by Frankel and Romer (1999), the extensions by Alcala and Ciccone (2004) and Noguer and Siscart (2005), and the critique by Rodrik (2000) and Rodriguez and Rodrik (2001). In addition we note an influential early contribution by Sachs and Warner (1995) who analyze the effect of trade policies (over the period 1965-1990) on economic growth. A number of papers focus on the role of country size in determining the income gains from trade openness: Alesina and Spolaore (1997, 2003), and Alesina, Spolaore and Wacziarg (2000, 2005). There are comparatively fewer empirical studies on the effects of international migration on income per person from a cross-country perspective. While several

[6] For excellent textbook treatments of openness and economic growth, see Acemoglu (2009, chapters 18 and 19) on the roles of knowledge diffusion and trade, Barro and Sala-i-Martin (2004, chapter 8) on technology diffusion and endogenous growth, or Weil (2005, chapter 11) on the relationship between economic growth and openness.

studies analyze the determinants of bilateral migration flows using a gravity equation (such as Adsera and Pytlikova (2012), Beine et al. (2011), Bertoli and Fernandez-Huertas (2011), Clark et al. (2008), Grogger and Hanson (2011), Llull (2011), Mayda (2007, 2010), or Pedersen et al. (2004), to name a few), very few empirical studies have focused on the causal effects of migration on income.[7] In a similar vein, there have been several historical studies of the First Globalization Era (from the middle of the XIX century to World War I). O'Rourke and Williamson (1994), Williamson (1996) and O'Rourke, Taylor and Williamson (1997) analyze the role of mass migration and the other dimensions of globalization on cross-country convergence in terms of income per capita convergence and factor prices.

Finally, our work is also related to the strand of literature studying the role of institutions and early development on economic growth. According to Hall and Jones (1999) and Acemoglu et al (2001, 2002), the main reason why geography appears to be a crucial determinant of cross-country differences in income per capita is that geography was decisive in determining a country's history of colonization, which set the foundations for the existing institutional arrangements in many countries.[8] In particular, good early institutions may have allowed for policies aimed at sustaining free markets, democracy, checks and balances and well-functioning legal and judicial systems. Current cross-country income differences are also closely related to differences in early development several centuries back (Diamond 1997, Comin et al 2010, among others). This point has been recently stressed by Putterman and Weil (2010). These authors show that most existing measures of a country's level of development in the distant past experienced an increase in their explanatory power over current income differences when we take into account the levels of early development of the countries of origin of the *ancestors* of the current population. We shall discuss in detail later on how our results relate to the findings in Putterman and Weil (2010).

3 Theoretical Framework

We present a simple model to derive our main empirical specification for the effect of openness to international trade and migration on income per person. The model builds heavily on Alesina, Spolaore and Wacziarg (2000).[9] Consider N regions in the world, indexed by $i = 1, 2...N$. These regions are partitioned into C countries. The size of each country, S_c, is given by the number of regions it encompasses. Each region i is endowed with human capital (workers) H_i and physical capital K_i. Each region's capital stock is used to produce a differentiated intermediate good, one unit for one unit. Human capital is also differentiated by country of origin. All regions produce a common final good (used as numeraire) by means of the following aggregate production function:

[7] Peri (2012) looks at the long-run effect of immigration on productivity and income per person across US states.
[8] We discuss this point in greater depth below.
[9] As these authors show, this static model can be interpreted as the steady state of a growth model. Hence, we stress that our predictions relate openness to long-run income levels.

$$Y_i = A_i \left(\sum_{j=1}^{N} H_{ij}^{\alpha} \right) \left(\sum_{j=1}^{N} X_{ij}^{1-\alpha} \right), \tag{1}$$

where $0 < \alpha < 1$. Expression (1) implies that producers in any region i have access to a full range of varieties for intermediate goods and human capital. H_{ij} denotes the units of human capital of variety j used in production of good i. Likewise, X_{ij} are units of intermediate good j used in region i.

Intermediate goods and workers are geographically mobile but subject to iceberg-type costs. Intermediate goods are shipped costlessly across regions within the same country. However, when Z^X units of intermediate good X are shipped to a foreign region only $(1-\gamma^X)Z^X$ units reach the destination, where $0 \leq \gamma^X \leq 1$ denotes the cost of shipping internationally as share of the goods' value. We denote by p_i the price charged by the producer of intermediate good i to ship one unit. The shipping costs (zero for domestic shipments) are paid by the buyer. Likewise, there are costs associated to hiring a foreign-born worker. These costs can be thought of as the additional costs of recruiting abroad, sponsoring an immigrants or training costs paid by the employer to help adapt foreign skills to the host economy. When Z^H foreign workers are hired by a firm, only $(1-\gamma^H)Z^H$ units are available for production of the final good, where $0 \leq \gamma^H \leq 1$ is the immigration cost per unit of human capital. Factors are paid their marginal products. For tractability we impose the following symmetry assumptions: $A_i = A$, $K_i = K$, and $H_i = H$, for all regions $i = 1, 2, ..., N$.

Let us now characterize the demand for domestic and foreign factors of production for a given region i. The marginal product of a unit of intermediate good purchased from a domestic producer from region j *in the same country* is

$$\frac{\partial Y_i}{\partial Z_{ij}^X} = A_i(1-\alpha) \left(Z_{ij}^X \right)^{-\alpha} \left(\sum_{k=1}^{N} H_{ik}^{\alpha} \right). \tag{2}$$

Let us now compute the marginal product of a unit of intermediate good purchased from *foreign* producer j', keeping in mind that only $X_{ij'} = (1-\gamma^X)Z_{ij'}^X$ units are available for production when $Z_{ij'}$ units are purchased. Then

$$\frac{\partial Y_i}{\partial Z_{ij'}^X} = A_i(1-\alpha)(1-\gamma^X)^{(1-\alpha)} \left(Z_{ij'}^X \right)^{-\alpha} \left(\sum_{k=1}^{N} H_{ik}^{\alpha} \right). \tag{3}$$

In a symmetric equilibrium all producers charge equal prices to all destinations (net of shipping costs), that is, $p_j = p_{j'}$. As a result each region purchases equal amounts of all domestically produced varieties (Z_D^X) and equal amounts of all foreign varieties (Z_F^X).[10] Equal prices (net of shipping costs) and profit maximization imply that the marginal products of domestic and imported intermediate capital goods will be equalized:

$$Z_F^X = \theta^X Z_D^X \tag{4}$$

[10] To guarantee existence of a symmetric equilibrium we also need to assume that all countries are equally sized in terms of the number of regions.

where $\theta^X = (1-\gamma^X)^{1-\alpha/\alpha}$.[11]

In similar fashion, wages (net of training costs) will be equal across regions in a symmetric equilibrium. Thus profit maximization will lead to equalization of the marginal products of domestic (Z_D^H) and foreign workers (Z_F^H). Thus

$$Z_F^H = \theta^H Z_D^H \tag{5}$$

where $\theta^H = (1-\gamma^H)^{\alpha/(1-\alpha)}$. Let us now turn to the resource constraints for intermediate goods and workers. The stock of capital in a region is used to produce its own variety of intermediate good. Then Z_D^X units are shipped to each region within the same country and Z_F^X are shipped to each region in another country. Similarly, Z_D^H workers will migrate to each domestic region and Z_F^H will migrate to each foreign region. The resulting resource constraints for each variety of human capital and intermediate input satisfy

$$SZ_D^X + (N-S)Z_F^X = K \tag{6}$$

$$SZ_D^H + (N-S)Z_F^H = H. \tag{7}$$

We can use these equations to derive closed-form solutions:

$$Z_D^X = \frac{K}{S+(N-S)\theta^X} \tag{8}$$

$$Z_D^H = \frac{H}{S+(N-S)\theta^H} \tag{9}$$

plus $Z_F^X = \theta^X Z_D^X$, and $Z_F^H = \theta^H Z_D^H$.

Let us now use these expressions to derive the measures of openness to international trade and migration that we will employ in the empirical section. Let us define the trade to GDP ratio, for short, the *trade share* (TSH) as the sum of exports plus imports as a share of GDP in country i:[12]

$$TSH_i = 2(1-\alpha)\frac{Z_F^X(N-S)}{Z_F^X(N-S)+Z_D^X S} = 2(1-\alpha)\frac{\theta^X(N-S)}{\theta^X(N-S)+S}. \tag{10}$$

Clearly, given country size, an increase in trade openness, θ^X, would increase the trade share. And an increase in size of the country, S, for a given degree of trade openness (as long as $\theta^X < 1$), will reduce the trade share. Expression (10) shows that the trade share also depends on the elasticity of final output to intermediates $(1-\alpha)$ and on the overall size of the world economy N. Similarly, we define the migration share (MSH) as the

[11] Note that $Z_F^X < Z_D^X$ as long as $\gamma^X > 0$.

[12] Imports are equal to exports in this model. In symmetric allocations the price of intermediate goods is the same (net of shipping costs) for all regions. Hence, the value of imports plus exports relative to the total value of intermediate goods is equal to twice the ratio of exported quantities relative to total quantities. Coefficient $(1-\alpha)$ is the share of capital in total income in symmetric allocations.

foreign-born share in the population.[13] That is, for country i,

$$MSH_i = \frac{Z_F^H(N-S)}{Z_D^H S + Z_F^H(N-S)} = \frac{\theta^H(N-S)}{S+\theta^H(N-S)}. \tag{11}$$

It is easy to see that, for a given country size S, the migration share depends positively on openness to immigration. Conversely, for given openness θ^H, the migration share depends negatively on the size of the country. Finally, substituting (8) and (9) into (1), we can express real GDP in country c, which is constant across regions within a country, as:

$$Y_c = A[S_c(1-(\theta^X)^{1-\alpha}) + N(\theta^X)^{1-\alpha}][S_c(1-(\theta^H)^{\alpha}) + N(\theta^H)^{\alpha}]H^{\alpha}K^{1-\alpha}. \tag{12}$$

Dividing by the initial population in the region, H_c, we can now compute GDP per capita:

$$y_c = TFP\left(A_c, \theta_c^H, \theta_c^X, S_c\right)\left(\frac{K_c}{H_c}\right)^{1-\alpha}, \tag{13}$$

where it has been convenient to reintroduce the country subindices for all variables. The first term collects all the determinants of total factor productivity in this model and the second is the factor intensity (capital-labor ratio). The previous expressions make clear that openness to migration θ^H and openness to trade θ^X affect positively TFP and income per person.[14] Similarly, for a given degree of openness, an increase in the size of the country, S, also increases productivity. We note also that this expression allows other factors to also affect TFP, such as government policies, institutions or social norms, which are contained in the term A_c.

4 Estimation

4.1 Empirical Specification and Instruments

Taking a log linear approximation of expression (13) we obtain the following relationship between income per person and openness to trade and immigration for country c:

$$\ln y_c = \beta_0 + \beta_1 \ln \theta_c^X + \beta_2 \ln \theta_c^H + \beta_3 \ln S_c + \beta_4 \mathbf{X}_c + \varepsilon_c \tag{14}$$

Coefficients β_1 and β_2 represent the long-run elasticity of income per person to trade and migration openness, respectively. β_3 represents the long run elasticity of income per person to country size, while $\mathbf{X}_c$ is a vector that includes other determinants of long-run output per person, such as productivity ($\ln \widetilde{A}_c$), human capital

[13]Note that the TSH and MSH are constant across regions within a country.
[14]Recall that N, the number of regions in the world, is obviously larger than S, the number of regions in a given country.

$(\ln \widetilde{h}_c)$ and physical capital $(\ln \widetilde{k}_c)$. The zero-mean term ε_c allows for idiosyncratic deviations of $\ln y_c$ from its steady state and is uncorrelated with the other explanatory variables $\mathbf{X}_c$. In equation (14) we allow the trade and migration costs to be country specific (indexed by c). This equation, though, cannot be estimated because openness to trade and migration (θ_c^X and θ_c^H) are not observable. Hence, we will rely on the empirical counterparts derived in equations (11) and (10). In particular we linearize expression (11) around the means and obtain:

$$MSH_c \approx \mathbf{A} + a_1 \ln \theta_c^H - a_2 S_c + \mathbf{a}\Xi_c^M. \tag{15}$$

Scalar $\mathbf{A}$ is a function of the average values of the parameters, while $\mathbf{a}$ is a vector of partial derivatives of MSH_i with respect to H, N and other potential determinants of the migration share. Coefficients a_1 and a_2 are the semi-elasticities of the migration share to migration openness and to country size, respectively. Residual term Ξ_M is a vector containing H, N, and other observable and unobservable determinants of MSH.

Similarly we can linearize expression (10) to obtain:

$$TSH_c \approx \mathbf{B} + b_1 \ln \theta_c^X - b_2 S_c + \mathbf{b}\Xi_{T,c}. \tag{16}$$

Scalar $\mathbf{B}$ collects the average values of the parameters affecting trade. Coefficients b_1 and b_2 are the semi-elasticities of the trade share to openness and to country size, respectively. The remaining terms are analogous to the previous equation. We can now solve equation (15) with respect to $\ln \theta_c^H$, equation (16) with respect to $\ln \theta_c^X$, and substitute these into equation (14) to obtain:

$$\ln y_c = \beta_0 + \beta_M MSH_c + \beta_T TSH_c + \beta_S \ln S_c + \beta_4 \mathbf{X}_c + \beta_5 \Xi_c + \varepsilon_c. \tag{17}$$

In expression (17) β_T is equal to β_1/a_1, β_M is equal to β_2/a_2, and $\beta_S = \beta_3 + \beta_1/b_1 + \beta_2/b_2$. Importantly, the term $\beta_5\Xi_c$ is a linear combination of the residual determinants of trade, $\mathbf{b}\Xi_{T,c}$, and immigration, $\mathbf{a}\Xi_{M,c}$. Hence it is *correlated* with MSH_c and TSH_c by construction. To the extent that this term contains some unobservable factors it will bias the OLS estimates of the following regression model:

$$\ln y_c = \beta_0 + \beta_M MSH_c + \beta_T TSH_c + \beta_S \ln S_c + \beta_C \mathbf{Controls} + u_c, \tag{18}$$

where the term $\beta_C\mathbf{Controls}$ includes $\beta_4\mathbf{X}_c$ and the observable components of $\beta_5\Xi_c$.

4.2 Gravity regressions

In order to obtain consistent estimates of β_M and β_T we exploit the cost-driven determinants of trade and migration for country c. Namely we use expressions (15) and (16), and isolate factors that only affect trade costs (θ^X) and migration costs (θ^H). Building on Frankel and Romer (1999), we begin by building a predictor for bilateral trade and migration shares:

$$
\begin{aligned}
\ln x_{cj} \quad = \quad & \gamma_1 \ln(Dist)_{cj} + \gamma_2 \ln(Pop)_c + \gamma_3 \ln(Pop)_j + \gamma_4 \ln(Area)_c + \gamma_5 \ln(Area)_j \\
& \gamma_6 (Landlocked)_c + \gamma_7 (Border)_{cj} + \gamma_8 (ComLang)_{cj} + \gamma_9 (Colony)_{cj} + \\
& \gamma_{10} \ln(Dist)_{cj}(Border)_{cj} + \gamma_{11} \ln(Pop)_c (Border)_{cj} + \gamma_{12} \ln(Pop)_j (Border)_{cj} \\
& + \gamma_{13} \ln(Area)_c (Border)_{cj} + \gamma_{14} \ln(Area)_j (Border)_{cj} + \gamma_{14} \ln(Landlocked)_j (Border)_{cj} + u_{cj}.
\end{aligned} \tag{19}
$$

The dependent variable x_{cj} is either MSH_{cj}, the stock of immigrants from country j to country c relative to the population of country c, or TSH_{cj}, the total value of trade between country c and j divided by the GDP of country c. The explanatory variables are the distance between the two countries, the population and area of each country, dummies for country c being landlocked, a dummy for country c and j sharing a border, a dummy for speaking a common language and a dummy for sharing a colonial past.[15] Finally, the interactions of the border dummies with the distance, population area, and landlocked dummies are also included. In one specification we include origin and destination dummy variables, which absorb the origin-specific and the destination-specific regressors.[16] We stress that bilateral distance (in a geographic, cultural, and institutional sense) is meant to proxy for bilateral trade and migration costs.

Once we have estimated the gravity regressions (19) we aggregate them across destinations j to obtain the predicted trade and migration shares for each country c. More specifically, define Z_{cj} to be the vector of explanatory variables included in (19) and γ_M to be the vector of coefficients in the regression for migration flows, while γ_T is the vector of coefficients in the bilateral trade regression. Then we define the cost-driven trade share for country c as:

$$\widehat{TSH}_c = \sum_{j \neq c} \exp(\widehat{\gamma}_T Z_{cj}). \tag{20}$$

Similarly we define the cost-driven migration share in country c as:

[15] The role of language in shaping international migration flows has been firmly established by Adsera and Pytlikova (2012). Their findings also show that sharing a common language matters more for non-English-speaking destinations.

[16] In order to safeguard exogeneity, the predictions are based solely on the regressors included explicitly. That is, the estimated fixed effects are not used in the construction of the predicted bilateral trade and migration shares that are then aggregated to build the instruments.

$$\widehat{MSH}_c = \sum_{j \neq c} \exp(\widehat{\gamma}_M Z_{cj}). \tag{21}$$

These predictors reflect the variation in bilateral trade and migration flows driven by country size and bilateral costs, arising from a country's geography and its cultural make out vis-a-vis other countries. Hence, once we control for the size of the country, variation in the predicted values of $\widehat{TSH}_c$ and $\widehat{MSH}_c$ is fundamentally driven by pairwise comparisons in geography and culture that affect trade costs θ_c^X and migration costs θ_c^H.

The trade and migration literature have estimated gravity equations like (19) repeatedly and our goal is not to have a structural interpretation of the coefficients $\widehat{\gamma}_T$ and $\widehat{\gamma}_M$ but rather to use the predictors (20) and (21) as instruments for the trade and migration shares. Nevertheless, we note that the more recent model-based implementations of the gravity equation to predict trade (e.g. Anderson and Van Wincoop 2003) and migration (e.g. Ortega and Peri 2009, 2012) include a full set of country of origin and of country of destination fixed effects. These are needed to capture the effect of "multilateral resistance" and not including them is thought to introduce omitted-variable bias. Hence, in our empirical implementation we also estimate (19) augmented by a set of country of origin and country of destination fixed effects, but do not use those estimated effects in our predictors, which are still calculated exactly as in (20) and (21).[17]

4.3 Discussion of the Identification Strategy

As described above our gravity-based identification strategy is an extension of Frankel and Romer (1999). The literature has emphasized two potential challenges to this strategy. The first challenge is that geography may play a direct role in determining a country's current income (Rodriguez and Rodrik 2001). This would invalidate the exclusion restriction and question the validity of the instruments. The second challenge arises from the potential role of institutional quality on income, and its dependence on geography via a country's colonization history (Hall and Jones 1999, Acemoglu et al. 2001). This would also invalidate our instrumental-variables strategy. We discuss each of these challenges in detail.

As noted by Rodriguez and Rodrik (2001), a country's remoteness (that is, its average distance from other countries) may be correlated with its geographic location. For example, European countries tend to be small and at a short distance from each other, in contrast with the country configuration in the American continent. Clearly, geographic location is very likely to have a "direct" effect on income per person through a variety of channels: soil quality and agricultural productivity (Comin et al 2010), disease environment (Weil 2007), or the quality of institutions (Hall and Jones 1999).[18] As a result, the instruments' exclusion restriction has been

[17] The reason for not using the estimated dummies in prediction is that these dummies absorb all sources of cross-country variation, potentially including factors that may not be considered exogenous to income per capita.

[18] In particular, latitude influenced the probability of being settled by population from a Western European country (Hall and Jones 1999). Geographic location also influences the disease environment, which has been argued affected the kind (and quality) of institutions established by colonizers (Acemoglu, Johnson and Robinson 2001). Putterman and Weil (2010) have emphasized

called into question.

In order to deal with these concerns we use two approaches. First, we follow Noguer and Siscart (2005) and include several controls in our basic model: distance to the equator, regional dummy variables (sub-Saharan Africa, Latin America, and East Asia), other geographic controls (namely the percent of land in the tropics, a landlocked dummy, and average distance to the coast), weather controls (average temperature and average humidity), morbidity variables (incidence of malaria and yellow fever) and colonial-history controls (former French colony, former English colony). All these variables can be considered as exogenous regressors and absorb the effect of geography, relaxing the assumptions required for the validity of the instrument.[19]

Second, besides controlling for geography explicitly, we also experiment with an econometric specification that explicitly includes institutional quality as an additional (endogenous)regressor. In the literature this has been one of the main channels through which geography is thought to affect income. Of course estimating the role of institutions raises additional endogeneity issues that need to be addressed.[20] The importance of explicitly accounting for institutional quality can be illustrated as follows. Good institutions, such as protection of property rights, granting balance of powers and ensuring economic freedom, are certainly a key determinant of a country's productivity and overall economic success. At the same time, because of high persistence, the quality of a country's current institutions may be closely related to its institutions in the past, which in turn has been partially determined by that country's latitude or distance from Europe. Failing to account for the direct effect of institutions on income may bias our estimates of the effects of openness to trade and migration.

Several authors have recently emphasized the extreme persistence of early development and its role in accounting for current income per capita through a variety of channels and, in particular, by leading to high degrees of openness to trade and migration (Chanda and Putterman (2005, 2007), Tabellini (2010), Putterman and Weil (2010)). Taking pre-1500 political institutions as pre-determined to current income (as done by Putterman and Weil (2010)), we address this point by controlling directly for early political development.

We address these challenges by implementing two distinct, though related, identification strategies. Our first strategy employs a specification that explicitly controls for the role of geography (including location, climate, disease, and soil quality regressors), and uses gravity-based predictors for the trade and migration shares as instruments. The second identification strategy is based on a specification that includes a measure of current institutional quality as an additional (endogenous) regressor. Here we follow Alcala and Ciccone (2004) and expand our vector of instruments by including distance to the equator and some other variables that capture the degree of influence of European colonial powers on good institutions. It is also important to keep in mind that our empirical measures of trade and migration openness, namely the trade and migration shares, are

the importance of the quality of the institutions of a country's ancestors, as opposed to the past institutions of a given geographic location.

[19]Our vector of controls contains the main variables included by Noguer and Siscart (2005) plus several additional variables.

[20]This was the strategy pursued by Alcala and Ciccone (2004) in their analysis of the effects of trade openness and institutional quality on income per person.

imperfect proxies for the underlying openness highlighted in our theoretical framework. Our instrumental-variables estimates may also be helpful in addressing this issue of measurement error.

5 Data and Summary Statistics

Our bilateral trade data is from the NBER-UN dataset.[21] This database uses National Accounts in order to obtain bilateral trade data and checks the importing as well as the exporting country statistics in order to improve on accuracy. We also cross-examined these data with the International Trade database (BACI) available at CEPII.[22] The UN-NBER database has a slightly larger coverage, filling some missing values, especially for smaller bilateral trade values. This dataset has information on imports for over thirty thousand bilateral pairs for the year 2000. We then replace missing values with zeros. We note that this will have no effect on our linear-in-logs predictors since the zero values will be dropped anyway. However, it will allow us to increase the number of observations in the non-linear estimation (Poisson pseudo-maximum likelihood). We build the trade flow for each country pair by adding imports and exports.

The bilateral migration data are from Docquier et al. (2010) and are described there in greater detail. They measure the number of people (older than 25) born in each of 194 world countries and residing in any of these countries in 2000. The original sources of these data are national censuses conducted around the year 2000. Specifically, for 194 countries we have their working-age population broken down by country of birth and education (with or without college education). There are 38,031 bilateral cells, none of which have missing values, however a large fraction contain zeros, corresponding to the fact that there are no migrants between many country pairs. We complement the bilateral dataset with data on geography (bilateral distance, a dummy for sharing a border, and the number of landlocked countries in the pair), country size (in terms of population and area), language (common languages), and colonial ties. These data are from the BACI dataset, provided by CEPII and described in Head, Mayer and Ries (2010). The resulting dataset has over 33,000 bilateral observations for trade and migration flows, around 24,000 of which have nonzero observations for trade flows, and about 8,000 have nonzero observations for migration flows (see Table 2). In comparison FR had 3,220 bilateral trade flows and Noguer and Siscart (2005) had 8,906, hence the coverage of our trade data is significantly larger than in the previous studies.

We now turn to our country-level dataset, which spans 188 countries, 146 of which were present in the FR dataset. To maintain comparability we estimate our main models on this sub-sample. The remaining 42 countries tend to be low-income and small in size, which raises some issues about the quality of their data.

[21] We thank Rob Feenstra for sharing these data with us. The data are available and described in detail at this website http://cid.econ.ucdavis.edu/.

[22] The correlation coefficient with the CEPII bilateral trade data for year 2000 is 0.99 when restricting to the same country pairs. These data can be downloaded at http://www.cepii.fr/anglaisgraph/bdd/baci.htm

However, we made a significant effort to extend the coverage for most variables, and thus also present results for the full sample.[23] Our main variables of interest are real GDP per person (PPP-adjusted), a measure of income inequality (Gini coefficient), the trade share in GDP (defined as imports plus exports over PPP-adjusted GDP), real trade openness (as in Alcala and Ciccone (2000)), the foreign-born share (both in terms of population and of human capital), an index of institutional quality and a measure of patents per person. The GDP and trade shares are from the Penn World Tables (version 7.0), the foreign-born share is calculated using the Docquier (2010) data. Along the lines of Hall and Jones (1999) and Alcala and Ciccone (2005) we build a measure of institutional quality. Our index of institutional quality is based on data in Acemoglu, Johnson and Robinson (2001) and is built as a simple average of an index of average protection against expropriation risk and an index of constraints on the executive (around year 1990).[24] Acemoglu, Johnson and Robinson (2001) is also our source for several additional variables that measure absolute geography, disease environment, climate, institutional characteristics and cultural traits. We use the database from Alesina et al. (2005) for ethnic, linguistic and religious fractionalization.

Table 1 reports some basic descriptive statistics and the source for the main variables of the paper. The mean real GDP per person is $10,682, with a standard deviation that is twenty percent larger than the mean. The mean Gini coefficient (from the UNU-WIDER dataset) is 41.53 (standard deviation 11.04). The mean trade share is 90%, with a standard deviation of 50 percentage points. The average degree of *real* trade openness is 0.50 (with a standard deviation of 0.42).[25] The correlation coefficient between the two variables is 0.76. The foreign-born share, defined as the foreign-born population over the total population in the country has a mean of 0.04 (standard deviation 0.08), and ranges from virtually zero to 0.52. When we build the migration share in terms of human capital (as opposed to population), we rely on estimates of Mincerian returns and the share of college-educated. The resulting migration share (in terms of human capital) is 0.09 on average (standard deviation 0.15), and ranges from zero to 0.80. These figures reflect the fact that immigrants are more educated than natives in many countries. As one would expect, the correlation coefficient between the two definitions of the migration share is very high (0.96).

We obtained two important control variables from Putterman and Weil (2010). The first is an index of early political development (the so-called *Statehist* variable). This index characterizes the level of sophistication of the sociopolitical institutions in the countries of origin of the ancestors around year 1500 of the current population for each country. This index is available for 160 of the countries in our sample. We also use their data to compute the share of the current population (year 2000) in each country whose ancestors in year 1500 lived in a different country. This is a measure of openness to international migration over the very long run. The average

[23] We have performed most of the regressions also on the full dataset, with very similar findings (available upon request).

[24] For more details see page 1397 in Acemoglu, Johnson and Robinson (2001).

[25] Following footnote 4 in Alcala and Ciccone (2004), real trade openness is defined as (nominal) openness times the price level, which undoes the dependence on relative nontradeable goods prices.

value is 0.24, with a large standard deviation (0.32), and ranges from zero to 100 percent. In addition the Table reports descriptive statistics on some of our main control variables (population, area, percent of the population speaking European languages), measures of income inequality (used as dependent variables used later in the analysis), and a series of variations on our gravity-based predictors for the trade share (TSH) and migration share (MSH), which are the core of our instrumental-variables strategy. We discuss their construction in detail below.

6 The Effect of Trade Openness on Income

We begin our empirical analysis by presenting the estimates of the gravity models for bilateral trade flows, and by reproducing the results of the previous literature that focused on the effect of trade openness on income.

6.1 Gravity Estimates for Trade Flows

Table 2 (left panel) reports the estimates of the gravity model for bilateral trade flows, based on equation (19) where the dependent variable is the log of the bilateral trade share. Column 1 reports the estimates of a linear-in-logs model. Column 2 reports the estimates of a similar model that includes country of origin and country of destination dummy variables. This specification will be helpful in assessing if the standard predictor (column 1) suffers from omitted-variable bias. We also note that the fixed-effects specification is better motivated theoretically (see Anderson and van Wincoop (2003) regarding trade flows, and Ortega and Peri (2009) and Bertoli and Fernandez-Huertas (2011) in the context of international migration).[26] Finally, in column 3, we follow Santos-Silva and Tenreyro (2008) and adopt a non-linear estimation method (Poisson pseudo-maximum likelihood). As argued by these authors, the latter estimation method addresses important heteroskedasticity issues and also boosts the sample size because it can naturally accommodate observations with zero bilateral values.

Qualitatively, the point estimates are similar across the three columns and have the expected signs: geographical distance is associated with lower bilateral trade shares, while destination size (in terms of population), sharing a common language and having colonial ties are all associated to larger bilateral trade shares. In particular, we note that the coefficient on log distance is very similar in the first two columns. This suggests that the vector of explanatory variables included in the first column is large enough to help identify the crucial role of bilateral distance in determining trade flows.[27] We also note that the point estimates of origin population are much smaller (even negative) than the corresponding destination size coefficients. This reflects the construction

[26] It is important to keep in mind that our goal here is not to identify the structural parameters of the underlying model for trade and migration flows. Our aim is to build predictors of these flows that can be considered plausibly exogenous.

[27] The same is true regarding bilateral migration flows (the right panel). We note though that the coefficient on log distance in column 6 is very similar to those in columns 4 and 5, while this is not the case for trade flows (column 3). This suggests that our estimates for migration flows may be more robust than the estimates for trade flows.

of trade shares where the denominator is the origin-country economic size, as measured by its GDP. The goodness of fit is obviously substantially higher in the specification including the fixed effects (column 2). Compared to the original exercise performed by Frankel and Romer (1999) our gravity model includes information on past colonial ties, along the lines of Head, Mayer and Ries (2010), which increases the explanatory power of the model and the resulting strength of the predictor for the trade share.

As explained earlier, we use our estimates of the vector of coefficients γ_T, obtained from specifications (1), (2) or (3) in Table 2, to build predicted values for all bilateral country pairs (not just those pairs used in the estimation). We then aggregate these predicted values following equation (20) to obtain the predicted trade share for each country. The right panel of the Table reports the estimates for the migration gravity regressions. For now it suffices to note that the overall pattern of coefficient signs is similar to that obtained for bilateral trade flows. We will return to the migration gravity regressions later in Section 7 below.

6.2 Replication of the Literature

The benchmark of our replication is the initial work of Frankel and Romer (1999), and a more updated version of the same exercise by Noguer and Siscart (2005) that explicitly deals with the criticism raised by Rodriguez and Rodrik (2001). More specifically, in this section we focus on a restricted model where the only dimension of openness is the trade share in GDP:

$$\ln y_c = \beta_0 + \beta_{TR} \left(TRsh_c \right) + \beta_P \ln Pop_c + \beta_A \ln Area_c + \beta_C \mathbf{Controls} + u_c \tag{22}$$

In equation (22) the dependent variable is income per person in country c measured in 2000 US Dollars, corrected for PPP as in the Penn World Tables. We include as explanatory variables the logarithm of area and population to capture the effect of country size. As an instrument for the trade share we use the gravity-based predictor proposed by FR and constructed using the estimates of Table 2 (column 1) described above.[28]

Table 3 reports the two-stage least-squares estimates for equation (22). Columns 1-4 of Table 3 report the estimates of the basic model, which includes only controls for country size (logs of area and population). The dependent variable in columns 1 through 3 is the log of income per person, whereas we use the log of income per worker in column 4 (as done in Alcala and Ciccone 2004). Our main sample is the one used by FR and contains 146 countries. We also report results with the largest sample that we could assemble (181 countries, column 3). Column 1 reproduces the finding in FR, where the trade share appears to have a positive and significant effect on income per person. Specifically, increasing the trade share by one percentage point is associated with a 2.5% increase in income per person. This estimate falls well within the range obtained by FR, who report estimates

[28] Our model contains an extended vector of controls, compared to FR. The main findings remain unaffected when using exactly their set of regressors (but standard errors are larger in that case).

between 1.97 and 2.96.[29] Notice that the log-linear gravity-based predictor for the trade share is a relatively strong instrument (the F-statistic for the excluded instruments in the first-stage regression is 13.71, which is close to the critical value of the most stringent test (16.38) tabulated by Stock and Yogo (2005)). Column 2 reports estimates for the same specification but where we have estimated the gravity model using a non-linear estimator (Poisson). In this case the instrument is very weak, which is reflected in a non-significant and very imprecise estimate for the coefficient on the trade share, signalling a lack of robustness. In order to give the FR strategy the best shot, in the following specifications we use the predicted trade share based on the linear-in-logs estimates of the bilateral trade shares. Columns 3 and 4 illustrate that the FR result also holds when using the whole sample of countries of replacing the dependent variable by income per worker.

Columns 5 through 7 include further controls, and represent the essence of the Rodrik and Rodriguez (2001) critique, which argues that geography may play a direct role in determining income, over and beyond its effect through trade costs and the resulting trade openness. Column 5 includes distance from the equator as a control. This variable is highly significant, confirming the results in Hall and Jones (1999). Importantly, the coefficient on the trade share falls dramatically and becomes statistically insignificant. Column 6 examines an alternative way of controlling for geographic location, by including a vector of regional dummies (sub-Saharan Africa, East Asia, and Latin America). The results confirm the findings in the previous column: the initial effect of the trade share is greatly diminished by controlling directly for geographical location. Column 7 includes additional variables, to control for geography, climate, soil quality, disease environment, and the colonial past.[30] The point estimate of the trade share coefficient remains very small and insignificant. The reason for the insignificant coefficient, however, is not only that the instruments are weak. As illustrated by the OLS estimates reported in column 9, once we include the geography and colonial controls, even the partial correlation between trade share and income falls to zero.[31] Finally, in column 8 (fixed effects predicted trade share), we do not obtain any significant effect of trade, with clear signs of a problem of weak instruments.[32]

In conclusion, our findings provide support for the critique raised by Rodriguez and Rodrik (2001) regarding the FR strategy as applied to bilateral trade flows. Underlying this finding we think that the overall trade share is too rough a measure of the increasingly complex trade flows between countries. For example it does not

[29] See the Table 3, page 387 of Frankel and Romer (1999). Our prediction for the trade share includes a few additional regressors not used by FR. If we use exactly their regressors the point estimate is 3.24, with standard error 1.40.

[30] The variables included besides the three regional dummies are: distance from equator, average temperature, humidity, an index of soil quality, an index of disease environment, a dummy for former English or French colonies and the share of European settlers and their descendants measured as of 1900.

[31] Our results differ from those of Noguer and Siscart (2005), who find that the positive effect of trade openness on income is robust to the inclusion of the geographic controls. We use different (more complete and updated) data, which accounts for the disparity in results. At minimum our results suggest that the effect of trade openness uncovered by these authors using the Frankel and Romer methodology is sensitive to the data used in the estimation. It is also possible that the trade to GDP ratio has become an increasingly worse proxy of openness to trade.

[32] We note that our fixed-effects gravity predictor does not include all the regressors that would be needed to provide a structural interpretation to the estimates. In addition the predictions we build on the basis of the estimated model do not make use of the estimated fixed effects in order to obtain a more exogenous predictor. Obviously, this reduces severely the prediction power (relevance) of this instrument.

distinguish between trade in final or intermediate goods, which have played an increasingly large role over the last few decades. Nevertheless some interesting questions remain: can the FR methodology be used successfully to identify the effect of openness to migration on income? To what extent was the finding in FR driven by openness to migration rather than to trade? We tackle these questions next.

7 The Effect of Openness to Trade and Migration on Income

The empirical growth literature has almost exclusively focused on a single dimension of economic openness, namely, openness to international trade.[33] This viewpoint neglects the well established fact that migration has played a very important role historically in spreading ideas.[34] It is certainly possible that migration may have played at least as large a role on long-run income as trade, if not larger. Economic research on immigration has taken a narrow focus, stressing the identification of the labor market effects of immigration. As pointed out recently by Hanson (2009), a more general approach is needed to carry out a comprehensive analysis of the aggregate economic effects of migration. From the point of view of economic growth, the study of the effects of migration on the stock of productive ideas, and on skill and product variety remains severely under-researched. All of these are potentially crucial channels in determining long-run income per capita (see Jones and Romer (2010) for a summary). The remainder of the paper focuses on the estimation of the long-run effects of openness to migration on income.

Let us first inspect some simple correlations. Figure 2 shows that there exists a robust positive partial correlation between the migration share and the logarithm of income per person across countries, after controlling for population and area. Figure 2A plots log income per person against the foreign-born share.[35] Figure 2B plots the gravity-predicted migration share (adjusted for country size) and income per capita.[36] In both cases the correlation is robust to dropping outliers. It is also not driven by the US, Canada, or Australia – countries that are both highly economically developed and have a high foreign-born share. This is particularly clear for the predicted migration share (Figure 2B). These countries' large size and relative remoteness lead to relatively low predicted immigration shares.

Yet these correlations might be driven by the confounding effect of trade, since the determinants of bilateral trade flows also influence bilateral migration quite strongly. We now examine the joint effects of openness to trade and migration on income in a more formal regression setting. Building on the basic FR specification, Table 4 includes openness to migration (measured by the migration share) as an additional explanatory variable. We estimate equation (18), treating both the trade and migration shares as *endogenous* regressors. We use the

[33]See, for instance, the review in the textbook by Weil (2007).
[34]See, for instance, Acemoglu et al. 2001, Comin et al. 2010, Diamond 1997, and more recently, Putterman and Weil 2010.
[35]The regression coefficient is 6.5 with a standard error of 1.18.
[36]The regression coefficient is 15.7 with a standard error of 3.95.

gravity-based predictors for the trade and migration shares as instrumental variables. Table 2 (columns 4 through 6) reports the estimates of the gravity migration model. The explanatory variables are the same as for the gravity trade regression and are meant to account for the determinants of migration costs and the gravitational force of origin and destination country sizes. The signs of the coefficients are largely as expected, and the point estimates for some of the main regressors are also roughly similar across specifications (notably, log distance, log destination population, common language, and colonial ties). As was the case with trade flows, bilateral distance reduces bilateral migration, while sharing a common language and colonial ties appear to significantly increase migration. While not dramatic, there are some noticeable differences between the marginal roles played by some variables in accounting for trade and migration flows. On the basis of the non-linear Poisson predictors (columns 3 and 6), which take into account the multiple zero bilateral entries, bilateral distance and sharing a border seem to play a much larger role in accounting for migration flows than for trade. On the contrary, past colonial ties appear to be a stronger predictor of trade rather than migration flows. Thus while the predictors for the overall trade and migration shares are based on the same underlying variables, the weights assigned to each of these variables differ in meaningful ways, which potentially can help us separately identify the causal effects of trade and migration on income per capita.

Before turning to the estimates, it is useful to visually examine the relationship between actual and predicted trade and migration shares. Figure 3 displays the corresponding scatterplots. Clearly, the predicted migration share is strongly correlated with the actual data (Figure 3A). This correlation is large, statistically significant, and not driven by outliers (Figure 3B). This is in stark contrast with the ability of the predicted trade share to account for the actual data (Figures 3C and 3D). In this case the positive correlation between predicted and actual values depends strongly on few influential observations. When the observations for Ireland, Luxembourg and Singapore are omitted, the correlation is weakened substantially and loses its statistical significance.[37] This is another sign of the relatively larger success of the gravity-based predictor in accounting for migration relative to trade flows.

Table 4 reports the 2SLS joint estimates of the effect of trade and migration openness on income per person. In columns 1-3 we only include controls for country size. In this case both the trade and immigration shares appear to be significant determinants of income per person. A one-percentage point increase in the trade share is associated with an increase in income per person of approximately 2.3-3%. The coefficient on the migration share is significant in columns 2 and 3, where we use the non-linear predictor for the migration share while we use the linear predictor for the trade share.[38] In this case a one-percentage point increase in the immigration share increases income per person by about 6% (column 3).[39]

[37]The role of influential observations in the prediction power of the gravity-based trade shares had already been noticed in the previous literature (see Figure 1 and following discussion in Frankel and Romer 1999).

[38]The difference in functional forms seems to aid in separately identifying the two coefficients.

[39]In the remainder we proceed with these combination of predictors, as they seem to maximize the strength of the instruments.

However, as we include, progressively, distance from the equator (specification 4), regional dummies and controls for geography and climate (specification 5), and for the colonial past (specification 6), the effect of trade vanishes while the effect of the migration share remains statistically significant, with a large and robust coefficient, ranging between 6.4 and 8.4. That is, the migration share appears to be a more robust determinant of income per person across countries. One could argue that these results simply reflect that the variables used to predict bilateral flows are more relevant for migration than for trade flows and, as a result, our instruments are weaker in predicting the trade shares than the migration shares. Notice, incidentally, that the instruments are rather weak in general in predicting jointly trade and migration, as indicated the Wald F-test of the first stage – only columns 3 and 5 allow for a relatively clear rejection of the null of weak instruments. To address this point column 7 reports OLS estimates of the regression model featuring migration and trade shares jointly and the whole set of controls. Again, the migration share seems to have a positive effect, while the coefficient on the trade share is very small and statistically non-significant. We also note that the OLS estimate for the effect of migration is only moderately lower than the corresponding instrumental-variables estimate (column 6), which may reflect some attenuation bias in OLS. Columns 8 and 9 focus on the role of the migration share alone. In column 8 we use the non-linear predictor for the migration share and in column 9 the fixed-effects predictor. In both cases the point estimate of the migration share remains stable, with a value ranging between 6.1 and 7.4. We note though that the non-linear predictor (column 8) delivers a much stronger first-stage regression.[40]

Let us illustrate here the magnitude of the effect of openness to immigration on income per person using a point estimate of 6.4 (the median estimate in Table 4) as a reference value. In our cross-country sample there is a 10 percentage-point difference in the foreign-born share, between countries at the tenth and ninetieth percentiles of income per person.[41] Based on our estimates, and assuming that the coefficient identifies the causal effect of openness to immigration, this would imply a difference in long-run income per capita by a factor of 1.87. By way of comparison, the difference in income per person between these two groups of countries attributed to differences in schooling (Hall and Jones 1999) was around 3.

8 The roles of Institutional Quality and Early Development

In our previous empirical model we included a fairly complete set of geographic controls. According to Hall and Jones (1999) and Acemoglu et al (2001), the main reason why these variables are relevant (in particular, distance from the equator) is that geography was decisive in determining a country's history of colonization. They argue that those initially less-developed countries that were colonized by a Western European power

Occasionally, we report results with the other predictors. The full set of results with all predictors can be obtained upon request.

[40] We note that the Stock and Yogo (2005) critical values are only strictly appropriate under homoskedasticity. We report heteroskedasticity-robust standard errors, which in our application tend to be higher than those obtain under the assumption of homoskedasticity.

[41] This also corresponds to the increase in the foreign-born share experienced, for instance, by Spain between 1998 and 2008.

through long-term settlements were endowed with good institutions. Since good institutions beget further good institutions, those countries are likely to enjoy high institutional quality today and the resulting high income per capita. On the contrary, countries that were colonized but not settled by Europeans experienced "exploitative" early institutions, a persistent burden on their economic development. In fact several authors have argued that institutional quality is the main determinant of long-run economic success (Hall and Jones 1999, Acemoglu et al (2001, 2002), Acemoglu and Johnson (2005), La Porta et al (1999), among others). Specifically, good early institutions may have allowed for policies aimed at sustaining free markets, democracy, checks and balances and well-functioning legal and judicial systems.

It is also plausible that good early institutions may have led to sustained openness to international trade and migration. Since our predictors for the trade and migration shares are based on geographical information, which has also influenced a country's history of colonization and resulting institutional quality, it is important to attempt to separately identify the roles of economic openness and good institutions on income. Our measure of institutional quality follows Acemoglu et al. (2001) and is the average between their indices for "protection against expropriation risk" and "constraints on the executive". Both are measured over the period 1975-85. These indices capture some fundamental aspects of protection of private property rights and the limitation of the power of government, which have been found to be crucial for an institutional setting conducive to economic growth.[42] Of course, institutions are likely to be endogenous to economic development. Following Hall and Jones (1999) and Alcala and Ciccone (2004), we complement the gravity-based predictors for openness (to trade and migration) with plausibly exogenous determinants of early institutions. Namely, distance to the equator and the share of the current population of European descent. The former has been shown to affect the odds of having been colonized and settled by a European power. The latter provides a measure of the resulting degree of social, economic and cultural influence.

Table 5 reports our 2SLS estimates. Column 1 considers the role of the migration share and our index of institutional quality on income per person, considering both as endogenous regressors. We include regional dummies and controls for geography, climate and disease environment in all specifications. Compared to the previous section, here we do not control for distance to the equator (as it is used as an excluded instrument) or the colonial past (as its influence is mainly through institutions). Both the migration share and institutional quality are highly significant, with coefficients around 9 and 0.4, respectively. In our sample the difference in the institutional quality index between the 90th and 10th percentiles is around 6. Based on our point estimate, the resulting income difference explained by institutions is equal to a factor of 14. In comparison, the migration share accounts for a factor of 3.5 in the income gap between the 90-10 percentiles. Hence, while institutions still appear to be the main determinant of income per capita disparities, openness to immigrants also

[42] The value of this index ranges between 0 and 8.

has an important and distinct contribution.[43] Column 2 confirms the previous findings using the fixed-effects predicted migration share, but the instrument is weaker now. Column 3 reports the estimates when we include the exogenous predictors of institutions directly as regressors (rather than using them as instruments). The positive effect of the migration share on income per capita remains.

Columns 4 and 5 estimate the specification in Alcala and Ciccone (2004). That is, trade openness is measured by the log of *real* trade share. As they found, the estimated coefficient associated to this variable is highly significant in column 4, with a point estimate of 1.97. However, when we introduce controls for geography, climate, and the disease environment, we cannot reject the null of a zero effect even on this more sophisticated measure of trade openness (column 5). Column 6 features institutional quality along with the trade and migration shares, all included together as endogenous variables. We use the predicted migration and trade shares, the distance from the equator, and the share of the population of European origin as instruments. This is a very demanding specification with three endogenous variables and four instruments (critical values not available). The point estimates suggest once again that the migration share and institutional quality have a positive effect on income per person, unlike the trade share.

Putterman and Weil (2010) emphasize the role of measures of early development, particularly political and administrative institutions around year 1500, in explaining current income levels. They argue that adjusting these measures of early political development to take into account the level of development of the countries of origin of the ancestors of the current population greatly increases their explanatory power. The idea is that migrants may have brought to the host country the ideas and social norms that prevailed in their countries of origin at the time of migration, acting as pollination agents. Thus migrants from more developed origin countries may have had a larger positive effect. Putterman and Weil (2010) also offer suggestive evidence indicating that greater variety in the composition by origin of migration flows may have had an additional positive effect. Their findings suggest that migration has played a crucial role in economic development as a vehicle for the dissemination of institutions. While related, the channel highlighted in our theoretical framework differs from theirs. Our measure of openness to migration refers to current levels of migration, which are unlikely to affect the quality of current institutions but rather affect income by increasing the varieties of skills available for production. We attempt to distinguish our mechanism from theirs in the following manner. First, we use the data in Putterman and Weil (2010) to directly control for the long-run effect of migration through institutions. Namely, we include the ancestor-adjusted quality of political institutions before year 1500 (the so-called *Statehist* variable) and the share of the current population whose ancestors around year 1500 were born abroad.[44] Columns 7 and 8 in Table 5 introduce these two controls. In column 7 we see that the positive effects

[43] We are able to reject the null of weak instruments when using the less stringent Stock and Yogo (2005) critical value (5.45), but typically not when using the more demanding one (13.43).

[44] The raw *Statehist*" variable is an index, ranging between 0 and 1, capturing the (discounted) length of time prior to year 1500 since the country had developed a supra-tribal government. The ancestor-adjusted variable, say for the US, is a weighted average

of the migration share and institutional quality are unaffected by introducing these controls. In column 8 we drop current institutional quality. In this case the migration share still displays a significant effect and we also reproduce the finding in Putterman and Weil (2010). In the previous column the quality of current institutions absorbed the effect of early institutions highlighted by these authors. Table A.1 further reinforces this point. Columns 3 and 4 in that table show that early institutional development plays a significant role in accounting for the quality of current institutions. The estimates in this columns also suggest that the current migration share does not have a direct effect on current institutional quality. In conclusion, the findings in Putterman and Weil (2010) suggest that migration had a historically important role in economic development by shaping a country's institutions. Our analysis shows that contemporary levels of migration appear to increase income through channels other than institutions. In the next section we test the specific channel highlighted in the theory section.

Thus far our analysis has ignored educational differences among natives and immigrants. It is certainly possible that migrants with high education have a larger contribution than those with lower education levels. On the other hand, several authors have argued that migrants' formal education is only a rough measure of the productive skills of immigrants (Dustmann et al, forthcoming). To investigate this question we now exploit our data further by distinguishing between individuals with a college degree and those without. We use this information to compute the share in the human capital of a country that is accounted for by its foreign-born population. Specifically, we assume that the average skilled worker has higher efficiency units of labor than the average unskilled worker, where the ratio of these units is given by the skilled-unskilled wage ratio. Following Hall and Jones (1999) we assume that the wage return to each additional year of education is 6.8% in all countries. Assuming that the average gap in years of schooling between skilled (college educated) and unskilled (non-college educated) workers is 6 years of schooling, we obtain that the efficiency units of skilled workers are 1.503 times the units for unskilled workers.[45] In column 9 of Table 5 we estimate the same specification as in column 1, but we now measure the migration share in terms of human capital, rather than in terms of population. The estimates confirm our previous finding: both the migration share and institutional quality have positive and significant effects on income per capita. We also note that the precision of the estimated effect of the migration share has increased when taking educational attainment into account. Let us examine the role of these estimates in accounting for observed cross-country differences in income per capita. Immigration as a share of human capital is around 2% for countries in the bottom decile of the income per capita distribution. To the contrary immigrants account for 22% of the human capital among the countries in the top decile. Based on an estimated coefficient of 4.8, if a country in the first decile by income per capita experienced an increase in

of *Statehist* across all countries in the world, where the weights correspond to the shares by country of origin of the ancestors of the current US population around year 1500. The exact definition can be found in pages 1640 and 1641 of Putterman and Weil (2010).

[45] Specifically, we define country c's stock of human capital as $H_c = U_c + 1.503 * S_c$, where S_c and U_c denote the number of college graduates and non-college graduates in the population, respectively.

its migration share (in terms of human capital) equal to twenty percentage points, its income per capita would increase by a factor of 2.61 in the long run.

9 Openness, Productivity, and Diversity

9.1 The Components of Income

The previous sections have shown that openness to migration has a robust, positive effect on income per person. This is certainly consistent with our model that considers openness to immigration as increasing variety in terms of skills in the receiving country, and thereby income per person. We now explore the channels behind this reduced-form relationship. We begin by performing a decomposition following Hall and Jones (1999) and Alcala and Ciccone (2004). We postulate a simple Cobb-Douglas aggregate production function in which output is produced using human capital and physical capital. Income per worker (as opposed to per person) can be broken down into: physical capital intensity, human capital intensity, and total factor productivity. Specifically,

$$\ln y_c = \frac{\alpha}{1-\alpha} \ln \frac{K_c}{Y_c} + \ln h_c + \ln TFP_c. \tag{23}$$

In expression 23, parameter α is the labor share in income (following the literature we set this equal to 0.33), K_c/Y_c is the capital-output ratio, $h_c = \exp(\gamma s_c)$ is the average human capital per person, calculated as the exponential of average years of schooling times their Mincerian return. Finally, TFP_c is the total factor productivity calculated as a Solow residual. The data on physical capital and output per worker can be obtained from the Penn World Tables while the data on average schooling are from the Barro and Lee (2011) and the Cohen and Soto (2007) databases.[46]

Table 6 reports the 2SLS estimates for a series of models where the dependent variables are the log of income per worker, the log of the capital-output ratio, the log of human capital per person, and the log of TFP. Our main regressor of interest is the migration share and the other regressors are log population, log area, and the set of controls for geography, climate, disease environment, and colonial past defined in Table 4. As in Table 4 (specification 9), we consider immigration as endogenous an endogenous regressor. In the left panel of the Table the migration share is computed as a share of the total resident population in the country. In the right panel it is computed as a share of the human capital in the country, which is a weighted sum of individuals that accounts for differences in educational attainment. In both sets of estimates the pattern that emerges is very similar. The migration share has a positive effect on income per worker, and this effect operates by and large

[46] Where available the data on years of schooling have been obtained from the most recent version of the Barro and Lee (2011) database. For a dozen countries for which the information is not available in that database we rely on Soto and Cohen's (2007) data, available at their personal website. Following Hall and Jones (1999) all dependent variables have been normalized by the US value.

through total factor productivity. This pattern is totally consistent with our theoretical framework. We find no evidence of an effect of immigration on capital intensity, consistent with the prediction of the neoclassical growth model stating that sporadic changes in the size of the labor force will not affect capital per worker in steady state. We find weak evidence that migration may stimulate investments in human capital, suggesting that the native population may seek to differentiate itself from migrants by acquiring more education.

The findings here beg another question. Why exactly does immigration lead to higher TFP? Our model suggests that this is because it increases the diversity of the skills in the labor force. Alternative interpretations are that immigration increases the variety of ideas, product variety and innovation. The analysis in the following sections tries to explore further the nature of the relationship between immigration and productivity.

9.2 Diversity in Trade and Migration Flows

If immigration affects productivity by increasing the variety of skills and ideas in a country then a highly diverse immigrant flow should be particularly beneficial. Several authors have provided evidence in partial support of this hypothesis. In the context of cities in the US, Ottaviano and Peri (2006) find that not only the share of foreign-born, but also the associated index of diversity (by country of origin) leads to higher productivity at the city level. Similarly, Broda and Weinstein (2006) find productivity gains of trading different goods with a variety of foreign countries. Peri and Sparber (2009) find that immigration affects the supply of tasks and induces task-specialization that produces efficiency gains through comparative advantage. One way to test the diversity of skills hypothesis is to assume that immigrants from different origin countries are endowed with different skill varieties. For instance, it is plausible to expect that a country's social norms and culture may shape the skills and production-relevant ideas among the workforce. The assumption of variety tied to country of origin was implicit in the theory we presented earlier and is akin to an Armington assumption in the context of international trade theory. From the empirical viewpoint, the findings by Putterman and Weil (2010) provide evidence in favor of this assumption.

In this section we analyze whether diversity by country of origin in migration and trade flows has an effect on income per capita. Specifically, we construct the so-called "fractionalization index" by country of origin:[47]

$$Div_c = 1 - \sum_j (M_{jc})^2 . \tag{24}$$

In expression (24), M_{jc} is the share of foreign-born from country j in the total foreign-born population residing in country c. As the index approaches zero, it indicates that one source country accounts for most of the immigration flows, hence diversity is very low. When it approaches one it describes a situation where

[47] Here we define the index for migration but we build it identically for trade flows. Our index is defined as one minus the corresponding Herfindahl-Hirschman concentration index.

immigration is balanced, with all countries of origin having equal shares, and hence diversity is very large. We also build a similar diversity index for trade flows using the share of trade with a specific country relative to total trade. Let us first comment briefly on some features of these indices, beginning with the migration diversity index. The value of this index for the US is 0.91, which will be a useful benchmark. This value indicates that migration flows into the US are fairly diverse.[48] Several countries attain higher values: Israel (0.94), Spain (0.94), the UK (0.96), Denmark (0.96), and Canada (0.96), to name just a few. Many countries display lower values: Bangladesh (0.06), Pakistan (0.09), India (0.60), Greece (0.70), or Japan (0.75). Turning now to the diversity of trade flows, we again use the US as the benchmark, with a relatively high value of 0.92. Several rich countries have more diverse trade flows, such as France (0.93), the UK (0.94), or Germany (0.95), reflecting the low trade costs within Europe. However, the countries with the highest values tend to be low income (Pakistan, India, Kenya or Tanzania are all in the top 10). At the other extreme, Mexico (0.39) and Canada (0.43) display very low values of the trade diversity index, reflecting the dominant position of the US as their main trading partner.

Columns 1 through 4 in Table 7 present two-stage least-squares estimates of the effects of trade and migration shares (both instrumented) on the log of GDP per capita, controlling for country size, distance to the equator and the complete vector of controls for region, geography, climate, disease environment, and colonial past. The first column reproduces the earlier findings: the migration share has a positive and significant effect on income, while the point estimate of the trade share is very low (even negative) and statistically insignificant. Column 2 adds the migration diversity index, treating it as an exogenous regressor. The coefficient is positive and highly significant and reduces the point estimate for the migration share only slightly. This provides suggestive evidence of a separate, positive effect of skill diversity. Column 3 adds the diversity index for trade. The point estimates for the migration share and the migration diversity index remain largely unaffected. The point estimate for diversity in trade flows is negative. Before reading too much into this finding we note that it is possible that the diversity of immigrants or trade flows by country of origin may be correlated with other determinants of income and certainly contains substantial measurement error. Ideally, one would like to treat them as endogenous regressors. In practice the gravity predictors for the diversity indices perform poorly, thwarting estimation by instrumental variables.[49]

To address this shortcoming we adopt a reduced-form approach and use the gravity predictor for the diversity indices as control variables. In column 4 we include gravity predictors for the diversity indices for immigration and trade flows as control variables in the regression. The share of immigrants and trade are still considered as endogenous and instrumented. The migration share still appears to have a positive effect. Moreover, immigrant

[48] Mexico plays a clearly dominant role in US immigration, however it is important to note that the US also host immigrants originating in virtually all other countries in the world, and the shares of these countries in the total immigrant population in the US are fairly balanced.

[49] The problem is particularly severe when we attempt to instrument for both the migration (or trade) share and the corresponding diversity index.

diversity seems to have an additional positive and significant effect on income per capita.[50] The income effect of immigrant diversity is large. An increase in the diversity of migrants from 0.05 (the value for Sri Lanka, whose immigrants are essentially all from India) to 0.95 (the value for the UK) implies a corresponding increase in output per person by a factor of 3.5. Our estimates also suggest a negative effect of predicted diversity in trade flows. While we have no intuition for this result we note that the higher values for the diversity index for trade flows were attained by low income countries, such as Pakistan, India or Kenya. In a nutshell, while the limited strength of the geographic instruments to predict migration diversity does not allow us to find conclusive proof, our results are suggestive of a positive effect of immigrant diversity on income. In comparison, we do not find evidence of a positive effect of trade openness or of trade diversity on income.

9.3 Fractionalization

The previous results suggest that large and diverse migration flows increase long-run income per capita. However, there may also be a negative byproduct associated to large and diverse migration flows. In particular, it may lead to ethnic or linguistic fractionalization, which in some cases may entail conflict with high social and economic costs. There is an extensive literature identifying the negative effect of ethnic fractionalization on institutional quality and local provision of public goods.[51] Alesina, Baqir and Easterly (1999) provide evidence indicating that ethnic or linguistic fractionalization increases conflict and reduces solidarity, leading to a reduction in the provision of public goods. However, Alesina et al. (2003) examine the consequences of different types of fractionalization (ethnic, linguistic and religious) for economic growth and several other economic outcomes. While they find effects of ethnic and linguistic fractionalization on some economic outcomes (corruption, political rights), they report that these effects appear to be sensitive, and find no consistent effect on economic growth.

Let us then examine if immigration has had an effect on fractionalization.[52] Columns 5 and 6 in Table 7 examine the effects of migration (and trade) openness on indices of ethnic and linguistic fractionalization. Both measures are taken from Alesina et al. (2003). In both cases we find that a higher migration share is associated with increases in fractionalization. This is reasonable as the current fractionalization of a country is, in part, driven by the inflow of ethnically and linguistically different immigrants. Trade shares, instead, are not associated to higher fractionalization.[53]

While not surprising, this finding begs an important economic question. Does increased fractionalization

[50] We have also estimated a specification where we treat migration diversity as an endogenous regressor, instrumented with its gravity predictor. When omitting the migration share from the regression the first-stage regression allows us to reject the null of weak instruments. In that case (not reported in Table 7) we obtain a point estimate for migration diversity equal to 3.87, significant at the 10% level.

[51] See Alesina and La Ferrara (2005) for an overview.

[52] The relationship between place of birth diversity and ethnic and linguistic fractionalization and their relation to income is also the focus of the analysis in Alesina, Hanoss and Rapoport (2012). They independently developed an analysis focusing on the effect of diversity by birthplace on income and productivity.

[53] We have also estimated an analogous specification where the dependent variable is an index of religious fractionalization. We did not find a significant effect of migration.

lead to lower long-run income levels? One would expect this to be the case if the opportunities for meaningful economic interactions are reduced by fractionalization and the ensuing residential or social segregation. It might even be the case that these costs swamp the gains from increased diversity discussed earlier. Columns 7 and 8 report the estimates of a specification where we explicitly control for ethnic and linguistic fractionalization, together with the shares of Catholic, Muslim and Protestant in the country's population as of 1980. In column 7 we include both the trade and migration shares, while in column 8 only the migration share. When controlling for ethnic and linguistic fractionalization, the point estimate for the migration share ranges from 8.85 to 10.11, statistically significant in both cases. Note also that linguistic fractionalization is negatively associated to income per person, while ethnic fractionalization does not seem to have an additional negative effect, confirming Alesina et al. (2003). In column 9 we do not control for ethnic or linguistic fractionalization. The point estimate for the migration share is smaller than in the previous column, but still highly significant and positive. Hence, the total effect of immigration on income appears to be positive, suggesting that the productivity gains of higher skill diversity are larger than the costs arising from increased linguistic fractionalization.[54]

10 Additional Results

This section gathers a number of additional results. First, we examine if trade and migration openness leads to higher income inequality. Second, we analyze if our main finding of an effect of immigration on income varies across subgroups of countries. Third, we explore the sensitivity of our findings to our definition of openness to migration. Finally, we conclude by examining an alternative channel through which migration may affect long-run income: patenting and innovation.

10.1 Income Inequality

There is an abundance of literature on the effects of international trade on income inequality. The debate has been reignited by the rise of trade flows with China and the public debate on the pros and cons of globalization.[55] There is also a similar ongoing debate surrounding the effects of immigration on the income distribution of the host country. Most of the work in this area is based on individual-level data or uses regional variation within a single country. In their review of the US literature, Raphael and Ronconi (2007) conclude that the effects of immigration on the wages and employment of native workers appear to be very small, although the academic debate is still ongoing. Few studies have examined the role of both international trade and migration. One influential contribution using US data is Borjas, Freeman and Katz (1992, 1997).[56] These authors conclude that

[54]It is certainly possible that immigration may have further effects on the type and level of public goods. While very interesting we leave this question to future research.

[55]See Richardson (1995) for a survey and the recent studies by Autor et al. (2011) and Levchenko and Zhang (2012).

[56]See also the replies by John DiNardo and John Abowd to the 1997 article.

the effects of immigration and trade flows on relative skill supplies have not been substantial enough to account for more than a small proportion of the overall widening of the wage inequality over the 1980s and 1990s.

We use our data to analyze the long-run effects of trade and immigration on income inequality exploiting cross-country variation. Before turning to empirical issues we note that in our theory the main economic effect of immigration is to enlarge the set of available productive skills. Recall also that our earlier results found no evidence of an effect of migration on capital intensity, suggesting no long-run effects on the rental rate of capital or the average wage. We also found partial evidence suggesting that immigration may have encouraged educational investments on the native population, which may lead to reduced wage inequality (if it lowers the skill wage premium).

We focus on a commonly used measure of income inequality at the national level: the Gini coefficient, which ranges between zero and one and is readily available for a large set of countries (UNU-WIDER database, version WIID 2C).[57] Table 8 reports our two-stage least-squares estimates. We consider the roles of the migration share (specifications 1 and 3) and migration and trade shares jointly (specifications 2 and 4) as potential determinants of inequality. In all regressions we also include country size, distance from the equator and regional, geography, climate, disease and colonial controls. The first two specifications use the fixed effects gravity predictor for the immigrant share of the population, while specifications 3 and 4 are based on the basic gravity predictor. In all specifications neither the trade nor the migration share have significant effects on the Gini coefficient, suggesting that while immigration may have had a positive effect on average income per person it does not seem to systematically have affected the income distribution. This is in line with the idea that immigrants bring a set of skills that complement those already existing in the host country.

10.2 Heteregenous effects across countries

Here we explore whether countries differ in their ability to employ productively the new skills brought by immigrants. It is possible that countries with better institutions, more open to trade, or with a more educated population provide a more conducive environment that leads to a better use of the new skills.

To explore these question we split our sample across three dimensions. We use the variable *Statehist* to separate countries with high (above the median) or low (below the median) level of pre-1500 state development. We use the trade share to distinguish between countries with high (above the median) or low (below the median) trade openness. We also use average human capital per person (relative to the US) to distinguish between countries with high (above the median) and low human capital levels. Then we estimate the effect of immigration on income interacted with the "high" and the "low" dummy for each of these dimensions, instrumenting with the predicted share of immigrants (interacted with the same dummies) and including all the

[57] This database collects studies based on population surveys and census data for a very large number of countries. We select countries with available data around year 2000.

controls as in Table 4 (specification 9).

The estimates of the coefficients are reported in Table 9. First, we note that all coefficients are positive and five out of six are significant. In most sub-samples immigration has a strong positive effect on income. Let us now examine the coefficients more carefully. Immigration seems to have similar effects on countries with high or low early development (in year 1500). This makes sense since in many countries, though not all, early advantage has been substantially eroded. In contrast, we find that immigration appears to have a larger, beneficial effect on income per capita in host countries characterized by good current institutions and high human capital. This is also intuitive since countries with good institutions provide a better economic environment for natives and immigrants to thrive economically. In addition, the skills possessed by highly educated workers are likely to be more specialized, reducing the potential for direct competition between natives and immigrants.

10.3 Alternative Measures of Openness

As we argued earlier the immigration and trade shares are noisy functions of the true underlying openness parameters. An alternative approach is to try to uncover the effects of the global latent *economic openness*. Specifically, we extract the first principal component of migration and trade shares and use it as an explanatory variable in regressions that are otherwise similar to those of Table 5. Our vector of instruments contains both the gravity predicted trade and migration shares. Columns 1 and 2 in Table 10 present the results. In both cases the coefficient of openness is positive and significant, and the point estimate is almost the same regardless of whether we include institutional quality (column 2) or not (column 1). These findings provide additional robustness to our earlier results.

Next, we turn to our definition of openness to international migration. Throughout the paper we have defined this variable as the foreign-born share in a country's population. We now examine whether *emigration* has an additional effect on income. It is possible that emigration has negative effects but the converse may also be true since remittances or human capital associated to return migration may compensate for the loss of workers.[58] To examine these issues columns 3 and 4 in Table 10 add the emigration share. The point estimates for the immigration and trade shares remain largely unaffected and the emigration share is never significant.[59] Next, we use the data on emigration to build the net immigration share, defined as stock of immigrants minus the stock of emigrants divided by the country's total population, and use it as our measure of migration openness. Columns 5 and 6 report the results. The results show that net immigration, relative to the population, has a positive effect on income, which provides additional robustness on our main findings.

[58] di Giovanni et al (2012) argue that the gains from remittances more than compensate for the loss of labor associated to emigration.

[59] We are aware that this regressor is likely to be endogenous. Unfortunately, if we treat the emigration share as an endogenous regressor, the first-stage regression becomes very weak and we obtain uninformative estimates with very large standard errors. At any rate, since we do not find any indication of an effect either way we have not explored the issue further.

10.4 Patents and Innovation

An alternative channel through which immigration might affect long-run income is through its effect on the rate of innovation. While not considered in our theoretical framework, it is possible that the degree of diversity in the skills of a country may spur creativity and innovation. Besides total factor productivity, a common way to try to measure innovation is by using data on patents. While not all innovations are patented and the patenting rate of innovations depends on the field and sector of discovery, statistics on patents have long been used as a measure of the innovation output of a country, region or sector.[60] Here we follow this approach. We are not the first in examining the relationship between immigration and their direct effect on innovation and entrepreneurship. Important contributions to this literature are Gauthier-Loiselle and Hunt (2010) and Hunt (2011). These studies provide evidence of high rates of patenting activity among the immigrant population, compared to natives with similar educational attainment. Similarly, some recent studies link openness to trade to technology adoption and innovation (Bloom et al, 2011). The World Intellectual Property Organization (WIPO) collects and distributes data on patents granted by any patent office in the world, to inventors residing in 108 countries between 1995 and 2010.[61] We then construct the average yearly number of patent per million inhabitants. The average number of patent per million people is 91 in our sample. The country with highest patenting per person counts 227 patent per million people, the country with the lowest value counts 0.01.

Table 11 reports the estimated effects of the migrant share and the trade share on the logarithm of patents per (million) person, using data on 108 countries. We include alternatively as explanatory variable MSH (column 1), TSH (column 2), their first principal component (column 3), and in column 4 we control for institutional quality. In all specifications we include regional dummies, geographic, climatic and disease variables and colonial origin variables as controls and we use the linear gravity predictor as an instrument. The results show a positive and marginally significant effect (at the 10% level) of the share of immigrants on patenting activity, both controlling (column 4) or not (column 1) for institutional quality. In comparison trade openness appears to have no effect on innovation. The first principal component of openness to immigration and trade also appears to have a (marginally significant) positive effect on patenting.

Taken together the findings here provide suggestive evidence of a potential link between immigration and innovation activity, as measured by patents per person. This may provide an additional channel, besides increasing the diversity of productive skills, through which immigration may affect productivity and long-run income.

[60] See for instance the book by Jaffe and Trajtenberg (2002).

[61] The data are available at the website http://www.wipo.int/ipstats/en/statistics/patents/.

11 Conclusions

This paper uses cross-country data to explore the relationship between income per person and economic openness, both in terms of openness to international trade and to immigration. To address endogeneity concerns we follow the instrumental-variables strategy introduced by Frankel and Romer (1999), initially used to estimate the impact of trade openness on income per person.

Our main finding is that openness to migration plays an important role in accounting for cross-country differences in income per capita. This finding is obtained using econometric specifications that include a comprehensive set of variables that control for geography, climate, disease environment, and colonial past. We also show that the effect of migration openness is distinct from the role played by institutional quality. Regarding the role of trade openness, we are able to reproduce the positive effect of trade on income found by Frankel and Romer (1999). However, the size and significance of its effect falls sharply when including our demanding set of controls. Second, we show that the main effect of migration operates through total factor productivity, consistent with the simple theoretical model that we propose. In this model the main channel by which immigration affects productivity and income is by enlarging the set of skills available for production. We also provide some more direct evidence of this channel by showing that diversity in immigration flows (by country of origin) appears to have an additional positive effect on income. As a byproduct of increased diversity we show that there is also an increase in linguistic fractionalization, which reduces but does not eliminate, the positive economic effects of greater variety in skills. Finally, we also provide partial evidence of a positive effect of immigration on innovation activity, which provides another channel through which immigration may affect long-run productivity an income.

Our analysis does not reveal a consistent effect of trade openness on income. In our view this reflects the fact that the trade to GDP ratio, which is the most common measure of trade openness, may increasingly be a poor measure of a country's degree of trade integration. While trade flows are fairly well measured, they measure the total value of imports and exports rather than the value added abroad or domestically. In contrast, GDP is measured as value added. Recently, trade economists have started paying a closer scrutiny to these issues (see Bems et al (2009) or Johnson and Noguer (2009)). Thus several countries that partake in a lot of processing trade and "transit" trade (Singapore, Hong Kong, Ireland) have deceptively large trade shares, but may have much smaller trade in value added. Using more accurate measures of trade openness may provide a more fruitful route to uncover the different channels through which trade openness matters for income.

References

Acemoglu Daron (2009) "Modern Economic Growth" Princeton University Press, Princeton NJ.

Acemoglu, Daron, Simon Johnson, and James Robinson (2001), "The Colonial Origins of Comparative Development: An Empirical Investigation," American Economic Review, 91, 1369–1401.

Acemoglu, Daron, Simon Johnson, and James Robinson (2002), "Reversal of Fortunes: Geography and Institutions in the Making of the Modern World Income Distribution," Quarterly Journal of Economics, 117 (2002), 1231–1294.

Adsera, Alicia and Mariola Pytlikova, 2012. "The role of language in shaping international migration." IZA Working Paper 6333.

Alcalá, Francisco and Antonio Ciccone (2004). "Trade and Productivity," The Quarterly Journal of Economics, MIT Press, vol. 119(2), pages 612-645, May.

Alesina Alberto, Reza Baqir and William Easterly, (1999). "Public Goods And Ethnic Divisions," The Quarterly Journal of Economics, MIT Press, vol. 114(4), pages 1243-1284, November.

Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat and Romain Wacziarg, (2003). " Fractionalization," Journal of Economic Growth, Springer, vol. 8(2), pages 155-94, June.

Alesina, Alberto, Johann Harnoss and Hillel Rapoport (2012) "Birthplace Diversity and Economic Growth" Manuscript.

Alesina, Alberto and Eliana La Ferrara, (2005). "Ethnic Diversity and Economic Performance," *Journal of Economic Literature*, American Economic Association, vol. 43(3), pages 762-800, September.

Alesina, Alberto and Spolaore, Enrico, (1997). "On the Number and Size of Nations," The Quarterly Journal of Economics, MIT Press, vol. 112(4), pages 1027-56, November.

Alesina Alberto, Enrico Spolaore and Romain Wacziarg, (2000). "Economic Integration and Political Disintegration," American Economic Review, American Economic Association, vol. 90(5), pages 1276-1296, December.

Alesina, Alberto, Spolaore, Enrico and Wacziarg, Romain, (2005). "Trade, Growth and the Size of Countries," Handbook of Economic Growth, in: Philippe Aghion & Steven Durlauf (ed.), Handbook of Economic Growth, edition 1, volume 1, chapter 23, pages 1499-1542 Elsevier.

Altonji Joseph G. and David Card, (1991). "The Effects of Immigration on the Labor Market Outcomes of Less-skilled Natives," NBER Chapters, in: Immigration, Trade and the Labor Market, pages 201-234 National Bureau of Economic Research, Inc.

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

Autor, David, Dorn, David, and Gordon Hanson (2011). ”The China Syndrome: Local Labor Market Effects of Import Competition in the United States.” Manuscript.

Barro Robert and Jong-Waa Lee (2011) ”Educational Attainment Dataset”, http://www.barrolee.com/.

Barro R. and Sala i Martin X. (2004) ”Economic Growth” Second Edition MIT Press, Cambridge Ma.

Beine, M., F. Docquier, and C. Ozden (2011): “Diasporas.” Journal of Development Economics, 95(1), 3041.

Bems, Rudolfs, Johnson, Robert, and Yi, Kei-Mu (2009). “The Role of Vertical Linkages in the Propagation of the Global Downturn of 2008.” Mimeo.

Bertoli, Simone, Fernandez-Huertas Moraga, Jesus, 2011. “Multilateral Resistance to Migration,” IZA Discussion Papers 5958, Institute for the Study of Labor (IZA).

Bloom, Nick, Draca, Mirko, and van Reenen, John (2011). “Trade induced technical change? The impact of Chinese imports on innovation and Information Technology.” CEP Discussion Paper No. 1000.

Borjas George J., (2003). ”The Labor Demand Curve Is Downward Sloping: Reexamining The Impact Of Immigration On The Labor Market,” The Quarterly Journal of Economics, MIT Press, vol. 118(4), pages 1335-1374, November.

Borjas George J. , Richard B. Freeman and Lawrence F. Katz, (1992). ”On the Labor Market Effects of Immigration and Trade,” NBER Chapters, in: Immigration and the Workforce: Economic Consequences for the United States and Source Areas, pages 213-244 National Bureau of Economic Research, Inc.

Borjas, George; Richard B. Freeman; Lawrence F. Katz (1997), “How Much Do Immigration and Trade Affect Labor Market Outcomes?” Brookings Papers on Economic Activity, Vol. 1997, No. 1. (1997), pp. 1-90.

Broda Christian and David E. Weinstein, (2006). ”Globalization and the Gains from Variety,” The Quarterly Journal of Economics, MIT Press, vol. 121(2), pages 541-585, May.

Card David, John DiNardo and Eugena Estes, (2000). ”The More Things Change: Immigrants and the Children of Immigrants in the 1940s, the 1970s, and the 1990s,” NBER Chapters, in: Issues in the Economics of Immigration, pages 227-270 National Bureau of Economic Research, Inc.

Chanda, Areendam, and Louis Putterman (2005), ”State Effectiveness, Economic Growth, and the Age of States”, in States and Development: Historical Antecedents of Stagnation and Advance, Matthew Lange and Dietrich Rueschemeyer, eds. (Basingstoke, UK: Palgrave MacMillan, 2005).

Chanda, Areendam, and Louis Putterman (2007), "Early Starts, Reversals and Catch-Up in the Process of Economic Development", *Scandinavian Journal of Economics*, 109 (2007), 387-413.

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

Cohen Daniel and Marcelo Soto, (2007). "Growth and human capital: good data, good results," Journal of Economic Growth, Springer, vol. 12(1), pages 51-76, March.

Comin, Easterly and Gong, (2010) "Was the Wealth of Nations Determined in 1000 BC?" American Economic Journal, 2 (2010), 69–57.

Di Giovanni, Julian, Levchenko, Andrei and Francesc Ortega, 2012. "A Global View of Cross-Border Migration," CReAM Discussion Paper Series 1218.

Diamond, Jared, Guns, Germs and Steel (New York: Norton, 1997).

Docquier, Frédéric Çaglar Özden and Giovanni Peri, (2010). "The Wage Effects of Immigration and Emigration," NBER Working Papers 16646, National Bureau of Economic Research, Inc.

Christian Dustmann, Tommaso Frattini, and Ian Preston, forthcoming. "The Effect of Immigration along the Distribution of Wages," Review of Economic Studies.

Eaton Jonathan and Samuel Kortum (1996) "Trade in Ideas: patenting and productivityin the OECD". Journal of International Economics 40 (May) 251-278

Frankel, Jeffrey A. and David Romer, (1999). "Does Trade Cause Growth?," American Economic Review, American Economic Association, vol. 89(3), pages 379-399, June.

Galor, Oded and Andrew Mountford, (2008). "Trading Population for Productivity: Theory and Evidence," Review of Economic Studies, Wiley Blackwell, vol. 75(4), pages 1143-1179, October.

Grogger, Jeffrey and Hanson, Gordon H., (2011). "Income maximization and the selection and sorting of international migrants," Journal of Development Economics, Elsevier, vol. 95(1), pages 42-57, May.

Gauthier-Loiselle, Marjolaine, and Hunt, Jennifer (2010). How Much Does Immigration Boost Innovation? American Economic Journal: Macroeconomics April 2010.

Grossman, Gene M. and Helpman, Elhanan, (1991). "Trade, knowledge spillovers, and growth," European Economic Review, Elsevier, vol. 35(2-3), pages 517-526, April.

Guiso Luigi, Paola Sapienza and Luigi Zingales, (2006). "Does Culture Affect Economic Outcomes?," Journal of Economic Perspectives, American Economic Association, vol. 20(2), pages 23-48, Spring.

Hall Robert E. and Charles I. Jones, (1999). "Why Do Some Countries Produce So Much More Output Per Worker Than Others?," The Quarterly Journal of Economics, MIT Press, vol. 114(1), pages 83-116, February.

Hanson, Gordon H. (2009) "The Economic Consequences of the International Migration of Labor," Annual Review of Economics, Annual Reviews, vol. 1(1), pages 179-208, 05.

Head, Keith, Mayer, Thierry and Ries, John, (2010). "The erosion of colonial trade linkages after independence," Journal of International Economics, Elsevier, vol. 81(1), pages 1-14, May.

Hunt, Jennifer (2011). Which Immigrants Are Most Innovative and Entrepreneurial? Distinctions by Entry Visa. Journal of Labor Economics July 2011.

Hunt, Jennifer (2012). The Impact of Immigration on the Educational Attainment of Natives." NBER Working Paper No. 18047.

Jaffe Adam and Trajtenberg Manuel (2002) "Patents, Citations and Innovations: a Window on the Knowledge Economy"MIT Press, Cambridge, Ma 2002.

Johnson, Robert C., and Guillermo Noguera. 2009. "Accounting for Intermediates: Production Sharing and Trade in Value Added." Unpublished Manuscript, Dartmouth College.

Jones Charles I. and Paul M. Romer, (2010). "The New Kaldor Facts: Ideas, Institutions, Population, and Human Capital," American Economic Journal: Macroeconomics, vol. 2(1), pages 224-45.

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

La Porta, Rafael , Lopez-de-Silanes, Florencio, Shleifer, Andrei and Vishny, Robert, (1999). "The Quality of Government," Journal of Law, Economics and Organization, OUP, vol. 15(1), pages 222-79.

Levchenko, Andrei and Jing Zhang, 2012. "Comparative Advantage and the Welfare Impact of European Integration," NBER Working Papers 18061, National Bureau of Economic Research, Inc.

Llull, Joan, 2011. "Understanding International Migration: Evidence from a New Dataset of Bilateral Stocks (1960-2000)." Manuscript.

Lucas, Robert E. (2009). "Trade and the Diffusion of the Industrial Revolution," American Economic Journal: Macroeconomics, American Economic Association, vol. 1(1), pages 1-25, January.

Matsuyama, Kiminori, (1992). "A Simple Model of Sectoral Adjustment," Review of Economic Studies, Wiley Blackwell, vol. 59(2), pages 375-88, April.

Mayda, Anna Maria (2007) "International Migration Flows: An Analysis of the forces and Constraints at Work" Manuscript Georgetown University.

Mayda, Anna Maria (2010). "International migration: a panel data analysis of the determinants of bilateral flows," Journal of Population Economics, Springer, vol. 23(4), pages 1249-1274, September.

McCleary, Rachel M. & Robert J. Barro, (2006). "Religion and Economy," Journal of Economic Perspectives, American Economic Association, vol. 20(2), pages 49-72, Spring.

Noguer, M., and M. Siscart (2005): "Trade raises income: a precise and robust result," Journal of International Economics, 65, 447–460.

Ortega, Francesc, and Giovanni Peri (2009). "The Causes and Effects of International Migrations: Evidence from OECD Countries 1980-2005," NBER Working Papers 14833.

Ortega, Francesc, and Giovanni Peri (2012). "The Role of Income and Immigration Policies in Attracting International Migrants." IZA Discussion Paper 6655.

O'Rourke, Kevin and Williamson, Jeffrey G., (1994). "Late Nineteenth-Century Anglo-American Factor-Price Convergence: Were Heckscher and Ohlin Right?," The Journal of Economic History, vol. 54(04), pages 892-916.

O'Rourke, Kevin H, Taylor, Alan M and Williamson, Jeffrey G, (1996). "Factor Price Convergence in the Late Nineteenth Century," International Economic Review, vol. 37(3), pages 499-530.

Ottaviano Gianmarco I.P. and Giovanni Peri, (2006). "The economic value of cultural diversity: evidence from US cities," Journal of Economic Geography, Oxford University Press, vol. 6(1), pages 9-44, January.

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 and Chad Sparber, 2009. "Task Specialization, Immigration, and Wages," American Economic Journal: Applied Economics, American Economic Association, vol. 1(3), pages 135-69, July.

Peri, Giovanni (2012). "The Effect Of Immigration On Productivity: Evidence From U.S. States," The Review of Economics and Statistics, MIT Press, vol. 94(1), pages 348-358, 07.

Putterman Louis, David N. Weil, (2010). "Post-1500 Population Flows and the Long-Run Determinants of Economic Growth and Inequality," The Quarterly Journal of Economics, vol. 125(4), pages 1627-1682.

Raphael, Steven, and Lucas Ronconi (2007). “The Effect of Labor Market Competition with Immigrants on the Wages and Employment of Natives: What Does Existing Research Tell Us?” Manuscript.

Richardson, J. David (1995). “Income Inequality and Trade: How to Think, What to Conclude.” The Journal of Economic Perspectives, Vol. 9, No. 3 (Summer, 1995), pp. 33-55.

Rivera-Batiz, Luis and Romer, Paul M, (1994). ”Economic Integration and Endogenous Growth: An Addendum,” The Quarterly Journal of Economics, MIT Press, vol. 109(1), pages 307-08, February.

Rodriguez, Francisco and Dani Rodrik, (2001). ”Trade Policy and Economic Growth: A Skeptic’s Guide to the Cross-National Evidence,” NBER Chapters, in: NBER Macroeconomics Annual 2000, Vol. 15, pages 261-338.

Rodrik, Dani (2000). ”How Far Will International Economic Integration Go?,” Journal of Economic Perspectives, American Economic Association, vol. 14(1), pages 177-186, Winter.

Sachs, Jeffrey D. Andrew Warner, (1995). ”Economic Reform and the Process of Global Integration,” Brookings Papers on Economic Activity, vol. 26(1, 25th A), pages 1-118.

Santos Silva J. M. C. and Silvana Tenreyro, (2006). ”The Log of Gravity,” The Review of Economics and Statistics, MIT Press, vol. 88(4), pages 641-658, 09.

Tabellini, Guido (2008). ”Presidential Address Institutions and Culture,” *Journal of the European Economic Association*, MIT Press, vol. 6(2-3), pages 255-294, 04-05.

Taylor, A. M., and J. G. Williamson. Convergence in the Age of Mass Migration. European Review of Economic History 1 (April 1997): 2763.

Taylor, A. M. Peopling the Pampa: On the Impact of Mass Migration to the River Plate, 18701914. Explorations in Economic History 34 (January 1997): 100132.

Taylor, A. M. Growth and Convergence in the Asia-Pacific Region: On the Role of Openness, Trade and Migration. In International Trade and Migration in the APEC Region, edited by P. J. Lloyd and L. S. Williams. Oxford: Oxford University Press, 1997.

Weil David N. (2005) ”Economic Growth” Pearson Addison Wesley, New York.

Weil, David N. (2007). ”Accounting for The Effect of Health on Economic Growth,” The Quarterly Journal of Economics, MIT Press, vol. 122(3), pages 1265-1306, 08.

Williamson, Jeffrey G, (1997). ”Globalization and Inequality, Past and Present,” World Bank Research Observer, Oxford University Press, vol. 12(2), pages 117-35, August.

Tables

Table 1: Descriptive Statistics and Data Sources for the main variables

Variable	Obs	Mean	Std. Dev.	Min	Max	Source
Dummy Frankel and Romer sample	188	0.78				Frankel and Romer (1999)
Real GDP per person in 2000 (PPP, chain-weighted 2005 USD)	184	10682	12881	117	74162	PWT, 7.2
TSH = Trade Flows / GDP	184	0.90	0.50	0.02	3.78	PWT, 7.2
Real TSH	184	0.50	0.42	0.01	2.72	Alcala and Ciccone (2004), PWT 7.2
MSH = Foreign-Born/Resident Pop.	188	0.04	0.08	0.00	0.52	Docquier et al (2010)
Emigrated/ Resident Population	188	0.06	0.09	0.00	0.49	Docquier et al (2010)
MSH in terms of human capital	175	0.09	0.15	0.00	0.80	Docquier et al (2010)
Institutional Quality Index	157	5.45	2.01	1.00	8.50	Acemoglu et al (2001)
Diversity index Immigration	168	0.70	0.22	0.02	0.96	Own calculations
Diversity index Trade flows	168	0.87	0.10	0.39	0.96	Own calculations
Logarithm of Population	183	1.71	2.01	-3.12	7.14	PWT, 7.2
Logarithm of Area	186	11.34	2.68	3.22	16.65	BACI dataset
Distance to equator	187	25.07	17.00	0.00	67.47	BACI dataset
Share of tropical land	153	0.49	0.48	0.00	1.00	BACI dataset
Pct. Euro. descent in 1900	153	28.38	40.97	0.00	100.00	Acemoglu et al (2001)
PW Share of foreign ancestors	188	0.24	0.32	0.00	1.00	Putterman and Weil (2010)
PW Early political dev. (Statehist)	160	0.48	0.23	0.00	0.96	Putterman and Weil (2010)
Pct. population speaking a European Language in 1975	149	31.01	43.01	0.00	100.00	Acemoglu et al (2001)
Gini Coefficient	130	41.53	11.04	21.80	76.60	UNU-WIDER
90-10 income ratio	71	11.57	11.21	3.16	67.58	UNU-WIDER
Predicted TSH (FR specification)	188	0.16	0.11	0.00	0.69	Own calculations
Predicted TSH (linear specification)	188	0.27	0.30	0.00	2.43	Own calculations
Predicted TSH (Non-linear spec.)	188	0.85	0.42	0.00	2.14	Own calculations
Predicted TSH (linear FE)	188	0.00	0.00	0.00	0.01	Own calculations
Predicted MSH (FR specification)	188	0.01	0.01	0.00	0.04	Own calculations
Predicted MSH (linear specification)	188	0.01	0.01	0.00	0.06	Own calculations
Predicted MSH (non-linear spec.)	188	0.04	0.03	0.00	0.16	Own calculations
Predicted MSH (linear FE)	188	0.01	0.01	0.00	0.03	Own calculations

Table 2. Gravity Models for Bilateral Trade Share (TSH) and Migration Share (MSH).

	1	2	3	4	5	6
Estimation	OLS	FE	Poisson	OLS	FE	Poisson
Dep. Var.	ln bilateral TSH	ln bilateral TSH	ln bilateral TSH	ln bilateral MSH	ln bilateral MSH	ln bilateral MSH
Ln distance	-1.82***	-1.71***	-0.87***	-1.38***	-1.37***	-1.46***
	[0.04]	[0.03]	[0.08]	[0.04]	[0.04]	[0.08]
Ln pop. origin	0.02		-0.21***	-0.40***		-0.30***
	[0.01]		[0.03]	[0.02]		[0.04]
Ln pop. dest.	1.08***		0.83***	0.63***		0.74***
	[0.01]		[0.04]	[0.02]		[0.07]
Ln area origin	-0.07***		0.04	0.20***		0.15***
	[0.01]		[0.03]	[0.02]		[0.04]
Ln area dest.	-0.25***		-0.21***	-0.08***		-0.08
	[0.01]		[0.05]	[0.02]		[0.05]
Sum landlocked	-0.82***	0.05	-0.64***	-0.25***	-2.50***	-0.67***
	[0.03]	[0.45]	[0.07]	[0.05]	[0.95]	[0.14]
Border	-4.71***	-7.64***	-1.95	-1.01	-1.45	-2.49**
	[1.00]	[0.95]	[1.25]	[0.94]	[1.09]	[1.19]
Border*(ln dist.)	0.69***	-0.04	0.23	-0.07	0.11	0.97***
	[0.21]	[0.20]	[0.39]	[0.23]	[0.24]	[0.36]
Border*(ln pop origin)	-0.32***	-0.49***	0.01	-0.21**	-0.06	-0.08
	[0.08]	[0.07]	[0.09]	[0.09]	[0.10]	[0.11]
Border*(ln pop dest.)	-0.34***	-0.54***	-0.28***	-0.25***	-0.35***	-0.58***
	[0.08]	[0.07]	[0.10]	[0.09]	[0.09]	[0.12]
Border*(ln area origin)	0.05	0.41***	-0.11	-0.06	-0.04	-0.34***
	[0.09]	[0.08]	[0.13]	[0.10]	[0.11]	[0.12]
Border*(ln area dest.)	0.11	0.45***	0.21	0.31***	0.25**	0.20
	[0.09]	[0.08]	[0.22]	[0.10]	[0.10]	[0.15]
Border*landlocked	0.81***	0.80***	0.83***	0.32**	0.06	0.49**
	[0.11]	[0.11]	[0.14]	[0.13]	[0.14]	[0.20]
Common language	0.60***	0.21***	1.00***	0.88***	0.50***	0.85***
	[0.08]	[0.07]	[0.26]	[0.10]	[0.10]	[0.19]
Common official lang.	0.01	0.69***	-0.38	0.47***	0.64***	0.13
	[0.08]	[0.07]	[0.27]	[0.10]	[0.09]	[0.20]
Time zone diff.	0.13***	0.01	0.02	0.09***	0.02*	0.02
	[0.01]	[0.01]	[0.03]	[0.01]	[0.01]	[0.03]
Colonial ties	3.09***	0.94***	1.43***	1.27***	1.49***	1.02***
	[0.13]	[0.09]	[0.13]	[0.17]	[0.11]	[0.22]
Origin hegemon	-2.23***		-1.78***	1.02***		0.53*
	[0.18]		[0.23]	[0.22]		[0.30]
Observations	24,627	24,627	33,108	8,022	8,022	34,782
R-squared	0.40	0.71	0.22	0.42	0.70	0.23

Note: All models contain an intercept (not shown here). The bilateral trade flows are from the NBER-UN dataset. The bilateral migration flows are from Docquier et al (2010). The trade share (TSH) is defined as the sum of bilateral imports and exports, over GDP of the receiving country, the migration share (MSH) is the number of foreign-born in the country over the total resident population in the country. The fixed-effects estimator includes a full set of origin and destination dummy variables (not reported). The estimated fixed effects are not used in building the predictors for TSH and MSH. In parenthesis we report the heteroskedasticity-robust standard errors. *,**,*** significant at 10%, 5% and 1% confidence level.

Table 3: The Effect of Trade openness on Income per Person. 2SLS estimates.

Specification:	(1) Linear predictor	(2) Nonlinear predictor	(3) Full sample	(4) GDP per worker	(5) Geography 1	(6) Geography 2	(7) Geography and Colonial	(8) FE predictor	(9) OLS
Dep.var.	ln GDP/Pop	ln GDP/Pop	ln GDP/Pop	ln GDP/Emp	ln GDP/Pop	ln GDP/Pop	ln GDP/Pop	ln GDP/Pop	ln GDP/Pop
TSH	2.53***	9.01	3.09**	2.45**	-0.58	-0.14	-0.40	8.30	0.33
	[0.95]	[9.79]	[1.21]	[0.97]	[0.77]	[0.75]	[0.96]	[20.91]	[0.22]
ln (Population)	0.10	0.09	0.10	0.08	0.02	-0.12*	-0.20*	0.06	-0.13*
	[0.11]	[0.35]	[0.12]	[0.11]	[0.09]	[0.08]	[0.11]	[0.32]	[0.07]
ln (Area)	0.11	0.86	0.11	0.12	-0.25**	-0.03	0.03	0.78	0.07
	[0.15]	[1.28]	[0.15]	[0.16]	[0.12]	[0.09]	[0.09]	[2.52]	[0.07]
Dist. equator					0.05***		0.03**	0.02	0.02***
					[0.01]		[0.01]	[0.08]	[0.01]
Pct. Land tropics							0.45		
							[0.49]		
Observations	146	146	181	146	146	146	122	146	122
Controls									
Region	No	no	no	no	no	yes	yes	no	yes
Geo/Climate/Disease	No	no	no	no	no	no	yes	no	yes
Colonial Origin	No	no	no	no	no	no	yes	no	yes
First-stage reg.									
Wald F test	13.71	0.7	9.31	13.71	9.38	7.65	7.3	0.12	.
Instruments	pred. TSH	pred. TSH NL	pred. TSH	pred. TSH	pred. TSH	pred. TSH	pred. TSH	pred. TSH FE	.

Note: All regressions include an intercept. **Regional Dummies** for sub-Saharan Africa, East Asia, and Latin America. **Geography, Climate and Disease controls** include the percentage of land in the tropics, a landlocked dummy, average distance to the coast, average yearly temperature, average yearly humidity, an index of soil quality, an index of the incidence of malaria, and an index of the incidence of yellow fever. **Colonial Controls** includes dummy variables for former French colony, former English colony, and a dummy for the 4 rich "young" countries (US, Canada, Australia and New Zealand), and the share of 1900 population of European origin. For columns 1 through 8 (one endogenous variable and one excluded instrument) the Stock and Yogo (2005) critical values (maximal IV size) range from 5.53 to 16.38, respectively, from the less stringent to the most stringent test (the 25% to 15% maximal IV size). Predictors based on fixed-effects gravity regression do not use the estimated fixed effects. In parenthesis we report the heteroskedasticity-robust standard errors. *,**,*** significant at 10%, 5% and 1% confidence level.

Table 4: The Effect of Trade and Migration Openness on Income. 2SLS estimates.

	(1) Basic	(2) NL Pred. MSH	(3) Full sample	(4) Dist. equator	(5) Region & Geo.	(6) All controls	(7) OLS	(8) Only MSH	(9) Only MSH, FE pred.
Dep. Var.	ln GDP/Pop	ln GDP/Pop	ln GDP/Pop	ln GDP/Pop	ln GDP/Pop	ln GDP/Pop	ln GDP/Pop	ln GDP/Pop	ln GDP/Pop
TSH	2.36**	2.31**	2.91**	-0.50	-0.33	-0.46	0.16		
	[0.94]	[0.91]	[1.22]	[0.63]	[0.62]	[1.18]	[0.19]		
MSH	4.56	5.89**	6.54*	6.40***	7.03***	8.37***	6.31***	7.41***	6.13**
	[4.41]	[2.98]	[3.88]	[1.86]	[2.17]	[2.51]	[1.04]	[1.80]	[3.09]
ln pop.	0.11	0.12	0.10	0.05	-0.07	-0.09	-0.06	-0.06	-0.08
	[0.12]	[0.11]	[0.12]	[0.07]	[0.07]	[0.10]	[0.06]	[0.05]	[0.05]
ln area	0.11	0.11	0.14	-0.22**	0.11	0.11	0.12**	0.12**	0.11*
	[0.16]	[0.16]	[0.16]	[0.09]	[0.07]	[0.07]	[0.06]	[0.06]	[0.06]
Dist. equator				0.05***	0.03***	0.02	0.01	0.02	0.01
				[0.01]	[0.01]	[0.01]	[0.01]	[0.01]	[0.01]
Observations	146	146	181	146	123	121	121	121	121
Controls									
Region	no	no	no	no	yes	yes	yes	yes	yes
Geo/Climate/Disease	no	no	no	no	yes	yes	yes	yes	yes
Colonial past	no	no	no	no	no	yes	yes	yes	yes
First stage reg.									
Wald F test exclusion	8.31	8.17	5.6	4.26	6.83	1.52	.	10.64	5.8
Instruments	pred. TSH	pred. TSH	pred. TSH	pred. TSH	pred. TSH	pred. TSH	.		
	pred. MSH	NL pred. MSH	pred. MSH	NL pred. MSH	NL pred. MSH	NL pred. MSH	.	NL pred. MSH	FE pred. MSH
SY 10% max IV size	7.03	7.03	7.03	7.03	7.03	7.03	.	16.38	16.38
SY 25% max IV size	3.63	3.63	3.63	3.63	3.63	3.63	.	5.53	5.53

Note: Unless noted otherwise the predicted TSH is based on the linear-in-logs gravity estimates and the predicted MSH is based on the NL gravity model. All regressions include an intercept. **Regional Dummies** for sub-Saharan Africa, East Asia, and Latin America. **Geography, Climate and Disease controls** include the percentage of land in the tropics, a landlocked dummy, average distance to the coast, average yearly temperature, average yearly humidity, an index of soil quality, an index of the incidence of malaria, and an index of the incidence of yellow fever. **Colonial Controls** includes dummy variables for former French colony, former English colony, and a dummy for the 4 rich "young" countries (US, Canada, Australia and New Zealand), and the share of 1900 population of European origin. For columns 1 through 7 (two endogenous regressors and two excluded instruments) the Stock and Yogo (2005) critical values (maximal IV size) are between 3.63 and 7.03, for the less stringent to the most stringent test (the 25% to 15% maximal IV size). Predictors based on fixed-effects gravity regression do not use the estimated fixed effects. In parenthesis we report the heteroskedasticity-robust standard errors. *,**,*** significant at 10%, 5% and 1% confidence level.

Table 5: Accounting for Institutional Quality and Early Development. 2SLS.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Main	FE instrum.	Exog. Det.	Alcala-Ciccone	AC controls	TSH, MSH, IQ	PW	PW, no IQ	MSH HK
Dep. var.	ln GDP/Pop	ln GDP/Pop	ln GDP/Pop	ln GDP/Emp	ln GDP/Emp	ln GDP/Pop	ln GDP/Pop	ln GDP/Pop	ln GDP/Pop
MSH	9.04***	10.20***	6.78***			8.75***	10.50***	7.02***	4.85***
	[2.04]	[3.79]	[2.01]			[1.92]	[2.72]	[2.52]	[1.17]
ln Real TSH				1.97***	-0.16				
				[0.70]	[0.93]				
TSH						0.11			
						[0.57]			
Instit. Quality	0.44***	0.43***		0.26*	0.39**	0.44***	0.50***		0.49***
	[0.10]	[0.10]		[0.15]	[0.19]	[0.10]	[0.12]		[0.11]
Dist. equator			0.01					0.04***	
			[0.01]					[0.00]	
Sh. Euro. Descent 1975			0.01***						
			[0.00]						
PW Statehist							-0.11	1.23***	
							[0.46]	[0.40]	
PW sh. fgn. ancestors							-0.62	0.69**	
							[0.50]	[0.28]	
Observations	120	120	122	128	120	120	117	133	120
R-squared	0.77	0.75	0.83	0.26	0.75	0.77	0.72	0.61	0.7
Region dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes
Geo/Climate/Disease	yes	yes	yes	yes	yes	yes	yes	yes	yes
Wald F test exclusion	7.8	2.97	8.06	2.06	0.99	3.04	6.13	7.32	7.24
Instruments	NL prd. MSH Euro1975 Dis. Equa.	FE prd. MSH Euro1975 Dis. Equa.	NL prd. MSH	NL prd. MSH Euro1975 Dis. Equa.	NL prd. MSH Euro1975 Dis. Equa.	NL prd. MSH Euro1975 Dis. Equa. Pred. TSH	NL prd. MSH Euro1975 Dis. Equa.	NL prd. MSH Euro1975 Dis. Equa.	NL prd. MSH Euro1975 Dis. Equa.
SY 10% max. IV	13.43	13.43	16.38	13.43	13.43	n.a.	13.43	13.43	13.43
SY 25% max. IV	5.45	5.45	5.53	5.45	5.45	n.a.	5.45	5.45	5.45

NOTE: All regressions include an intercept and controls for log population and log area (not shown). Unless otherwise noted the predicted TSH is based on the linear gravity model and the MSH on the NL gravity model. **Regional** for sub-Saharan Africa, East Asia, and Latin America. **Controls for Geography, Climate and the Disease environment** include the percentage of land in the tropics, a landlocked dummy, average distance to the coast, average yearly temperature, average yearly humidity, an index of soil quality, an index of the incidence of malaria, and an index of the incidence of yellow fever. **Colonial Controls includes d**ummy variables for former French colony, former English colony, and a dummy for the group of 4 young rich countries (US, Canada, Australia and New Zealand). In specification 9 the migration share has been defined in terms of human capital, with each skilled (college-graduate) worker has 1.5 times the efficiency units of unskilled workers. In parenthesis we report the heteroskedasticity-robust standard errors. *,**,*** significant at 10%, 5% and 1% confidence level.

Table 6: Effect of Migration on the Components of Income per Person. 2SLS estimates.

MSH in Dep. var.	(1) Pop. ln Y/L	(2) Pop. (α/1-α)*ln K/Y	(3) Pop. ln H/L	(4) Pop. ln TFP	(5) Human Cap. ln Y/L	(6) Human Cap. (α/1-α)*ln K/Y	(7) Human Cap. ln H/L	(8) Human Cap. ln TFP
MSH	7.68***	0.94	1.43*	5.31**	3.84***	0.47	0.72*	2.66**
	[2.23]	[0.90]	[0.77]	[2.15]	[1.15]	[0.44]	[0.38]	[1.13]
Observations	98	98	98	98	98	98	98	98
R-squared	0.85	0.34	0.78	0.79	0.83	0.33	0.76	0.78
Controls								
Region	yes	yes	yes	yes	yes	yes	yes	yes
Geo/Climate/Disease	yes	yes	yes	yes	yes	yes	yes	yes
Colonial past	yes	yes	yes	yes	yes	yes	yes	yes
First stage reg.								
Wald F test exclusion	9.06	9.06	9.06	9.06	10.69	10.69	10.69	10.69
Instruments	NL pred. MSH	NL pred. MSH	NL pred. MSH	NL pred. MSH	NL pred. MSH	NL pred. MSH	NL pred. MSH	NL pred. MSH
SY 10% max IV size	16.38	16.38	16.38	16.38	16.38	16.38	16.38	16.38
SY 25% max IV size	5.53	5.53	5.53	5.53	5.53	5.53	5.53	5.53

Note: Dependent variables normalized by the US value. Coefficient "α" is the capital share in the Cobb-Douglas production function underlying this decomposition (Hall and Jones 1999). We have assumed a value α=0.33. All regression models include an intercept and control for log population and log area (not shown). **Regional Dummies** for sub-Saharan Africa, East Asia, and Latin America. **Geography, Climate and Disease controls** include the percentage of land in the tropics, a landlocked dummy, average distance to the coast, average yearly temperature, average yearly humidity, an index of soil quality, an index of the incidence of malaria, and an index of the incidence of yellow fever. **Colonial Controls** includes dummy variables for former French colony, former English colony, and a dummy for the 4 rich "young" countries (US, Canada, Australia and New Zealand), and the share of 1900 population of European origin. The predicted migration share is based on the non-linear Poisson pseudo-ML estimator. In parenthesis we report the heteroskedasticity-robust standard errors. *,**,*** significant at 10%, 5% and 1% confidence level.

Table 7: Diversity and Fractionalization. 2SLS estimates.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dep. Var.	ln GDP/Pop	ln GDP/Pop	ln GDP/Pop	ln GDP/Pop	Ethnic Fractionaliz.	Ling. Fractionaliz.	ln GDP/Pop	ln GDP/Pop	lngdppop
TSH	-0.46	-1.06	-1.02	-0.42	0.34	0.19	1.59		
	[1.18]	[1.03]	[0.93]	[0.76]	[0.28]	[0.36]	[1.77]		
MSH	8.37***	6.79**	7.28***	7.83***	1.15*	1.86**	8.85***	10.11***	9.37***
	[2.51]	[2.84]	[2.81]	[1.69]	[0.65]	[0.87]	[2.42]	[2.15]	[1.86]
Diversity M.		1.16**	1.02**						
		[0.47]	[0.48]						
Diversity T.			-2.15***						
			[0.81]						
Pred. Div. M.				1.19**					
				[0.58]					
Pred. Div. T.				-2.56***					
				[0.59]					
Ethnic Frac.							0.44	0.45	
Ling. Frac.							-1.06	-0.61**	
Catholic sh.							0.00	0.00	0.00
Muslim sh.							-0.01*	-0.00	-0.00
Protestant sh.							0.01***	0.01***	0.01***
Observations	121	119	119	119	120	117	115	115	120
Controls									
Region	yes	yes	yes	yes	yes	yes	yes	yes	yes
Geo/Climate/Dis.	yes	yes	yes	yes	yes	yes	yes	yes	yes
Colonial past	yes	yes	yes	yes	yes	yes	yes	yes	yes
Instruments									
	pred. TSH	pred. TSH	pred. TSH	pred. TSH	pred. TSH	pred. TSH	pred. TSH		
	NL prd. MSH	NL prd. MSH	NL prd. MSH	NL prd. MSH	NL prd. MSH	NL prd. MSH	NL prd. MSH	NL prd. MSH	NL prd. MSH

Note: The predicted TSH is based on the linear in logs gravity estimates and the predicted MSH is based on the non-linear Poisson-ML. Predicted values for the TSH and the Diversity (Fractionalization) index for Trade flows are based on the linear-in-logs gravity estimates and the analogous variables for Migration are based on the non-linear gravity estimates. All regressions include an intercept, log population, log area, and distance to the equator (not shown). **Regional Dummies** for sub-Saharan Africa, East Asia, and Latin America. **Other Geography, Climate and Disease controls** include the percentage of land in the tropics, a landlocked dummy, average distance to the coast, average yearly temperature, average yearly humidity, an index of soil quality, an index of the incidence of malaria, and an index of the incidence of yellow fever. **Colonial Controls includes** dummy variables for former French colony, former English colony, and a dummy for the group of 4 young, rich countries (US, Canada, Australia and New Zealand), and the share of 1900 population of European origin. Standard errors for some control variables have been omitted for lack of space. In parenthesis we report the heteroskedasticity-robust standard errors. *,**,***= significant at 10%, 5% and 1% confidence level.

Table 8: Economic Openness and Income Inequality. 2SLS

	(1)	(2)	(3)	(4)
Dep. Var.	Gini coeff.	Gini coeff.	90-10	90-10
TSH		-0.19		-2.61
		[0.23]		[7.79]
MSH	-0.30	0.66	-67.65	-50.95
	[0.41]	[1.08]	[48.25]	[55.52]
Obs.	104	104	59	59
R-squared	0.57	0.40	0.51	0.54
Controls				
Region	yes	yes	yes	yes
Geo/Climate/Dis.	yes	yes	yes	yes
Colonial past	yes	yes	yes	yes

Note: The dependent variables are the Gini coefficient and the 90-10 ratio of income percentiles. All regressions include an intercept and controls for log population, log area and distance to the equator (not shown). The predicted TSH is based on the linear in logs gravity estimates and the predicted MSH is based on the non-linear Poisson-ML. **Regional Dummies** for sub-Saharan Africa, East Asia, and Latin America. **Geography, Climate and Disease controls** include the percentage of land in the tropics, a landlocked dummy, average distance to the coast, average yearly temperature, average yearly humidity, an index of soil quality, an index of the incidence of malaria, and an index of the incidence of yellow fever. **Colonial Controls includes** dummy variables for former French colony, former English colony, and a dummy for the group of 4 young, rich countries (US, Canada, Australia and New Zealand), and the share of 1900 population of European origin. In parenthesis we report the heteroskedasticity-robust standard errors. *,**,***= significant at 10%, 5% and 1% confidence level.

Table 9: Heterogeneous Effects of Immigration. 2SLS.

Dep. Var.: ln GDP/Pop	(1) By Development in year 1500	(2) By Current Instit. Quality	(3) By Current Human Capital
MSH*High	7.38**	9.08***	8.12***
	[3.17]	[2.67]	[1.79]
MSH*Low	7.42***	6.16***	3.57
	[1.66]	[2.24]	[4.66]
Observations	121	121	121
Controls			
Region dummies	yes	yes	yes
Geo/Climate/Diseae	yes	yes	yes
Colonial Ties	yes	yes	yes
First-stage regression			
Wald F test exclusion	5.51	4.74	2.66
Instruments	NL prd. MSH*High	NL prd. MSH*High	NL prd. MSH*High
	NL prd. MSH*Low	NL prd. MSH*Low	NL prd. MSH*Low
SY Critical values for maximal IV size			
10%	7.03	7.03	7.03
25%	3.63	3.63	3.63

NOTE: The dependent variable is log of income per person.

The instruments are the (non-linear Poisson) gravity-predictor of share of foreign-born interacted with the dummy "high" and "low" for the considered dimension. Countries are classified as having a high (low) value if they are above (below) the median. All regressions include an intercept, logarithm of population, logarithm of area, distance to the equator, and the following three sets of dummies. **Regional Dummies** for sub-Saharan Africa, East Asia, and Latin America. **Geography, Climate and Disease controls** include the percentage of land in the tropics, a landlocked dummy, average distance to the coast, average yearly temperature, average yearly humidity, an index of soil quality, an index of the incidence of malaria, and an index of the incidence of yellow fever. **Colonial Controls includes** dummy variables for former French colony, former English colony, and a dummy for the group of 4 young, rich countries (US, Canada, Australia and New Zealand), and the share of 1900 population of European origin. In parenthesis we report the heteroskedasticity-robust standard errors. *,**,***= significant at 10%, 5% and 1% confidence level.

Table 10: Alternative Measures of Openness . 2SLS.

Dep. Var.	(1) ln GDP/Pop	(2) ln GDP/Pop	(3) ln GDP/Pop	(4) ln GDP/Pop	(5) ln GDP/Pop	(6) ln GDP/Pop
Principal Comp. Open	0.74***	0.82***				
	[0.20]	[0.20]				
Instit. Quality		0.42***		0.42***		0.54***
		[0.10]		[0.09]		[0.12]
MSH			8.56***	9.06***		
			[2.60]	[2.06]		
TSH			-0.41	0.04	-0.92	0.58
			[1.13]	[0.60]	[1.42]	[0.65]
Emig/Pop			0.85	1.05		
			[0.92]	[0.90]		
Net Immig./Pop					6.85***	6.75***
					[2.52]	[1.74]
Dist. Equator	0.01		0.02		0.02	
	[0.01]		[0.01]		[0.01]	
Observations	121	120	121	120	121	120
Controls						
Region	yes	yes	yes	yes	yes	yes
Geo/Climate/Disease	yes	yes	yes	yes	yes	yes
Colonial past	yes	no	yes	no	yes	no
Instruments						
	NL prd. MSH pred. TSH	NL prd. MSH pred. TSH Euro descent Dist. equator	NL prd. MSH pred. TSH	NL prd. MSH pred. TSH Euro descent Dist. equator	NL prd. MSH pred. TSH	NL prd. MSH pred. TSH Euro descnt. Dist. equator

NOTE: Predicted values for the TSH are based on the linear-in-logs gravity estimates and the analogous variables for Migration are based on the non-linear gravity estimates. All regressions include an intercept, log population, and log area. **Regional Dummies** for sub-Saharan Africa, East Asia, and Latin America. **Other Geography, Climate and Disease controls** include the percentage of land in the tropics, a landlocked dummy, average distance to the coast, average yearly temperature, average yearly humidity, an index of soil quality, an index of the incidence of malaria, and an index of the incidence of yellow fever. **Colonial Controls includes** dummy variables for former French colony, former English colony, and a dummy for the group of 4 young, rich countries (US, Canada, Australia and New Zealand), and the share of 1900 population of European origin. In parenthesis we report the heteroskedasticity-robust standard errors. *,**,***= significant at 10%, 5% and 1% confidence level.

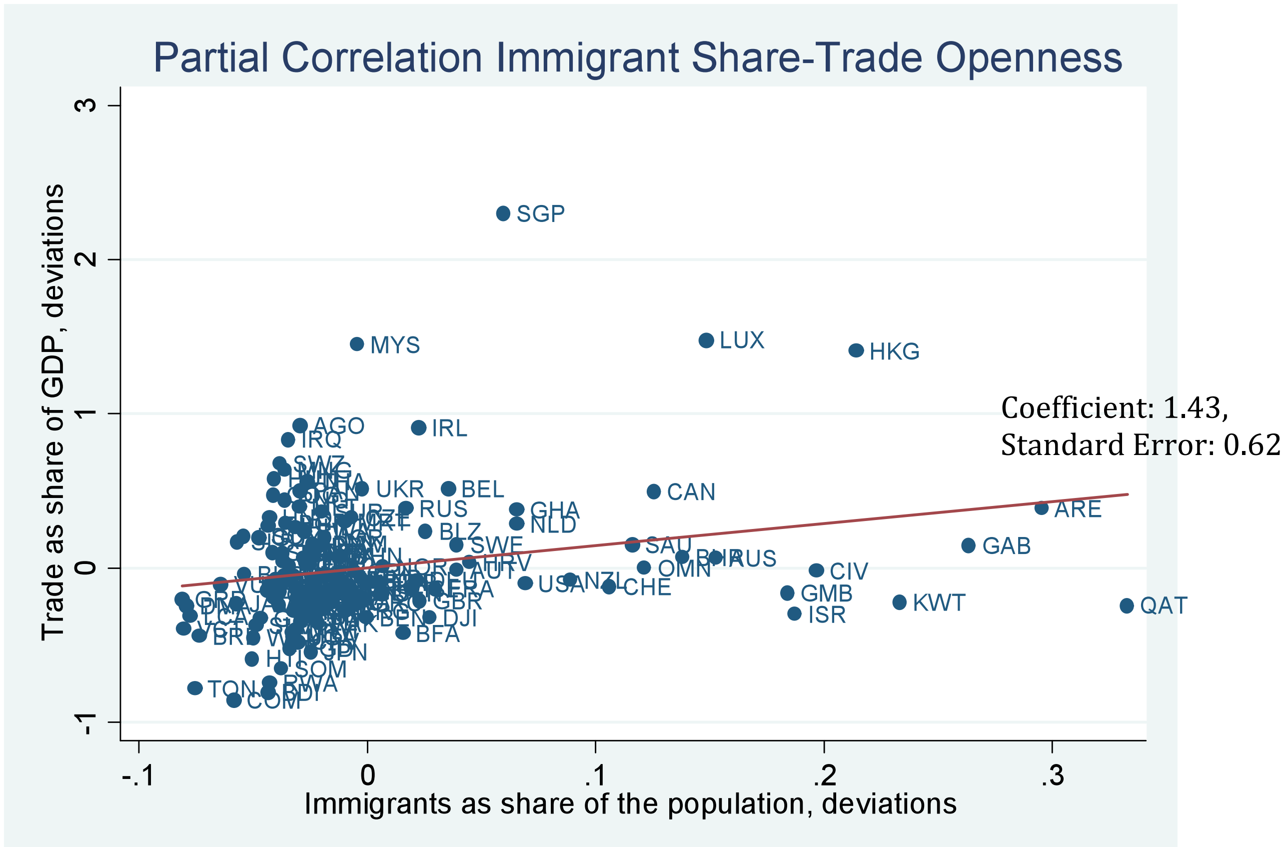
Table 11: Patenting Activity. 2SLS.

	(1)	(2)	(3)	(4)
Dep. Var	Ln Patents / Pop	Ln Patents / Pop	Ln Patents / Pop	Ln Patents / Pop
MSH	11.67*			10.05*
	[6.34]			[5.17]
TSH		1.32		
		[1.57]		
Princip. Comp. Open			1.04*	
			[0.56]	
Institutional Quality				1.57***
				[0.44]
Observations	108	108	108	108
Controls				
Region dummies	yes	yes	yes	yes
Geo/Climate/Disease	yes	yes	yes	yes
Colonial Ties	yes	yes	yes	yes
First-stage regression				
Wald F test exclusion	10.87	6.77	5.90	6.37
Instruments	Pred. MSH	Pred. TSH	Pred. MSH Pred. TSH	Pred. MSH Euro descent in1975 Distance equator

NOTES: The dependent variable is the logarithm of the average number of yearly patents (1995-2010) granted to applicants residing in the country by any patent office in the world, per million inhabitants. All regressions include an intercept, log population, and log area. **Regional Dummies** for sub-Saharan Africa, East Asia, and Latin America. **Other Geography, Climate and Disease controls** include the percentage of land in the tropics, a landlocked dummy, average distance to the coast, average yearly temperature, average yearly humidity, an index of soil quality, an index of the incidence of malaria, and an index of the incidence of yellow fever. **Colonial Controls includes** dummy variables for former French colony, former English colony, and a dummy for the group of 4 young, rich countries (US, Canada, Australia and New Zealand), and the share of 1900 population of European origin. In parenthesis we report the heteroskedasticity-robust standard errors. *,**,***= significant at 10%, 5% and 1% confidence level.

Figures

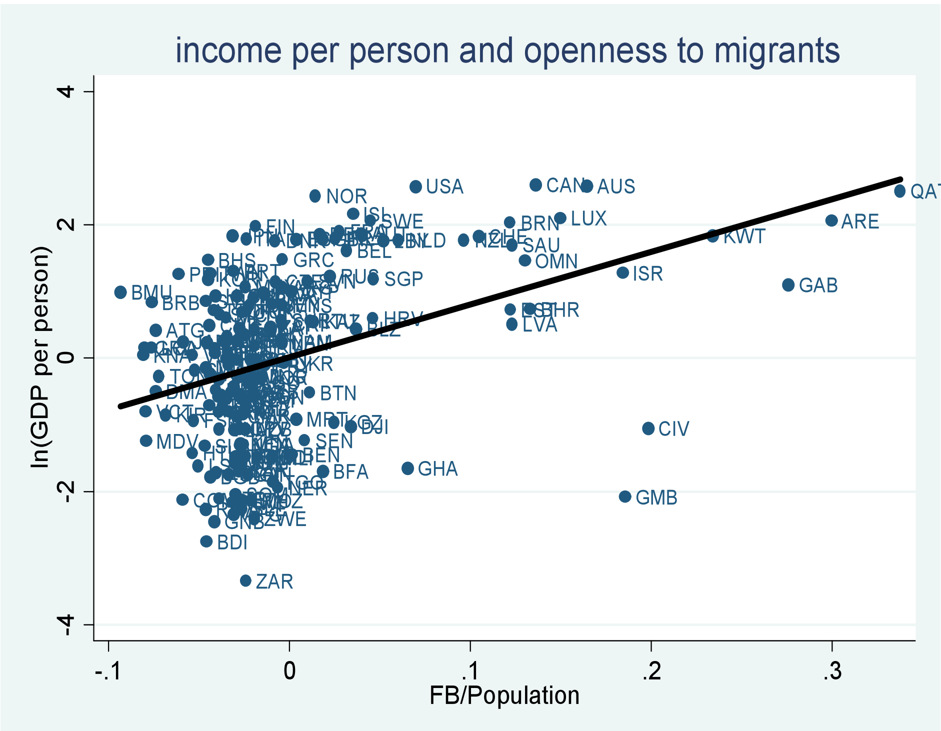
Figure 1: Migration Share and Trade Share



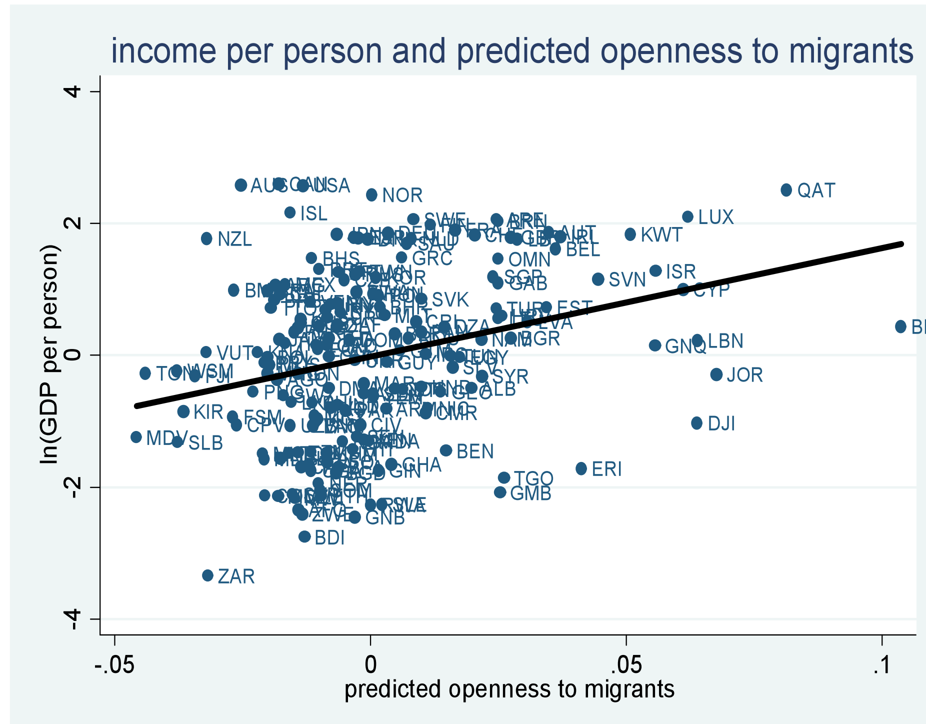
Note: The data are relative to 147 countries in year 2000. We plot the residuals after adjusting by log population and log area. The sources and construction of the trade as share of GDP and of the foreign-born share are described in the text.

Figure 2: Openness to Immigration (MSH) and GDP per person, adjusted for country size

2A: MSH and GDP per person



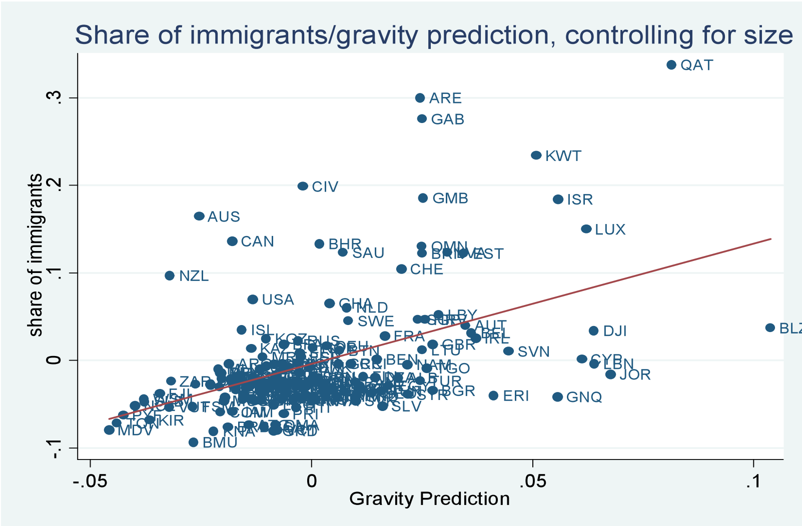
2B: Gravity-predicted MSH and GDP per person



Note: The scatterplot shows each variable after adjusting for logarithm of population and area. The predictor for immigration share used is the linear gravity predictor.

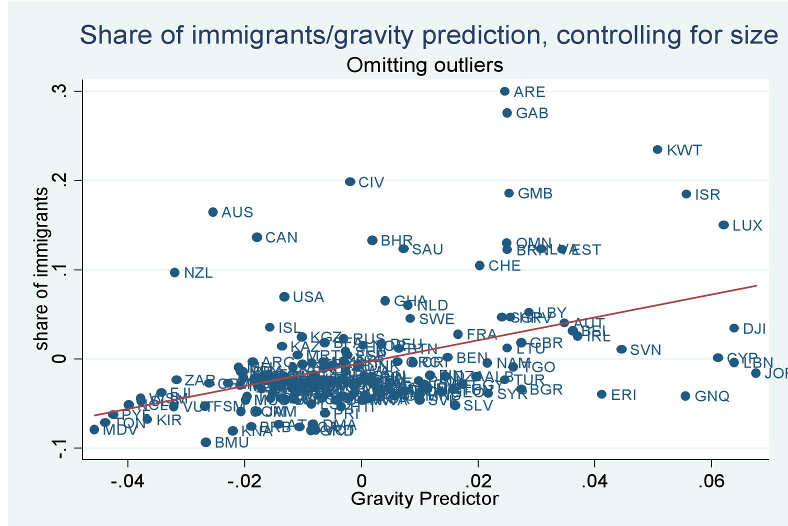
Figure 3

3A: Fit of the predicted migration share, adjusted for country size



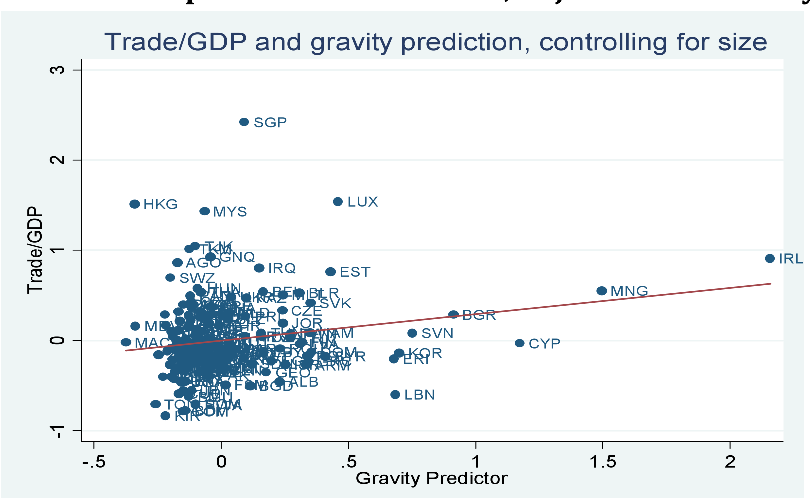
Slope: 1.37, standard error: 0.30 F-stat: 20.56

3B: Excluding 2 outliers



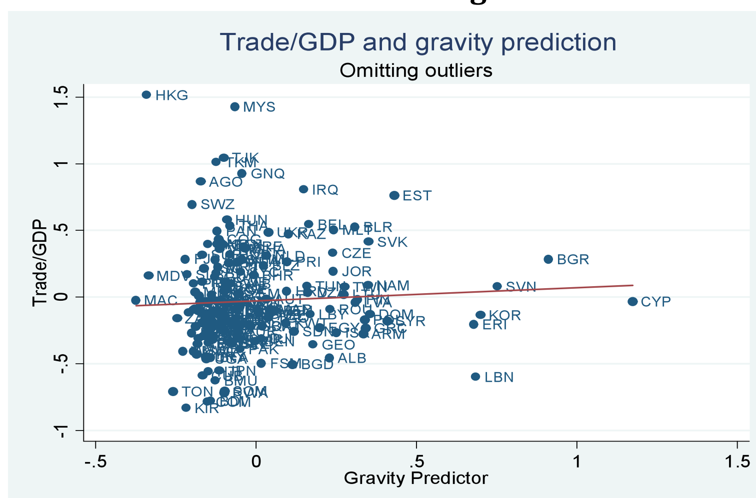
Slope: 1.28, standard error: 0.26 F-stat: 22.90

3C: Fit of the predicted trade share, adjusted for country size



Slope: 0.29 std. error 0.09, F-test 9.39

3D: Excluding 4 outliers



Slope: 0.09 std. error 0.11, F-test 0.39

Appendix

Table A.1

	Effect on Income	Effect on Income	Effect on Institutions	Effect on Institutions
	(1) ln GDP/Pop	(2) ln GDP/Pop	(3) Institutional Quality	(4) Institutional Quality
MSH	12.01*** [3.23]	9.49*** [2.49]	2.37 [4.93]	-2.89 [4.30]
Putterman-Weil Statehist	2.53*** [0.49]	1.06** [0.40]	3.56*** [0.74]	1.39** [0.70]
Putt.-Weil Share fgn. Ancestors	-0.04 [0.33]	0.62** [0.27]	0.31 [0.54]	1.03** [0.50]
Dist. equator		0.023*** [0.007]		0.02* [0.01]
European descent 1975		0.01*** [0.002]		0.025*** [0.005]
Observations	133	128	126	140
Wald F first stage	8.89	6.69	9.83	7.81
Instruments	pred. MSH	pred. MSH	pred. MSH	pred. MSH
SY 10% maxIV	16.38	16.38	16.38	16.38
SY 25% maxIV	5.53	5.53	5.53	5.53

Note: All regressions include a constant and ln (area) and ln (population) as explanatory variables. Heteroskedasticity-robust standard errors. *,**,***= significant at 10%, 5% and 1% confidence level.