



DEPARTMENT OF ECONOMICS

Working Paper Series

The Effect of Trade and Migration on Income

Francesc Ortega
Queens College CUNY
Giovanni Peri
UC Davis

June 19, 2012

Paper # 12-13

This paper explores the relationship between economic openness and income per person using cross-country data. To address endogeneity concerns we extend the instrumental-variables strategy first used by Frankel and Romer (1999). First, we show that bilateral geographic characteristics of countries such as distance, contiguity and commonality of language are successful in predicting openness to immigration and to trade. Equipped with these instruments we then establish a robust, positive effect of openness to immigration on long-run income per capita across countries using econometric specifications that include a comprehensive set of variables controlling for geography, climate, disease environment, and colonial past. In contrast the positive effect of trade openness on income vanishes once the control variables are included in the specification. Our main finding is robust to explicitly including institutional quality as an (endogenous) regressor, controlling for measures of early economic development, and measuring the share of immigrants in terms of efficiency units of labor. We also show that the main effect of migration operates through total factor productivity but not through institutional quality. This is consistent with the idea that immigration increases the variety of skills and ideas available for production. We also provide some more direct evidence of this channel by building an index of the degree of diversity in immigration flows by country of origin. Immigration also increases linguistic fractionalization which, in turn, has a negative effect on income per capita, however the direct gains from greater skill diversity are much larger than the costs arising from increased fractionalization due to immigrants. We do not find evidence of increased income inequality due to openness to immigration or trade. Finally we find evidence that immigration benefits the amount of innovation produced in a country as measured by patents.

Department of Economics
One Shields Avenue
Davis, CA 95616
(530)752-0741

http://www.econ.ucdavis.edu/working_search.cfm

The Effect of Trade and Migration on Income

Francesc Ortega (Queens College CUNY)

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

This Draft: June 18th, 2012

Abstract

This paper explores the relationship between economic openness and income per person using cross-country data. To address endogeneity concerns we extend the instrumental-variables strategy first used by Frankel and Romer (1999). First, we show that bilateral geographic characteristics of countries such as distance, contiguity and commonality of language are successful in predicting openness to immigration and to trade. Equipped with these instruments we then establish a robust, positive effect of openness to immigration on long-run income per capita across countries using econometric specifications that include a comprehensive set of variables controlling for geography, climate, disease environment, and colonial past. In contrast the positive effect of trade openness on income vanishes once the control variables are included in the specification. Our main finding is robust to explicitly including institutional quality as an (endogenous) regressor, controlling for measures of early economic development, and measuring the share of immigrants in terms of efficiency units of labor. We also show that the main effect of migration operates through total factor productivity but not through institutional quality. This is consistent with the idea that immigration increases the variety of skills and ideas available for production. We also provide some more direct evidence of this channel by building an index of the degree of diversity in immigration flows by country of origin. Immigration also increases linguistic fractionalization which, in turn, has a negative effect on income per capita, however the direct gains from greater skill diversity are much larger than the costs arising from increased fractionalization due to immigrants. We do not find evidence of increased income inequality due to openness to immigration or trade. Finally we find evidence that immigration benefits the amount of innovation produced in a country as measured by patents.

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

Openness to interactions and exchange with the rest of the world may be one of the most powerful channels of economic development, technological change and social progress for a country. This has been recognized and studied in depth by economists and historians. As interactions and exchanges with other people and other groups are at the very foundation of economic gains (from variety, specialization and spillovers) openness to the rest of the world should be relevant for growth, especially for relatively small countries (Alesina, Spolaore and Wacziarg (2000, 2005), Frankel and Romer (1999)).

Empirical tests of this proposition in the existing literature have focused on measures of trade openness. Measures of policy openness to international trade (Sachs and Warner (1995), Lucas (2010)), or measures of trade as a share of GDP (Frankel and Romer (1999), Rodrik (2000), Rodriguez and Rodrik (2001), Noguer and Siscart (2005), Alcalá and Ciccone (2004)) have been analyzed as determinants of long-run economic performance of countries.

An important challenge in this exercise stems from the potentially severe problems of endogeneity and omitted variable bias. Both trade policies and trade volumes in a country are affected by its level of development. Furthermore, trade shares are an imperfect measure of economic openness. In an influential paper Frankel and Romer (1999) suggested using the geography driven variation in bilateral trade costs as the basis for identification of the effect of trade openness on income. The idea behind this identification approach is that the relative location of a country, in geographical space, influences its income per person only by affecting the extent of its trade flows with the rest of the world.

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 in the world. The latter has a *direct* effect on income through climate, soil quality, agricultural productivity, disease environment, and institutional quality (through a country's history of colonization as argued in Hall and Jones (1999) and Acemoglu et al. (2001)). Thus the validity of the Frankel and Romer (1999) strategy requires controlling for a country's "absolute" geographic location, and its geographic and climatic characteristics. Subsequent authors, such as Noguer and Siscart (2005), Alcalá and Ciccone (2004), and many others, have also emphasized the importance of using better data, improved measures of trade openness, and controlling for institutional quality in studying the effect of openness on income per person.

In this paper we re-examine the relationship between economic openness and economic development, emphasizing a different dimension of openness. During the last century (especially before World War I and again since the 1980's), a crucial aspect of openness has operated through international migration flows. This channel remained neglected in the literature because of the limited size of migration flows during the 1930-1970 period. However, migration flows have played a key role in the historical and economic development of many countries,

particularly those in the New World. Migrants brought ideas, skills, and labor, providing an important channel for the international diffusion of knowledge. As recently emphasized by Putterman and Weil (2010) these migration flows have had extremely persistent effects.

Thus a complete analysis of the effects of economic openness on development requires an analysis of openness to migration in addition to openness to trade. The empirical challenges are analogous to those faced by the literature on the effects of trade on income since migrants are attracted to countries with high income and good institutions. Still, migration costs also influence migrants' decisions and therefore the relative location of a country should also affect its openness to immigrants. Specifically, individuals tend to migrate to countries that are at a short distance, both geographically and culturally. Consequently, we propose to apply the identification strategy of Frankel and Romer (1999) to the analysis of openness to migration. **We use measures of geographic and linguistic distance between countries in a basic gravity framework to predict bilateral migration and trade flows. We then aggregate these flows by receiving country to predict aggregate migration and trade, and use these as instruments for actual openness to migration and trade.

Since trade and migration depend, to some extent, on the same set of variables which may affect bilateral costs, the empirical literature estimating the effect of trade openness on income may suffer from omitted variable bias. Figure 1 reports the partial correlation between trade as a share of GDP and immigrants (foreign-born) as a share of the resident population for 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. A positive and significant correlation is evident from Figure 1.

This paper offers several novel contributions. First, using new data on the bilateral stocks of migrants in year 2000, we include the share of immigrant population as an additional measure of economic openness. The response of trade and migration to geographical variables is differentiated enough to allow us to separately identify their effects on income per person. This is an extension of the Frankel and Romer (1999) strategy. Second, our analysis includes a very extensive set of controls (geography, climate, disease and colonial ties) and we pay particular attention to the role of institutional quality and pre-industrial history as additional determinants of income per capita. Third, we consider the effect of trade and migration on other country-level outcomes such as productivity, human capital accumulation, inequality and innovation. We also explore the role of immigrant diversity on productivity and income per person and find that diversity of immigrants by country of origin may be responsible for a large fraction of the positive effect of openness to immigration.¹

Our results show that openness to immigrants, instrumented by its geography-based gravity predictors in a cross-country regression of 147 countries, has a significantly positive and large effect on income per person

¹A 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.

in year 2000. A 10% 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%) in the long-run across countries. We then consider whether the effect of immigration on income is robust to a series of geographical, climate, institutional and historical controls. In particular we check that the effect of immigration is not simply driven by institutional quality differences instrumented using distance to the equator (Hall and Jones (1999)) and the share of people with European origin (Alcala and Ciccone (2004) and Acemoglu, Johnson and Robinson (2001)). The results are also very robust to controlling for the pre-1500 history of development of countries' ancestors that Putterman and Weil (2010) have emphasized as an important determinant of current development. In contrast, the effect of openness to trade, while large and significant in the basic specification (as found in previous studies), does not withstand the inclusion of geographic controls and becomes insignificantly different from 0.

Overall our findings strongly suggest that openness to foreigners plays an important role in determining the long-run economic success of countries and is possibly even more important than openness to trade. Even after controlling for institutions, policies, pre 1500 development, colonial history and geographic characteristics, our estimates imply that immigration flows have a large and significant effect on income per person. Once we account for immigration, the role played by trade openness is not significant. Our results also suggest the effect of immigration on long-run income per capita is roughly orthogonal to the effect of institutional quality. Specifically, we find no evidence that openness to foreigners leads to better institutions, but given institutional quality, higher openness to immigrants generates more productive economic systems. Our empirical findings are consistent with a model where immigration leads to greater skill/ideas diversity, which in turn increases total factor productivity (TFP) and income per capita in the long run. We also explore further the relationship between immigration and receiving country skills, and find evidence that openness to immigration is more beneficial to income in countries with high human capital.

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 a simple theoretical framework that justifies the main empirical specification used in this paper. 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 the effect of openness to immigration. Section 8 explores whether the effect of immigration survives the inclusion of controls for institutions, culture, historical events and other checks. In section 9 we show the effects of immigration on each component of income per person. We also explore the effect of diversity of immigrants on income per person, linguistic fractionalization, inequality, and innovation, and we perform several other checks. Section 10 concludes.

2 Literature Review

There is a vast theoretical literature linking several aspects of openness to income per capita levels and growth.² A series of models (such as Grossman and Helpman (1991), Rivera-Batiz and Romer (1994), Eaton and Kortum (1996), Lucas (2010)) emphasize the role of openness to trade in promoting innovation, technological diffusion and catch-up. Depending on the details of the model such processes may affect the income of countries relative to the technological frontier or, in the case of endogenous growth, their growth rates. The effect of market size on innovation and growth spurred by trade has also been analyzed as a driver of economic growth. The size of the market can affect the speed and type of technological adoption (Acemoglu (2003)), or may encourage specialization and learning by doing (e.g. Matsuyama (1992), Galor and Mountford (2008)). Other studies emphasize the role of openness in increasing the efficiency of production of domestic firms subject to international competition (e.g. Weil (2005)).

More closely related to this paper are studies that empirically analyze the relationship between trade and income per capita across countries, controlling for other determinants of growth and instrumenting openness with policies or geographical determinants. A very influential early paper by Sachs and Warner (1995) analyzes the effect of trade policies (classified for a group of countries over the period 1965-1990) on economic growth. Frankel and Romer (1999) (FR from now on) introduce an identification strategy based on the fact that geographic determinants of bilateral trade can be used as instruments for trade if relative geography affects income only via trade flows. They use this approach to estimate the impact of trade on income per person. Following FR, Rodrik (2000) and Rodriguez and Rodrik (2001) question the validity of the results arguing that the initial findings of a positive and significant effect of trade on income per person disappear when one controls for some simple geographical features such as the distance from the equator or the share of population in tropical area. Noguer and Siscart (2005) and Alcalá and Ciccone (2004) show that openness to trade still matters, even after controlling for institutional quality, using a larger database on trade flows and measures of “real openness”. Also closely related to this literature are papers by Alesina and Spolaore (1997, 2003), and Alesina, Spolaore and Wacziarg (2000, 2005), which focus on the effect of trade openness on growth, controlling for the size of countries. They further develop the analysis by allowing the size of a country to be endogenously determined by a trade off between economic benefits and political costs of size.

In comparison to the very large literature on trade and income per capita there are almost no studies on the long-run effect of migration on income per person. While there are several papers on the analysis of cross-country determinants of migrations using a gravity equation (e.g. Clark et al. (2008), Pedersen et al.

²There are some textbook treatments of theories of openness and growth. Chapter 18 and 19 of the textbook by Acemoglu (2009) present and overview of the effect of knowledge diffusion and trade on income per person. Chapter 8 of the textbook by Barro and Sala i Martin (2004) is devoted to the analysis of the diffusion of technology in models with endogenous growth. Chapter 11 of the textbook by David Weil (2005) analyzes the relationship between growth and openness.

(2004), and Mayda (2010)) very few studies have looked at the opposite direction of causality from migration to income growth in the long run. Peri (2012) looks at the long-run effect of immigration on productivity and income per person across US states. Related to our paper is also Putterman and Weil (2010) who relate the level of development in a country to the history of the ancestors of that country's inhabitants by constructing a matrix of international migration flows since 1500. While the authors consider immigration a key element for economic development, it only affects the past origins of current inhabitants. They do not consider the effect of immigration on growth, although they argue that immigrants from advanced countries bring human capital, culture, and institutions, which increases the economic success of receiving countries. In section 8 we consider the Putterman and Weil (2010) story in greater detail, and borrow their long-run migration matrix to construct the degree of historical openness to immigration. However, we more directly analyze the contribution of immigration as carrying benefits from varieties of skills/human capital not mediated by institutions.

Finally there have been several historical studies focusing on the increased openness of migration and trade during the 1800 and their effects on growth. Among others, O'Rourke and Williamson (1994), Williamson (1996) and O'Rourke, Taylor and Williamson (1997) show the important role of mass migration and globalization on income per capita convergence in the late nineteenth century.

3 The Theoretical Framework

In this section we describe a simple model that justifies the basic regression specification we use in section 7 to analyze the effect of openness to trade and openness to migration on income per person. The model is based on Alesina, Spolaore and Wacziarg (2000). We consider the determinants of income per capita using a simple static endowment economy. This can be interpreted as the steady state of a long-run growth model, hence our predictions are about the long-run effects of openness on income. The goal of the model is simply to show how income per person, trade as a share of GDP and immigration as a share of population depend on the size and openness of a country.

Consider N identical regions in the world, indexed by $i = 1, 2, \dots, N$. grouped into to C countries. Each country contains S_c regions, $c = 1, 2, \dots, C$ so that S_c can be considered the size of the country (both as area and population). Each region is endowed with a given amount of labor (unskilled workers) L_i , physical capital K_i , and human capital (skilled labor) H_i . Labor is homogeneous across regions, while physical and human capital are differentiated and each region is endowed with a unique variety. Output Y , homogeneous across regions, is produced using the following production function:

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

Expression (1) implies that producers in region i combine different varieties of skilled labor. H_{ij} represents the human capital input of region-specific variety j used in region i . They also combine different varieties of an intermediate good to be used in region i , X_{ij} , produced using physical capital specific to region j according to a linear function $X_{ij} = K_{ij}$. For simplicity we impose a strongly symmetric structure across regions. We assume that all regions have the same productivity level $A_i = 1$, capital endowment k , number of unskilled workers (equal to 1), and number of skilled workers (equal to h). We also assume that all countries are the same size so that $N = S * C$.

Assuming output as the numeraire, each region demands skilled workers, unskilled workers and intermediate goods to produce the final good. We assume that there is a cost when using an intermediate good produced in a region of a different country and in hiring a skilled worker from a different country, which we model as an iceberg-type cost. When employing a skilled worker from a different country only a fraction $(1 - \gamma)$ of her human capital can be used in production because of downgrading skills, or due to time spent away from production to fulfill immigration requirements. When using an intermediate input from another country only a fraction $(1 - \delta)$ can be employed in production after paying the trade costs. The parameters γ and δ (between 0 and 1) can therefore be considered as measures of migration and trade costs. When using intermediate goods and human capital from different regions within the same country there are no migration or trade costs.

Intermediate goods and skilled workers have an international market, hence their price/wage must be equal across regions. Unskilled workers are homogeneous across regions, hence in this perfectly symmetric world, if there is any migration cost they do not migrate. Considering the production function (1) in region i of country c we can derive the demand for skilled workers from other regions of the same country (D_{ii} as in "Domestic" workers) and for skilled workers from regions of other countries (F_{ij} as in "Foreign-born" workers). Using the symmetry across regions and types we have that the wage paid to any skilled worker has to be the same so that $w = w_j = w_i$ and we have:

$$w = \alpha \left(\sum_{i=1}^N X_{ij}^\beta \right) (D_{ii})^{1-\alpha} = \alpha \left(\sum_{i=1}^N X_{ij}^\beta \right) (1 - \gamma)^\alpha (F_{ij})^{1-\alpha} \quad (2)$$

Similarly for intermediate goods, given their international price p (equal across varieties because of symmetry), we can derive the demand for the varieties produced domestically, X_{ii} , and for those produced abroad, X_{ij} , as follows:

$$p = \beta \left(\sum_{i=1}^N H_{ij}^\alpha \right) (X_{ii})^{1-\beta} = \beta \left(\sum_{i=1}^N H_{ij}^\alpha \right) (1 - \delta)^\beta (X_{ij})^{1-\beta} \quad (3)$$

Therefore there is the following relationship between the number of skilled workers from each region working in another region of the home country and those working in a region abroad: $F_{ij} = \theta D_{ii}$, where $\theta = (1 - \gamma)^{\frac{\alpha}{1-\alpha}}$

is between 0 and 1 and measures how "free" migration is. A similar relationship exists between the amount of intermediate good produced for use in a domestic region and the amount produced for use in a foreign region: $X_{ij} = \phi X_{ii}$ where $\phi = (1 - \delta)^{\frac{\beta}{1-\beta}}$ is between 0 and 1 and measures how "free" trade is. The market clearing conditions for each type of human capital and for each type of physical capital are as follows:

$$SD_{ii} + (C - 1)S\theta D_{ii} = h \quad (4)$$

$$SX_{ii} + (C - 1)S\phi X_{ii} = k \quad (5)$$

Equations (4) and (5) and the conditions $F_{ij} = \theta D_{ii}$ and $X_{ij} = \phi X_{ii}$ can be used to obtain:

$$F_{ij} = \theta D_{ii} = \theta \frac{h}{S(1 - \theta) + \theta N} \quad (6)$$

$$X_{ij} = \phi X_{ii} = \phi \frac{k}{S(1 - \phi) + \phi N} \quad (7)$$

We can now derive the migration share of the population (MSH_i) and the value of trade as share of GDP (TSH_i) in each region i .³ These are the two most commonly used measures of immigration and trade intensity. Immigrants are the sum of highly skilled from all regions in the rest of the world ($N - S$), while natives are the sum of the highly skilled from country S who work in region i plus the local less skilled, whose measure is standardized to one:

$$MSH_i = \frac{F_{ij}(N - S)}{F_{ij}(N - S) + D_{ii}S + 1} = \left(\frac{h}{1 + h} \right) \frac{\theta(N - S)}{\theta(N - S) + S} \quad (8)$$

It is easy to show, for a given country size S , the migration share depends positively on the openness to immigration, $[\partial(MSH_i)/\partial \ln \theta] \theta = a_1 > 0$. Conversely for given openness θ the migration share depends negatively on the size of the country, $\partial(MSH_i)/\partial S = -a_2 < 0$. Expression 8 also shows other determinants of the share of foreign born, such as the ratio of skilled to unskilled workers (h) and the size of the world economy (N).

Similarly trade (equal to imports+exports)⁴ as a share of GDP in region i is given by:

$$TSH_i = 2\beta \frac{X_{ij}(N - S)}{X_{ij}(N - S) + X_{ii}S} = 2\beta \frac{\phi(N - S)}{\phi(N - S) + S} \quad (9)$$

³Those shares in the home country will be the same as in region i .

⁴Imports are equal to exports in this model. Recall that the share of income, β , goes to pay for intermediate goods, whose price is the same. Hence the value of imports plus exports relative to the total value of intermediate goods used is equal to the ratio of their quantities.

For given country size an increase in openness, ϕ , would increase trade as a share of GDP, $[\partial(TSH_i)/\partial \ln \phi] \phi = b_1 > 0$, while an increase in size of the country, for a given openness (as long as $\phi < 1$), will decrease trade, $\partial(TSH_i)/\partial S = -b_2 < 0$. Expression (9) shows that the share of trade also depends on the elasticity of final output to intermediates β and on the overall size of the world economy N .

Finally, substituting (6) and (7) into (1) and dividing by population in the region, $(1+h)$, we can express output per person in region i , and hence in country C as:

$$y_i = \left(\frac{A_i^{1-\alpha-\beta} h_i^\alpha k_i^\beta}{1+h} \right) [S(1-\theta) + N\theta]^{1-\alpha} [S(1-\phi) + N\phi]^{1-\beta} \quad (10)$$

It is easy to see that openness to migration θ and openness to trade ϕ both positively affect income per person (i.e. $\partial y_i / \partial \theta > 0$, $\partial y_i / \partial \phi > 0$) through an increase in the second and third terms in brackets which affects the productivity of factors. Similarly, for a given openness, an increase in the size of the country, S , increases productivity (i.e. $\partial y_i / \partial S > 0$). Hence, expression (10) that shows income per capita in region i of country c can be re-written as:

$$y_c = TFP(\tilde{A}_c, \phi, \theta, S_c) \tilde{h}_c^\alpha \tilde{k}_c^\beta \quad (11)$$

In (11) we substituted the country subscript and we use the notation $\tilde{h}$ and $\tilde{k}$ to denote human capital and physical capita per person. Hence TFP is a function of the country size, economic openness (in terms of trade and migration), and other factors such as institutions, social norms and government policies (embodied in $\tilde{A}_c$).

Our model has two main implications that we test. First, for a given size, economies that are more open to international trade and migration enjoy greater input variety, higher TFP and higher income per person. Second, the share of trade relative to GDP and the share of immigration relative to the total population are increasing functions of trade and migration openness, respectively. These shares can be directly measured in our data.

4 Estimation

4.1 Empirical Specification and Instruments

Taking a log linear approximation of expression (10) we obtain the following relationship between income per person and openness to trade and immigration for country c :

$$\ln y_c = \beta_0 + \beta_1 \ln \theta_c + \beta_2 \ln \phi_c + \beta_3 \ln S_c + \beta_4 \mathbf{X}_c + \varepsilon_c \quad (12)$$

The 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 ($\ln \tilde{A}_c$), human capital ($\ln \tilde{h}_c$) and physical capital ($\ln \tilde{k}_c$). 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 $\mathbf{X}_c$. A difference from (11) is that in (12) we allow the trade and migration costs to be country specific (indexed by c). However we cannot estimate (12) because openness to migration and trade, θ_c and ϕ_c , are not observable. Hence, we will use 8 and 9, and rely on the dependence of migration and trade on their respective costs. In particular we can linearize the expression (8) around the average values of the parameters and obtain:

$$MSH_c \approx \mathbf{A} + a_1 \ln \theta_c - a_2 S_c + \mathbf{a} \Xi_{M,c} \quad (13)$$

The 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 share of foreign-born. The coefficients a_1 and a_2 are defined in section 3 as the semi-elasticities of the migration share to migration openness and to country size, respectively. The term Ξ_M is a vector containing h , N and other observable and unobservable determinants of MSH .

Similarly we can linearize expression (9) to obtain:

$$TSH_c \approx \mathbf{B} + b_1 \ln \phi_c - b_2 S_c + \mathbf{b} \Xi_{T,c} \quad (14)$$

The scalar $\mathbf{B}$ collects the average values of the parameters affecting trade. The coefficients b_1 and b_2 are defined in section 3 as the semi-elasticities of the trade share to openness and to country size, respectively. The term $\mathbf{b}$ is a vector of partial derivatives of TSH_i with respect to β , N and other potential determinants of trade. Ξ_T is the vector containing β , N and all other determinants of TSH .

We can now solve (13) with respect to $\ln \theta_c$, (14) with respect to $\ln \phi_c$ and substitute these into (12) to obtain:

$$\ln y_c = \beta_0 + \beta_M MSH_c + \beta_T TSH_c + \beta_S \ln S_c + \beta_4 \mathbf{X}_c + \beta_5 \Xi_c + \varepsilon_c \quad (15)$$

In expression (15) 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 . If there are some unobservable factors in Ξ_c they will be confounded in the residual and will be correlated with the explanatory variables. In particular, if these factors are omitted variables that affect both income and trade/immigration they will cause an upward bias in the OLS estimator. If they are factors that only affect the correlation between openness and actual trade/immigration, they will

act as measurement error and bias the OLS estimates down. These two effects would cause inconsistency in the OLS estimates of β_M and β_T in the following regression:

$$\ln y_c = \beta_0 + \beta_M MSH_c + \beta_T TSH_c + \beta_S \ln S_c + \beta_C \mathbf{Controls} + u_c \quad (16)$$

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

In order to obtain consistent estimates of β_M and β_T we identify the cost-driven determinants of trade and migration for country c . Namely we use expressions (13) and (14), and isolate factors that only affect trade costs (ϕ) and migration costs (θ) for each country. We follow Frankel and Romer (1999) who use the gravity regression to predict bilateral flows. In particular, using a bilateral gravity model we first run the following regression to predict the logarithm of bilateral trade and migration shares:

$$\begin{aligned} \ln x_{cj} = & \gamma_1 \ln(Dist)_{cj} + \gamma_2 \ln(Pop)_c + \gamma_3 \ln(Pop)_j + \gamma_4 \ln(Area)_c + \gamma_5 \ln(Area)_j \\ & \gamma_6 (Landlocked)_c + \gamma_7 (Border)_{cj} + \gamma_8 (ComLang)_{cj} + \gamma_9 (Colony)_{cj} + \\ & \gamma_{10} \ln(Dist)_{cj} (Border)_{cj} + \gamma_{11} \ln(Pop)_c (Border)_{cj} + \gamma_{12} \ln(Pop)_j (Border)_{cj} \\ & + \gamma_{13} \ln(Area)_c (Border)_{cj} + \gamma_{14} \ln(Area)_j (Border)_{cj} + \\ & + \gamma_{14} \ln(Landlocked)_j (Border)_{cj} + u_{cj} \end{aligned} \quad (17)$$

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. Finally the interactions of the border dummies with the distance, population area, and landlocked dummies are also included. These factors are determined by the relative geography and the relative spatial and cultural location of countries.

Once we have estimated the gravity regressions (17) we use them to obtain the predicted bilateral share of migrants and trade, and then we aggregate across j to obtain the predicted trade and immigration intensity for each country c . In particular, we define Z_{cj} to be the vector of explanatory variables included in (17) 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 measure of (bilateral) cost-driven trade as a share of GDP for country c as follows:

$$\widehat{TSH}_c = \sum_{j \neq c} \exp(\widehat{\gamma}_T Z_{cj}) \quad (18)$$

Similarly we define the measure of gravity-predicted migration share in country c as follows:

$$\widehat{MSH}_c = \sum_{j \neq c} \exp(\hat{\gamma}_M Z_{cj}) \quad (19)$$

We consider the gravity predictions as identifying the part of (bilateral) trade and migration driven by size and bilateral physical and cultural geography of a country vis-a-vis the rest of the world. Hence, once we control for the size of the country, the predicted $\widehat{TSH}_c$ and $\widehat{MSH}_c$ are functions of trade costs θ_c and migration costs ϕ_c determined by relative geography only. The trade and migration literature have estimated equations like 17 several times in the past and our goal is not to have a structural interpretation of the coefficients $\hat{\gamma}_T$ and $\hat{\gamma}_M$ but rather to use the formulas (18) and (19) to predict the part of migration and trade uncorrelated with income per person. Nevertheless, the more recent and model-based implementation of the gravity equation to predict trade (e.g. Anderson and Van Wincoop (2003)) and migration (Ortega and Peri (2009)) include a set of country of origin and of country of destination fixed effects. These capture the effect of "multilateral resistance" on trade and migration and are needed to obtain consistent estimates of the effect of geography on trade/immigration flows. Hence, in our empirical implementation we also estimate (17) including a set of country of origin and country of destination fixed effects. We do not include the estimates of those fixed effects in the geography-predicted flows that are still calculated exactly as in (18) and (19). This method allows us to consistently estimate the effects of geography on flows, and then use these consistently estimated coefficients in the prediction.

4.2 Discussion of the Identification Strategy

As described above our identification strategy, based on relative geography, is an extension of Frenkel 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 institutions in determining both income levels and the degree of openness to trade and migration (Hall and Jones (1999)). Moreover if institutional quality is related to geography (as argued in Acemoglu et al. (2001)) this may also invalidate the 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)).⁵ Clearly the exclusion restriction required for the validity of this instrumental variables approach has been called into question.

In order to deal with these concerns we use the following approach. 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 are considered as exogenous regressors and absorb the effect of geography through other omitted variables strengthening the assumption behind the exclusion restriction for the instruments.⁶

Second, we explicitly include institutional quality as an additional regressor, which raises additional endogeneity issues that need to be dealt with.⁷ The importance of explicitly accounting for institutional quality can be illustrated as follows. Good institutions, protecting property rights, granting balance of powers and ensuring economic freedom are certainly a key to (probably the main determinants of) the economic success of a country, high income per person, and high productivity. At the same time these institutions might have been favored by latitude or distance from Europe. Failing to account for the direct effect of institutions on income may bias our estimates of the effects of (trade and migration) openness. Along the same lines it is also plausible that good early political institutions may have produced both high income levels today (Chanda and Putterman (2005, 2007), Tabellini (2010), Putterman and Weil (2010)) and higher degrees of (trade and migration) openness. 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 institutional development.

Given these challenges in estimation we provide results based on two different, 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 additional endogenous regressor. Under this specification we use distance to the equator and other variables that capture the degree of influence of European powers on good institutions as additional instruments for institution, while we maintain the gravity-based instrument for openness to immigration.

⁵In particular, latitude influenced the probability of being settled by population from a Western European country (Hall and Jones (1999)). Geographic location also influences the country's disease environment, which affected the strategy followed by colonizers (Acemoglu, Johnson and Robinson (2001)). Putterman and Weil (2010) have emphasized the importance of the quality of the institutions of a country's ancestors, as opposed to the institutions in that actual location.

⁶Our vector of controls contains the main variables included in Noguer and Siscart (2005) plus several additional variables.

⁷This was the strategy pursued by Alcalá and Ciccone (2004) in their analysis of the effects of trade openness and institutional quality on income per person.

Let us also emphasize, that the instrumental variable approach helps us deal with the fact that we are using the trade and immigrant shares as proxies for the economic openness of a country. The latent variable affecting both measures, and also the economic development of a country, is the openness to economic interactions with people, capital and goods from other countries. Hence the instruments address the issue of measurement error of the latent variable. In the presence of measurement error and omitted variable bias, that could be both severe, it is hard to even predict the direction of the OLS bias.

5 Data and Summary Statistics

Our bilateral trade data is from the NBER-UN dataset.⁸ 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 obtain the most accurate values on bilateral trade. We also checked this database with the International Trade database (BACI) available at CEPII⁹ and the data for year 2000 are essentially identical, where they both exist.¹⁰ 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 source of these data is the national censuses of countries, 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 8000

⁸We 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/>.

⁹Downloadable at <http://www.cepii.fr/anglaisgraph/bdd/baci.htm>.

¹⁰The correlation coefficient with the CEPII bilateral trade data is 0.99 when restricting to the same country pairs.

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. However, we made a significant effort to extend the coverage for most variables, and thus also present results for the full sample, to confirm their robustness.¹¹

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 Alcalá 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 as share of the population are calculated using the Docquier (2010) data. Along the lines of Hall and Jones (1999) and Alcalá 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).¹² Acemoglu, Johnson and Robinson (2001) is also the source of several 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).¹³ 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 measure immigrants in terms of human capital, by using Mincerian returns and the share of college-educated, the foreign-born share 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 variables is very high (0.96).

¹¹We have performed (without reporting them) most of the regression on the full dataset. There is never a significant difference with the results found using the FR dataset only.

¹²For more details see page 1397 in Acemoglu, Johnson and Robinson (2001).

¹³Following footnote 4 in Alcalá and Ciccone (2004), real trade openness is defined as (nominal) openness times the price level, which undoes the dependence on relative nontradeable goods prices.

We also obtained two variables from Putterman and Weil (2010) that are important controls. The first is an index of early 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 of the current population of each country around year 1500. 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 value is 0.24, with a large standard deviation(0.32), and ranges from zero to 100 percent.

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. Next we reproduce the results of the previous literature that focused on the effect of trade openness on income.

6.1 Gravity estimates

Table 2 reports the estimates of the gravity model for bilateral trade flows, following equation (17). Column 1 reports the estimates of a linear-in-logs model. Column 2 reports the estimates of a similar model that includes sets of country of origin and country of destination fixed effects. Finally in column 3, following Santos-Silva and Tenreyro (2008), we use a non-linear estimation method (Poisson pseudo-maximum likelihood), which allows us to include zero bilateral trade flows. Accordingly, the sample size is larger for column 3. The goodness of fit is obviously substantially higher in the specification including the fixed effects (column 2), since they capture all the country-specific determinants of immigration and emigration flows. However, in most cases, the coefficients on the geographical/linguistic and colonial variables have similar signs and significance across specifications. Many of the estimated coefficients are highly significant, with the expected signs. Bilateral distance significantly reduces trade flows, while size (in terms of population), sharing a common language, and having a shared colonial past all increase trade flows.¹⁴ As explained earlier, we use our estimates of the vector of coefficients γ_T , obtained from specifications (1), (2) or (3) of 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 in 18 to obtain the predicted trade share for each country, which form our instrumental variables. We will discuss the estimates for the migration gravity regressions (also reported in Table 2, columns 4-6) in section 7 below.

¹⁴Departing from Frankel and Romer (1999) our gravity model includes information on past colonial ties, along the lines of Head, Mayer and Ries (2010). This increases the explanatory power (and the resulting strength of the predictor for the trade share).

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:

$$\ln y_c = \beta_0 + \beta_{TR}(TRsh_c) + \beta_P \ln Pop_c + \beta_A \ln Area_c + \beta_C \mathbf{Controls} + u_c \quad (20)$$

In (20) the dependent variable is income per person in country c measured in 2000 US Dollars, corrected for PPP as in Penn World Table, 7.0. 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 version of the gravity-based predictor proposed by FR and constructed using the estimates of Table 2 described above.

Table 3 reports the two-stage least-squares estimates for the model (20). Columns 1-4 of Table 3 report the estimates of the basic model, which includes only controls for country size (log 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 Ciccone and Alcala, 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 is exactly within the range obtained by FR, who show estimates between 1.97 and 2.96.¹⁵ 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 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 pseudo-maximum likelihood, in Column 3 of Table 2) to include the zero-observations. 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. This signals the problems with using trade flows as measure of openness. Still, to give this strategy the best shot at identifying an effect of trade on income, in the following specifications we use the predicted trade share based on the linear-in-logs estimates of the bilateral trade shares. Column 4 illustrates that the FR result also holds when using 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. Column 5 includes distance from the equator as a control. This variable is highly significant, confirming the results in Hall and Jones (1999).

¹⁵ See the Table 3, page 387 of FR.

Moreover the coefficient on the trade share falls dramatically and becomes statistically insignificant.

Column 6 examines an alternative way of controlling for geography—including a number of regional dummies (sub-Saharan Africa, East Asia, and Latin America)—and confirms the findings in the previous column: the effect of the trade share becomes small and insignificant. Column 7 includes additional variables, to control for geography, climate, soil quality, disease environment, and the colonial past.¹⁶ 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 seen in the first-stage F test). 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.¹⁷ Finally, even when we use the theoretically more appropriate fixed-effect gravity estimator to predict the trade share, in Column 8, we do not obtain any significant effect of trade. The estimates (with very large standard errors) exhibit the symptoms of weak instrument bias. Our interpretation is that the trade share is a rather imprecise measure of the economically relevant interactions between countries. Alternatively openness to trade may not capture some key aspects of the economic openness of a country. This leads us to analyze whether openness to people, as measured by migrations, has a more robust effect on income.

7 The effect of Openness to Trade and Openness to Migration on Income

The economic growth literature has almost exclusively considered openness to trade as measure of economic openness.¹⁸ However, migration has always been a very important channel for people and ideas to interact. Compared to international trade, which may involve only a limited number of agents (firms), long-term migration may better facilitate economic interactions encouraging the exchange of skills and ideas. Economists, as pointed out by Gordon Hanson (2009), have mostly studied the partial effect of immigration on employment and wages of specific groups, thus ignoring the long-run potential effect of immigrants on competition, scale, variety of skills, and new ideas. These are all crucial channels in determining income per capita, as emphasized in the growth literature (see Jones and Romer (2010) for a summary). Hence it seems only natural to analyze the effect of openness to immigration in addition to openness to international trade. The remainder of the paper deals with this analysis.

Let us first inspect some simple correlations. Figure 2 shows that there exists a robust positive partial

¹⁶The 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.

¹⁷Our 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. Clearly, since our data differ from Noguer and Siscart's, our estimates are bound to differ from theirs. 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.

¹⁸See for instance the summary in the textbook by Weil (2007).

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.¹⁹ Figure 2B plots the gravity-predicted migration share (adjusted for country size) and income per capita.²⁰ In both cases the correlation is robust to dropping outliers. Moreover the positive correlation is strong and it is not driven by the US, Canada, or Australia— countries that are both highly economically developed and have a high share of foreign-born. This is particularly true for the predicted migration share (Figure 2B), given the large size and relative remoteness of these countries which generates relatively low imputed immigration.

Let us now examine the joint relationship between trade and migration openness on income in a more formal regression setting. In Table 4 we include openness to migration (measured by the share of foreign-born in the total resident population) as an additional explanatory variable. We estimate specification (16), treating both the trade and migration shares as *endogenous* regressors. We use the gravity models to estimate bilateral immigration and trade flows. To instrument the migration share we use, alternatively, a linear in logs predictor (whose coefficients are reported in column 4 of Table 2), a linear in log predictor with country of origin and destination fixed effects (in column 5 of Table 2), or a non-linear Poisson maximum-likelihood predictor (column 6 of Table 2). The bilateral explanatory variables are the same as for the trade regression and the estimated coefficient for the migration gravity regression are shown in Table 2, columns (4) to (6). The signs of the coefficients are largely as expected if we interpret these variables as proxying for migration costs, and they usually share the same sign and level of significance across specifications. Specifically, bilateral distance reduces bilateral migration, while sharing a common language and colonial ties appear to significantly increase migration. Notice that several geographic variables have significantly different impact on trade or migration flows.²¹ These differences enable us to predict each flow separately.

Before going through the estimates of Table 4 it is useful to visually examine the fit of the predicted trade and migration shares. Figure 3 displays the scatterplots of the predicted and actual trade and migration shares. It is clear that the predicted migration share fits the actual data fairly well (Figure 3A). In fact the two are highly correlated and the correlation is not driven by outliers. Omitting the two observations with highest predicted migration shares (Belize and Qatar) leaves the correlation unchanged (Figure 3B). The picture, however, is different for the predicted trade share (Figures 3C and 3D). In this case the positive correlation between predicted and actual values depends strongly on few influential observations, namely Ireland, Luxembourg and Singapore. Omitting these (Figure 3D), the correlation is weakened and loses its statistical significance.²²

Table 4 reports the 2SLS joint estimates of the effect of trade and migration openness on income per person.

¹⁹The regression coefficient is 6.5 with a standard error of 1.18.

²⁰The regression coefficient is 15.7 with a standard error of 3.95.

²¹For instance Common colonial ties seem to affect trade more than migration while common language seem to affect migration more than trade.

²²The 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)).

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 2.3-2.9%. Likewise a one-percentage point increase in the immigration share increases income per person by about 6%. However, as we include, progressively, distance from the Equator (specification 4), regional dummies and controls for geography and climate (specification 5), and control for the colonial origin of countries (former English or French colonies, share of European settlers and their descendants measured as of 1900) in 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 5.9 and 7.03. The migration share appears to be a more robust determinant of income per person across countries, suggesting that migration plays a very important economic role. In specification 7 we use the theoretically correct, fixed-effects gravity estimator to predict the migration share (while keeping the linear estimator for the trade share). Even in this case the effect of immigration remains large and statistically significant.

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 the Wald F-test of the first stage is at most 2.79, which is below the Stock and Yogo (2005) less stringent critical value. Column 8, however, reports OLS estimates of the regression model featuring migration and trade shares jointly. The estimates here reveal that when controlling for geography, climate, and countries' colonial past the partial correlation of the migration share with income per capita remains positive while that for the trade share is quantitatively small and statistically not significant. We also note that the OLS estimate for the effect of migration is not far from the corresponding instrumental-variables estimate (column 6). This suggests that the endogeneity bias and the measurement error bias (both likely to be present) are of similar size. We will analyze explicitly several possible omitted variables in the remaining sections.

Finally, column 9 reports a specification where the trade share has been dropped. As we can see, the point estimate of the migration share remains stable, with a value around 6.3, and an associated standard error around 1. In this case, however, the Wald F-test of the excluded instrument is above 62, much larger than the most stringent Stock and Yogo (2005) critical value. As the effect of openness to trade is never significant in specifications including geographic controls, we exclude this regressor in several subsequent specifications to improve the power of the instruments.

Let us illustrate, here, the magnitude of the effect of openness to immigration on income per person using a point estimate of 6.3 (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 of residents, between countries at the tenth and ninetieth percentiles of income per person.²³ 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 History

In our previous empirical model we included a fairly complete set of geographic controls. According to Hall and Jones (1999) and Acemoglu, Johnson and Robinson (2001), the main reason why these variables are relevant (in particular the distance from the equator) is that geography was decisive in determining if, by whom, and how a country was colonized. Acemoglu, Johnson and Robinson (2001) argue that those initially less-developed countries that were colonized by a Western European power through long-term settlements were endowed with good institutions, with extremely persistent effects on economic development in the long-run. On the contrary, countries that were colonized but not settled by Europeans experienced “exploitative” institutions, with persistent negative long-term effects.

The institutional quality of a country has been considered the main determinant of its long-run economic success by a series of very influential papers.²⁴ We have already alluded to the role of European Colonization in bringing different types of institutions, such as free markets, democracy, checks and balances and the legal system. Economists have argued that differences in institutions could depend on the colonizing country (e.g. the distinction between French and English legal systems made by La Porta et al. (1999)) or on the type of colonization (for settlement or for exploitation) as argued by Acemoglu et al. (2001). It is also likely that countries with good institutions encouraged international trade and facilitated immigration. This might create a spurious correlation between openness and income. In this section we consider the role of institutional quality explicitly in our empirical specification. Our measure of institutional quality follows Acemoglu et al. (2001) and is the average between the “Protection against expropriation risk” and the “Constraint 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. The value of this index ranges between 0 and 8.

Of course, institutions are likely to be endogenous to economic development. Following Hall and Jones (1999) and Alcalá and Ciccone (2004), we complement the gravity-based predictors for openness to trade (and migration) with arguably exogenous determinants of initial institutions. Namely, distance to the equator and

²³This also corresponds to the increase in the foreign-born share experienced, for instance, by Spain between 1998 and 2008.

²⁴Hall and Jones (1999), Acemoglu et al (2001, 2002), Acemoglu and Johnson (2005), La Porta et al (1999), among others.

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 variable provides a measure of the social, economic and cultural influence of being colonized and settled by Europeans. Table 5 reports our estimates. Column 1 considers the role of the migration share and our index of institutional quality on income per person. 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. Both are treated as endogenous and instrumented. 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 90th and 10th percentiles. Hence while institutions are still the main determinant of income differences, as indicated by the previous literature, openness to immigrants adds a very significant long-run effect. Notice, moreover that the instruments are jointly not too weak as the Wald test is higher than the less stringent Stock and Yogo (2005) critical value.

Column 2 shows the estimated effect of immigration and institutions when we instrument the first endogenous variable using the fixed-effects gravity predictor. While the point estimate of the effect is somewhat smaller (6.30) the effect is still very significant, precisely estimated and very close to the estimates of Table 4. The instrument is also similarly strong.

Columns 3 and 4 show the estimates of the effect of immigration, when we include the exogenous predictors of institutions directly as regressors (rather than using them as instruments). Specifically we include the distance from the equator and the share of the population of European Descent as controls, in addition to the regional dummies and the other geographic, climatic and disease-environment controls. Both the point estimates of the coefficient on the migrant share and its significance confirm the results of the previous two columns. In addition, when we only include the share of migrants as an endogenous regressor, the fixed-effects gravity estimator (specification 4) turns out to be a very strong instrument (Wald F-test equal to 52), suggesting that the specification with fixed effects is theoretically sound and empirically very effective.

Column 5 estimates the specification in Alcalá and Ciccone (2004). That is, trade openness is measured by the log of *real* trade share (in which trade values, that is, are expressed in PPP). As in their paper, the estimated coefficient associated to this variable is highly significant. However, when we introduce controls for geography, climate, and the disease environment, we cannot reject the null of a zero effect even on this new measure of trade openness that better captures its real scale. Column 7 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 3 endogenous variables and 4 instruments. While the instruments are somewhat weak, the estimates confirm once again that the migration share and institutional quality have a positive and highly statistically significant effect on income per person, whereas the coefficient on trade share does not have any significant effect.

Putterman and Weil (2010) emphasize the role of measures of early development, particularly of pre-1500 political and administrative institutions, in explaining current income levels. They also argue that adjusting these measures to take into account the countries of origin of the ancestors of the current population greatly increases the magnitude of the effects. In a way their analysis also puts migration at the center-stage of cross-country development. However, they focus on a story rather different from ours. Immigration of people from countries with a long history of developed state and political institutions is likely to bring better culture and institutions to a country. This would generate long-run benefits for countries that received immigrants from long-civilized European populations. We explore differences between their story and ours in two ways. First, using their data we include two variables in our regressions: the ancestor-adjusted quality of political institutions before year 1500 (the so-called *Statehist* variable) and the share of the population with pre-1500 foreign ancestors.²⁵ Column 8 introduces these two controls. The coefficients on the migration share and the institutional quality index are even larger than in previous specifications, and the new controls based on the Putterman and Weil (2010) analysis are not significant. This however, is unfair to their theory as they argue that development brought by highly civilized ancestors mainly work through institutions. Hence in table A.1 of the appendix we explore the differences in the effects of migration openness and of Putterman and Weil variables on income and on institutional quality separately. When we control for distance from the equator and for the share of population of European descent and not for institutions (specification 2 and 4 in Table A.1) we find that openness to immigration has a very significant effect on income, and also that *Statehist* and the share of foreign ancestors have a significant positive effect. Such an effect of *Statehist* and foreign ancestors must come through institutions (in fact, when controlling for those, as we do in specification 8 of Table 5, the variables from Putterman and Weil do not affect income any longer). Regression (4) shows that *Statehist* and the share of foreign ancestors have a positive and significant effect on institutional quality, even when controlling for distance from equator and European ancestry. To the contrary openness to immigrants does not matter for the institutional quality of a country. Hence, in the Putterman and Weil analysis historical flows of people affect income across countries mainly by affecting their institutions, while our analysis finds that recent immigration

²⁵The raw *Statehist* variable is an index, ranging between 0 and 1, capturing the length of time prior to year 1500 since the country developed a supra-tribal government and discounting the distant past. The ancestor-adjusted variable, say for the US, is a weighted average of *Statehist* across all countries in the world, where the weights correspond to the shares of the current US population whose ancestors in year 1500 lived in each country in the world. More plainly, this adjustment accounts for the fact that migrants may have brought ideas, skills and social norms whose quality correlates with the level of development of their societies of birth in year 1500. The exact definition of the index can be found at page 1640 and 1641 of Putterman and Weil (2010).

has an additional and positive effect on income, controlling for institutions.

In all our results so far we have measured the migration share without accounting for the educational attainment of natives and immigrants. However, our data allow us to disaggregate the native and immigrant populations in two groups: those with a college education 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.²⁶

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 (effectiveness) units, 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 standard errors associated with the migration share variable have fallen by almost half in column 9.

The magnitude of the effect implied by this coefficient is comparable with what we obtained in column 1. 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 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 Additional Results

9.1 Effects on income components

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 diversity in terms of ideas, skills, and abilities, in the receiving country. This translates into higher long-run income per person. We now explore the channels behind this reduced-form relationship. We begin by breaking down this effect into the effects on income components. Following Hall and Jones (1999) and Alcalá and Ciccone (2004), we postulate a simple Cobb-Douglas aggregate production function in which output is produced using human

²⁶Specifically, 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.

capital and physical capital. We derive a decomposition of the effects of migration on income per worker (as opposed to per person) along three main channels: physical capital intensity, human capital intensity, and TFP. Specifically,

$$\ln y_c = \frac{\alpha}{1 - \alpha} \ln(K_c/Y_c) + \ln(h_c) + \ln(TFP_c) \quad (21)$$

In expression 21, the 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 are from the Penn World Table 7.0 while the data on average schooling are from the Barro and Lee (2011) and the Cohen and Soto (2007) databases.²⁷

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 and instrument it. The difference between specifications (1)-(4) and (5)-(8) is that in the first set of estimates we use the fixed-effects gravity predictor for the foreign-born share of the population, while in the second set of estimates we use the non linear gravity equation predictor for the foreign-born share.²⁸

While the first set of estimates exhibit somewhat smaller effects, the pattern and significance of the coefficients is the same in the two specifications. The main channel that accounts for the effect of migration on income per capita is TFP, which is consistent with our theoretical model. We find no evidence of an effect of immigration on capital intensity, and a marginally significant effect operating through human capital when we use the non linear gravity equation predictor as an instrument. Overall immigration does not seem to affect much the capital intensity of the receiving economy, it somewhat increases its human capital intensity, and it substantially increases its productivity. Why does immigration lead to higher TFP? Our model suggests that this is because it increases the diversity of the skills in the labor force. An alternative interpretation is that immigration increases the variety of ideas, enhancing problem-solving and innovation potentials. The next section provides further evidence suggestive of this relation between immigration-diversity and productivity. Section 9.7 considers more directly the role of immigration in generating innovation as measured by patents.

²⁷Where 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 use the database gathered by Marcelo Soto and hosted at his website.

²⁸Following Hall and Jones (1999) all dependent variables have been normalized by the US value.

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. By spanning the set of abilities, skills and ideas, immigration may induce specialization along comparative advantage, and increase the varieties of goods and services available. 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) increased the efficiency and productivity of US cities. Similarly Broda and Weinstein (2006) quantify the productivity gains of trading different goods with a variety of foreign countries. While it is not easy to capture the dimension of diversity of immigrants that are relevant for productivity, we think that it is useful and reasonable to begin analyzing the effect of diversity with respect to country of origin. On one hand different countries of origin affect culture, skills and language spoken, and these in turn are likely to affect the varieties of productive ideas. On the other hand, the trade literature (e.g. Broda and Weinstein (2006)) considers the country dimension as a crucial reason for differentiated goods.

In this section we analyze whether diversity in migration and trade flows has an effect on income per capita. Specifically, we construct the so-called “fractionalization index” by country of origin:²⁹

$$Div_c = 1 - \sum_j (M_{jc})^2 \quad (22)$$

In expression (22), M_{jc} is the share of foreign-born from country j in the total of foreign-born 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 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 in formula (22).

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.³⁰ 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).

Let us now turn briefly to the diversity index for trade flows. Again, let us take 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). This is likely to be driven by the low trade costs for intra-Europe trade. However, the countries with the highest values tend to be low income (Pakistan, India, Kenya or Tanzania are

²⁹Here we define the index for migration but it is analogous for trade flows. Our index is defined as one minus the corresponding Herfindahl-Hirschman concentration index.

³⁰Mexico 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.

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 very dominant position of the US as their main trading partner.

In principle the diversity of immigrants by country of origin can be correlated with other determinants of income and certainly contains substantial measurement error. Ideally, we would like to use our gravity predictors for bilateral flows to build an instrument for the diversity index as well. In practice the gravity model delivers a predicted diversity index that is a poor fit for the index calculated using actual data.³¹ To address this shortcoming we begin by including the diversity index as an exogenous regressor, while still treating the share of immigrants (and the trade share) as endogenous. Later we replace the diversity index with its gravity-based prediction, which is a function solely of geographical location and hence not affected by income and institutions.

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 ties. The first column reproduces the earlier findings, using the fixed-effects gravity predictor as an instrument, and shows that the migration share has a positive and significant effect on income, while the point estimate of the trade share is very low and statistically insignificant. Column 2 adds the diversity index for migration, treating it as an exogenous regressor. The coefficient associated with this variable is positive and highly significant. In addition it cuts the coefficient on the migration share almost by half, rendering it non-significant. Column 3 adds the diversity index for trade, which has an insignificant negative effect. The coefficients on the migration share and the diversity index for migration remain practically unchanged.

Column 4 reports the point estimates when we include the gravity-predicted diversity index for immigrants and trade as controls. The share of immigrants and trade are still considered as endogenous and instrumented by the predicted shares. The estimated effects of diversity are similar to those in specification (3), while the effect of openness to immigration is larger in this specification. Based on column (4), the point estimate for the effect of the migration share on income per person is around 5 and is highly significant. The point estimates for the migration and trade diversity indices are 1.25 and -1.97 , respectively, and both are significantly different from zero.

The income effects 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. When considering diversity as endogenous, and instrumenting it with the predicted gravity-based diversity, we need to omit the share of immigrants for the instruments to have power. In that case (not reported in Table 7) we obtain a larger coefficient of 3.87, significant at the 10% level, but with a relatively large standard error. While the limited power of the geographic

³¹The problem is particularly severe when we attempt to instrument for both the migration (or trade) share and the corresponding diversity index.

instruments for the diversity variable does not allow us to find conclusive proof, there seem to be reasonable evidence that one channel through which immigration increases income per person is via its diversity. This may correspond to variety of ideas, skills, cultures and a selection of human capital types, very important to enhance productivity and efficiency.

Taken together these estimates suggest that both the size of migration, relative to the population, and its diversity, in terms of origin countries, have a positive effect on income per capita. In contrast, 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 negative effects associated to large and diverse migration flows. In particular, it may lead to ethnic or linguistic fractionalization, which may have 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.³² 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 also find these effects to be sensitive to the specification. Moreover they rarely find any effect on growth of income per person. As immigration is likely to increase ethnic and linguistic diversity of a country we explore if this channel matters in our regressions.³³

We focus on this issue in the second part of Table 7. Columns 5 and 6 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. This is not surprising at all but it begs the question of whether the potentially negative economic consequences of increased linguistic fractionalization associated with immigration (such as reducing the scope for productive economic interactions) can be compared to the gains of greater diversity on income per person found in the previous sections.³⁴

³²See Alesina and La Ferrara (2005) for an overview.

³³The 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 a similar analysis looking at the effect of place of birth diversity on income and productivity across countries.

³⁴We have also estimated an analogous specification where the dependent variable is an index of religious fractionalization. We did not find a significant effect associated to the migration share.

Column 7 and 8 of Table 7 report the estimates of a specification where we explicitly control for ethnic and linguistic fractionalization, and for the share of Catholic, Muslim and Protestant in the country's population as of 1980. In specification 7 we use the log linear gravity predictor for the share of immigrants while in specification 8 we use the fixed effects gravity predictor. First, we find that there is a significant, negative effect of linguistic fractionalization on income per person but no effect from ethnic fractionalization. This is exactly consistent with the growth regressions of Alesina et al. (2003).³⁵ Second, the coefficient of the migration share is large and very significant. In particular, when using the fixed effects predictor (Column 8), the migration share has a coefficient of 7. This is similar to the effect obtained when we do not include ethnic and linguistic fractionalization (and leave the rest of the regression unchanged), which is reported in Column (9) and is equal to 7.22. Hence the positive effect of immigration on income is much larger than its indirect effect through increased linguistic fractionalization. Let us notice that increased immigration (and the consequent increased fractionalization) may have effects on other economic outcomes, especially on the provision of public goods or on the generosity of the public sectors. Analyzing those specific effect is interesting and we leave it for further research. In the next section we consider the impact of immigration on income inequality.

9.4 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.³⁶ 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).³⁷ These authors conclude that 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. In particular we note that to the extent that immigration may increase variety in skills and ideas, as emphasized by our model, it may not affect the degree of income inequality, as it mainly produces horizontal differentiation of skills.³⁸

³⁵We do not include religious fractionalization in the regression as it is never significant.

³⁶See Richardson (1995) for a survey of the literature and the recent study by Autor, Dorn and Hanson (2011).

³⁷See also the replies by John DiNardo and John Abowd to the 1997 article.

³⁸We already showed in section 9.1 that immigration does not affect a country's capital intensity in the long run, and it may even

We consider a common measures of income inequality at the national level: the Gini coefficient, which ranges between zero and one and is available for a large set of countries. This measure is available from the UNU-WIDER database, version WIID 2C.³⁹ Table 8 reports our two-stage least-squares estimates. We consider, alternatively immigrants as share of the population (specifications 1 and 3) or both immigrants and trade as share of GDP (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 specification (3) and (4) use the non linear gravity predictor. In all specifications neither the trade nor the migration share have significant effects on the Gini coefficient, revealing that while immigration has a positive effect on average income per person it does not seem to systematically affect the income distribution. This is in line with the idea that immigrants bring skill complementarity to a country that does not significantly alter the income differentials between the rich and poor. In conclusion the cross-country evidence does not indicate that economic openness, as measured by the trade share of GDP and by the migration share in the total population, leads to long-run increases in income inequality.

9.5 Heterogeneous effects across countries

The empirical analysis so far has shown a very significant and strong effect of openness to immigration on income per worker, that is robust to a series of geographic controls, controls for institutional quality and history, and the omission of outliers. It is reasonable to think, however, that the effect of immigrants on income and productivity may depend on the ability of the receiving country to best utilize the different skills and ideas that arrive. While we have seen that immigration has a positive effect on income, independent from the institutional quality and history of the receiving country, what we estimated, nevertheless, is an average effect. It may be the case that countries with better institutions or with more educated population are better at exploiting the inflow of skills and ideas brought by immigrants. Moreover, while we have seen that openness to trade does not seem to be correlated with income per person it may be the case that more open countries have a better environment for the skills of immigrants to thrive. Trade openness, that is, may be complementary to immigration in producing effects on income per person. 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. Finally we use the variable measuring average human capital per person (relative to the US) to differentiate countries with high (above the median) and low human capital levels. Then

increase human capital intensity. Both facts, if highly educated and capital owners are richer than workers, would tend to reduce wage inequality.

³⁹This 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.

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 controls as in Table 4 specification 9. The estimates of the coefficients are reported in Table 9, where we report two estimates (using the non-linear predictor or the fixed effects predictor as IV) for each specification. First notice that all coefficients are positive and nine out of twelve are significant. In most subsamples immigration has a strong positive effect on income. Focusing on the differences we notice that only for human capital levels do we obtain larger point estimates of the effect of immigration for countries endowed with high human capital values. Immigration has a significantly stronger positive effect on income in countries with higher human capital. For differences in early development (*Statehist*) and in trade openness the coefficients are rather similar. A formal test rejects equality of the coefficients only in specification 6, confirming the idea that high human capital may be important in enhancing the positive effect of immigration on income. This implies a positive complementarity between immigrants and local human capital, that is potentially stronger for highly educated countries. If this is correct, the largest benefit from immigration would accrue to countries that are likely to be better equipped to absorb them and to use their skills.

9.6 Other Measures of Openness

As we argued throughout the paper the immigration and trade shares are noisy functions of the underlying "openness" parameters ϕ and θ . If those parameters are similar, because the latent true measures of openness to trade and to immigrants coincide for a country, then we can use those two measures together to increase our precision and efficiency in identifying their effect. We do this in Table 10 by using the first principal component of migration and trade shares as an explanatory variable in regressions otherwise similar to those of Table 5. The advantage is that we now use both imputed trade and migration as instruments. Columns 1 and 2 in Table 10 present these 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). We notice that the strength of the instrument, however, has not increased. While combining the measures of trade and immigration to construct a measure of the global openness of a country is promising, it seems to add little to the simple measures taken individually.

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 want to examine whether *emigration* affects income. Openness, after all, would imply more immigration and also more emigration. One may imagine that an outflow of labor may reduce skill diversity and hence average worker productivity so that emigration may have a symmetric and negative effect on income, relative to immigration. However, there are also a number of channels that could compensate for this loss: remittances, enhanced informational networks for trade and

foreign direct investment, return migration after learning and so on. To keep things as simple as possible we add the share of emigrants in the population to our regression model. Columns 3 and 4 report our findings. The point estimates for the immigration and trade shares remain fairly unaffected and the emigration share is never significant—its point estimate is actually positive.⁴⁰ One good feature of our data is that since they are based on censuses of all countries in the world we can construct measures of emigration based on data of destination countries. These are much more precise than the statistics on emigration kept by sending countries. Next, we use the data on emigration to build the net immigration share, defined as stock of immigrants minus stock of emigrants divided by the country’s total population, and include it as our measure of migration openness. Columns 5 and 6 report the results, where we include this measure of net immigration relative to population. The results are similar to what we have obtained previously (e.g. relative to columns 3 and 4). Hence there seems to be a genuine positive effect of openness to immigration: the long-run benefits in income per person accrue to the receiving country, but there are no negative effects for the sending country. These “Pareto-improving” effects of migration in the World as a whole supports the idea that migration moves diverse and productive skills where they would create the most value. Productive countries attract immigrants and their productivity is enhanced by them. Sending countries are deprived of those skills but gain remittances, network and highly educated returnees that help offset the negative effects.

9.7 Effects on Innovation as measured by Patenting

The most important long-run effect immigration may have on a country is to increase its ability to innovate. Diversity in skills and openness in attitudes can be important inputs for creativity and innovation. While there is no good measure of economically relevant innovation, economists consider TFP as the best measure of the technological level of a country. In this section we consider another proxy for innovation, namely the amount of patenting done by inventors residing in the country. From the World Intellectual Property Organization (WIPO) we obtained data on patents granted by any patent office in the world, to inventors residing in 108 countries between 1995 and 2010.⁴¹ We then construct the average yearly number of patent per million inhabitant in each of those countries and we consider this as a measure of innovative ability of a country.⁴² 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 innovative output of a country, region or sector.⁴³ Table 11 shows the estimated effects of immigrant share and trade share on the logarithm of patent per (million)

⁴⁰We are aware that this regressor is likely to be endogenous. Unfortunately, if we try to include emigration as endogenous variable and instrument both immigration and emigration with the gravity instrument we obtain very weak first stage and uninformative estimates due to 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.

⁴¹The data are available at the website <http://www.wipo.int/ipstats/en/statistics/patents/>.

⁴²The 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.

⁴³See for instance the book by Jaffe and Trajtenmberg (2002).

person, using data on 108 countries. We include, alternatively as explanatory variable, *MSH* (column 1), *TSH* (column 2). the first principal component of both (column 3) and in column 4 we control for institutional quality treating it as an endogenous variable. 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 significant effect (significant at the 10% level) of the share of immigrants on innovation either controlling (Column 4) or not (Column 1) for institutional quality, while trade has no effect at all on innovation. The first principal component that combines openness to immigration and trade is also found to have a positive effect on innovation. While the patent measure is certainly a coarse proxy for innovation, this result, in line with the previous ones, confirm the important role of openness to immigration for the long-run idea-creation potential of a country. Furthermore, this empirical result suggests that openness to immigrants may be an even more important element for growth than openness to trade.

10 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. We use new and recently gathered data on the stock of immigrants in all countries in the world as of 2000 to measure openness to migration. To address endogeneity concerns we extend the instrumental-variables strategy, based on geography and relative country size, first used by Frankel and Romer (1999) to estimate the impact of trade openness on income per person. In essence we use geography-based determinants of trade and migration costs as the sources of variation in openness to identify the effect on income.

First, we show that this empirical strategy is successful in generating enough identifying variation, particularly in the case of openness to immigration. Equipped with these instruments we then establish a robust, positive effect of openness to immigration on long-run income per capita using econometric specifications that include a comprehensive set of variables that control for geography, climate, disease environment, and colonial past. In contrast the positive effect of trade openness on income vanishes once these control variables are included in the specification.

Our main finding is robust to explicitly including institutional quality as an (endogenous) regressor, to controlling for measures of early economic development, and to measuring the share of immigrants in terms of efficiency units of labor. We also show that the main effect of migration operates through total factor productivity, as predicted by our simple theoretical model, which postulates that immigration increases the variety of skills available for production. We also provide some more direct evidence of this channel by building an index of the diversity of immigrants by country of origin and find that countries with more human capital benefit more from immigration. We do not find evidence of increased income inequality due to openness to immigration

or trade. However, we find that immigration significantly increases ethnic and linguistic fractionalization. Our results indicate that fractionalization has a negative effect on income per capita, however the gains from greater skill diversity appear to more than offset the costs arising from fractionalization. Finally we also find evidence that countries with a larger share of foreign born have higher innovation per capita, controlling for institutional quality, geography and other historical variables. Trade, in contrast, is not associated with higher innovation rates.

Our analysis does not find any effect of trade openness on income, productivity and inequality, once we control for institutions and geography. While it is possible that openness to immigrants matters much more for income than openness to trade, we believe that trade openness may be poorly measured relative to immigration openness. While trade data are good, they measure total value of imports and exports rather than the value added abroad or domestically. GDP on the other hand is measured as value added. Thus several countries that partake in a lot of processing trade and "transit" trade (Singapore, Hong Kong, Ireland) have deceptively large measures of trade shares (in our data), but may have much smaller trade in value added. Thinking more carefully about the measure of trade openness that is relevant for a country's income and how to correct statistics to measure only import and export value added is needed and may help to better identify these effects in the future.

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 (2001), 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.
- 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.
- 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.

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

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.

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.

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.

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

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

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

- 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 (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.
- 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.
- 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, National Bureau of Economic Research, Inc.
- 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*, Cambridge University Press, vol. 54(04), pages 892-916, December.
- O'Rourke, Kevin H, Taylor, Alan M and Williamson, Jeffrey G, (1996). "Factor Price Convergence in the Late Nineteenth Century," *International Economic Review*, Department of Economics, University of Pennsylvania and Osaka University Institute of Social and Economic Research Association, vol. 37(3), pages 499-530, August.
- 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 (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*, MIT Press, vol. 125(4), pages 1627-1682, November.

- 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*, Volume 15, pages 261-338 National Bureau of Economic Research, Inc.
- Rodrik, Dani (2000). "How Far Will International Economic Integration Go?," *Journal of Economic Perspectives*, American Economic Association, vol. 14(1), pages 177-186, Winter.
- 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 available at <http://gsppi.berkeley.edu/faculty/sraphael/du-bois-review-january-2007.pdf>.
- Sachs, Jeffrey D. Andrew Warner, (1995). "Economic Reform and the Process of Global Integration," *Brookings Papers on Economic Activity*, Economic Studies Program, The Brookings Institution, 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.
- 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	Our calculations
Diversity index Trade flows	168	0.87	0.10	0.39	0.96	Our 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. of European descendants in 1900	153	28.38	40.97	0.00	100	Acemoglu et al (2001)
Share of foreign Ancestors	188	0.24	0.32	0.00	1.00	Putterman- Weil (2010)
Early political development (Statehist)	160	0.48	0.23	0.00	0.96	Putterman- 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 plus one spec.)	188	0.31	0.30	0.00	1.85	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 plus one spec.)	188	0.01	0.04	0.00	0.44	Own calculations

Table 2. Gravity models for bilateral trade shares and foreign-born shares.

Estimation Method	1 OLS	2 FE	3 Poisson PML	4 OLS	5 FE	6 Poisson PML
Dependent Variable	In bilateral TSH	In bilateral TSH	In bilateral TSH	In bilateral MSH	In bilateral MSH	In bilateral MSH
In (distance)	-1.37*** [0.02]	-1.68*** [0.02]	-0.81*** [0.04]	-1.12*** [0.03]	-1.31*** [0.03]	-1.41*** [0.07]
In (pop. Origin)	0.03** [0.01]	-1.04*** [0.11]	-0.21*** [0.03]	-0.40*** [0.02]	-0.68*** [0.17]	-0.30*** [0.04]
In (pop. Dest.)	1.09*** [0.01]	0.59*** [0.06]	0.83*** [0.04]	0.65*** [0.02]	-0.30* [0.17]	0.74*** [0.07]
In (area origin)	-0.09*** [0.01]	0.12** [0.05]	0.04 [0.03]	0.21*** [0.02]	-0.50*** [0.11]	0.15*** [0.04]
In (area dest.)	-0.26*** [0.01]	-0.26*** [0.06]	-0.21*** [0.05]	-0.09*** [0.02]	0.86*** [0.15]	-0.08* [0.05]
sum landlocked	-0.87*** [0.03]	1.43*** [0.26]	-0.64*** [0.07]	-0.26*** [0.05]	-2.95*** [0.42]	-0.67*** [0.14]
Border	-2.13** [0.98]	-7.47*** [0.94]	-1.68 [1.28]	0.49 [0.92]	-1.14 [1.08]	-2.34** [1.18]
Border* In (dist.)	0.32 [0.20]	-0.07 [0.19]	0.17 [0.39]	-0.28 [0.23]	0.05 [0.24]	0.93** [0.36]
Border* In (pop origin)	-0.32*** [0.08]	-0.49*** [0.07]	0.02 [0.09]	-0.21** [0.09]	-0.07 [0.09]	-0.08 [0.11]
Border* In (pop dest.)	-0.34*** [0.08]	-0.54*** [0.07]	-0.29*** [0.10]	-0.26*** [0.09]	-0.34*** [0.09]	-0.58*** [0.12]
Border* In (area origin)	0.06 [0.10]	0.41*** [0.08]	-0.12 [0.13]	-0.06 [0.10]	-0.04 [0.11]	-0.34*** [0.12]
Border* In (area dest.)	0.12 [0.09]	0.45*** [0.08]	0.23 [0.22]	0.32*** [0.10]	0.27*** [0.10]	0.21 [0.15]
Border*landlocked	0.84*** [0.11]	0.80*** [0.11]	0.83*** [0.14]	0.32** [0.13]	0.05 [0.14]	0.50** [0.20]
Common language	0.61*** [0.08]	0.21*** [0.07]	1.00*** [0.26]	0.89*** [0.10]	0.49*** [0.09]	0.85*** [0.19]
Common official lang.	-0.00 [0.08]	0.69*** [0.07]	-0.38 [0.26]	0.43*** [0.10]	0.63*** [0.09]	0.13 [0.19]
Colonial ties	3.07*** [0.12]	0.92*** [0.12]	1.43*** [0.13]	1.25*** [0.17]	0.74*** [0.16]	1.02*** [0.23]
Origin Hegemon	-2.24*** [0.18]	0.02 [0.17]	-1.77*** [0.23]	1.03*** [0.22]	1.32*** [0.21]	0.53* [0.30]
Observations	24,627	24,627	33,108	8,022	8,022	34,782
R-squared	0.39	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. 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)	(6)	(7)	(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*** [0.95]	9.01 [9.79]	3.09** [1.21]	2.45** [0.97]	-0.58 [0.77]	-0.14 [0.75]	-0.40 [0.96]	18.54 [59.02]	0.33 [0.22]
ln (Population)	0.10 [0.11]	0.09 [0.35]	0.10 [0.12]	0.08 [0.11]	0.02 [0.09]	-0.12* [0.08]	-0.20* [0.11]	1.64 [5.78]	-0.13* [0.07]
ln (Area)	0.11 [0.15]	0.86 [1.28]	0.11 [0.15]	0.12 [0.16]	-0.25** [0.12]	-0.03 [0.09]	0.03 [0.09]	0.98 [2.69]	0.07 [0.07]
Dist. equator					0.05*** [0.01]		0.03** [0.01]	-0.11 [0.43]	0.02*** [0.01]
Pct. Land tropics							0.45 [0.49]	-4.97 [17.02]	
Observations	146	146	181	146	146	146	122	122	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	0.07	.
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). 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). 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, FR sample	(2) Basic, Full sample	(3) Income per Worker	(4) Includes Distance from equator	(5) With Region and geographic controls	(6) With all controls	(7) Using FE Predictor With all controls	(8) OLS	(9) Includes only MSH
Dep. Var.	ln (GDP/Pop)	ln (GDP/Pop)	ln (GDP/Emp)	ln (GDP/Pop)	ln (GDP/Pop)	ln (GDP/Pop)	ln (GDP/Pop)	ln (GDP/Pop)	ln (GDP/Pop)
TSH	2.31** [0.91]	2.91** [1.22]	2.23** [0.94]	-0.50 [0.63]	-0.33 [0.62]	-0.08 [0.85]	-0.75 [1.65]	0.16 [0.19]	
MSH	5.89** [2.98]	6.54* [3.88]	6.01** [2.93]	6.40*** [1.86]	7.03*** [2.17]	5.91*** [1.73]	5.40** [2.41]	6.31*** [1.04]	6.34*** [1.15]
ln (Population)	0.12 [0.11]	0.10 [0.12]	0.10 [0.11]	0.05 [0.07]	-0.07 [0.07]	-0.08 [0.11]	-0.16 [0.20]	-0.06 [0.06]	-0.07 [0.06]
ln (Area)	0.11 [0.16]	0.14 [0.16]	0.12 [0.16]	-0.22** [0.09]	0.11 [0.07]	0.11 [0.08]	0.07 [0.12]	0.12** [0.06]	0.11** [0.05]
Dist. equator				0.05*** [0.01]	0.03*** [0.01]	0.03*** [0.01]	0.02 [0.01]	0.01 [0.01]	0.01 [0.01]
Observations	146	181	146	146	123	122	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 Origin	no	no	no	no	no	yes	yes	yes	yes
First stage reg.									
Wald F test	0.09	1.47	0.09	0.21	2.79	2.64	0.84	.	62.09
exclusion									
Instruments	pred. TSH pred. MSH	pred. TSH pred. MSH	pred. TSH pred. MSH	pred. TSH pred. MSH	pred. TSH pred. MSH	pred. TSH pred. MSH	pred. TSH pred. MSH FE	.	pred. MSH

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). 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). 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.

Specification	(1) Main	(2) FE predictor.	(3) Exogenous Determinant.	(4) As .3 with FE pred.	(5) Alcala- Ciccone	(6) AC plus controls	(7) TSH, MSH, IQ	(8) Putterman- Weil	(9) Human capital
Dependent variable: ln (GDP/Pop)	ln (GDP/Pop)	ln (GDP/Pop)	ln (GDP/Pop)	ln (GDP/Pop)	ln (GDP/Emp)	ln (GDP/Emp)	ln (GDP/Pop)	ln (GDP/Pop)	ln (GDP/Pop)
MSH	9.04*** [2.04]	6.30*** [1.43]	6.78*** [2.01]	5.76*** [1.12]			8.75*** [1.92]	10.50*** [2.72]	4.85*** [1.17]
Institutional Quality	0.44*** [0.10]	0.46*** [0.10]			0.26* [0.15]	0.39** [0.19]	0.44*** [0.10]	0.50*** [0.12]	0.49*** [0.11]
ln (Real TSH)					1.97*** [0.70]	-0.16 [0.93]			
TSH							0.11 [0.57]		
Dist. Equator			0.01 [0.01]	0.01 [0.01]					
Share of Euro. descent 1975			0.01*** [0.00]	0.01*** [0.00]					
PW Statehist								-0.11 [0.46]	
PW Share of foreign ancestors								-0.62 [0.50]	
Observations	120	107	119	122	128	120	120	117	120
Region dummies	yes	yes	yes	yes	no	yes	yes	yes	yes
Geo/Climate/Disease	yes	yes	yes	yes	no	yes	yes	yes	yes
First-stage reg. Wald F test exclusion	7.79	7.47	8.05	52.04	2.06	0.99	3.44	6.13	7.24
Instruments	pred. MSH	FE pred. MSH	pred. MSH	FE pred. MSH	ln pred. TSH	ln pred. TSH	pred. MSH pred. TSH	pred. MSH	pred. MSH
	EU descent distance equator	EU descent distance equator			EU descent distance equator	EU descent distance equator	EU descent distance equator	EU descent distance equator	EU descent distance equator
SY 10% max. IV	13.43	13.43	13.43	13.43	13.43	13.43	n.a.	13.43	13.43
SY 25% max. IV	5.45	5.45	5.45	5.45	5.45	5.45	n.a.	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 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). 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. Two-stage least-squares.

Dep. Var.	Migration share predicted using FE gravity				Migration predicted using simple gravity			
	(1) ln (Y/L)	(2) $(\alpha/1-\alpha)*\ln (K/Y)$	(3) ln (H/L)	(4) ln (TFP)	(5) ln (Y/L)	(6) $(\alpha/1-\alpha)*\ln (K/Y)$	(7) ln (H/L)	(8) ln (TFP)
MSH (Pop)	4.78** [1.97]	0.47 [0.47]	0.18 [0.46]	4.13*** [1.50]	7.68*** [2.23]	0.94 [0.90]	1.43* [0.77]	5.31** [2.15]
Ln(Population)	0.02 [0.06]	-0.07*** [0.02]	-0.00 [0.02]	0.09* [0.05]	0.05 [0.05]	-0.07*** [0.02]	0.01 [0.02]	0.10** [0.05]
Ln(area)	0.04 [0.05]	0.02 [0.02]	0.02 [0.01]	-0.00 [0.05]	0.06 [0.06]	0.03 [0.02]	0.03 [0.02]	0.01 [0.05]
Observations	98	98	98	98	98	98	98	98
Regional dummies	yes	yes	yes	yes	yes	yes	yes	yes
Geo/Climate/Disease	yes	yes	yes	yes	yes	yes	yes	yes
Colonial ties	yes	yes	yes	yes	yes	yes	yes	yes
First-stage reg. Wald F test exclusion Instrument:	70.34	70.34	70.34	70.34	9.057	9.057	9.057	9.057
	FE predicted MSH	FE predicted MSH	FE predicted MSH	FE predicted MSH	predicted MSH	predicted MSH	predicted MSH	predicted MSH
SY 10% max. IV	16.38	16.38	16.38	16.38	16.38	16.38	16.38	16.38
SY 25% max. IV	5.53	5.53	5.53	5.53	5.53	5.53	5.53	5.53

Not: 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 $\alpha=0.33$. All regression models include an intercept, regional dummies, and the set of geographic, climate and disease controls defined earlier. The instrument is the predicted migration share (based on the fixed-effects gravity model, where the estimated fixed effects have not been used in the prediction of the NL gravity model). In parenthesis we report the heteroskedasticity-robust standard errors. *, **, *** significant at 10%, 5% and 1% confidence level.

Table 7: Immigrant Diversity and Linguistic Fractionalization. Two-stage least-squares.

Specification:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	9
Dependent Variable	ln (GDP/Pop)	ln (GDP/Pop)	ln (GDP/Pop)	ln (GDP/Pop)	Ethnic Fractionalization	Linguistic Fractionalization	ln (GDP/Pop)	ln (GDP/Pop)	ln (GDP/Pop)
TSH	-0.08 [0.85]	-0.20 [1.03]	-0.22 [1.04]	-0.22 [0.75]	0.03 [0.22]	-0.07 [0.19]			
MSH	5.91*** [1.73]	3.11 [2.32]	3.18 [2.14]	5.10*** [1.60]	1.66*** [0.50]	1.22* [0.69]	10.01*** [2.19]	7.03*** [1.11]	7.22*** [1.00]
Diversity M.		1.54*** [0.45]	1.50*** [0.43]						
Diversity T.			-1.40 [0.86]						
Predicted Diversity M.				1.25** [0.56]					
Predicted Diversity T.				-1.97*** [0.41]					
Ethnic Frac.							0.42	0.53	
Ling. Frac.							-0.62**	-0.45*	
Catholic sh.							0.01*	0.01**	0.01**
Muslim sh.							-0.01**	0.00	0.00
Protestant sh.							0.01***	0.01***	0.01***
Observations	122	120	120	120	121	118	117	116	121
Instrument:	Predicted TSH; FE predicted MSH	Predicted TSH; FE predicted MSH	Predicted TSH; FE predicted MSH	Predicted TSH; FE predicted MSH	Predicted TSH; FE predicted MSH	Predicted TSH; FE predicted MSH	Predicted MSH	FE predicted MSH	FE predicted MSH

Note: The predicted TSH is based on the linear in logs gravity estimates and the predicted MSH is based on the fixed-effects gravity model (not using the estimated fixed effects in the prediction). Predicted values for the TSH and the Diversity 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). 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

Dep. Var.	(1) Gini Coefficient	(2) Gini Coefficient	(3) Gini Coefficient	(4) Gini Coefficient
TSH		-0.10 [0.09]		-0.10 [0.08]
MSH	0.23 [0.44]	0.05 [0.46]	-0.32 [0.42]	0.22 [0.50]
ln (Population)	-0.00 [0.01]	-0.01 [0.01]	-0.00 [0.01]	-0.01 [0.01]
ln (Area)	0.01 [0.01]	-0.00 [0.01]	-0.00 [0.01]	-0.00 [0.01]
Distance from Equator	-0.00** [0.00]	-0.00 [0.00]	-0.00* [0.00]	-0.00 [0.00]
Observations	115	105	105	115
Controls	yes	yes	yes	yes
Region	yes	yes	yes	yes
Geo/Climate/Disease	yes	yes	yes	yes
Colonial Origin	yes	yes	yes	yes
First stage Wald Statistics	3.98	2.32	11.13	3.67
Instrument:	FE Predicted MSH	FE Predicted MSH Predicted TSH	Predicted MSH	predicted MSH, Predicted TSH
SY 10% max. IV	26.38	7.03	16.38	7.03
SY 25% max. IV	5.53	3.63	5.53	3.63

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 young, rich countries (US, Canada, Australia and New Zealand). In parenthesis we report the heteroskedasticity-robust standard errors. *, **, *** = significant at 10%, 5% and 1% confidence level.

Table 9: Heterogeneity of the Effects of Immigration, 2SLS

	By Development in year 1500		By trade openness		By Current Human Capital	
Dep. Var.: ln GDP/Pop	(1) 2SLS, Predicted MSH	(2) 2SLS, FE Predicted MSH	(3) 2SLS, Predicted MSH	(4) 2SLS, FE Predicted MSH	(5) 2SLS, Predicted MSH	(6) 2SLS, FE Predicted MSH
MSH*High	7.38** [3.17]	7.01** [3.47]	7.39*** [1.78]	6.26*** [1.16]	8.12*** [1.79]	6.76*** [1.02]
MSH*Low	7.42*** [1.66]	6.34*** [1.16]	7.55** [2.96]	4.51 [3.21]	3.57 [4.66]	1.41 [2.46]
Observations	121	121	121	121	121	121
Controls						
Region dummies	yes	yes	yes	yes	yes	yes
Geo/Climate/Disease	yes	yes	yes	yes	yes	yes
Colonial Ties	yes	yes	yes	yes	yes	yes
First-stage regression						
Wald F test exclusion	5.51	6.16	1.63	5.64	2.65	14.34
Instruments	pred. MSH*High	FE pred. MSH*High	pred. MSH*High	FE pred. MSH*High	pred. MSH*High	FE pred. MSH*High
	pred. MSH*Low	FE pred. MSH*Low	pred. MSH*Low	FE pred. MSH*Low	pred. MSH*Low	FE pred. MSH*Low

NOTE: The dependent variable is log of income per person. Estimation method is 2SLS, the instruments are the (fixed-effects) 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 and logarithm of area plus 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 4 young, rich countries (US, Canada, Australia and New Zealand). 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)
First Principal Component of TSH and MSH	0.91*** [0.23]	0.85*** [0.27]				
Institutional Quality		0.42*** [0.10]		0.46*** [0.10]		0.54*** [0.12]
MSH			5.46** [2.53]	6.24*** [1.75]		
TSH			-0.74 [1.66]	-0.09 [0.71]	-1.06 [1.64]	0.43 [0.85]
Emigration/Pop			0.21 [0.85]	0.42 [0.74]		
Net Immigration/Pop					4.11** [2.04]	5.99*** [1.69]
Dist. Equator	0.01 [0.01]		0.02 [0.01]		0.02 [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 ties	yes	no	yes	no	yes	no
First stage						
Wald F test	11.34	6.81	0.82	1.68	0.97	2.13

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 fixed-effects 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 4 young rich countries (US, Canada, Australia and New Zealand). In parenthesis we report the heteroskedasticity-robust standard errors. *, **, ***= significant at 10%, 5% and 1% confidence level.

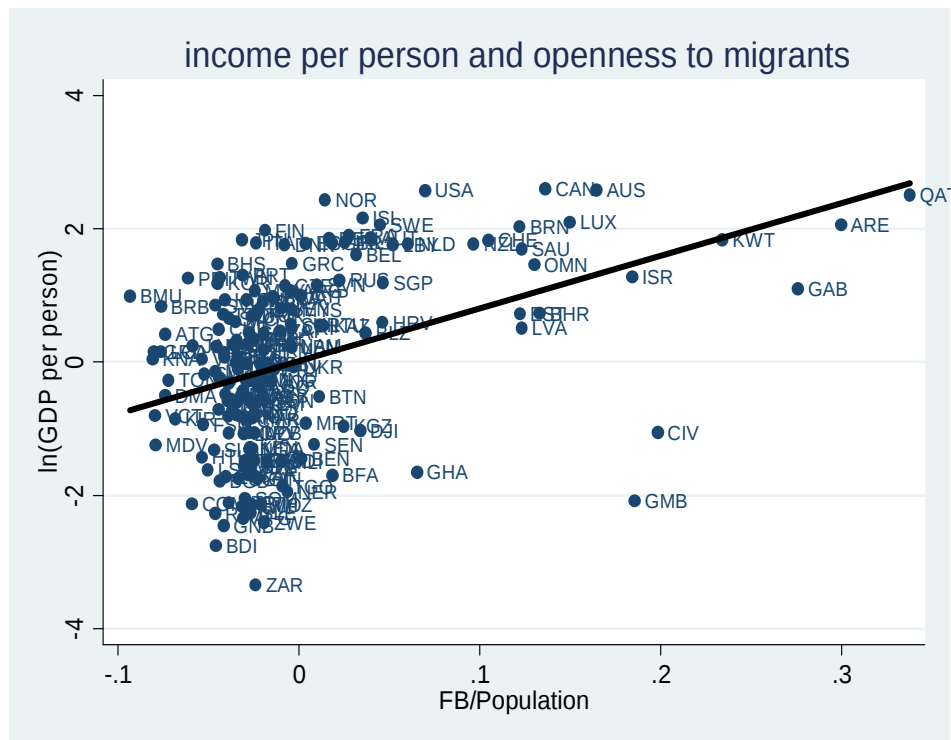
Table 11: Effects of Openness on Patents

Dep. Var	(1) ln (patent/Population)	(2) ln (patent/Population)	(3) ln (patent/Population)	(4) ln (patent/Population)
MSH	11.67* [6.34]			10.05* [5.17]
TSH		1.32 [1.57]		
Principal Component Openness			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 EU descent in 1975 distance equator

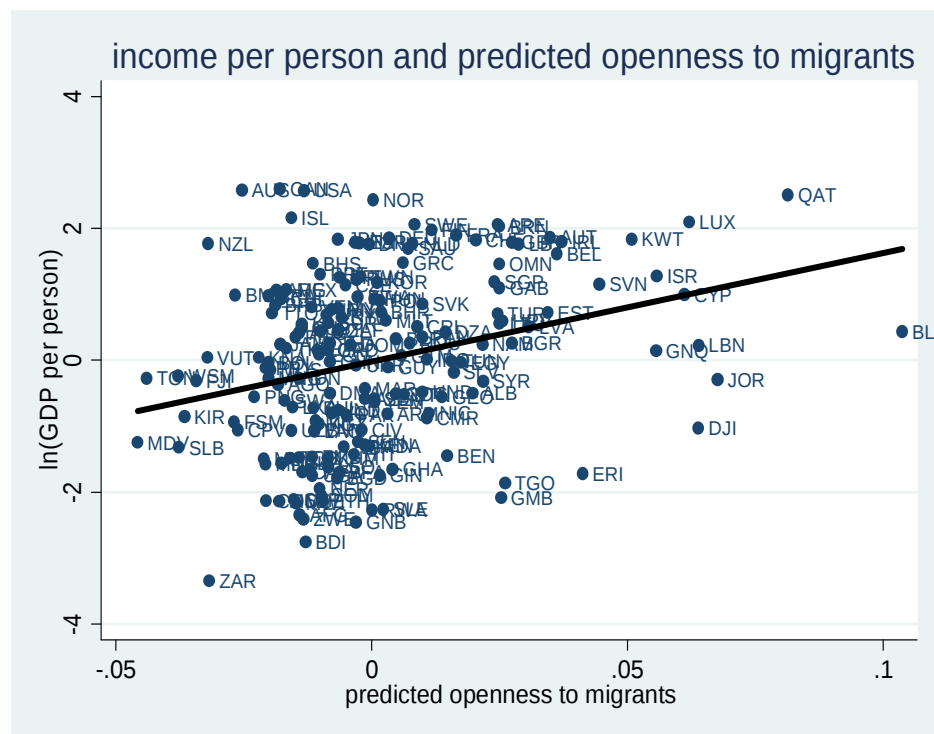
NOTES: The dependent variable is the logarithm of the average number of yearly patents (1995-2010) granted to applicants residing in the country by any patent office in the world, per million inhabitants. All regressions include an intercept, log population, and log area. **Regional Dummies** for sub-Saharan 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 4 young rich countries (US, Canada, Australia and New Zealand). In parenthesis we report the heteroskedasticity-robust standard errors. *, **, ***= significant at 10%, 5% and 1% confidence level.

Figure 2: Openness to immigrants and GDP per person, adjusted for country size

2A: Foreign-born share and GDP per person



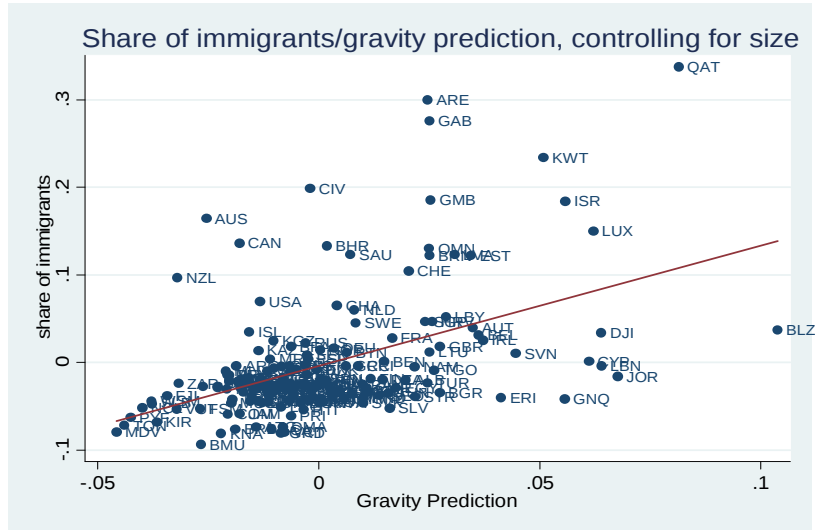
2B: Gravity-predicted foreign-born share and GDP per person



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

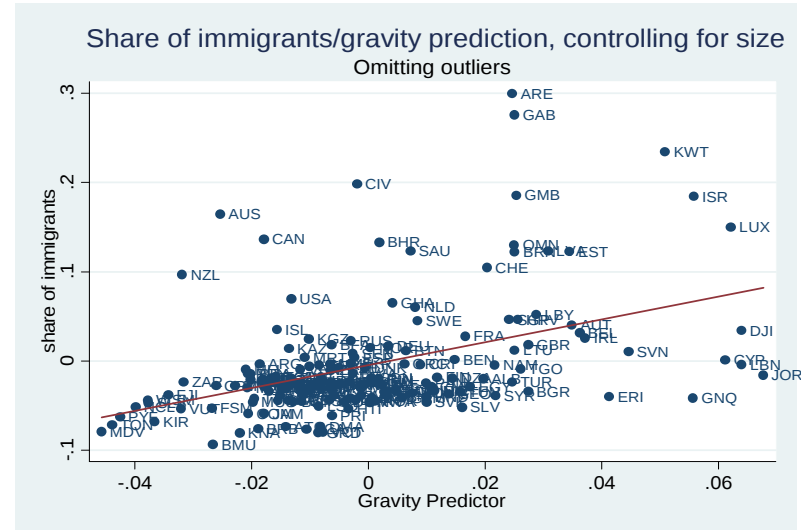
Figure 3

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



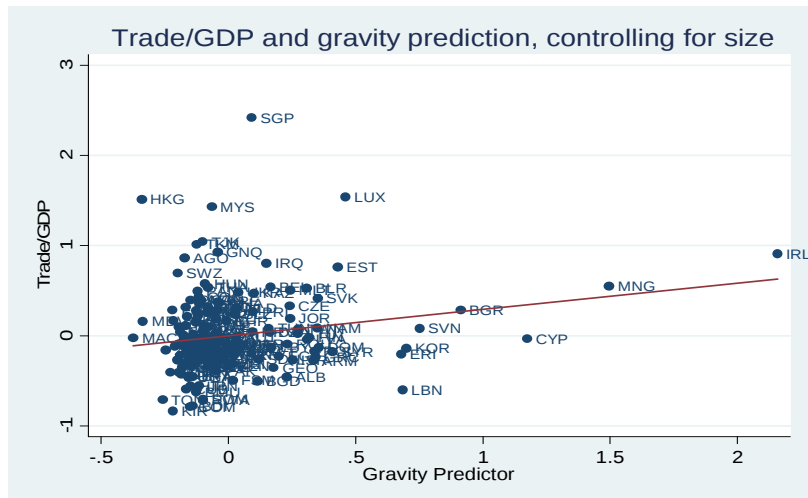
Slope: 1.37, standard error: 0.30 F-stat: 20.56

3B: Excluding 2 outliers



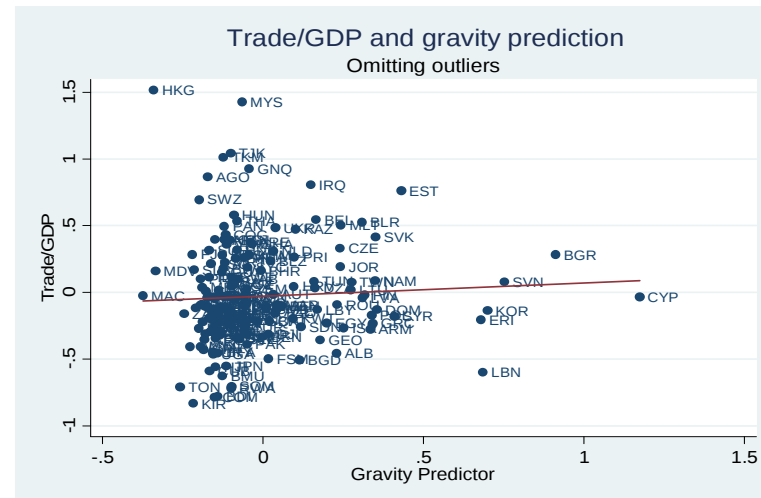
Slope: 1.28, standard error: 0.26 F-stat: 22.90

3C: Fit of the predicted trade share, controlling for 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 (1) ln (income/Po pulation)	Effect on Income (2) ln (income/Po pulation)	Effect on Institutions (3) Institutional Quality	Effect on Institutions (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]
Putterman-Weil "Share foreign. 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.