



International Migration: A Panel Data Analysis of the Determinants of Bilateral Flows

Anna Maria Mayda

I INTRODUCTION

International migration patterns vary considerably over time, and across destination and origin countries. Some OECD countries have experienced a decrease in the size of the annual immigrant inflow between 1980 and 1995. Over the same years, the number of immigrants per year has increased in several other OECD countries.¹ The percentage change of the annual immigrant inflow from 1980 to 1995 ranges between negative 42% (in Japan) and positive 48% (in Canada). For all destinations, such changes

¹There has been a decrease in France, Japan, Netherlands, and the United Kingdom. There has been an increase in Belgium, Canada, Germany, Luxembourg, Norway, Switzerland, and the United States (OECD, 1997).

A. M. Mayda (✉)

Department of Economics and School of Foreign Service, Georgetown
University, Washington, DC, USA
e-mail: amm223@georgetown.edu

© The Author(s), under exclusive license to Springer Nature
Switzerland AG 2022

P. Goulart et al. (eds.), *Global Labour in Distress, Volume I*, Palgrave
Readers in Economics,

https://doi.org/10.1007/978-3-030-89258-6_3

are anything but monotonic (OECD, 1997). The variation in terms of origin countries is remarkable as well (OECD, 1997).

Several factors are likely to influence the size, origin, and destination of labor movements at each point in time and contribute to the variation observed in the data. However, very few empirical works in the literature have tried to understand what drives international migration, perhaps due to past unavailability of cross-country data.

In turn, international migration has recently received a great deal of attention in light of research showing its beneficial effects from an economic-development point of view. For example, the recent literature has pointed out repeatedly the potential of free migration to produce large benefits—most likely greater than the gains from liberalizing existing trade barriers (Rodrik, 2002). To fully understand these and other effects, it is important to identify the forces and constraints that shape international migration movements.

In this paper, I empirically investigate the determinants—economic, geographic, cultural, and demographic—of bilateral immigration flows. My analysis is based on the predictions of a simple theoretical framework that focuses on both supply and demand factors. I use yearly data on immigrant inflows into fourteen OECD countries by country of origin, between 1980 and 1995. The source of this data is the International Migration Statistics for OECD countries (OECD, 1997), based on the OECD's Continuous Reporting System on Migration (SOPEMI).

My paper is related to a vast literature on the determinants of migration. Clark et al. (2007) and Karemera et al. (2000) both focus on the fundamentals explaining immigrant inflows into the United States by country of origin in the last decades. Other papers in the literature that analyze the determinants of migration to the U.S. are Borjas (1987) and Borjas and Bratsberg (1996). Hatton (2005) investigates trends in UK net migration in the last decades. Finally, Helliwell (1998) sheds light on factors affecting labor movements in his investigation of the magnitude of immigration border effects, using data on Canadian interprovincial, US interstate, and US-Canada cross-border immigration.

This paper makes three contributions to the literature. First, my analysis puts greater emphasis than previous works on the demand side of international migration, namely destination countries' migration policies. This change of perspective is important, given restrictive immigration policies in the vast majority of host countries. Second, my work is the first one I am aware of to use the OECD (1997) data on international migration to

systematically investigate the drivers of international flows of migrants. Previous works have either used country cross-sections (Borjas 1987, Yang 1995), or have focused on a single destination country over time (Borjas & Bratsberg, 1996; Brücker et al., 2003; Clark et al., 2007; Karemera et al., 2000) or a single origin country over time (Yang, 2003). By extending the focus of the analysis to a multitude of origin and destination countries and taking advantage of both the time-series and cross-country variation in the data, I can test the robustness and broader validity of the results found in earlier works.² Third, this paper carefully reviews and proposes solutions to various econometric issues that arise in the estimation, such as endogeneity and reverse causality. These econometric complications have not all been addressed in the previous literature.³ Once I deal with them (e.g., by controlling for destination and origin countries' fixed effects and for year effects), my analysis both delivers estimates broadly consistent with the predictions of the international migration model and generates empirical puzzles.

According to the international migration model, pull and push factors have either similar-sized effects (with opposite signs), when migration quotas are not binding, or they both have no (or a small) effect on emigration rates, when migration quotas are binding. It is not clear, *ex ante*, which one of the two scenarios characterizes actual flows. Migration policies in the majority of destination countries are very restrictive, which should imply binding constraints on the number of migrants. On the other hand, even countries with binding official immigration quotas often accept unwanted (legal) immigration.⁴ Restrictive immigration policies are often characterized by loopholes, that leave room for potential migrants to take advantage of economic incentives. For example, immigration to Western European countries still took place after the late 1970s, despite the official closed-door policy. Family-reunification and asylum-seekers policies can explain continuing migration inflows to Western Europe (Joppke, 1998).

²Since I began working on this paper, I have become aware of other related, but independent papers analyzing cross-country migration patterns: Alvarez-Plata et al. (2003); Pedersen et al. (2004) Pedersen et al. (2006). I discuss these very recent contributions to the literature below, in relation to the data I use and results I find.

³See also Alvarez-Plata et al. (2003) for an excellent discussion of the properties of different estimators of the determinants of migration flows.

⁴Notice that the data set I use only covers legal migration.

My empirical results are puzzling because they are in part consistent with the first scenario and in part with the second one. I find that pull factors—proxied by the per worker GDP in the destination country—significantly increase the size of emigration rates. This result is robust to changes in the specification of the empirical model. Both absolute and relative pull factors matter. That is, the emigration rate to a given destination is an increasing function of that country's per worker GDP and a decreasing function of the average per worker GDP of all the other host countries in the sample⁵ (each weighted by the inverse of distance from the origin country). On the other hand, the impact of push factors—proxied by the per worker GDP in the origin country—is seldom negative as theory suggests would be the case with not-binding migration quotas and, when it is, the size of the effect is smaller than for pull factors and insignificant. Therefore my analysis finds evidence of an asymmetric impact of pull and push factors on emigration rates.⁶

The asymmetry is a familiar puzzle. For example, it has been documented in several works in the literature on internal migration (see, e.g., Hunt (2006) and the papers referenced in its footnote 4). Based on the existing literature, there might be numerous reasons for the asymmetry and possibly different ones operating across borders versus within country borders. At the national level, where migration quotas do not exist, Hunt (2006) provides an explanation of the asymmetry by breaking down data by age group: Origin region's unemployment rates (push factor) have an insignificant impact on migration flows because the insignificant effect for the young—who are not as sensitive to their own layoffs as the old—dominates the significant positive effect for the old. This explanation cannot be investigated at the international level because of data unavailability.

Another interpretation in the literature of the asymmetry is that migration quotas are effectively not binding but the impact of income

⁵ Since the host countries in the sample receive a large fraction of immigrants in the world, it is not overly restrictive to focus on them. For example, according to the United Nations (2004), the list of leading host countries of international migrants in 2000—as measured by the percentage of the world's migrant stock in each of these countries—includes the United States (20%), Germany (4.2%), France (3.6%), Canada (3.3%), Australia (2.7%), United Kingdom (2.3%), Switzerland (1%), Japan (0.9%), and the Netherlands (0.9%) (see Table ii.3, p.30). These countries all belong to my sample.

⁶ By asymmetry between pull and push factors, I mean that the coefficient of economic conditions in the source region does not have the expected sign, while the coefficient of economic conditions in the destination region is, as expected, positive and significant.

opportunities in the origin country is affected by poverty constraints, due to fixed costs of migration and credit-market imperfections (Lopez & Schiff, 1998; Yang, 2003). Since lower levels of per worker GDP in the source country both strengthen incentives to leave and make it more difficult to overcome poverty constraints, the net effect might be close to zero. In the empirical analysis I investigate this possibility and I find very weak evidence that my result on push factors is driven by poverty constraints in the origin country.

Yet an alternative explanation of my findings is that the asymmetric effect I estimate for pull and push factors is explained by the demand side of international migration—namely, migration policies—and not by the supply side as is often assumed in the previous literature. Changes in mean income opportunities in the destination country not only affect migrants' incentive to move there but also impact the political process behind the formation of migration policies. For example, in periods of economic booms, policymakers are better able to overcome political opposition to and accommodate increasing migration inflows.⁷ *If migration quotas are binding*, the latter political-economy channel will be at work while the determinants on the supply side will have no (or a small) impact. This would explain the asymmetric effect I estimate for pull and push factors. While I do not investigate this interpretation directly,⁸ I find evidence which is consistent with migration policy playing a constraining role. In the empirical analysis, I differentiate the effect of pull and push factors according to changes in destination countries' migration policy. I find that the effect of pull factors becomes more positive and the impact of push factors turns negative in those years when a host country's immigration laws become less restrictive. This is also true for the impact of other supply-side determinants such as geography and demographics (see below). In sum, my results suggest that migration quotas matter as they mitigate supply-side effects.⁹

⁷Hanson and Spilimbergo (2001) focus on US border enforcement and show that enforcement softens when the sectors that use illegal immigrants expand, which is evidence that migration policy is affected by changes in economic conditions in the destination country.

⁸This interpretation goes beyond the theoretical model in this paper, which assumes exogenous migration quotas. The empirical analysis of the endogenous determination of migration policy and its role in explaining the asymmetric effect of pull and push factors is outside the scope of this paper.

⁹This result is consistent with the findings in Hatton (2004) where emigration from Britain in the era of free migration (before 1914) is compared to emigration in 1950

My empirical analysis also finds that inequality in the source and host economies is related to the size of emigration rates as predicted by Borjas (1987) selection model. An increase in the origin country's relative inequality has a non-monotonic effect on the size of the emigration rate: The impact is estimated to be positive if there is positive selection, negative if there is negative selection. Among the variables affecting the costs of migration, distance between destination and origin countries appears to be the most important one: Its effect is negative, significant, and steady across specifications. On the other hand, there is no evidence that cultural variables related to each country pair play a significant role. Demographics—in particular, the share of the origin country's population who is young—shape bilateral flows as predicted by the theory. Since the effect of geography and demographics works through the supply side of the model, their impact should be even stronger when migration quotas are relaxed, which is what I find in the data.

Finally, I empirically investigate the importance of network effects. Since immigrants are likely to receive support from other immigrants from the same origin country already established in the host country, they will have an incentive to choose destinations with larger communities of fellow citizens. Network effects imply that bilateral migration flows are highly correlated over time, which is what the data shows. However, it is not clear how to interpret this result. While it is consistent with supply factors (i.e., network effects), it could also be driven by demand factors (family reunification policies, for example).

The rest of this paper is organized as follows. Section 2 presents a simple model of international migration. In Sect. 3 I describe the data sets used, while in Sect. 4 I discuss the estimating equations and some econometric issues that complicate the analysis. Finally, I present the main empirical results and additional results in, respectively, Sects. 5 and 6. Section 7 concludes.

onwards, when immigration policies were in place in the four main host countries of British migrants. The paper finds that, from the mid-1960s, the impact of economic and demographic forces “became less powerful as they were increasingly inhibited by immigration policies in the principal destination countries.” (p.1).

2 THEORETICAL FRAMEWORK

Both supply and demand factors affect international migration flows. Migrants' decisions to move, according to economic and non-economic incentives, shape the supply side of labor movements. The host country's immigration policy represents the demand side, namely the demand for immigrants in the destination country. The theoretical framework in this paper is closely related to the previous literature (Borjas, 1999; Clark et al., 2007), the main difference being the greater emphasis in my model on destination countries' immigration policy. I consider two countries: country 0, which is the origin of immigrant flows and country 1, which is the destination. I first focus on the supply side of immigration and look at the probability that an individual chosen randomly from the population of country 0 will migrate to country 1. In each country, wages are a function of the individual skill level (s_i). The wages that individual i receives in country 0 and would receive if he migrated to country 1 are respectively equal to $w_{0i} = \alpha_0 + \theta_0 \cdot s_i + \epsilon_{0i}$ and $w_{1i} = \alpha_1 + \theta_1 \cdot s_i + \epsilon_{1i}$, where the two disturbances have zero means over the origin country's population. In light of the empirical analysis below, based on aggregate data, it is helpful to rewrite individual i 's wages in the two locations as a function of first and second moments of the income distributions (of the origin country's population) at home and abroad respectively:

$$w_{0i} = \mu_0 + v_{0i}, \text{ where } v_{0i} \sim N(0, \sigma_0^2), \quad (1)$$

$$w_{1i} = \mu_1^0 + v_{1i}, \text{ where } v_{1i} \sim N(0, \sigma_1^2), \quad (2)$$

where the correlation coefficient between v_{0i} and v_{1i} equals ρ_{01} , μ_0 equals $\alpha_0 + \theta_0 \cdot \bar{s}_0$ and μ_1^0 equals $\alpha_1 + \theta_1 \cdot \bar{s}_0$ (\bar{s}_0 is the mean skill level of the origin country's population).

Notice that μ_1^0 , which is equal to the mean wage of the origin country's population if it all migrated to country 1, is different from $\mu_1 = \alpha_1 + \theta_1 \cdot \bar{s}_1$, which is equal to the mean wage of the destination country's population in country 1 (\bar{s}_1 represents the mean skill level of the destination country's population). This point will be relevant in one of the robustness checks in the empirical analysis.

I assume that each individual has Cobb Douglas preferences for the two goods produced in the world (x_A and x_B), which implies an indirect utility (function) from having an income y given by $v(p_A, p_B; y) = \bar{A}(p_A, p_B) \cdot y$. I assume that each country is a small open economy characterized by free trade with the rest of the world:¹⁰ therefore goods' prices p_A and p_B , as well as $\bar{A}(p_A, p_B)$, are given and equal across countries.¹¹ An individual in country 0 will migrate to country 1 if the utility of moving is greater than the utility of staying at home that is, given the assumptions above, if the expected income in country 1 net of migration costs is greater than the expected income in country 0. Following the literature, I can define an index I_i that measures the net benefit of moving relative to staying at home for a risk-neutral individual i :

$$I_i = \eta_{01} \cdot w_{1i} - C_i - w_{0i}, \quad (3)$$

where η_{01} is the probability that the migrant from country 0 will be allowed to stay in country 1, and $C_i = \mu_C + v_i^C$, with $v_i^C \sim N(0, \sigma_C^2)$, represents the level of individual migration costs.¹² The correlation coefficients between v_i^C and (v_{0i}, v_{1i}) are equal to (ρ_{0C}, ρ_{1C}) . The implicit assumption in (3) is that, if the migrant moves to but is not allowed to stay in the destination country, he still incurs the migration costs C_i and gives up the home wage w_{0i} . In other words, the individual migrates to the host country before knowing whether he will be able to stay (for a longer period of time) and gain the income w_{1i} .¹³ Immigrants may not be able to stay in the host country because of quotas due to a restrictive immigration policy.

The probability that an individual chosen randomly from the population of the origin country will migrate from country 0 to country 1 therefore equals:

¹⁰ Given free trade, what explains the difference in rates of return to labor across countries? The answer is that, besides free trade, the other conditions for factor-price-equalization are not satisfied: for example, if international productivity differences exist, then only *adjusted* factor-price-equalization holds.

¹¹ In the empirical analysis I adjust for international differences in goods' prices, using PPP income levels.

¹² I assume that each individual knows the wage levels w_{1i} and w_{0i} he would get in each location, the migration costs C_i , and the probability η_{01} .

¹³ This assumption is consistent with the evidence that immigrants often arrive to a destination country with temporary tourist or student visas with the hope of being able to stay.

$$P = \Pr[I_i > 0] = \Pr[\eta_{01} \times (\mu_1^0 + v_{1i}) - (\mu_c + v_i^c) - (\mu_0 + v_{0i}) > 0], \quad (4)$$

which can be rewritten as $P = 1 - \Phi(z)$, where $z = -\frac{(\eta_{01} \cdot \mu_1^0 - \mu_0 - \mu_c)}{\sigma_v}$, σ_v is the standard deviation of $(\eta_{01} \cdot v_{1i} - v_{0i} - v_i^c)$, and $\Phi(\cdot)$ is the cumulative distribution function of a standard normal. The probability in (4) is the *supply* emigration rate $\frac{I_{01}^S}{P_0}$, where I_{01}^S represents the size of the migration flow as determined by the supply side of the model and P_0 the population in the origin country.

Next, I assume that the destination country's immigration policy sets quantity constraints for immigrants coming from each origin country. Let I_{01}^D be the maximum number of migrants from country 0 allowed each year into country 1. These immigration quotas, which represent country 1's demand for immigrants from country 0, may or may not be binding. Only in the latter case does the emigration rate we observe in the data ($\frac{I_{01}}{P_0}$) equals the *supply* emigration rate $\frac{I_{01}^S}{P_0}$ defined above. On the other hand, if quantity constraints are binding, $\frac{I_{01}}{P_0}$ will be less than $\frac{I_{01}^S}{P_0}$. In

general, the emigration rate we observe in the data is equal to the minimum of $\frac{I_{01}^S}{P_0}$ and $\frac{I_{01}^D}{P_0}$, and is represented in Fig. 1 by the heavy lines, as a

function of μ_1^0 , μ_0 and μ_c . The figure assumes that quotas I_{01}^D are exogenous, which means that they are not affected by μ_1^0 nor by μ_0 nor by μ_c . This is a strong assumption that is questioned in the interpretation of the empirical results.

I assume that the probability η_{01} that the migrant from country 0 will be allowed to stay in country 1 is equal to $\min\{1, \frac{I_{01}^D}{I_{01}^S}\}$. It is then possible to derive testable predictions for the impact of μ_1^0 , μ_0 , and μ_c on the emigration rate from country 0 to country 1:

$$\frac{d(\frac{I_{01}}{P_0})}{d\mu_1^0} = \begin{cases} \frac{\phi(z)}{\sigma_v} > 0, & \text{if } \frac{I_{01}^S}{P_0} < \frac{I_{01}^D}{P_0}; \\ 0, & \text{if } \frac{I_{01}^S}{P_0} \geq \frac{I_{01}^D}{P_0} \end{cases} \quad (5)$$

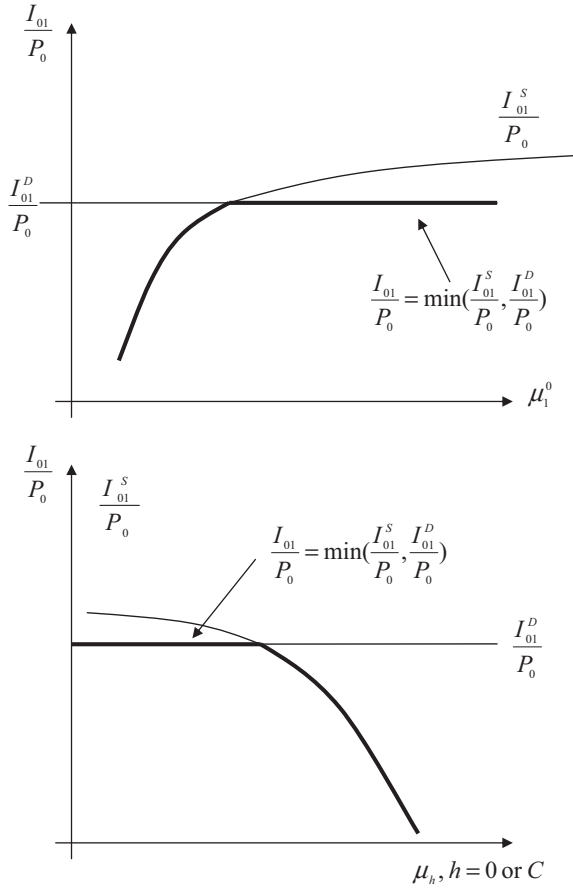


Fig. 1 The actual emigration rate as a function of mean income opportunities in the destination and origin country and of mean moving costs

$$\frac{d\left(\frac{I_{01}}{P_0}\right)}{d\mu_h} = \begin{cases} -\frac{\phi(z)}{\sigma_v} < 0, & \text{if } \frac{I_{01}^S}{P_0} \leq \frac{I_{01}^D}{P_0}; \\ 0, & \text{if } \frac{I_{01}^S}{P_0} > \frac{I_{01}^D}{P_0} \end{cases} \quad (6)$$

where $\phi(\cdot)$ is the density function of a standard normal and $h=0$, C. According to (5) pull effects (namely, improvements in the mean income opportunities in the destination country) are positive and strongest when restrictions are not binding neither ex-ante nor ex-post, they are positive but smaller in size when the quota is binding ex-post but not ex-ante and, finally, they are equal to zero in a quantity-constrained world. A parallel interpretation explains the comparative-static results in (6), which describes push effects (changes of μ_0 , that is mean income opportunities in the origin country) and the impact of mean migration costs (changes of μ_C), according to the immigration-policy regime.

Thus, according to this simple model, pull and push factors have either similar-sized effects (with opposite signs), when quotas are not binding, or they both have no (or a small) effect on emigration rates, when quotas are binding. In the empirical analysis I will not be able to control for whether migration quotas are binding for a country pair in a given year (since I do not have data on I_{01}^D). Therefore I will estimate an average effect across country pairs with different degrees of restrictiveness. However, I will be able to use information on changes in I_{01}^D : I should find that pull (push) effects are more positive (negative) than average, for a given destination country, if that country's migration policy becomes less restrictive.¹⁴

Focusing for simplicity on the region where immigration quotas are not binding, it is straightforward to derive predictions for the impact of second moments of the income distributions (of the origin country's population) at home and abroad respectively. In particular, assuming that $\sigma_C=0$, we obtain the following expressions, where $k = \phi(z)(\sigma_1^2 + \sigma_0^2 - 2\rho_{01}\sigma_0\sigma_1)^{-\frac{1}{2}}(-\frac{1}{\sigma_v^2}) < 0$ (Borjas, 1987):

$$\frac{d\left(\frac{I_{01}}{P_0}\right)}{d\sigma_1} = k \times (\mu_1^0 - \mu_0 - \mu_C) \cdot (\sigma_1 - \rho_{01}\sigma_0), \quad (7)$$

¹⁴The reason is that, with higher I_{01}^D , the range of μ_1^0 (μ_0) for which the effect is strictly positive (negative) is wider (see Fig. 2).

$$\frac{d\left(\frac{I_{01}}{P_0}\right)}{d\sigma_0} = k \times (\mu_1^0 - \mu_0 - \mu_c) \times (\sigma_0 - \rho_{01}\sigma_1). \quad (8)$$

In my discussion I will assume that $(\mu_1^0 - \mu_0 - \mu_c) > 0$ so that, based on first-moments considerations, on average immigrants have an incentive to migrate. The results in (7) and (8) imply that, if $\frac{\sigma_0}{\sigma_1} < 1$ and ρ_{01} is sufficiently high ($\rho_{01} > \frac{\sigma_0}{\sigma_1}$), then $d\sigma_0 > 0$ or $d\sigma_1 < 0$ (i.e., an increase in the relative inequality $\frac{\sigma_0}{\sigma_1}$ will increase the emigration rate. Similarly, if $\frac{\sigma_0}{\sigma_1} > 1$ and ρ_{01} is sufficiently high ($\rho_{01} > \frac{\sigma_1}{\sigma_0}$), then $d\sigma_0 > 0$ or $d\sigma_1 < 0$ (i.e., an increase in the relative inequality $\frac{\sigma_0}{\sigma_1}$) will decrease the emigration rate.

3 DATA

In this paper, I merge data from an international migration panel with macroeconomic and other information on the origin and destination countries of immigrant flows. Data on immigration comes from the International Migration Statistics (IMS) data set for OECD countries (OECD, 1997), which provides information on bilateral immigrant flows based on the OECD's Continuous Reporting System on Migration (SOPEMI).¹⁵ In particular, I use data on yearly immigrant inflows into fourteen OECD countries by country of origin, in the period 1980–1995. The IMS data only covers legal immigration; population registers and residence and work permits are the main sources of these statistics.¹⁶ Based on this dataset, labor movements to the fourteen OECD countries appear to be both South-North and North-North flows. The sample includes seventy-nine origin countries with per worker GDP levels ranging from

¹⁵Alvarez-Plata et al. (2003) and Pedersen et al. (2004) use different international-migration data sets: the former paper uses the Eurostat Labor Force Survey which covers all destination countries within the EU-15 over nine years; the latter paper uses a dataset constructed by the authors after contacting the statistical bureaus in 27 selected destination countries (this data set covers the years between 1990 and 2000).

¹⁶Although the migration data is not perfectly comparable across OECD countries (some countries in the OECD (1997) data set define immigrants based on country of birth, while others based on citizenship), it is reasonable to think that changes over time can be compared.

approximately \$1000 to \$55,000 (PPP-adjusted) on average in the period considered.

The quality of the IMS data is high even though the coverage is not complete. The data set is supposed to cover immigrant inflows into each of the fourteen destination countries from all over the world. However, the sum by country of origin of the IMS numbers is not equal to 100% of the total flow into each destination country. The percentage of the total immigrant inflow covered by the disaggregate data ranges between 45% (Belgium) and 84% (United States). Put differently, the data set includes zero flows in correspondence of some country pairs (immigrant inflows from Italy to the United States, for example): some of these observations are likely to correspond to very small flows rather than zero flows. If very small flows are recorded as zeros in the disaggregate data set, there will be a discrepancy between total flows and the sum of flows by origin country. In the empirical analysis I will keep zero-flows observations in the data set and will investigate the robustness of my results to using a Tobit model.

Summary statistics and data sources for the other regressors used in the empirical model are documented in Appendix. Data on macroeconomic variables comes from various sources: the 2001 World Development Indicators data set (World Bank 2001) and the Penn World Tables (versions 5.6 and 6.1). Geographic and cultural information, such as on great-circle distance,¹⁷ land border, common language, and colonial ties, comes from Glick and Rose's (2002) data set on gravity-model variables. I also use statistics on the average number of schooling years in the total population of destination and origin countries (over age 15) from Barro and Lee's (2000) data set. Data on Gini coefficients of destination and origin countries, used to construct the origin country's relative inequality variable, comes from Deininger and Squire (1996) data set (I only use so-called high-quality observations).¹⁸ Finally, information on origin countries' share of young population comes from the United Nations.

Figure 2 shows that many destination countries in the sample are characterized by substantial volatility of immigrant inflows year after year. An important cause of variation over time in the number of immigrants to a given destination country is changes in that country's migration policy. For example, the United States' graph in Fig. 2 displays a peak around the

¹⁷Distance is calculated with the great circle formula using each capital city's latitude and longitude data.

¹⁸I linearly extrapolate data on schooling years and Gini coefficients for the years in which it is not available, based on the values for other years for the same country.

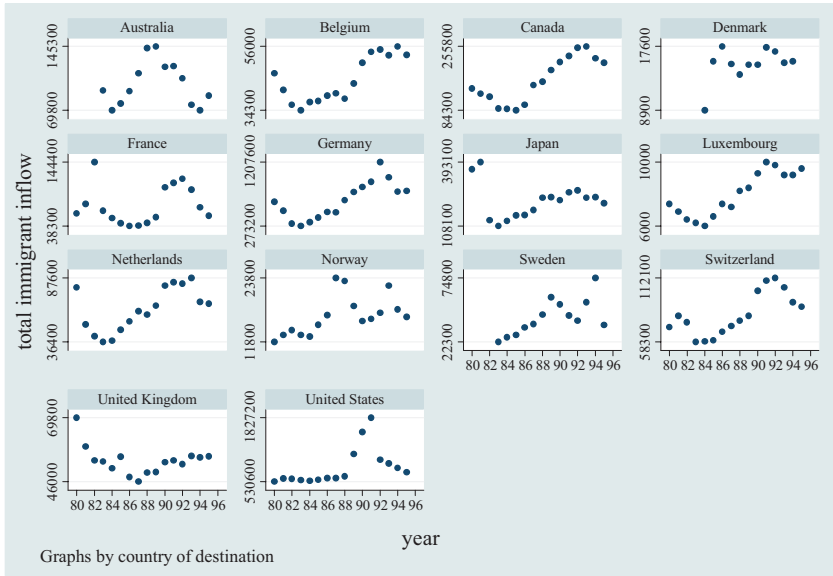


Fig. 2 Total immigrant inflow by destination country

year 1990. This is not surprising given that an amnesty law, the Immigration Reform and Control Act, was passed in 1986 and put in effect in the following years, with the bulk of the legalizations taking place in 1989–1991. The graph for Japan, on the other hand, displays a sudden decrease in the total immigrant inflow around the year 1982, which is when the Immigration Control and Refugee Recognition Act was passed. A separate Appendix to this paper documents the main characteristics of the migration policies of the destination countries in the sample and the timing (after 1980) of changes in their legislations (Mayda and Patel, 2004). A data set of destination countries' migration-policy changes, between 1980 and 1995, was constructed on the basis of the information in this Appendix and used in the empirical analysis.¹⁹

¹⁹In particular, the information in the Appendix (and in the background papers listed in the References) was used to identify: first, the timing of immigration-policy changes taking place in each destination country (the years in which migration policy laws were passed or enforced); second, the direction of the change in the case of substantial changes (loosening vs. tightening), based on a qualitative assessment of the laws (we mainly focused on aspects of migration policies related to the size of immigration flows, as opposed to, for example, issues of citizenship).

4 EMPIRICAL MODEL

According to the theoretical framework in Sect. 2, the estimating equation should include the emigration rate as the dependent variable and, among the explanatory variables, the mean wage of the origin country's population in, respectively, the origin and destination countries. As approximations for the latter two variables, I use the (log) level of per worker GDP, PPP-adjusted (constant 1996 international dollars) in the two countries.²⁰ Based on the theoretical model, I expect pull and push effects to be, respectively, positive and negative on average, if migration quotas are not binding, and both zero (or small) otherwise.

Another determinant of bilateral immigration flows implied by the model of Sect. 2 is the physical distance between the two locations, which affects migration costs C_i . The further away the two countries are, the higher the monetary travel costs for the initial move, as well as for visits back home. Remote destinations may also discourage migration because they require longer travel time and thus higher foregone earnings. Another explanation as to why distance may negatively affect migration is that it is more costly to acquire information *ex-ante* about far-away countries. Besides distance, I introduce additional variables that affect the level of migration costs C_i . A common land border is likely to encourage migration flows, since land travel is usually less expensive than air travel. Linguistic and cultural similarity are also likely to reduce the magnitude of migration costs, for example by improving the transferability of individual skill from one place to the other. Past colonial relationships should increase emigration rates, to the extent that they translate into similar institutions and stronger political ties between the two countries, thus decreasing the level of migration costs C_i .

Finally, I introduce the share of the origin country's population who is young (between 15 and 29 years old) as a demographic determinant of migration flows. Consider an extension of the basic model in Sect. 2 to a multi-period setting. In this set-up, the individual cares not only about current wage differentials net of moving costs, but about future ones too. This implies that a potential migrant from country 0 will have a bigger incentive to migrate the younger he is, as the present discounted value of net benefits will be higher the longer the remaining work life time is (for

²⁰ Unfortunately, wage data cannot be used because wage income series are not available for all countries (especially origin ones) in the sample. Since per worker GDP is not a direct measure of the mean wage of the origin country's population at home and abroad, I run robustness checks to test whether it is a good proxy for it.

positive I_i in each year). We would then expect the share of the young population in the origin country to positively affect the emigration rate out of that country.

In a cross-country analysis, such as in this paper, unobserved country-specific effects could result in biased estimates. For example, the estimate of the coefficient on the destination country's per worker GDP may be positive. Based on this result, it is not clear whether immigrants go to countries with higher wages or, alternatively, whether countries with higher wages have other characteristics that attract immigrants. Along the same lines, a negative coefficient on income at home leaves open the question of whether immigrants leave countries with lower wages or, alternatively, whether countries with lower wages have certain features that push immigrants to leave. To (partly) get around this problem, I exploit the panel structure of the data set and I introduce dummy variables for both destination and origin countries. This allows me to control for unobserved country-specific effects which are additive and time-invariant.²¹ All the regressions also have year effects, to account for common time shocks, and robust standard errors clustered by country pair, to address heteroscedasticity and allow for correlation over time of country-pair observations. Notice that destination countries' fixed effects also allow me to control for features of their immigration policy which are time-invariant and common across origin countries. In order to capture the effect of *changes* in destination countries' migration policies, I introduce two interaction terms of an indicator variable of such changes with pull and push factors, respectively. According to the theory, if the migration policy of a destination country becomes less restrictive, the effect of pull (push) factors should turn more positive (negative).

The basic empirical specification thus looks as follows:

$$\begin{aligned} \frac{flow_{ijt}}{P_{it}} = & \beta + \beta_0 p w g d p_{it-1} + \beta_1 p w g d p_{jt-1} + \beta_2 dist_{ij} + \beta_3 border_{ij} \\ & + \beta_4 comlang_{ij} + \beta_5 colony_{ij} + \beta_6 youngpop_{it-1} \\ & + \beta_7 p w g d p_{it-1} \times immigpol_{jt} + \beta_8 p w g d p_{jt-1} \times immigpol_{jt} \\ & + \delta_i I_i + \delta_j I_j + \delta_t I_t + \varepsilon_{ijt} \end{aligned} \quad (9)$$

²¹ In one robustness check, I control for country-pair fixed effects. In all the other regressions, I include separate destination and origin countries' fixed effects.

where i is the origin country, j the destination country, and t time.

$\frac{flow_{ijt}}{P_{it}}$ is the emigration rate from i to j at time t ($flow_{ijt}$ is the inflow into country j from country i at time t , P_{it} is the population of the origin country at time t). $pwgdp$ is the (log) per worker GDP, PPP-adjusted (constant 1996 international dollars) and $dist$ measures the (log) great-circle distance between the two countries. The variable $border$ equals one if the two countries in the pair share a land border. $comlang$ and $colony$ are two dummy variables equal to one, respectively, if a common language is spoken in the two locations, and for pairs of countries which were, at some point in the past, in a colonial relationship. The variable $youngpop$ is the share of the population in the origin country aged 15–29 years old. The variable $immigpol$ increases by one (decreases by one) if in that year the destination country's immigration policy became less (more) restrictive, zero otherwise. In other words, a change in policy is modelled as leading to a lasting effect (i.e., in the year when the policy change occurred and in the following years). Finally, the basic empirical specification also includes destination and origin countries' fixed effects (I_j and I_i) and year effects (I_t). According to the model in Sect. 2, I expect that $\beta_0 \leq 0$, $\beta_1 \geq 0$, $\beta_2 \leq 0$, $\beta_3 \geq 0$, $\beta_4 \geq 0$, $\beta_5 \geq 0$, $\beta_6 \geq 0$, $\beta_7 < 0$, and $\beta_8 > 0$.

An econometric complication is the possibility of reverse causality and, more in general, of endogeneity in the time-series dimension of the analysis. For example, the theoretical model in Sect. 2 predicts that, if migration quotas are not binding, better (worse) income opportunities in the destination (origin) country increase emigration rates. However, a positive β_1 (negative β_0) may just reflect causation in the opposite direction, that is the impact of immigrant flows on wages in host and source countries. After all, this channel is the main focus of analysis in many labor-economics papers (see Friedberg & Hunt, 1995 for a survey of this literature). More broadly, other time-variant third factors may drive contemporaneous wages and immigrant flows.

As for reverse causality, notice that it is likely to bias the estimates toward zero. The reason is that, if anything, immigrant inflows are likely to decrease wages in the destination country and outflows are likely to increase wages in the origin country. While the opposite signs are a theoretical possibility (e.g., in the economic-geography literature, because of economies of scale), the empirical evidence in the labor-economics

literature is that immigrant inflows have a negative or zero impact on the destination country's wages (Friedberg & Hunt, 1995; Borjas, 2003) and that immigrant outflows have a positive impact on the origin country's wages (Mishra, 2007).

Although reverse causality may not be an issue, it is still important to address endogeneity. Thus, I relate *current* emigration rates to *lagged* values of (log) per worker GDP, at home and abroad (and to *lagged* values of all the other time-varying regressors). While it is unrealistic to claim that wages at home and abroad are strictly exogenous, it is plausible to assume that they are predetermined, in the sense that immigrant inflows—and third factors in the error term—can only affect contemporaneous and future wages.²²

5 EMPIRICAL RESULTS

Table 1 presents the results from estimation of Eq. (9). The estimates show a systematic pattern, broadly consistent with the theoretical predictions of the international migration model. The analysis also generates empirical puzzles.

First, the emigration rate is positively related to the destination country's (log) per worker GDP. According to the estimate in regression (1), a 10% increase in the level of per worker GDP in the destination country increases emigration by 2.6 emigrants per 100,000 individuals of the origin country's population (significant at the 5% level). In other words, a 10% increase in the host country's per worker GDP implies a 20% increase in the emigration rate (as the mean of the dependent variable is, in regression (1), 13 emigrants per 100,000 individuals). This result would suggest that migration quotas are not binding on average across destination countries. However, the impact on the emigration rate of a change in the income opportunities at home is not consistent with this interpretation. Push effects are estimated to be insignificantly different from zero in Table 1 (and often of the wrong sign). One possibility is that, in practice, migration quotas are not binding, but push factors are zero due to the

²² Strict exogeneity of an explanatory variable implies $E[X_{it}\varepsilon_{it}] = 0$, for $\forall s, t$, while predeterminedness implies $E[X_{it}\varepsilon_{is}] = 0$, for $\forall s > t$. In one of the following specifications, I also control for lagged values of the emigration rate, since if the emigration rate is autocorrelated, predeterminedness of the regressors does not guarantee consistency of the estimates.

Table 1 Determinants of bilateral immigrant flows

<i>Equation</i>	1	2	3	4	5	6	7	8	9
<i>Dependent variable</i>	<i>Emigration rate</i>								
log per worker gdp (destination)	25.57	25.55	25.81	23.39	37.8	25.74	189.66	22.05	19.33
	10.79*	10.76*	10.79*	11.36*	15.56*	10.83*	76.66*	10.97*	14.55
log per worker gdp (origin)	2.05	1.78	4.82	6.01	-20.59	5.45	5.62	-7.97	9.74
	7.95	7.8	8.52	10.31	14.53	8.74	23.91	6.16	9.19
log distance	-40.99	-40.63	-40.64	-40.84	-31.94	-37.92	-148.76		-41.91
	9.51**	9.09**	9.09**	9.33**	10.51**	8.00**	4.46**		8.42**
Land border	-28.14	-36.94	-36.93	-38.88	-17.48				
	19.69	23.24	23.28	24.7	27.68				
Common language	22.05	22.03	23.78	23.78	16.72				
	15.88	15.88	17.24	16.61					
Colony	3.02	2.9	1.96	11.1					
	16.91	16.94	18.03	18.75					
Share of young population (origin)			198.46	164.38	-10.2	204.45	573.14	124.97	223.43
			103.78+	104.73	112.51	106.00+	311.80+	82.85	109.42*
log yrs schooling (destination)				-29.76	-32.96				
				13.65*	23.83				
log yrs schooling (origin)				17.36	46.25				
				13.91	35.81				

(continued)

Table 1 (continued)

Equation	1	2	3	4	5	6	7	8	9
<i>Dependent variable</i>	<i>Emigration rate</i>								
log capital per worker (destination)					-21.49				
					21.38				
log capital per worker (origin)					11.77				
					15.57				
Per worker gdp (destination)*immig policy change									17.63
per worker gdp (origin)*immig policy change									6.12** -3.24
log distance*immig policy change									1.44* -10.4
Share of young population (origin)*immig policy change									151.49
immig policy change									48.73** -109.19 73.18
Number of observations	8010	8010	8010	7313	4103	8010	8010	8010	8010
R-squared	0.24	0.25	0.25	0.26	0.26	0.24		0.85	0.27

OLS estimates except for regression (7), which is a Tobit. All regressions include year effects. Destination and origin countries' dummy variables are included in specifications (1)–(7) and (9). Regression (8) includes country-pair fixed effects. Standard errors, clustered by country pairs, are presented under each estimated coefficient. Constant not shown. + significant at 10%; * significant at 5%; ** significant at 1%. See Appendix for data sources.

The emigration rate (immigrant inflow from origin to destination country (multiplied by 100,000), divided by origin country's population) gives the number of incoming immigrants per 100,000 individuals in the origin country's population. `per worker gdp` is the level of per worker GDP, PPP-adjusted (constant 1996 international dollars), lagged by one year. `distance` is the great-circle distance. `land border` equals one if the destination and origin countries share a land border.

`common language` equals one if a common language is spoken in both host and origin countries. `colony` equals one for pairs of countries ever in a colonial relationship. `share of young population (origin)` is the share of the population in the origin country aged 15–29, lagged by one year. `log yrs schooling` is the log of the average schooling years in the population over age 15, lagged by one year. `log capital per worker` is non-residential capital stock per worker (1985 intl. prices), lagged by one year. `immig policy change` increases by one if in that year the host country's immigration policy became less restrictive, decreases by one if it became more restrictive, zero if there was no change.

effect of poverty constraints in the origin country. I will investigate this hypothesis in Table 2.

In regressions (1)–(3), Table 1, I also explore the role played by geographic (*log distance* and *land border*), cultural (*common language* and *colony*), and demographic (*share of young population (origin)*) determinants, respectively. The picture that emerges from my results is one in which geography and demographics are the most important among this set of drivers of migration flows. According to the estimate in column (1), doubling the great-circle distance between the source and host country decreases the number of emigrants by 41 per 100,000 individuals in the origin country (significant at the 1% level). On the other hand, a common land border does not appear to play a significant role. The impact of a common language, though of the right sign, is not statistically significant and, surprisingly, past colonial relationships do not appear to affect migration rates (this is true whether *common language* and *colony* are entered in

Table 2 Economic determinants more in detail

<i>Equation</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>
<i>Dependent variable</i>	<i>Emigration rate</i>					
log per worker gdp (destination)	6.64	26.42	25.51	17.1	23.02	37.78
	20.5	27.11	10.86*	12.1	11.15*	12.31**
log per worker gdp (origin)	4.12	4.5	75.7	8.87	6.73	1.68
	17.64	17.93	53.08	15.26	8.58	7.15
Square of log per worker gdp (origin)			-3.84			
			3.07			
Origin country's relative inequality	74.89	80.08				
	38.66+	44.06+				
Square of relative inequality	-28.11	-29.29				
	13.15*	14.51*				
Unemployment rate (destination)				-0.35		
				0.4		
Unemployment rate (origin)				0.96		
				0.91		
Multilateral pull					-9.33	
					5.05+	
Emigration rate ($t-1$)						0.66
						0.02**

(continued)

Table 2 (continued)

<i>Equation</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>
<i>Dependent variable</i>	<i>Emigration rate</i>					
log distance			-36.82 7.57**	-30.43 8.28**	-36.16 7.36**	
Common language			18.86 11.85	19.16 11.55+	18.97 11.86	
Share of young population (origin)			186.06	-35.59	195.65	
Constant	-143.56 274.49	-349.04 333.02	-327.12 234.08	-24.72 220.43	-4.54 186.64	-0.59 0.24*
Number of observations	4028	3350	8010	5010	8010	6429
R-squared	0.18	0.18	0.25	0.23	0.25	

OLS estimates with year effects, except for regression (6) (see below). Destination and origin countries' dummy variables are included in each specification (except in regression (6)). Standard errors, clustered by country pairs, are presented under each estimated coefficient. + significant at 10%; * significant at 5%; ** significant at 1%. See the end of Table 1 for definitions of the main variables used.

The emigration rate (immigrant inflow from origin to destination country (multiplied by 100,000), divided by origin country's population) gives the number of incoming immigrants per 100,000 individuals in the origin country's population.

multilateral pull gives, for each destination/origin country pair, the average of (log per worker gdp (destination)-log distance) over all the other destination countries. origin country's relative inequality gives a measure of the inequality in the origin country relative to the destination country (it equals the Gini coefficient in the origin country divided by the Gini coefficient in the destination country).

In regression (2), I only include observations characterized by a positive difference between the per capita GDP levels of destination and origin countries in any given year. In Eq.(6), I include as regressors the emigration rate lagged by one, two, three, and four years (the coefficients on the latter three lags are not shown). Only by introducing all these lags, I don't reject the null of zero autocovariance in residuals of order 2 (which is one of the requirements of the Arellano and Bond estimator). Column (6): Arellano-Bond test that average autocovariance in residuals of order 1 is 0: H_0 : no autocorrelation $z = -55.05$ $Pr > z = 0.0000$. Column (6): Arellano-Bond test that average autocovariance in residuals of order 2 is 0: H_0 : no autocorrelation $z = -0.35$ $Pr > z = 0.7269$. See Appendix for data sources.

the regression together or one at a time). Finally, the share of the origin country's population who is young has a positive and significant impact on emigration rates. A ten percentage point increase in the origin country's 15–29 years old population raises the emigration rate by 20 emigrants per 100,000 individuals (regression (3)).

Next, I investigate whether per worker GDP (PPP-adjusted) of origin and destination countries is a good proxy for mean income opportunities of migrant workers at home and abroad. Per worker GDP is not a direct measure of wages of a potential migrant, since it depends on rates of return

to both capital and labor and on endowments of each factor. For example, a higher per worker GDP in the destination country does not necessarily mean better income opportunities on average for an immigrant worker, since it could be due to a higher capital-labor ratio or to a more skilled labor force in the destination country's population. To address this concern, I run a robustness check where I control for the mean skill level and per worker capital endowment in destination and origin countries (columns (4)–(5)).²³ I first control for the average schooling level in both countries in regression (4). I still estimate pull effects which are positive and significant (at the 1% level). The results on push effects are the same as in previous estimates as well. In line with the theoretical predictions, the average skill level in the population of the destination (origin) country has a negative (positive) impact on the emigration rate. In regression (5) I control for the per worker endowments of both skill and capital and find that their coefficients are of the right sign (although not significant). Most importantly, my prior findings on pull and push factors are robust.

In column (6), out of all the geographic, cultural, and demographic determinants, I only include the ones which are significant based on regressions (1)–(3), that is *log distance* and *share of young population (origin)*. I find evidence consistent with my previous results. Using a specification with these variables, I test how robust the results are—in particular, in terms of the asymmetry between pull and push factors—to using a Tobit specification (regression (7)). The estimates are again in line with the picture based on OLS regressions but they are larger in magnitude.

In the next regression (column (8)) I only exploit the variation over time within country pairs, by introducing fixed effects for each combination of origin and destination countries.²⁴ These country-pairs dummy variables allow me to control for time-invariant features of the destination country's immigration policy which are specific for each origin country. The results from this specification confirm that push and pull factors have an asymmetric effect in terms of magnitudes and significance levels.²⁵

²³ Since capital is assumed to be internationally mobile, there are no international differences in rates of return to capital.

²⁴ Therefore I do not include the regressors *log distance*, *land border*, *common language*, and *colony* since they are constant within country pairs and, therefore, would be perfectly collinear with the country-pair dummy variables.

²⁵ If country pairs differ in terms of out-migration and return migration rates, net migration flows can be very different from gross flows. Since out-migration and return migration are likely to characterize specific country pairs, they are partially accounted for by including country-pair fixed effects.

Next, I investigate the interaction between changes in destination countries' migration policies and, respectively, pull and push factors (column (9), Table 1). Consistent with the theoretical predictions, positive pull factors are bigger than average for a destination country whose migration policy becomes less restrictive. Setting aside the average effect, push factors turn negative and significant once migration restrictions are relaxed. The opposite is true when policy becomes more protectionist. In the same regression I also add the interaction of the indicator variable of changes in destination countries' migration policy with, respectively, *log distance* and *share of young population (origin)*. I find that the effect of the latter two variables is more pronounced (more negative and more positive, respectively) when a host country's immigration laws turn less restrictive. The opposite is true when policy becomes more protectionist. Notice that I also include the linear effect of immigration policy changes, which is insignificant. Regression (9) represents the preferred specification of the model. It shows that migration restrictions matter by mitigating effects on the supply side of the model (pull and push factors, geography, and demographics).

6 ADDITIONAL RESULTS

In Table 2, I analyze economic determinants more in detail. First, I investigate the impact of the second moments of the income distributions in the origin and destination countries. According to the theory (formulas (7) and (8)), given low values of the origin country's relative inequality ($\frac{\sigma_0}{\sigma_1}$), if $\frac{\sigma_0}{\sigma_1}$ increases, the emigration rate will increase, while given high values of $\frac{\sigma_0}{\sigma_1}$, if $\frac{\sigma_0}{\sigma_1}$ increases, the emigration rate will decrease.²⁶ The

²⁶I assume that ρ_{01} is sufficiently high ($\rho_{01} > \max\{\frac{\sigma_0}{\sigma_1}, \frac{\sigma_1}{\sigma_0}\}$). The motivation for this assumption is explained in Borjas (1987): "It seems plausible to argue that for non-Communist countries, ρ_{01} is likely to be positive and large. After all, profit-maximizing employers are likely to value the same factors in any market economy" (p.534). I also assume that $(\mu_1^0 - \mu_0 - \mu_c) > 0$ so that, based on first-moments considerations, on average immigrants have an incentive to migrate. The motivation for the last assumption is that the data set mostly includes migration flows from lower to higher average-income countries: the average difference in per capita GDP levels of destination and origin countries is positive and

intuition for these results is straightforward. If income inequality in the origin country is lower than in the destination country ($\frac{\sigma_0}{\sigma_1} < 1$), there is positive selection of immigrants from country 0 to country 1: migrants are selected from the upper tail of the income distribution at home and end up in the upper tail of the income distribution abroad (in both cases, the relevant distribution is the origin country's population one). For example, consider potential migrants from Portugal to the United States. Given that income inequality is lower in Portugal than in the U.S., among Portuguese workers it is the better-off who have an incentive to migrate while those at the very low tail of the income curve have an incentive to stay. The reason is that the probability of both very high and very low incomes is higher in the U.S. than in Portugal. An increase in income inequality in Portugal will make the marginal individual (who is in the lower tail of the income distribution) relatively worse-off at home and will increase her incentive to leave. Similarly, if income is more dispersed at home than abroad ($\frac{\sigma_0}{\sigma_1} > 1$), then there is negative selection of immigrants from country 0 to country 1: migrants are selected from the lower tail of the income distribution at home and end up in the lower tail of the income distribution abroad. An example of this situation is migration from Brazil to the U.S., given that income inequality in the latter is lower than in the former.²⁷ An increase in income inequality in Brazil will lower the emigration rate because those who were not migrating beforehand, the better-off, will have even less incentive to do so afterwards. In order to test these predictions, I introduce in the estimating equation a measure of the origin country's relative inequality ($\frac{\sigma_0}{\sigma_1}$) both in linear and quadratic forms. As expected, I find that the coefficient on the linear term is positive and on the quadratic term is negative (both significant at conventional levels),

substantial (approximately \$20,600). I also add a robustness check (regression (2), Table 2) where I only include observations characterized by a positive difference between the per capita GDP levels of destination and origin countries in any given year.

²⁷The Gini coefficient for Portugal was 36.76 in 1990, while in the U.S. it was 37.8. The Gini coefficient for Brazil was 61.76 in 1985, while in the U.S. it was 37.26 (Deininger and Squire 1996).

which is consistent with Borjas (1987) selection model (regressions (1)–(2), Table 2).²⁸

The remaining specifications in Table 2 investigate empirically a few extensions of the theoretical framework of Sect. 2. First, it is possible to incorporate poverty constraints in the model, due to fixed costs of migration and credit market imperfections in the origin country. As Yang (2003) shows, these assumptions imply that the effect on emigration rates of income opportunities at home is non-monotonic, positive at very low levels of income and negative for higher levels. Accordingly, I extend the empirical model previously specified by introducing both a linear and a quadratic term in per worker GDP of the origin country. I find very weak evidence of poverty constraints in regression (3). The sign of the coefficients is consistent with the theory but the lack of significance of the estimates prevents me from reading too much support into them.²⁹ This result thus leaves open the question of why push and pull effects are different in size and, indirectly, lends support to the alternative hypothesis of binding (and endogenous) migration quotas.

Next, the theoretical model can be modified by taking into account uncertainty in finding a job in each place. This extension suggests using the unemployment rate (which is approximately equal to one minus the probability of finding a job) as a regressor in the estimating equation. My results in column (4) are not significant. In an additional extension (column (5)), I test whether workers choose among multiple destination countries. In the theoretical model, the choice is between the origin country and one particular destination country. In practice, however, potential migrants are likely to compare mean income opportunities in their origin country to those in the destination country considered *and in any other*

²⁸ I evaluate the effect of relative inequality over the relevant range of values. Based on the coefficient estimates in column (1), Table 2, the threshold value of relative inequality is approximately equal to 2.6642: if $\frac{\sigma_0}{\sigma_1}$ is below this value (which is the case almost always in

my sample, based on the summary statistics in Appendix), an increase in $\frac{\sigma_0}{\sigma_1}$ raises the emigration rate. This is consistent with positive selection taking place.

²⁹ In contrast with my results—which are not significant—Hatton and Williamson (2003) and Pedersen et al. (2004) find evidence of an inverted U-shaped effect on emigration of the origin country's economic conditions.

host country. For each pair of source and host economies, I construct and control for a *multilateral pull* term which is an average of per worker GDP levels of all the other destination countries in the sample, each weighted by the inverse of distance from the origin country. Regression (5) shows that third-country effects shape bilateral migration flows as expected, given that the coefficient of the *multilateral pull* term is indeed negative and significant (at the 10% level).³⁰

To conclude, I investigate the role of past migration flows to the destination country from the same origin country. Lagged emigration rates capture the impact of network effects, which are likely to reduce the cost C_i of migration. The introduction of the lagged emigration rate among the explanatory variables makes the model a dynamic one. I use Arellano and Bond's GMM estimator to deal with the incidental parameter problem that arises with fixed-effects estimation of such a dynamic equation.³¹ Emigration rates show considerable inertia in regression (6), where the coefficient of the lagged emigration rate is 0.66 (significant at the 1% level).³² However, outside the model of Sect. 2—which assumes exogenous migration quotas—it is unclear how to interpret this autocorrelation. While it is consistent with network effects on the supply side, it could also be driven by factors working on the demand side. In particular, through the latter channel, past migration flows can influence the emigration rate in two different ways: through family-reunification immigration policies and through political-economy factors (see, e.g., Goldin (1994) and Ortega (2005), where the votes of naturalized immigrants affect immigration-policy outcomes).

³⁰The multilateral pull term places migrants' decision to move in a multi-country framework. It is inspired by the multilateral trade resistance term in Anderson and van Wincoop (2003) (even though mine is an atheoretical measure).

³¹The Arellano and Bond estimator transforms into a difference the initial equation to remove the country-pair fixed effect and produces an equation that can be estimated with instrumental variables using a generalized method-of-moments estimator. The instruments include the lagged values of the dependent variable starting from $t-4-2$ (since the regression includes, as regressors, the emigration rate lagged by one, two, three, and four years).

³²Regression (6) includes, as regressors, the emigration rate lagged by one, two, three, and four years (the coefficients of the latter three lags are not shown in the table). The reason is that, only by introducing all these lags, I don't reject the null of zero autocovariance in residuals of order 2 (which is one of the requirements of the Arellano and Bond estimator).

7 CONCLUSIONS

In this paper, I empirically investigate the determinants of international bilateral migration flows. This analysis both delivers estimates consistent with the predictions of the international migration model and generates empirical puzzles.

In particular, I find evidence that pull factors, that is income opportunities in the destination country, significantly increase the size of emigration rates. This result is very robust to changes in the specification of the empirical model. On the other hand, the sign of the impact of push factors—that is, per worker GDP in the origin country—is seldom negative and, when it is, the size of the effect is smaller than for pull factors and insignificant. Therefore the evidence uncovered by the estimates is mixed in terms of the migration-policy regime that characterizes, on average, the destination countries in the sample: Push effects suggest that migration quotas are more binding than pull effects do. A possible explanation of the asymmetry between push and pull factors is the role played by the demand side of the model, that is destination countries' migration policies. While the theoretical framework of Sect. 2 assumes that migration quotas are exogenous, in practice they are not. Indeed migration policies can be thought of as the outcome of a political-economy model in which voters' attitudes toward immigrants, interest-groups pressure, policy-makers preferences, and the institutional structure of government interact with each other and give rise to a final immigration-policy outcome (Rodrik, 1995; Facchini & Willmann, 2005; Mayda, 2006). Binding and endogenous migration quotas can explain the asymmetric effect I estimate for pull and push factors. While I do not investigate the endogenous determination of migration policy, I find evidence consistent with the constraining role played by migration policies. In the empirical analysis, I interact an indicator variable of changes in destination countries' migration policies with pull and push factors, respectively. I find that pull effects become more positive and push effects turn negative in those years when a host country's immigration laws become less restrictive.

Among the variables affecting the costs of migration, distance appears to be the most important one. Its effect is negative, significant, and steady across specifications. Demographics, in particular the share of the origin country's population who is young, represent a significant determinant of emigration rates as well. I find that the effect of both variables is more pronounced in those years when a host country's immigration laws become less restrictive. In sum, my results suggest that migration quotas matter: They mitigate supply-side effects, that is pull and push factors, geography, and demographics.

The investigation of the determinants of international migration leads to other interesting research questions. The framework I have used in this paper to study migration flows is related to the gravity model of trade, which is used to analyze bilateral *trade* flows across countries. As a matter of fact, I have used several variables that appear frequently in the trade gravity literature (*log distance*, *land border*, *common language*, and *colony*). A common framework of empirical analysis for trade and migration makes it possible to combine the study of these two dimensions of international integration.

To conclude, by taking advantage of both the time-series and cross-country variation in an annual panel data set, this paper makes progress in explaining the determinants of international migration flows and in providing a framework for future analyses of migration relative to other dimensions of globalization.

Acknowledgments I would like to thank Alberto Alesina, Elhanan Helpman, Dani Rodrik, Klaus Zimmermann, and three anonymous referees for many insightful comments. For helpful suggestions, I am also grateful to Richard Adams, Marcos Chamon, Giovanni Facchini, Bryan Graham, Louise Grogan, Russell Hillberry, Arik Levinson, Lindsay Lowell, Rod Ludema, Lant Pritchett, Maurice Schiff, Tara Watson, Jeffrey Williamson, and participants at the International Workshop at Harvard University, at the 2003 NEUDC Conference at Yale University, at the International Trade Commission Workshop, at the World Bank International Trade Seminar, and at the IZA Annual Migration Meeting. I am grateful for research assistance provided by Krishna Patel and Pramod Khadka. All errors remain mine.

APPENDIX

See Table 3.

Table 3 Summary statistics (1980–1995) and data sources

<i>Variable</i>	<i>Obs</i>	<i>Mean</i>	<i>Std. Dev.</i>	<i>Min</i>	<i>Max</i>
Emigration rate	8010	13.2433	81.5410	0.0000	1568.9430
Per worker gdp (destination)	8010	40682	5895	25252	55361
Per worker gdp (origin)	8010	20061	14106	1027	55361
log distance	8010	8.1715	0.8694	5.0872	9.3836
Land border	8010	0.0268	0.1616	0	1
Common language	8010	0.1704	0.3760	0	1
Colony	8010	0.0385	0.1923	0	1
Share of young population (origin)	8010	0.2612	0.0303	0.1951	0.3152
Years schooling (destination)	4103	9.6403	1.3096	6.8370	11.8650
Years schooling (origin)	4103	7.0285	2.4659	2.7240	11.8650
Capital per worker (destination)	4103	36041	12167	16992	76733
Capital per worker (origin)	4103	19232	13290	822	48135
Unemployment rate (destination)	5010	6.7306	3.4646	0.5000	14.1000
Unemployment rate (origin)	5010	8.1476	5.2840	0.0800	27.6000
Origin country's relative inequality	4028	1.2123	0.3846	0.3861	2.6810

The emigration rate (immigrant inflow from origin to destination country (multiplied by 100,000), divided by origin country's population) is from the IMS data set (OECD, 1997). Per worker GDP, PPP-adjusted (constant 1996 international dollars) is from the Penn World Tables, version 6.1. Log distance, land border, common language, and colony (countries ever in a colonial relationship) are from Glick and Rose (2002). Years of schooling are from Barro and Lee (2000) data set. Capital per worker (Nonresidential Capital Stock per Worker, 1985 intl. prices) is from the Penn World Tables, version 5.6.

The share of young population (origin) is based on data from the United Nations. The unemployment rate is from the World Development Indicators (2001), World Bank. The origin country's relative inequality is based on data on Gini coefficients from Deininger and Squire (1996) data set (only high-quality observations were used). The data set on immigration policy changes was constructed by Mayda and Patel (2004). All time-varying variables (except the emigration rate) are lagged by one year.

Summary statistics for the emigration rate, per worker gdp (destination), per worker gdp (origin), log distance, land border, common language, colony, share of young population (origin) are based on the same observations as in regressions (1)–(3) and (6)–(9), Table 1. Summary statistics for years of schooling (destination), years of schooling (origin), capital per worker (destination), capital per worker (origin) are based on the same observations as in regression (5), Table 1. Summary statistics for unemployment rate (destination) and unemployment rate (origin) are based on the same observations as in regression (4), Table 2. Finally, summary statistics for the origin country's relative inequality are based on the same observations as in regression (1), Table 2.

REFERENCES

- Alvarez-Plata, P., Brucker, H., & Siliverstovs, B. (2003). Potential migration from Central and Eastern Europe into the EU-15—An update Report for the European Commission, DG Employment and Social Affairs.
- Anderson, J. E., & Wincoop, E. V. (2003). Gravity with gravitas: A solution to the border puzzle. *American Economic Review*, 93(1), 170–192.
- Barro, R., & Lee, J. (2000). International data on educational attainment. Data Set.
- Borjas, G., & Bratsberg, B. (1996). Who leaves? The outmigration of the foreign-born. *Review of Economics and Statistics*, 78(1), 165–176.
- Borjas, G. J. (1987). Self selection and the earnings of immigrants. *American Economic Review*, 77(4), 531–553.
- Borjas, G. J. (1999). The economic analysis of immigration. In O. Ashenfelter & D. Card (Eds.), *Handbook of labor economics* (chapter 28, pp. 1697–1760). The Netherlands: North-Holland Elsevier Science.
- Borjas, G. J. (2003). The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *Quarterly Journal of Economics*, 118(4), 1335–1374. Harvard University.
- Brücker, H., Siliverstovs, B., & Trübswetter, P. (2003). International migration to Germany: Estimation of a time-series model and inference in panel cointegration. DIW Discussion paper No. 391.
- Clark, X., Hatton, T. J., & Williamson, J. G. (2007). Explaining US immigration, 1971–1998. *Review of Economics and Statistics*, 89(2), 359–373.
- Facchini, G., & Willmann, G. (2005). The political economy of international factor mobility. *Journal of International Economics*, 67(1), 201–219.
- Friedberg, R., & Hunt, J. (1995). The impact of immigrants on host country wages, employment, and growth. *Journal of Economic Perspectives*, 9(2), 23–44.
- Glick, R., & Rose, A. (2002). Does a currency union affect trade? The time-series evidence. *European Economic Review*, 46(6), 1125–1151.
- Goldin, C. (1994). The political economy of immigration restriction in the United States, 1890 to 1921. In C. Goldin & G. Libecap (Eds.), *The regulated economy: A historical approach to political economy* (pp. 223–257). Chicago, IL: University of Chicago Press.
- Hanson, G. H., & Spilimbergo, A. (2001). Political economy, sectoral shocks, and border enforcement. *Canadian Journal of Economics*, 34(3), 612–638.
- Hatton, T. J. (2004). Emigration from the UK, 1870–1913 and 1950–1998. *European Review of Economic History*, 8(2), 149–171.
- Hatton, T. J. (2005). Explaining trends in UK immigration. *Journal of Population Economics*, 18(4), 719–740.
- Hatton, T. J., & Williamson, J. G. (2003). Demographic and economic pressure on emigration out of Africa. *Scandinavian Journal of Economics*, 105(3), 465–482.

- Helliwell, J. F. (1998). *How much do national borders matter?* (chapter 5, pp. 79–91) Washington, DC: Brookings Institution Press.
- Hunt, J. (2006). Staunching emigration from East Germany: Age and the determinants of migration. *Journal of the European Economic Association*, 4(5), 1014–1037.
- Joppke, L. (1998). Why liberal states accept unwanted immigration. *World Politics*, 50(2), 266–293.
- Karemera, D., Oguledo, V. I., & Davis, B. (2000). A gravity model analysis of international migration to North America. *Applied Economics*, 32(13), 1745–1755.
- Lopez, R., & Schiff, M. (1998). Migration and the skill composition of the labor force: The impact of trade liberalization in LDCs. *Canadian Journal of Economics*, 31(2), 318–336.
- Mayda, A. M. (2006). Who is against immigration? A cross-country investigation of individual attitudes toward immigrants. *Review of Economics and Statistics*, 88(3), 510–530.
- Mayda, A. M., & Patel, K. (2004). OECD countries migration policies changes. Appendix to International migration: A panel data analysis of the determinants of bilateral flows, by Anna Maria Mayda. <http://www.georgetown.edu/faculty/amm223/>
- Mishra, P. (2007). Emigration and wages in source countries: Evidence from Mexico. *Journal of Development Economics*, 82(1), 180–199.
- Nations, U. (2004). *World economic and social survey 2004. International migration*. New York: United Nations.
- OECD (1997). International migration statistics for OECD countries. Data set.
- Ortega, F. (2005). Immigration policy and skill upgrading. *Journal of Public Economics*, 89(9–10):1841–1863.
- Pedersen, P. J., Pytlikova, M., & Smith, N. (2004). Selection or network effects? Migration flows into 27 OECD countries, 1990–2000. IZA Discussion Paper 1104.
- Pedersen, P. J., Pytlikova, M., & Smith, N. (2006). Migration into OECD countries 1990–2000. In Parson & Smeeding (Eds.), *Immigration and the transformation of Europe*. Cambridge: Cambridge University Press.
- Rodrik, D. (1995). Political economy of trade policy. In G. Grossman & K. Rogoff (Eds.), *Handbook of international economics* (Vol. 3, chapter 28, pp. 1457–1494). The Netherlands: North Holland Elsevier Science.
- Rodrik, D. (2002). Comments at the conference on “Immigration policy and the welfare state” In T. Boeri, G. H. Hanson, & B. McCormick (Eds.), *Immigration policy and the welfare system*. Oxford: Oxford University Press.
- Yang, D. (2003). *Financing constraints, economic shocks, and international labor migration: Understanding the departure and return of Philippine overseas workers*. Harvard University, dissertation chapter.