Erschienen in: Kyklos. 2018, 71(1), pp. 86-109
https://doi.org/10.1111/kykl.12163

Migration, Diasporas and Culture: An Empirical Investigation

Paul Collier and Anke Hoeffler*

I. INTRODUCTION

The increase in international migration since 1945 is often considered as an inte gral part of globalization. However, this is belied by its composition. Migration among OECD countries has been broadly constant despite a vast expansion in both trade and capital flows between them. The expansion in global migration is entirely accounted for by flows from developing countries to the OECD. Developed countries account for 15 per cent of the global population, but are the destination of 56 per cent of all migrants. While the defining difference between developed and developing countries is income, there are also substantial differences in political freedom and in culture. Migration from developing to developed countries is likely to be driven by these three gaps and their interactions.

In this article, we analyze the flow of voluntary migration from developing to OECD countries for each decade between 1960 and 2000. We use recently released global panel data and build on the work by Docquier and Rapoport (2012) and Mayda (2010). We first estimate a core model based on the assump tion that migration is determined by push and pull factors. This gravity model of migration suggests that differences in income and political freedom between host and origin countries are important factors in the migration decision. This supports a conceptualization of the act of migration as an investment on which the migrant gets a subsequent return. Impediments to mobility, such as greater

*Paul Collier: Blavatnik School of Government Oxford University, and Centre for the Study of African Economies. Anke Hoeffler: Centre for the Study of African Economies, Department of Economics, Oxford University. Manor Road Building, Manor Road, Oxford OX1 3UQ, anke.hoeffler@economics.ox.ac.uk.
We would like to thank the anonymous referee, Michel Beine, Simone Bertoli, Stephen Fisher and Mariola Pytlikova for comments on an earlier draft and David Hendry and Jennifer Castle for help with Oxmetrics. We are also grateful to Çağlar Özden and Romain Wacziarg for their advice on the use of the migration and cultural distance data. All remaining errors are our own. Anke Hoeffler's participation was funded through the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 290752 (NOPOOR).

geographical distance, require increased investments. The investment costs of migration are reduced by support from prior migrants, also commonly referred to as the diaspora. Consistent with much of the migration literature we find that this effect of diasporas is powerful (e.g. Beine et al. 2011). As a result, migration has a tendency to accelerate: migration builds diasporas, and diasporas ease sub sequent migration. In addition to examining the effect of the diaspora we allow for a number of interaction effects between the diaspora, the push and pull factors as well as obstacles to migration. We also account for unobserved country effects and use an automated process to derive a more parsimonious model. We find that origin countries do not systematically influence the emigration decision while host countries regulate immigration effectively and they differ systematically in the degree of regulation from each other. The interaction effects as well as the country effects go some way towards accounting for omitted variables that otherwise bias the coefficient on the diaspora. We then develop this core model by examining the impact of culture. Potentially, one important impediment to mi gration is a difference in cultures: experimental methods of behavioral economics are now able to quantify substantial differences between cultures (Gächter and Schulz 2016) and genetic data has been used to proxy for differences in cultures (Spolaore and Wacziarg 2013). Surprisingly, we do not find that larger difference in cultures *per se* constitute an obstacle to migration. However, the interaction between cultural distance and diaspora is positive, suggesting that the diaspora introduces a further potential effect of culture. The cultural distance between home and host societies may affect the rate at which immigrants absorb into their host society, and hence the rate at which their connections with their country of origin decay. If a wide culture gap preserved the link between immigrants and their home societies, it would increase the gearing from the stock of immigrants onto the subsequent flow. A large difference in cultures could therefore potentially both impede migration and accelerate its increase. Thus, our article contributes to the migration literature in two ways. First, we improve on existing gravity models and second, we examine the impact of the cultural distance between sending and receiving countries.

Our article is structured in the following way. In Section II we estimate a gravity model of migration, using a set of economic and political explanatory variables, and show how the results support an investment return interpretation. We introduce a measure of cultural distance in Section III and show that it has powerful effects. Section IV concludes.

II. ESTIMATING A CORE MODEL

Migration has been examined in many different academic disciplines: history, economics, political science, geography, anthropology and sociology. The core assumption of economic models is that individuals migrate from low

to high income regions in order to improve their well being (Ravenstein's 1885, 1889 'laws of migration'). In the twentieth century a number of gravity based models expanded on the idea that migration is a function of spatial disequilibria (see for example Todaro 1969; Harris and Todaro 1970; Borjas 1989; Hatton and Williamson 1998). We follow this tradition and for mulate a gravity model to examine the impact of economic, political, demo graphic and environmental factors on the migration decision.

II.1. The dependent variable

In our model we analyze the flow of migrants from low and middle income coun tries to high income OECD countries over the period 1970 2000. The classifica tion is based on the World Bank's country definition for 2000, the most recent year for which comprehensive migration data are available. For a full list of the 20 high income OECD countries and the 101 low and middle income countries please refer to appendix. Özden et al. (2011) provide a panel database of migration, their impressive data effort covers migrant populations for 1960, 1970, 1980, 1990 and 2000 in 175 countries. Migrants are defined as people living in one country but having been born in another country. Thus, the concept of migration is based on where people are born and not on their nationality.

The following figures and tables provide some description of the Özden et al. (2011) data. Figures 1 3 show that while the high income OECD countries only account for 15 per cent of the global population, they are the destination of 56 per cent of all global migrants (originating from low, middle and high income countries).

Figure 1

Global Population, Year 2000. [Colour figure can be viewed at wileyonlinelibrary.com]

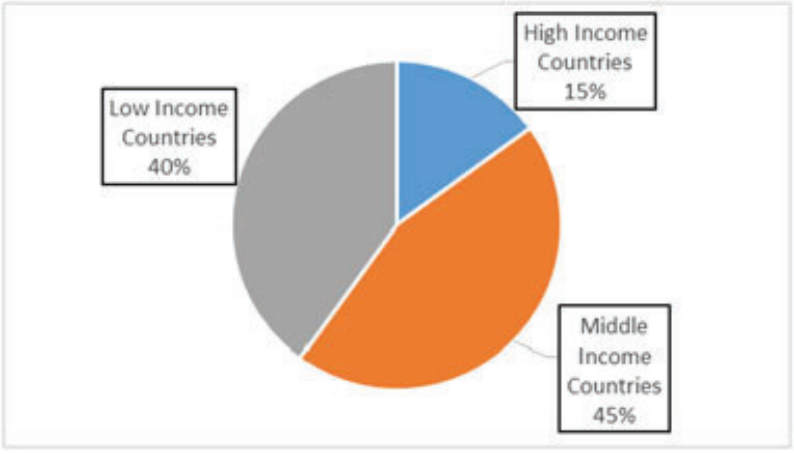


Figure 2

Global Migration by Destination, Year 2000. [Colour figure can be viewed at wileyonlinelibrary.com]

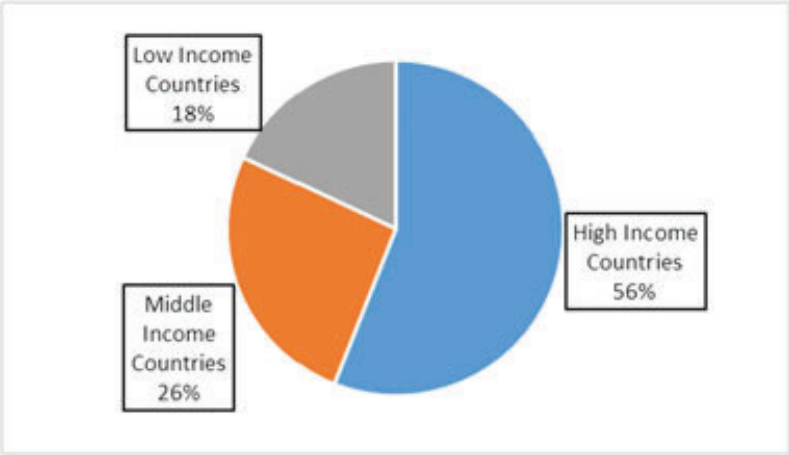
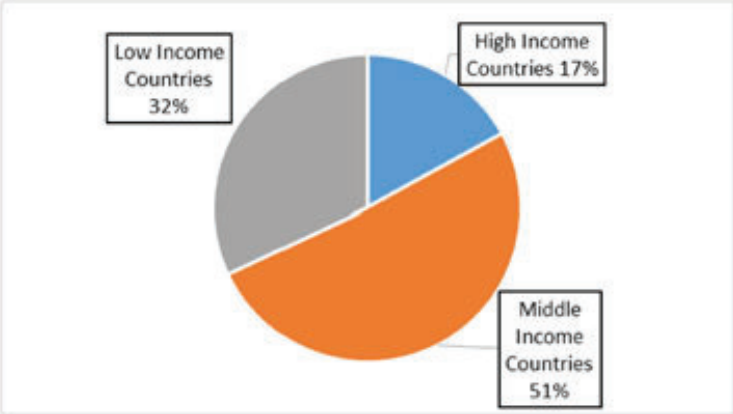


Figure 3

Global Migration by Origin, Year 2000. [Colour figure can be viewed at wileyonlinelibrary.com]



Source: *Özden et al. (2011), Penn World Tables, own calculations. The country classifications are based on the World Development Report 2000/2001, for a full list of countries see the appendix*

Table 1 provides a further breakdown of the origin and destination of migrants. In 2000 about 70 million migrants from low and middle income countries were living in the high income OECD countries. This accounts for about 42 per cent of all global migration. In our analysis we concentrate on the examination of the migrant flows from low and middle income countries to high income OECD countries.

Based on these migrant stock data we calculate the inflow of new migrants. The difference in the stocks between time *t 1* and *t* is not equal to the inflow of new migrants because some of the migrants present at time *t 1* will have

Table 1

Stocks of Migrants in the Year 2000

Origin down/ destination across	To Low income countries	To Middle income countries	To High income countries	Total Migrants (% of total population)	Total population
From Low income countries	21.2	17.9	**13.6**	52.7 (2.2%)	2380
From Middle income countries	8.19	19.7	**56.1**	83.99 (3.1%)	2710
From High income countries	0.71	5.09	22.4	28.2 (3.2%)	881
Total	30.1	42.69	92.1	164.89 (2.8%)	5971

Note: Numbers in millions Source: Özden et al. (2011), Penn World Tables, own calculations. The country classifications are based on the World Development Report 2000/2001.

died, returned or migrated to a different country by time *t*. We do not have sufficient information on return or onward migration but we do have some information on mortality. How many migrants survive from one round of censuses to the next depends on the age structure of the migrant population and the mortality rates of these migrants in their host country. In the Appendix we outline how we used data on the age structure of the migrant stock and host country mortality rates to adjust the flow of migrants. To illustrate our adjust ment we present migration flow data from India to the UK in Table 2, account ing for mortality within the Indian stock of migrants (or diaspora) the adjusted migration flows from India to the United Kingdom are larger than simply taking the difference between the Indian diaspora at the beginning and the end of the decade.

It is worth pointing out that the (adjusted) migrant flow data can take a value of zero and negative values. Many of the migrant stock figures in Özden et al. (2011) are zero. These are either 'true' zeros or are perhaps very small numbers, because in many countries published census data does not provide information on very small groups of people, because they could be identified. For example, there are no recorded migrants from the Central African Republic in Finland for the years 1960, 70, 80 and 90. In our regression sample about eleven per cent of all of the

Table 2

Migration Flows from India to the United Kingdom

	Unadjusted	**Adjusted**
1960-69	143,896	164,251
1970-79	63,202	98,509
1980-89	29,747	58,454
1990-99	92,507	112,270

flow data are based on zero stock data. The difference between the migrant stocks over time can also be negative if a large number of migrants left. One example are Romanian migrants to Germany. In 1990 the stock of migrants was three times higher than in 2000, resulting in an adjusted migration flow of 767,413 people during this period. This is the largest outflow of migrants in our sample.[1]

After this short discussion of the stock and flow data we provide a more formal description of our model. We use data on adjusted migration stocks, mad_{ij}, where i denotes the country of origin and j the destination country. These data are directed dyads because the migration stock for country pair ij is not the same as for country pair ji. Since we only consider cross border migration we do not consider any movements within countries and do not analyze mad_{ii}, mad_{jj} etc. In our migration model we model the change in the stock of migrants, or the flow of migrants, over each decade:

$$mad_{ijt} - mad_{ijt\ \ 1} = \alpha + \beta_0 m_{ijt\ \ 1} + \beta_1 z_{jt} + \beta_2 z_{ij} + \beta_3 w_{ij} + \beta_4 m_{ijt\ \ 1} \cdot z_{it} + \beta_5 m_{ijt\ \ 1} \cdot z_{jt} + \beta_6 m_{ijt\ \ 1} \cdot w_{jjt} + \eta_i + \eta_j + \tau_t + e_{ijt} \quad (1)$$

where i and j are the origin and destination country, respectively and t denotes time (1960, 1970, 1980, 1990 and 2000). The stock of migrants at time t is denoted by m_{ijt} and the difference between the adjusted stock of migrants at time t and $t\ 1$ approximates the flow of migrants ($mad_{ijt}\ \ mad_{ijt-1}$). Our dyadic dataset contains two types of explanatory variables. One set is specific to the country, be it the country of origin or destination, for example income (z_{it}, z_{jt}). Another set of variables is specific to the country pair, such as a common border or colonial history (w_{ij}). We expect that the effect of the country characteristics depends on the size of the existing migrant stock, or the diaspora. We therefore interact the country and country pair characteristics with the diaspora (e.g. $w_{ij} \cdot m_{ijt-1}$). We also include country fixed effects for the origin countries as well as destina tion countries, η_i and η_j, and a time effect, τ_t. We also note that the error terms e_{ij} are likely to be correlated across observations.

II.2. The explanatory variables

We regress the flow of migrants between country dyads on measures of that proxy the incentive to migrate and the cost of migration. The fundamental economic incentive to migrate is the absolute difference in per capita income between the host and home economies. However, while income in the host

[1]We provide additional discussion of the Romanian migrant stock in the appendix. Furthermore, we investigated the importance of outliers and found that excluding these did not substantively change our results.

country has a straightforward incentive effect, that in the country of origin has a dual role since it affects not only the incentive to migrate but the ability to finance the required investment. Hence, rather than use the income difference as an ex planatory variable, we enter incomes in the host and home economies separately. We measure the income variables at the beginning of the decade in order to re duce endogeneity.

While the potential migrant has no choice as to country of origin, the country of destination is a choice variable. The migrant determines where to migrate as well as whether to migrate. Bertoli and Fernández Huertas Moraga (2013) consider these alternative choices: migration between two countries does not depend solely on their relative attractiveness, but also on the one of alternative destinations. Using high frequency data they account for 'mulitlateral resistance' in their model of migration to Spain. With our large n, small t data set we cannot employ their estimation technique but include a relative income variable that captures the alternative choice. In estimating the rate of migration from country *i* to *j*, where *j* is some particular OECD country, we therefore investigate whether the income of *j* relative to other potential OECD destinations has significant effects. We again use lagged values.

Since the decision to migrate is intrinsically forward looking, the pertinent economic variable may be the expectation of future incomes. However, the growth of per capita income in home and host countries over the course of the decade is potentially endogenous to migration during the decade: for example, immigration may raise per capita growth. To control for this we proxy growth of the destination country's economy by the weighted growth rates of the countries which are the markets for its exports.

Analogous to the economic incentive to migrate, there is an incentive to societies that offer greater political freedom. We proxy political freedom by the commonly used polity measure from the Polity V dataset.[2] The measure ranges from +10 (full democracy) to 10 (complete autocracy) and while there is little variation among high income OECD societies, there is considerable variation both between low and middle income countries and over time within them. As with income, the extent of political freedom in countries of origin has ambiguous effects: some restrictions on freedom may directly or indirectly reduce emigra tion. Hence, as with income, we enter the degree of political freedom separately for host and home societies. We also include social disruption as proxied by periods of civil and international warfare.

The investment cost of migration is influenced by a range of variables. The most evident proxy variable is the distance between the country dyads. As noted above, diasporas lower the costs of migration. Relatives established in host

[2]For a full description please refer to Jaggers and Gurr (1995), data source: http://www.systemicpeace.org/polityproject.html, accessed 11 January 2017

countries provide finance for travel, hospitality on arrival, and networks for employment (see Collier 2013, chapter 6 for a discussion). They can also confer the right to migrate through sponsorship, marriage and political influence on regulations. We proxy the size of the diaspora by the stock of immigrants from the country of origin in the host society, measured as discussed above. We also include whether there was a former colonial relationship between the dyads. This may influence decisions through the ease of information flows, or more subtly through expectations set during education, as exemplified by the sense of Britain as the 'home country' for West Indians described in the writings of V.S. Naipaul.

Migration is subject to a range of regulations imposed by host societies. How effective these regulations are is discussed by Mayda (2010). She argues that despite the official 'closed door' policy of Western European countries since the late 1970s migrants have been able to gain entry by using rules on family reunification and political asylum and these rules differ across countries. There is also evidence that changes in immigration regulation have been effective in limiting migration. For example Ortega and Peri (2013) show that non EU migration decreased after the implementation of the Schengen Agreement. Closely related to migration regulation is the attitude towards immigrants. Gorinas and Pytliková (2015) show that attitudes vary across the OECD coun tries and that these variations have a significant impact on migration. We account for the differences in regulation and attitudes in the following way. First, we include an indicator for former colonial relationship, since this affects the attitudes and regulations imposed by host societies. Second, we include a dummy which takes a value of one if the country of origin is landlocked. Other than by direct flight, emigration from such countries requires passage through transit countries which are likely to impose their own regulations, so that the migrant is subject to multiple restrictions. Third, as discussed below, we introduce country fixed effects. These will capture the regulatory and attitudinal differences between OECD societies which are persistent and which are not idiosyncratic as to particular countries of origin.

Finally, we control for time effects by decade dummies, and for the size of country, through the (lagged) populations of host and origin societies. A well established result in the literature is that small countries, such as islands, have high rates of emigration (Brown et al. 2014; Briguglio 1995).

Because migration is a complex decision and subject to a wide array of potential influences, it is valuable to investigate the interactions between variables. However, even though our model is relatively parsimonious in the use of explanatory variables, there are 23 interaction terms. The inclusion of these interaction terms is likely to introduce multicollinearity. In order to decide which interaction terms to retain in our core model we used the automated selec tion process as suggested by Johansen and Nielsen (2009) and Doornik (2009).

This establishes which interaction terms are significant. The results are presented in Table 3. Column 1 is the General Unrestricted Model (GUM) with all interac tion terms included. The more manageable model that resulted from the automated selection process is presented in column 2, which includes 15 interac tion effects.

Our model may suffer from omitted variable bias and we investigated whether including country fixed effects changed our results. In order to examine this, we included all fixed effects for countries of origin and destination in our core model of column 2 and used the automated process to establish which country effects were significant. The resulting model is presented in column 3. It appears that the inclusion of the country fixed effects leaves the coefficient estimates on the explanatory variables and the interaction terms qualitatively unchanged. We found that most of the country of origin fixed effects were insignificant and that only the fixed effects for Hungary, Korea, Libya, the Philippines and Poland survive the automated selection process. In contrast, most destination country dummies are retained, migration differs across the 20 high income destination countries and these differences are not explained in our model. The most likely explanation for these destination country fixed effects is that immigration policies are significantly different across these coun tries. This interpretation is supported by Koopmans et al. (2012), they compare ten West European countries, concluding that immigrant rights differ signifi cantly across countries and that these differences are persistent. Unfortunately, the use of the Koopmans et al data would reduce our sample substantially and we are thus not able to test the impact of different immigration rights more formally.

One concern is that our regressions suffer from endogeneity bias. This bias may be due to omitted variables and/or from reverse causality (Beine et al. 2016 provide a discussion of endogeneity concerns and other issues arising in this type of analysis). Since we account for unobserved origin and destination country effects we are more concerned with the problem of reverse causality. Our migration stock is the accumulated flow of migrants and ideally we should employ an instrumental variable approach to address the endogeneity concerns. Beine et al. 2010 use historical guest worker legislation and historical (civil) wars as instruments for the initial stock of migrants. They find that their results are robust to this instrumentation. We decided not to follow their example because the focus of our analysis is on the interaction of the stock of migrants (the dias pora) and other variables. The application of the Johansen and Nielsen (2009) and Doornik (2009) automated selection process to determine which interactions should be included in the core model does currently not allow for instrumentation.

A further concern is how we should treat the standard errors in this regression. In our dyadic regressions the error terms are likely to be correlated across

Table 3

Analysis of Migration Flows

	(1) All Interaction Terms	(2) Core Model	(3) Core Model plus Country Dummies
Diaspora	4.930 (2.277)**	4.911 (2.127)**	5.192 (2.146)**
Colonial Ties	-436,185 (293,279)	-524,781 (288,616)*	-517,111 (282,860)*
ln Distance (km)	79,893 (54,619)	55,278 (23,047)**	75,351 (24,425)***
Landlocked (dummy)	-1,601 (554)***	-1,428 (525)***	249 (488)
ln Population origin	384 (314)	414 (301)	442 (281)
ln population destination	1,955 (437)***	1,968 (428)***	10,447 (2,742)***
ln GDP origin t-1	4,409 (8,325)	253 (379)	1,002 (412)**
ln GDP destination	64,035 (48,845)	46,405 (23,649)**	69,373 (27,980)**
Growth origin	6,082 (5,476)	422 (222)*	466 (238)*
Growth destination	4,576 (22,057)	1,260 (654)*	-3,536 (1,408)**
Relative GDP	-24,378 (62,618)	11,270 (16,293)	12,364 (21,062)
Polity origin	2,267 (2,041)	280 (72.2)***	309 (70.86)***
Polity destination	-295 (1,921)	229 (117)**	-145 (175)
Civil War	1,176 (1,133)	1,204 (1,150)	570 (1,157)
International War	2,213 (3,101)	2,289 (3,027)	2,841 (2,997)
Colonial·ln GDP origin	8,654 (3,745)**	8,276 (3,724)**	7,996 (3,647)**
Colonial·ln GDP destin.	37,028 (29,469)	47,239 (29,119)	47,241 (28,583)*
Colonial·Growth origin	-5,139 (3,729)	-4,839 (3,839)	-4,649 (3,725)
Colonial·Growth destin.	27,877 (14,075)**	27,272 (14,759)*	26,606 (14,311)*
Colonial·Relative GDP	-58,516 (40,085)	-56,274 (39,173)	-58,558 (38,469)
Colonial·Polity origin	-794 (544)	-779 (538)	-617 (538)
Colonial·Polity destin.	1,111 (1,209)		
ln Distance· lnGDP o	-462 (935)		
ln Distance· lnGDP d	-8,606 (5,533)	-6,247 (2,437)**	-8,640 (2,601)***
ln dist·Growth origin	-641 (609)		
ln dist·Growth destin.	-405 (2,434)		
ln dist·relative GDP	4,501 (6,912)		
ln dist·Polity origin	-224 (226)		
ln dist·Polity destin.	53.23 (215)		
Diaspora·ln GDP origin	-0.191 (0.086)**	-0.190 (0.087)**	-0.189 (0.087)**
Diaspora·ln GDP destin.	-0.413 (0.199)**	-0.415 (0.195)**	-0.439 (0.197)**

(Continues)

Table 3. (Continued)

	(1) All Interaction Terms	(2) Core Model	(3) Core Model plus Country Dummies
Diaspora·Growth origin	0.059 (0.103)	0.056 (0.065)	0.047 (0.066)
Diaspora·Growth destin.	-0.009 (0.134)		
Diaspora·relative GDP	0.820 (0.127)***	0.819 (0.129)***	0.829 (0.133)***
Diaspora·Polity origin	0.005 (0.010)	0.005 (0.009)	0.005 (0.009)
Diaspora·Polity destin.	-0.145 (0.066)**	-0.143 (0.065)**	-0.147 (0.064)**
Diaspora· ln Distance	0.206 (0.044)***	0.206 (0.046)***	0.204 (0.046)***
Diaspora·colonial ties	-0.225 (0.115)*	-0.224 (0.115)*	-0.231 (0.114)**
70s decade dummy	14,626 (9,729)	12,197 (11,070)	6,250 (14,718)
80s decade dummy	20,532 (12,180)*	17,497 (14,158)	6,033 (18,998)
90s decade dummy	17,613 (12,555)	14,558 (14,584)	2,623 (19,812)
Hungary			-20,622 (5,024)***
Korea			10,987 (6,106)*
Libya			-11,259 (2,858)***
Philippines			10,243 (4,172)**
Poland			-17,519 (16,209)
Australia			21,236 (5,469)***
Austria			16,664 (5,582)***
Canada			12,313 (3,759)***
Denmark			21,662 (6,625)***
Finland			23,615 (6,958)***
Greece			15,027 (4,829)***
Ireland			27,816 (8,174)***
Netherlands			11,037 (4,043)***
New Zealand			36,204 (9,158)***
Norway			23,342 (7,103)***
Portugal			15,650 (5,565)***
Sweden			18,508 (5,594)***
Switzerland			15,764 (6,253)**
Constant	-660,333 (481,855)	-469,627 (215,951)**	-803,757 (254,769)***
Observations	6379	6379	6379
R-squared	0.84	0.84	0.85

Notes Dependent variable is migration flow, robust standard errors in parentheses
*significant at 10%;
**significant at 5%;
***significant at 1%

observations.[3] Possible treatments include clustering the standard errors by destination, by origin or by dyad. We present these regressions in appendix table A3. Comparison across the different standard error treatments suggests that our results are robust to different ways of clustering. In particular the coefficient on diaspora and the interaction terms remain qualitatively unchanged.

II.3. Discussion

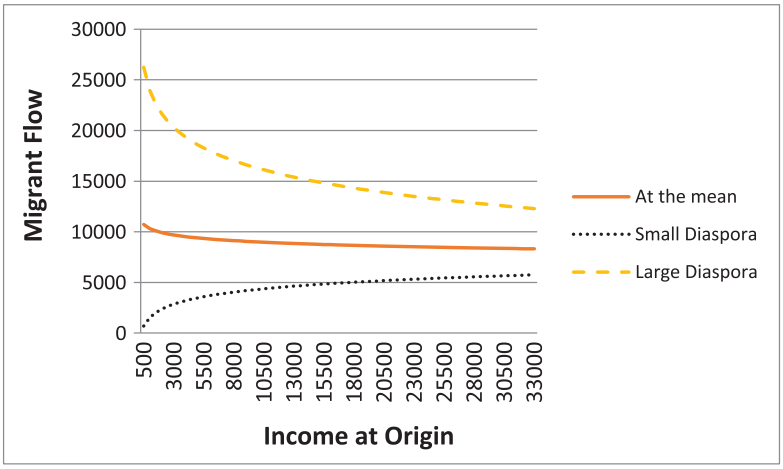
The overall influence of each variable is the sum of its direct and interaction effects. We find that interactions are central to the migration decision. Since it is difficult to discern from the regression tables the net effects of variables that operate through several interactions, we show the salient relationships through graphs.

As suggested by previous authors the effect of the diaspora on migration flows is large (e.g. Beine et al. 2011; Munshi 2003; Massey et al. 1993). However, its overall effect is determined through a series of interactions with other variables. Of the 15 interaction terms included in our core model, eight are interactions with the diaspora. Figures 4 to 6 bring out the salient net effects. In Figure 4 the flow of migration is shown as a function of the level of income in the country of origin. The three lines show this relationship for different sizes of diasporas. The flat black line shows the relationship at the mean of all variables: income in the country of origin has little effect. This confirms the suggestions by De Haas (2007) that the level of development has little impact on the migration. The upward sloping dotted line shows the relationship if the diaspora is small. Now, income in the country of origin is strongly supportive of migration: the ability to finance the investment cost evidently predominates over the incentive effect. Finally, the downward sloping dashed line shows the relationship if the diaspora is large. Now, higher income in the country of origin reduces migration: the incentive effect predominates over the enhanced ability to finance investment. Our interpretation of these substantial differences is that the diaspora is critical in reducing the investment cost that must be borne by the migrant, and so substitutes

[3]Consider for example individuals from destination country i that have decided to migrate to country k rather than to country l, this implies that the u_{ik} residual is negatively correlated with the u_{il} residual: if migrants choose k they cannot by assumption choose l. Since they are observationally equivalent, their probability of being chosen is the same, but they cannot both be chosen. By the same argument, it follows that u_{ik} is negatively correlated with all u_i residuals while u_{il} is positively correlated with all residuals other than k: since k was selected, all other alternative countries have not been chosen. The pattern of correlation across residuals is thus not the standard one encountered in fixed effects and random effect models, where it is usually acceptable to focus on positive correlation across residuals. For a detailed discussion see Fafchamps and Gubert (2007). They also suggest a correction of dyadic correlation but we cannot apply their method because we only consider migration from low and middle income countries to high income countries. For the Fafchamps and Gubert method we would have to consider all possible dyads, e.g. also migration from high income countries.

Figure 4

Income, Diaspora and Migrant Flow. [Colour figure can be viewed at wileyonlinelibrary.com]

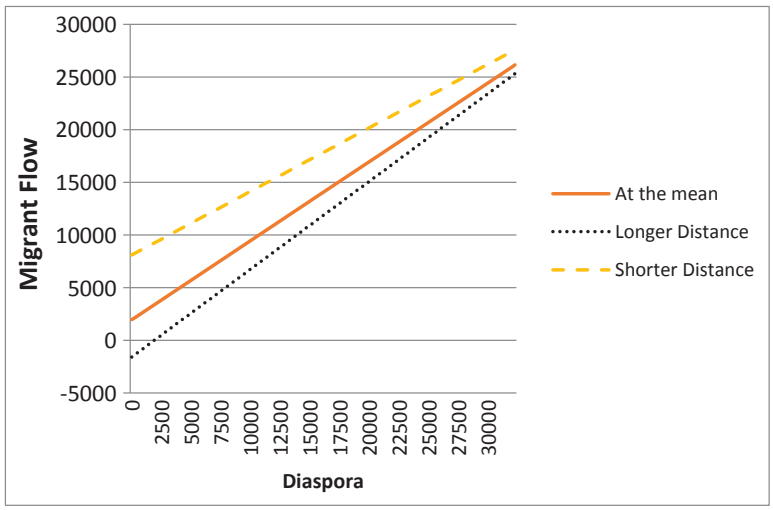


Note: Income per capita in constant USD, base year 2000, PPP adjusted

for income in the country of origin. A similar implication follows from Figure 5, which depicts the flow of migration as a function of the size of the diaspora. At the mean of all other variables (the black line) the diaspora substantially increases migration. The slope of this line is 0.749, suggesting that for an extra 1,000 persons in the diaspora, the flow in the following decade will increase by about 749 migrants. Although we use different data and employ a different model and estimation strategy, this diaspora effect is comparable to the one found by

Figure 5

Distance and Migrant Flow. [Colour figure can be viewed at wileyonlinelibrary.com]

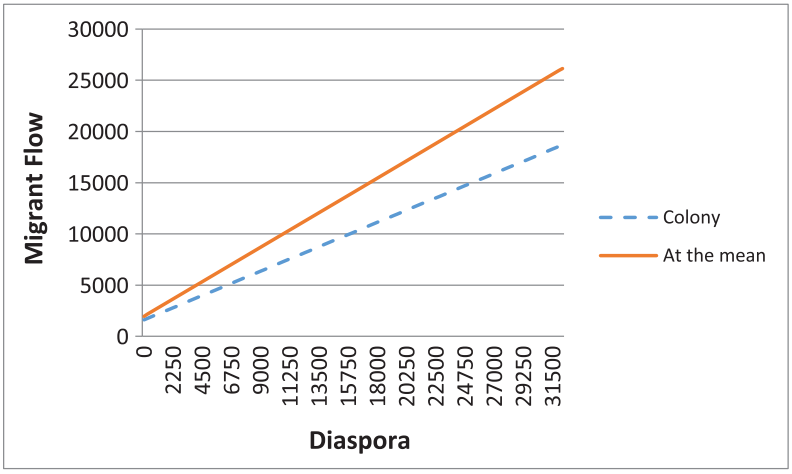


Beine et al. (2011). They regress the change in the stock of migration between 1990 and 2000 on the stock of migrants in 1990. The coefficient on this lagged stock of migrants takes values of 0.669 to 0.766 (Their Table 1, column 3 and ro bustness checks in Table 2). This is a very large effect and an obvious concern is that since the diaspora is measured as the accumulated stock of past migration it spuriously includes all those unobserved effects which are time invariant. This concern is mitigated because we are controlling for fixed effects in both the country of origin and the host country; because we are controlling for the past re lationship of the country dyad; because the effect of the diaspora is not direct but comes about only through interactions with other variables that have clear eco nomic interpretations; and because they are consistent with the recent literature. The effect of the diaspora is considerably more powerful if migration requires a long distance (the dotted line) rather than a short distance (the dashed line). Dis tance is an important impediment to migration, but since long distances increase the investment cost, diasporas have a more important role. Finally, Figure 6 depicts the same relationship but for former colonies (the dashed line). A former colonial relationship reduces the importance of diasporas, our interpretation being that it substitutes for the information flow provided by diasporas.

While a key attraction of OECD countries is evidently their level of income, the small variations between them and the small changes over time do not appear to be important. In our results, the significant and positive direct effect of the level of income is offset by interaction terms. Similarly, growth during the decade in the destination country has only one weakly significant effect, through an interaction with a former colonial relationship. Growth during the decade in the country of origin is also weakly significant, increasing migration. This is consistent with the effect on the ability to finance investment predominating over

Figure 6

Colonial Past and Migrant Flow. [Colour figure can be viewed at wileyonlinelibrary.com]



the incentive effect. However, it should be recalled that we proxy the destination economy's growth rate by the weighted growth rates of its export partners.

To summarize, the economic aspects of the migration decision appear to be determined by the tension between the prospect of reaping a large economic gain and the impediment of financing the required initial investment. Diasporas are critical in overcoming the investment costs and hence elevating the importance of the incentive constituted by the income gap.

Whereas the economic effects are powerful but complex, the effect of political openness is weak but straightforward. There are only small differences among OECD societies in measured political openness and these have small, though statistically significant effects. Political openness in countries of origin is more important, but greater openness increases emigration rather than reducing it. Evidently, variations in political rights include the right to emigrate and the right to information pertinent for the migration decision.

Turning to our proxies for variations in restrictions on migration, both a former colonial relationship and whether the country of origin is landlocked are significant. However, quantitatively the effects are minor, and that concerning whether the country of origin is landlocked is likely to be spurious since it appears to be driven entirely by Hungary.

III. MIGRATION AND CULTURAL DISTANCE

We now introduce the possible effects of cultural difference into our analysis. In the economic literature Spolaore and Wacziarg (2009, 2013) have pioneered the use of the objective measurement of cultural distance between pairs of countries by means of the proxy of neutral genetic divergence.[4] They (2013:18) define genetic distance as a "summary measure of differences in allele frequencies between populations across a range of neutral genes (chromosomal loci) … [it] captures the length of time since two populations became separated from each other." In less technical language, this describes the distance from some neutral genes that occur over time. Over time genetic mutations occur and are passed down throughout the generations. When part of the population migrates and set tles in different world regions, genetic differences develop within these different regional populations. The genetic distance measure is smaller if this population separation (through migration) occurred more recently. This distance measure can therefore be interpreted as a proxy for the relatedness between populations. The measured genetic differences are neutral in the sense that they have no direct bearing upon human performance, but they are reasonable proxies for the time during which different pairs of human cultures have been separated and so have presumably diverged. Like Spolaore and Wacziarg (2009, 2013) we concentrate

[4]The original genetic distance data were collected by Cavalli-Sforza, Menozzi, and Piazza (1994, pp. 75–76).

on the genetic distance that takes into account that a country's population is not homogeneous but made up of different groups. We therefore use the "weighted genetic distance," representing the *expected* genetic distance between two randomly chosen individuals, one from each country, using the genetic distances associated with their respective ancestor populations (Spolaore and Wacziarg 2013:18 Footnote 17).

One concern is that the populations of our twenty destination countries may be very similar in terms of genetic make up. The majority of these countries are European and may be genetically very close. Since many citizens in the USA, Australia and New Zealand have European ancestors, their genetic make up may also be very similar. It may therefore be difficult to examine genetic distance as a determinant of the migration decision. However, there is considerable vari ation in the data as Table A5 in the Appendix suggests. Comparing the 20 desti nation countries, it appears that Japan is very different from the other emigration destinations, followed by Finland and the USA. This variation suggests that ge netic distance may be a useful variable in adding explanation to our core model.

Spolaore and Wacziarg (2009) have used their measure of the cultural distance between countries to demonstrate that it can be a barrier to the flow of knowledge. While their concern is the adoption of new technologies, potentially the same knowledge barrier impedes the movement of people. We test this by introducing cultural distance as an explanatory variable. Our results are reported in Table 4. In column 1 we repeat our core model for ease of comparison. Our measure of the genetic distance between countries is not available for as many pairs of countries as our other data and so the number of observations is approx imately halved. However, it remains substantial at 3,022 and when our core regression is rerun on this reduced sample the results are very similar: the size of the diaspora continues to have a large and significant effect. We report these results in column (2). At the mean of all the variables our core model predicts a flow of 9,760 migrants whereas the results from the reduced sample in column (2) predict one thousand more migrants (10,760). Having established the differ ence of predicted outcomes given different sample sizes we then introduce the genetic distance variable in column (3). We find no significant direct effect of genetic distance on the rate of migration. Migration from poor countries to rich ones appears not to be impeded by the gap in cultures.

However, cultural distance can potentially affect migration not only as a direct barrier, but indirectly as a barrier to the integration of immigrants into their host society. Potentially this barrier can affect migration by reducing the pace at which the diaspora decays through integration into mainstream society. This is, for example, an implication of the Dunbar constant, which posits that the number of other people with which an individual interacts is approximately the same across societies (Dunbar 1992). Recall that we have found further evidence for the proposition of Beine et al. (2011) that diasporas are critical adjuncts to the

Table 4

Migration and Cultural Distance

	(1) Core Model	(2) Core Model Reduced Sample	(3)	(4)	(5)
Genetic Distance			1 170 (1 090)	1 686 (0 966)*	0 414 (0 816)
Diaspora Gen Dist			0 0005 (0 0002)**		0 001 (0 000)***
New Diaspora				18 570 (3 835)***	
New Dias GenDist				-0 010 (0 003)***	
Old Diaspora				2 692 (2 637)	
Old Dias GenDist				0 0004 (0 0002)**	
Linguistic Dist					-937 (355)***
Diaspora LingDist					0 003 (0 019)
Diaspora	5 192 (2 146)**	8 072 (2 485)***	8 169 (3 259)**		11 824 (2 926)***
Colonial Ties	-517,111 (282,860)*	-622,813 (283,780)**	-445,214 (318,847)	-305,005 (237,477)	-608,305 (282,663)**
ln Distance (km)	75,351 (24,425)***	79,086 (22,129)***	50,015 (32,101)	-9,964 (23,751)	49,006 (20,784)**
Landlocked	249	130	663	-234	455
(dummy)	(488)	(607)	(864)	(535)	(623)
ln GDP origin	1,002 (412)**	1,585 (451)***	2,687 (907)***	856 (492)*	1,516 (547)***
ln GDP destination	69,373 (27,979)**	48,605 (31,293)	36,866 (39,209)	9,552 (25,212)	20,377 (30,808)
ln Population orig	442 (281)	632 (301)**	1,042 (415)**	407 (246)*	721 (281)**
ln Population dest	10,447 (2,742)***	12,897 (2,475)***	12,502 (3,070)***	9,020 (3,221)***	10,818 (1,968)***
Polity origin	309 (70 86)***	384 (86 91)***	475 (117)***	230 (99)**	242 (79)***
Polity destination	-145 (175)	-233 (149)	-241 (187)	74 (175)	-56 (141)
Growth origin	466 (238)*	716 (305)**	630 (354)*	537 (278)*	692 (269)**
Growth destination	-3,536 (1,408)**	-6,773 (2,369)***	-7,680 (2,647)***	-4,585 (2,614)*	-4,909 (2,011)**
International War	2,841 (2,997)	-1,722 (1,602)	6,316 (5,551)	297 (2,120)	962 (1,707)

(Continues)

Table 4. (Continued)

	(1) Core Model	(2) Core Model Reduced Sample	(3)	(4)	(5)
Civil War	570 (1,157)	1,294 (1,536)	1,608 (2,154)	2,544 (1,524)*	1,826 (1,561)
Relative GDP	12,364 (21,062)	38,648 (32,062)	23,591 (29,109)	-21,101 (15,569)	35,645 (30,152)
Colonial GDP orig	7,996 (3,647)**	7,174 (3,554)**	4,235 (3,769)	2,831 (3,488)	5,076 (3,699)
Colonial GDP dest	47,241 (28,583)*	57,964 (28,448)**	43,836 (33,598)	29,209 (23,690)	56,877 (28,368)**
Colonial Growth o	-4,649 (3,724 751)	-4,720 (3,849)	-6,126 (5,078)	-5,173 (2,736)*	-3,910 (3,563)
Colonial Growth d	26,606 (14,311)*	28,641 (14,204)**	25,226 (18,103)	19,384 (11,781)*	28,339 (13,756)**
Colonial rel GDP	-58,558 (38,469)	-48,835 (38,421)	-61,595 (48,611)	-25,205 (34,581)	-41,106 (37,803)
Colonial Polity o	-617 (538)	-527 (494)	39 (648)	721 (488 413)	-315 (497)
ln Dist lnGDP d	-8,640 (2,601)***	-9,103 (2,365)***	-6,237 (3,500)*	714 (2,499)	-5,796 (2,203)***
ln Dist Growth o	0 047 (0 066)	-0 064 (0 065)	0 039 (0 091)	-0 030 (0 078)	-0 144 (0 067)**
Diaspora rel GDP	0 829 (0 133)***	0 847 (0 092)***	0 784 (0 113)***	0 531 (0 117)***	0 732 (0 086)***
Diaspora ln Dist	0 204 (0 046)***	0 182 (0 072)**	0 047 (0 071)	-0 018 (0 049)	0 043 (0 095)
Diaspora colonial	-0 231 (0 114)**	-0 425 (0 126)***	-0 301 (0 118)**	-0 238 (0 127)*	-0 497 (0 126)***
Diaspora Polity o	0 005 (0 009)	0 009 (0 007)	0 004 (0 009)	-0 003 (0 008)	0 012 (0 008)
Diaspora Polity d	-0 147 (0 064)**	-0 242 (0 080)***	-0 203 (0 059)***	-0 257 (0 093)***	-0 257 (0 081)***
Diaspora ln GDP o	-0 189 (0 087)**	-0 307 (0 065)***	-0 278 (0 103)***	-0 246 (0 083)***	-0 332 (0 073)***
Diaspora ln GDP d	-0 439 (0 197)**	-0 487 (0 208)**	-0 503 (0 269)*	0 145 (0 194)	-0 715 (0 244)***
Observations	6379	3022	3394	3274	3022
R-squared	0 85	0 96	0 88	0 95	0 96

Notes Dependent variable is migration flow, robust standard errors in parentheses All regressions include decade and country dummies
*significant at 10%;
**significant at 5%;
***significant at 1%

migration process. As the diaspora accumulates, migration becomes easier and so increases. While migration adds to the diaspora, there is also an offsetting process of integration of immigrants into the mainstream society of their host country. Gradually migrants lose their connections with their country of origin. We hypothesize that this offsetting effect of integration is lower when diasporas originate from culturally distant countries. This hypothesis was proposed in Collier (2013) but was not tested.

Here, we test the hypothesis through introducing an interaction term between the cultural distance between a host society and a country of origin and our previous estimate of the accumulated stock of migrants from that country. While we have referred to this cumulated stock as the diaspora, our measure has made no allowance for cultural integration.

As reported in Table 4 column 3, whereas the direct effect of cultural distance on migration is small and insignificant, its interaction with the diaspora has a large and significant effect. As postulated, the wider is the gap in cultures the more useful is the accumulated stock of immigrants to subsequent migrants. The effect is substantial. At the mean value of genetic distance (a value of 1,040) the predicted flow of migrants during a decade is 12,066. If we consider a smaller genetic distance (reducing the distance by one standard deviation, 599) this predicted value is 8,347 and if we consider a larger genetic distance (plus one standard deviation) this predicted value is 15,784. As we noted in the previous section, while the apparent direct effect of the diaspora might be contaminated by the effect of omitted variables, interaction effects are less exposed to this critique. In the present case, we have the same measured diaspora having powerfully different effects on migration due to observed differences in cultural distance. If cultural distance affects migration by slowing the rate of decay of the diaspora (due to less integration), this is potentially discernible by decomposing accumulated migration into vintages. In Table 4 column 4 we use the information on migration flows to split the diaspora into the stock of migrants who have been living in their country of destination 20 years or longer (‘old diaspora’) and the more recently arrived (‘new diaspora’). The interaction of cultural distance with the old diaspora is positive and statistically significant. We interpret this as further evidence that culturally distant diasporas are less likely to integrate into their host societies and thus still provide powerful networks for the latest migrants.

While the above table shows the single decade effect of cultural distance, the impact of the diaspora on migration cumulates over several decades. In our core model, adding 100 migrants to the stock in 1960 would cumulate by 2000 to 823 immigrants as the higher initial stock increases the flow in the next decade, which in turn increases the stock, in a process that echoes down the decades. Re estimated on the reduced sample for which it is feasible to include genetic distance this baseline effect is somewhat larger at 1247. Around this baseline if

we repeat the comparison of plus or minus one standard deviation variation in genetic distance the cumulated stocks of migrants by 2000 are respectively 2213 and 637. Thus, setting all other variables to the same values, a culturally distant country of origin will over the course of 40 years have sent approximately triple the number of immigrants to a host country as a culturally proximate one.

We now turn to robustness checks. We control for linguistic distance and use an alternative measure of genetic difference (which we present in the appendix[5]). The results remain qualitatively similar. In Table 4, column 5 we include a mea sure of linguistic distance. These data were also used by Spolaore and Wacziarg (2009) but were originally compiled by Fearon (2003). Using ethnographic data from Ethnologue, Fearon classifies languages into common families and describes the relatedness of world languages in a language tree.[6] Languages develop over time and space. If two languages share many common nodes in a historic language tree, these languages are more likely to trace their roots to a more recent common ancestor language. The number of common nodes in the linguistic tree is the measure of linguistic similarity. We converted this measure of similarity into a measure of genetic distance. We include this linguistic distance measure and its interaction with the diaspora (column 5). The correla tion between genetic and linguistic distance is relatively low at $\rho=0.24$, thus the two variables measure different things.[7]

Linguistic distance has quite different effects from cultural distance, mirroring some of the recent results by Adsera and Pytlikova (2015). Whereas a wide cul ture gap does not directly impede migration, a wide language gap is a significant direct impediment. There are several potential routes by which this may operate: for example, information flows may be more limited, or anticipated earning power in a host society may be reduced.[8] However, the width of the language gap does not significantly interact with the diaspora. Again this is in contrast to the effect of the culture gap.

The contrast between the effects of linguistic distance and genetic distance, the latter being used as a proxy for cultural distance, is striking. Based on the regres sion presented in column 5, a one standard deviation increase in genetic distance *increases* the migration flow by 27 per cent, whereas a one standard deviation

[5]Cavalli-Sforza et al. (1994) offer alternative measurements of genetic distance, mainly through weighting population groups differently.

[6]See https://www.ethnologue.com, accessed 31st March 2014.

[7]This is also confirmed in the recent work by Alesina et al. (2016) where they examine the differential effect of birthplace, ethnic (linguistic) and genetic diversity on output.

[8]Isphording and Otten (2013) and Isphording (2015) show that linguistic distance impedes migrants' language acquisition in the US, Germany and Spain. Dustman and van Soest (2002) provide evidence that language proficiency has a significant impact on the earnings of immigrants.

increase in linguistic distance *reduces* the migration flow by 16 per cent. Clearly, these distinctive effects warrant further research, specifically on whether genetic distance is a good proxy for cultural distance, and if so how more specifically cultural distance affects migration.

IV. CONCLUSION

Recent advances in global data sets have greatly enhanced the scope for the quantitative analysis of international migration. We have focused on migration from developing to developed countries. During the first half of the 20th Century such migration was negligible, but since 1945 it has continuously increased and is now substantial. By using global data since 1960 we are able to analyze this process of escalation. Developed and developing countries are characterized by three substantial differences: in incomes, in political rights, and in cultures. There are now objective measures for each of these differences and our analysis is able to integrate them. We find that migration is driven by interactions between a few key variables.

One powerful effect is that diasporas increase migration. Most plausibly this is through reducing the investment costs that migrants must surmount in order to access the markedly higher incomes that can be earned in OECD countries. For example, diasporas are more important if migration requires a long journey, and if incomes are low in the country of origin. As migration feeds diasporas and diasporas facilitate subsequent migration, this accounts for the observed tendency of migration to accelerate.

The other powerful effect stems from cultural differences. We find no evidence that a wider gap in cultural differences decreases migration. On the contrary, we find that a wider gap between the cultures of the host society and the country of origin *increases* migration. However, the migration enhancing effect is solely due to the interaction between cultural distance and the accumulated stock of migrants: the wider the gap the more powerfully does past migration facilitate further migration. This effect appears to be stronger for most recent migrants, they are most important to attracting new migrants. We have interpreted these results as being due to the effect that although cultural difference might make migration more difficult, recent migrants provide a more effective bridge into the new host country than migrants who arrived a long time ago. Over time migrants are absorbed into their host society and hence loosen their connections with their country of origin and are thus less useful for new migrants. This interpretation is amena ble to investigation in case studies. More generally, now that the scope for global aggregate analysis has opened up, there is a need for close integration with field studies.

REFERENCES

Adsera, Alicia, and Mariola Pytlikova (2015). The Role of Language in Shaping International Migration, *The Economic Journal.* 125: F49-F81.

Alesina, Alberto, Johann Harnoss, and Hillel Rapoport (2016). Birthplace diversity and economic prosperity, *Journal of Economic Growth* 21.2: 101-138.

Beine, Michel, Frédéric Docquier, and Çağlar Özden (2011). Diasporas, *Journal of Development Economics* 95: 30-41.

Beine, Michel, Frédéric Docquier, and Çağlar Özden (2010). Diaspora Effects in International Migration. Key Questions and Methodological Issues, *Swiss Journal of Economics and Statistics.* 146: 639-659.

Beine, Michel, Simone Bertoli, and Jesús Fernández-Huertas Moraga (2016). A practitioners' guide to gravity models of international migration, *The World Economy.* 39: 496-512.

Bertoli, Simone, and Jesús Fernández-Huertas Moraga (2013). Multilateral resistance to migration, *Journal of Development Economics* 102: 79 100.

Borjas, George J. (1989). Economic theory and international migration, *International Migration Review*: 457-485.

Brettell, Caroline B. and Hollifield, James F. (eds) (2008). *Migration Theory: Talking Across Disciplines,* 2nd edition. New York: Routledge.

Briguglio, Lino. (1995). Small island developing states and their economic vulnerabilities. *World Development* 23: 1615-1632.

Brown, Richard P., Connell, John and Eliana V. Jimenez-Soto (2014). Migrants' remittances, poverty and social protection in the South Pacific: Fiji and Tonga, *Population, Space and Place.* 20: 434-454.

Cavalli-Sforza, Luigi Luca, Paolo Menozzi, and Alberto Piazza (1994). *The History and Geography of Human Genes*. Princeton, NJ: Princeton University Press.

Collier, Paul (2013). *Exodus: Immigration and Multiculturalism in the 21st Century*. Oxford: Oxford University Press.

Constant, Amelie F., Liliya Gataullina, and Klaus F. Zimmermann (2009). Ethnosizing immigrants, *Journal of Economic Behavior and Organization* 69: 274-287.

Dunbar, Robin I.M. (1992). Neocortex size as a constraint on group size in primates, *Journal of Human Evolution* 22: 469-493.

De Haas, H. (2010). *Migration Transitions: A Theoretical and Empirical Enquiry into the Developmental Drivers of International Migration*. International Migration Institute, University of Oxford, IMI Working Paper 24/2010.

De Haas, Hein (2007). Turning the tide? Why development will not stop migration, *Development and Change* 38: 819-840.

Doornik, Juergen. 2009. Autometrics in Castle, Jennifer L. and Neil Shephard (eds.), *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry*. Oxford: Oxford University Press.

Docquier, Frédéric and Hillel Rapoport (2012). Globalisation, Brain Drain, and Development., *Journal of Economic Literature* 50: 681-730.

Dustmann, Christian and Arthur van Soest (2002). Language and the earnings of immigrants. *Industrial & Labor Relations Review* 55: 473-492.

Fearon, James D. (2003). Ethnic and Cultural Diversity by Country, *Journal of Economic Growth*. 8:195 222.

Fafchamps, Marcel and Flore Gubert (2007). The formation of risk sharing networks, *Journal of Development Economics* 83: 326 350.

Gächter, Simon and Jonathan F. Schulz (2016). Intrinsic honesty and the prevalence of rule violations across societies, *Nature* 531.7595: 496-499.

Gorinas, Cedric and Mariola Pytliková (2015). The Influence of Attitudes toward Immigrants on International Migration, *International Migration Review*. Fall: 1-36.

Harris, John R., and Michael P. Todaro (1970). Migration, unemployment and development: a two-sector analysis, *The American Economic Review.* 60: 126-142.

Hatton, Timothy J., and Jeffrey G. Williamson (1998). *The Age of Mass Migration: Causes and Economic Impact*. New York: Oxford University Press.

Isphording, Ingo Eduard and Sebastian Otten (2013). The costs of Babylon linguistic distance in applied economics, *Review of International Economics* 21: 354-369.

Isphording, Ingo Eduard (2015). What drives the language proficiency of immigrants?. *IZA World of Labor WP* 177.

Jaggers, Keith and Ted Robert Gurr (1995). 'Tracking democracy's third wave with the Polity III data', *Journal of Peace Research* 32: 469 482.

Johansen, Søren and Bent Nielsen (2009). An Analysis of the Indicator Saturation Estimator as a Robust Regression Estimator in Castle, Jennifer L. and Neil Shephard (eds.), *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry,* Oxford University Press, Oxford, UK.

Koopmans, Ruud, Ines Michalowski and Stine Waibel (2012). Citizenship Rights for Immigrants: National Political Processes and Cross-National Convergence in Western Europe, 1980-2008, *The American Journal of Sociology* 117:1202-1245.

Lee, Everett S. (1966). A theory of migration, *Demography* 3: 47 57.

Massey, Douglas S., Joaquin Arango, Graeme Hugo, Ali Kouaouci, Adela Pellegrino and J. Edward Taylor (1993). Theories of international migration: A review and appraisal, *Population and Development Review* 3:431-466.

Mayda, Anna Maria (2010). International migration: A panel data analysis of the determinants of bilateral flows, *Journal of Population Economics* 23: 1249-1274.

Meng, Xin and Robert G. Gregory (2005). Intermarriage and the economic assimilation of immigrants, *Journal of Labor Economics* 23: 135-174.

Munshi, Kaivan (2003). Networks in the modern economy: Mexican migrants in the US labor market, *The Quarterly Journal of Economics.* 118: 549-599.

Ortega, Francesc and Giovanni Peri (2013). The effect of income and immigration policies on international migration. *Migration Studies.* 1: 47-74.

Özden, Çağlar, Christopher R. Parsons, Maurice Schiff, and Terrie L. Walmsley (2011). Where on Earth is Everybody? The Evolution of Global Bilateral Migration 1960 2000, *World Bank Economic Review* 25: 12-56.

Ravenstein, Ernest George (1885). The laws of migration. *Journal of the Royal Statistical Society* 48: 167 227.

Ravenstein, Ernest George (1889). The laws of migration. *Journal of the Royal Statistical Society* 52: 214 301.

Spolaore, Enrico and Romain Wacziarg (2009). The Diffusion of Development, *Quarterly Journal of Economics* 124: 469 529.

Spolaore, Enrico, and Romain Wacziarg (2013). How deep are the Roots of Human Development?, *Journal of Economic Literature* 51: 1 45.

Todaro, Michael P. (1969). A model of labor migration and urban unemployment in less developed countries, *The American Economic Review.* 59: 138-148.

Năstu ă, Sebastian and Ancu a-Daniela Tompea (2009). Romania, chapter 9 in Fassmann, Heinz, Ursula Reeger, and Wiebke Sievers (eds). *Statistics and Reality: Concepts and Measurements of Migration in Europe.* Amsterdam: Amsterdam University Press.

Zimmermann, Laura, Klaus F. Zimmermann and Amelie Constant (2007). Ethnic Self-Identification of First-Generation Immigrants, *International Migration Review* 41: 769-781.

SUMMARY

Using a gravity model, we examine the dynamics of migration from developing to OECD countries. Origin and destination countries are characterized by substantial differences in incomes, political rights and cultures. Incentives as well as costs shape the decision to migrate. One powerful dynamic effect is that diasporas increase migration, mainly because they lower the cost of migration. Diasporas assist the next wave of migrants by overcoming the high cost of the emigration, in particular when the origin country is far away and poor. The interaction between the diaspora and cultural distance is also significant. Diasporas in culturally distant countries appear to be particularly useful in overcoming the cost of migration. We interpret this as evidence that culturally distant diasporas are less likely to integrate and maintain closer links with their country of origin, while diasporas from culturally similar countries are more likely to integrate and thus be less useful to potential new migrants.