CESifo WORKING PAPERS

7950
2019

November 2019

Does Birthplace Diversity Affect Economic Complexity? Cross-Country Evidence

Dany Bahar, Hillel Rapoport, Riccardo Turati

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 7950*
Category 6: Fiscal Policy, Macroeconomics and Growth

Does Birthplace Diversity Affect Economic Complexity? Cross-Country Evidence

Abstract

We empirically investigate the relationship between a country's economic complexity and the diversity in the birthplaces of its immigrants. Our cross-country analysis suggests that birthplace diversity is strongly and positively associated with economic complexity. This holds particularly for diversity among highly educated migrants and for countries at intermediate levels of economic complexity. The results are robust to accounting for previous trends in birthplace diversity as well as to using alternatives diversity measures. We address endogeneity concerns by instrumenting diversity through predicted stocks from a pseudo-gravity model as well as from a standard shift-share approach. Finally, we provide evidence suggesting that birthplace diversity boosts economic complexity by increasing the diversification of the host country's export basket.

JEL-Codes: F220, O310, O330.

Keywords: economic complexity, birthplace diversity, immigration, growth.

Dany Bahar
The Brookings Institution
USA - Washington DC 20001
dbahar@brookings.edu

Hillel Rapoport
Paris School of Economics
Université Paris 1 Panthéon-Sorbonne
France - 75014 Paris
hillel.rapoport@psemail.eu

Riccardo Turati
UCLouvain / IRES / LIDAM
Belgium – 01348 Louvain-la-Neuve
r.turati@uclouvain.be

November 2019
The usual disclaimer applies. Hillel Rapoport acknowledges support from Agence Nationale de la Recherche under the framework of the Investissements d'avenir program (ANR-17-EURE-001).

1 Introduction

Diversity of immigration (or immigrants' birthplace diversity) has been shown to boost economic performance at different levels of aggregation: countries (Ortega and Peri (2014), Alesina *et al.* (2016) and Bove and Elia (2016)), regions (Trax *et al.* (2015)), US metropolitan areas, counties and states (Ottaviano and Peri (2006), Ager and Brückner (2013), Fulford *et al.* (2017), Docquier *et al.* (2019)), or firms (e.g., Parrotta *et al.* (2014), Trax *et al.* (2015)).[1] The dominant interpretation for the positive association between birthplace diversity and economic performance has to do with the complementarity in skills and knowledge sets brought about by immigrants who grew up in different environments, went to different school systems, and learned different trades and skills as a result. As immigrants expand the set of skills a country can access to, by the same token they also expand its opportunities to become competitive in a broader set of economic activities. One way in which economists have recently measured the extent to which an economy is capable of becoming competitive in a more diverse set of industries is by quantifying its "economic complexity" (Hausmann and Klinger (2007), Hidalgo *et al.* (2007), Hidalgo and Hausmann (2009)). This paper empirically explores the relationship between a country's population birthplace diversity and its levels of economic complexity.

During the last decades, virtually all modern economies have become more demographically diverse. While the global share of immigrants has remained stable over time –at about three percent of the world population– the share of the foreign-born population in OECD countries has increased from seven to thirteen percent, on average, between 1970 and 2010. Among them, some countries experienced a dramatic growth of the share of foreign born over the 1970-2010 period. For example, Spain (from 1.1 to 15 percent), Greece (from 1 to 10 percent), and Portugal (from 1 to 9 percent).[2] Moreover, the composition of the foreign-born population in OECD countries has changed mostly due to immigration from developing countries, with the average share of non-OECD migrants growing by a factor of 8.7 over the 1970-2010 period against just 1.2 for OECD migrants over the same period.

Our focus on economic complexity is based on the findings of a burgeoning literature which has documented how more complex economies tend to enjoy higher levels of income and better growth prospects (e.g. Hidalgo and Hausmann (2009)). Exploring the different ways through which countries achieve higher complexity, however, is still an open question. A common interpretation of the drivers of economic complexity is the availability of accumulated unobserved capabilities, or productive know-how, which in turn is reflected in the composition of a country's export basket, namely its level of diversification and its uniqueness. Achieving an export basket that is more diversified and includes less ubiquitous products (and, hence, generates higher prospects of economic growth), conceptually, would imply drawing from a more diverse set of complementary skills in the economy. Following this line of thought, we investigate whether economic complexity is posi-

[1] At very low levels of aggregation such as teams, there are many instances of negative outcomes, especially when focusing on ethnic diversity in contexts of conflict (e.g., Hjort (2014), Lyons (2017)). In this paper we focus on birthplace diversity, rather than other measures of diversity used in the literature, such as ethnic (Alesina and La Ferrara (2005)) or linguistic (Desmet *et al.* (2012)) diversity.

[2] The OECD countries with the highest share of foreign-born in 2010 are Israel (36 percent), Luxembourg (33 percent), Australia (25 percent) and Canada (21 percent), according to the World Bank's Bilateral Migration Database (available here).

tively affected by higher birthplace diversity of the local labor force. In essence, the more diverse a country is in terms of its population's birthplaces, the more accumulated knowledge there is in terms of idiosyncratic individual characteristics to draw from. As long as these characteristics and skills are complementary one to the other (and to the native set of skills and knowledge), they should translate into higher levels of economic complexity, income per capita, and economic growth.

In this paper, we put these ideas to the test. In particular, we gather data on bilateral migration stocks between countries to construct measures of immigration diversity and investigate how these relate to measures of economic complexity (i.e., to the "Economic Complexity Index" of Hidalgo and Hausmann (2009), Hausmann and Hidalgo (2011)). Our baseline regressions use the Herfindahl index of diversity (a measure of the likelihood that two randomly drawn immigrants are from different countries), however in our robustness analysis we use alternative measures of diversity such as the Theil index and its components. We go beyond Alesina *et al.* (2016) as we look at a potential driver behind their results, showing that birthplace diversity affects complexity which in turns affects growth. Moreover, since economic complexity (as measured in Hidalgo and Hausmann (2009)) is a combination of countries' export diversification and abilities to produce more unique products, we are also able to decompose the relation between diversity and complexity, shedding light on the mechanisms through which diversity ultimately affects economic growth.

We find that, on average, countries with one standard deviation from our sample mean in terms of birthplace diversity exhibit higher levels of economic complexity by 0.18 standard deviations above the mean. These results are particularly strong when limiting our birthplace diversity measure to skilled migrants (i.e., to migrants with college education or more), consistently with the idea that the relationship between diversity and complexity is driven by skill-complementarity. Our results are robust to controlling for the diversification level of countries' export baskets and their income levels. In other words, countries with higher birthplace diversity of their populations tend to be more economically complex, regardless of their income level and the baseline composition of their export baskets. Furthermore our results are particularly strong among developing countries that are in the middle of the economic complexity distribution, suggesting that an increase in skill variety is particularly relevant for boosting countries development path. We also find that our results are particularly strong for inflows of immigrants from origin countries with little to no prior immigration in the host country of interest.

We further show that our results are robust to using the Theil index of diversity and its decomposition into a between and a within component. We find that most of the effect is driven by the between component, which can be interpreted as the extensive margin of diversity in the sense that it captures the contribution of new origin countries to the evolution of immigration diversity. This is consistent with our interpretation that the relation between diversity and complexity is fundamentally about expanding the set of skills and knowledge to be combined in production. Our results are robust to controlling for potential omitted factors related to immigration diversity such as transfers of technological norms from origin to destination countries with an origin-specific effect (Valette (2018)), or to controlling for birthplace diversity of immigrants in 1960 (thereby ruling out the possibility that the results are driven by the diversity of the second and third generations of immigrants). Finally, we are able to rule out our results being driven by time-invariant unobserved heterogeneity when looking at a short panel of high income countries.[3] We also test the robustness of our cross-country analysis once we control for time-invariant country-specific factors, showing that our estimates are still robust and unlikely explained by missing unobserved factors.

In an attempt to go beyond correlation and deal with endogeneity concerns, we propose two instrumental variable approaches. The first is based on the prediction of skill-specific bilateral stocks of immigrants through a gravity model (see Alesina *et al.* (2016) and Docquier *et al.* (2019)). Including a set of common gravity controls from the trade literature (such as distance, common language, sharing borders, etc.) and interactions between origin and year dummies to capture all the push-time variant factors, we then predict skill-specific bilateral stocks. With those predicted stocks we built our first set of instrumental variables, as skill-specific predicted birthplace diversity indexes. The second approach is based on the shift-share instrument (Card (2001) and Moriconi *et al.* (2018)) to predict the skill-specific bilateral immigration stocks based on the location of immigrants by origin countries in the 1970s and the aggregate skill-specific flows from origin countries. The predicted stocks are then determined by the previous "enclaves" of immigrants and not by recent economic/social factors. Therefore, we argue, our instruments based on this methodology are less prone to reverse causation bias. Moreover, since important historical and migration specific events happened which are orthogonal to the cross-country distribution of immigrants in the 1970s and produced exogenous variation of migration flows,[4] the 1970 distribution of immigrants can reasonably be less correlated with country persisting factors. Those exogenous shocks were indeed large and strengthen the reliability of our approach, since the correlation between our instrument and unobserved persistent countries factors is mitigated. Combining both approaches (gravity model and shift-share based instruments), our estimates support a causal interpretation for the effect of diversity on economic complexity, in particular among developing countries.

The rest of the paper is organized as follows. Section 2 discusses the data sources and presents some stylized facts on the evolution and distribution of Economic Complexity and birthplace diversity. Section 3 presents our empirical approach and our strategies to deal with endogeneity. Section 4 shows the main results of the analysis. In Section 5 we discuss the robustness of our results. Section 6 discusses candidate mechanisms. Finally, Section 7 concludes.

[3] The only data source that provides skill-specific bilateral stocks of migration is the ADOP (2015) dataset over the 1990-2000 period and the Database on Immigrants in OECD countries (DIOC) over the 2000-2010 period. Only high-income countries are available in both databases.

[4] Some examples are the fall of the Soviet Union and resulting opening of the Iron Curtain in 1989, the creation of the European Union and of the Schengen Area that allowed free movement of Europeans in 1995 and the recent implementation of the Hart-Cellar Act which changed the quota system in the US and became effective at the end of 1968.

2 Data and Stylized Facts

Our paper investigates the relation between the birthplace diversity of immigrants and the level of economic complexity of the receiving countries. Combining different sources, our main analysis covers a sample of 100 countries over the period 1990 to 2000. In Section 2.1 we describe our data on skill-specific immigration and various measures of birthplace diversity, and in Section 2.2 we describe the measures of economic complexity used. In Section 2.3 we present some stylized facts that include suggestive evidence of a relationship between birthplace diversity and complexity, as a preamble to our econometric estimations.

2.1 Immigration and Diversity Indexes

Our data source on migration stocks comes from Artuc et al. (2015) (henceforth ADOP (2015)). This dataset provides a square matrix of bilateral migration stocks for 195 countries for the years 1990 and 2000. It also contains migration stocks disaggregated by education level of migrants, splitting them between college and non-college educated, as well as disaggregated by gender. In the dataset, an immigrant is defined as a 25 year or older foreign-born individual. Due to the high coverage of destination and origin countries and the skill composition of the stocks, this data set is increasingly used in cross-country analysis (e.g., Docquier *et al.* (2015), Alesina *et al.* (2016), Bahar and Rapoport (2018)). In keeping with the literature on birthplace diversity reviewed in the introduction and following the decomposition of birthplace diversity over the total population suggested by Alesina *et al.* (2016), we compute a country-level fractionalization index based on the immigrant population by skill group.[5] Namely, given destination country d, origin country $o \in O = \{1, ..., 195\}$ and year t, our immigrant birthplace diversity index is:

$$Div^{Mig}_{d,t,s} = \sum_{o}^{O} \overline{m}_{d,o,t,s}(1 - \overline{m}_{d,o,t,s}) \tag{1}$$

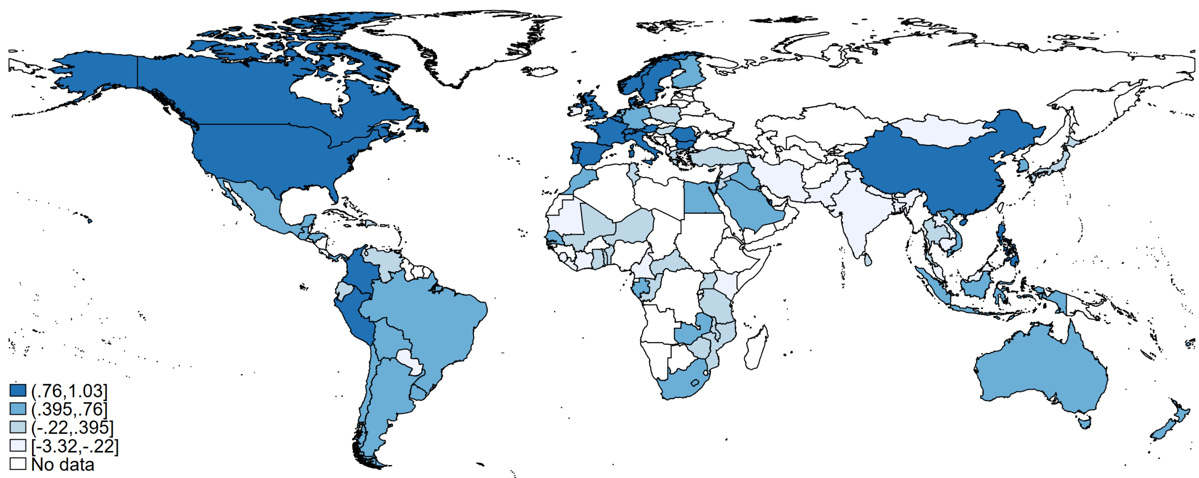
where $\overline{m}_{d,o,t,s}$ is the stock of immigrants from country of origin o in destination country d at year t with education $s \in E = \{All, HS, LS\}$ over the total stock of immigrants in country d with the same education level. Thank to the decomposition of the fractionalization index and including simultaneously in our empirical strategy both the immigration birthplace diversity index and the immigration share allows us to control both for the composition and the size of immigrants population. Moreover, as it is pointed out by the literature, the correlation between the immigration share and the same birthplace diversity index computed on the total country population rather then immigrant population is extremely high and close to one even in our sample (0.981).

By construction, our immigrant birthplace diversity index is between 0 and 1, and it measures the probability to draw randomly two individuals from the immigrants population that are born in different countries.

[5] Alesina *et al.* (2016) show that a birthplace fractionalization index computed over the total population can be decomposed in two components: the between component, which captures the diversity driven by the overall immigrants population and natives, and the within component, which captures the diversity within immigrants population. The correlation between the former component and the immigration share is extremely high and around 0.98.

A higher value for the index implies a more diverse immigrant population. To simplify further analysis, we standardized our measure of diversity with mean 0 and standard deviation equal to 1. Figure 1 plots the standardized level of birthplace diversity across countries. As can be seen, the birthplace diversity of immigrants can be quite high in countries with large foreign-born populations (e.g., in Western countries) or in countries where the share of foreign-born is small (e.g., China). In fact, there is little correlation between the share of immigrants in the host-country population and their diversity (see Alesina *et al.* (2016)).

Figure 1: Birthplace Migration Diversity 2000



Note: authors' calculations on ADOP (2015) data. The map plots the standardized value of the birthplace diversity index on the overall population of migrants in year 2000.

Computing a skill-specific fractionalization index over the immigrants population is not the only way to measure countries' diversity. An alternative and opposite approach is to measure countries polarization rather than diversity. The intuition is that a more polarized population can be associated with a reduction of social cohesion and public good provision. Since the highest level of polarization index is associated with a population characterized by two homogeneous groups-of-origin immigrants, a highly polarized society implies less diversity,and so a less diverse set of skills and competences. Following Montalvo and Reynal-Querol (2005) and Ager and Brückner (2013) we compute a skill-specific polarization index over the immigrants population which captures how far the population distribution is from a bimodal distribution, which is written:

$$Pol^{Mig}_{d,t,s} = 1 - \sum_{o}^{O} \left(\frac{0.5 - \overline{m}_{d,o,t,s}}{0.5} \right)^2 \overline{m}_{d,o,t,s}. \tag{2}$$

Since we compute it over the immigrant population, this index reaches its maximum level when a country's immigrant population is composed by two groups of equal size. Table B-1 shows that skill-specific fractionalization and polarization are negatively correlated. Moreover, if our prior is that the positive relation between population diversity and economic complexity is driven by the expansion of the set of skills and competences brought by different immigrants, we should find a negative relation between the birthplace polarization index and economic complexity. We will test this hypothesis and we will show it later on in our results.

Another alternative index of population diversity is the Theil index. Adopted in the trade literature (e.g., Cadot *et al.*, 2013), such index is used to compute measures export diversification using as inputs a vector of a country's per-product export shares. We can apply this index, too, to measure diversity by computing a skill-specific Theil index over the immigrant population as follow:

$$Theil^{Mig}_{d,t,s} = \frac{1}{N}\sum_{o}^{O}\left(\frac{MIG_{d,o,t,s}}{\mu_{d,t,s}}\right) ln\left(\frac{MIG_{d,o,t,s}}{\mu_{d,t,s}}\right) \tag{3}$$

where $MIG_{d,o,t,s}$ is the total stock of immigrants from origin country o in destination country d at year t with education s. The number of immigrants countries of origin is represented by $N = 195$, while $\mu_{d,t,s}$ is the average skill-specific size of immigrants group.[6] Compared to our benchmark fractionalization index computed in equation (1), the Theil index has two substantial differences. First, the Theil index is equal to zero when we have perfect diversification within the immigrants population, implying that immigrants are perfectly distributed among countries of origin. For this reason the Theil index is negatively related with our fractionalization index, as Table B-1 shows. Second, the Theil index can be decomposed along two dimensions, allowing us a better understanding of the relation between diversity and economic complexity. Following Cadot *et al.* (2011), we decompose the Theil index in two additive components: the between-origin component ($Theil^{B,Mig}_{d,t,s}$) and within-origin component ($Theil^{W,Mig}_{d,t,s}$). The former captures the extensive margin of the Theil Index, which implies a variation in the number of origin countries represented in the immigrants population, while the latter capture the intensive margin, which is driven by the change of balance among immigrants groups. Analyzing the relation between the Theil index (and its components) and economic complexity can help decompose the relationship we are studying by looking at whether economic complexity correlates differently with a more diverse migrant population in terms of its composition (e.g., the distribution of the shares of each group) or, rather, the number of origin countries the migrants are coming from.

2.2 The Economic Complexity Index

Measuring a country's capabilities to produce and export, and quantify its impact on future economic growth, is challenging. Measures of productivity, based on the seminal Solow residual (Solow, 1956) to more advanced methodologies that incorporate socioeconomic indicators to measure countries' growth capacity are still the subject of numerous studies.

Our paper focuses on one of those measures that has been at the core of a burgeoning literature related to economic growth: the Economic Complexity Index (ECI), developed by Hidalgo and Hausmann (2009). The ECI aims to capture a country's accumulated capabilities or productive know-how. Countries with highly diversified export baskets which in turn include industries exported by fewer countries have higher levels

[6] Namely, we compute $\mu_{d,t,s}$ as follow: $\mu_{d,t,s} = \frac{1}{N}\sum_{o}^{O} MIG_{d,o,t,s}$.

of economic complexity.Hidalgo and Hausmann (2009) document that higher levels of economic complexity positively relate to income per capita it has a particularly strong explanatory power in predicting future economic growth (as compared to other long-established determinants of economic growth).

For our exercise we collect yearly ECI data for 222 countries from 1962 to 2016, from the Atlas of Economic Complexity (Hausmann *et al.* (2014)).[7] As explained by Hausmann *et al.* (2014), the ECI is constructed using exports data for each country and year on nearly 800 tradable industries categorized using the SITC 4-digits classification.

As explained in Hidalgo and Hausmann (2009), the ECI is computed using product level export value data as its main input. It organizes the export data for every year in the form of a matrix M_{dp}, sized $d \times p$, where a cell is 1 if country d exports product p with comparative advantage, and 0 otherwise. To quantify whether a country exports a product with comparative advantage, the authors rely on the measure Revealed Comparative Advantage (RCA), originally suggested by Balassa (1965). The RCA for country d and product p for any year is computed according to the following formula:

$$RCA_{d,p} \equiv \frac{exp_{d,p} \sum\limits_{p} exp_{d,p}}{\sum\limits_{d} exp_{d,p} \sum\limits_{d} \sum\limits_{p} exp_{d,p}}$$

where $exp_{d,p}$ is the exported value of product p by country d. Thus, to create the M_{dp} matrix, it is assumed that country d exports product p with comparative advantage if its RCA is 1 or more. This threshold can be interpreted as that country exporting that product in higher relative proportion than the World as a whole.

As a first step to compute ECI, M_{dp} is used to quantify two indicators. First, for each country, the diversification of a country's export basket, measured as the number of products in which the country holds a RCA equal or greater than 1. Second, for each product, the number of countries that export such product with a RCA of unit or more. The indexes of country diversification ($K_{d,0}$) and product ubiquity ($K_{p,0}$) are defined as follow:

$$K_{d,0} = \sum_{p} M_{dp} \tag{4}$$

$$K_{p,0} = \sum_{d} M_{dp} \tag{5}$$

Equations (4) and (5) are the building blocks of the measure of economic complexity: countries with a more diverse export basket of less ubiquitous products are characterized by a high degree of economic complexity. The intuition is that both aspects (diversification and less ubiquitous products) requires more know-how and capabilities (thus, countries export more varieties) that, in turn, are more exclusive and rare (thus, the varieties a country exports, on average, are less common).

The calculation of the ECI as explained by Hidalgo and Hausmann (2009) relies on the application of

[7]Data are available from `atlas.cid.harvard.edu`

what they define as the *Method of Reflections*. The construction of ECI is an iterative process. The method starts with $K_{d,0}$ in its first iteration, as a measure of a country's complexity simply by counting the number of products it exports (in a given year). Then, it incorporates information of each one of those products using $K_{p,0}$, and looks at the number of products d exports weighted by the ubiquity of each one of them. It then incorporates the average diversification of the countries that export the same products as d, and iterates again and again, until converging. As explained by Hidalgo and Hausmann (2009), the nth iteration of this measure is:

$$K_{d,n} = \frac{1}{K_{d,0}} \sum_{p} M_{dp} \times K_{p,n-1} \tag{6}$$

Where,

$$K_{p,n-1} = \frac{1}{K_{p,0}} \sum_{d} M_{dp} \times K_{d,n-2} \tag{7}$$

Substituting (7) in (6), we have:

$$K_{d,n} = \frac{1}{K_{d,0}} \sum_{p} M_{dp} \times \frac{1}{K_{p,0}} \sum_{d'} M_{d'p} \times K_{d',n-2} \tag{8}$$

As shown by Hausmann *et al.* (2014), this expression can be rewritten as:

$$K_{d,n} = \sum_{d'} \widetilde{M_{dd'}} \times K_{d',n-2} \tag{9}$$

Where,

$$\widetilde{M_{dd'}} = \sum_{p} \frac{M_{d,p} M_{d',p}}{K_{d,0} K_{p,0}} \tag{10}$$

Hausmann *et al.* (2014) note that equation (9) is satisfied when $K_{d,n} = K_{d,n-2}$ (e.g., convergence is achieved). Using matricial algebra, this corresponds to the eigenvector of $M_{dd'}$ associated with the second largest eigenvalue (where captures most of the variance). Following their notation, let's call this eigenvector $\vec{K}$. Thus, the vector of all ECI values (one per country) is computed by standardizing $\vec{K}$:

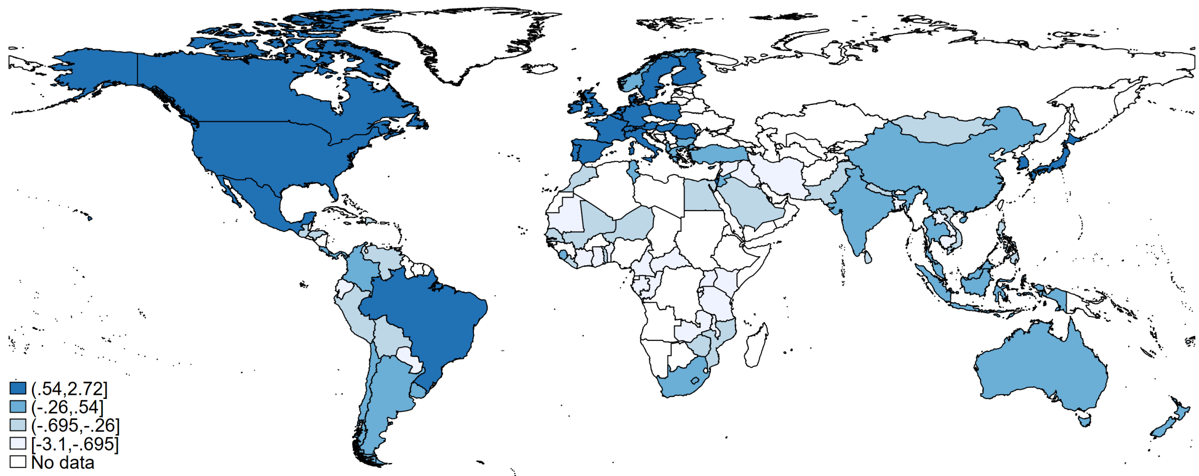
$$ECI = \frac{\vec{K} - < \vec{K} >}{stdev(\vec{K})} \tag{11}$$

Where $< . >$ and $stdev(.)$ represent average and standard deviation, respectively.

This calculation provides one ECI value for each country d in every year, using data on its export basket.

Figure 2 plots the distribution of ECI values per country for year 2000. We can clearly see that the geographic distribution of the intensity in economic complexity is rather similar to the one of birthplace diversity. Moreover, developed high-income societies are characterized by an high level of economic complexity, compared to developing/low-income countries.

Figure 2: Economic Complexity Index 2000



Note: authors' calculations using data from Hausmann *et al.* (2014). The map plot the standardized value of the Economic Complexity Index by country in year 2000.

2.3 Stylized facts

While there is a literature that has focused its attention to investigating the relation between birthplace diversity and several economic outcomes, such as income per capita and productivity, our focus is particularly on economic complexity for three main reasons.

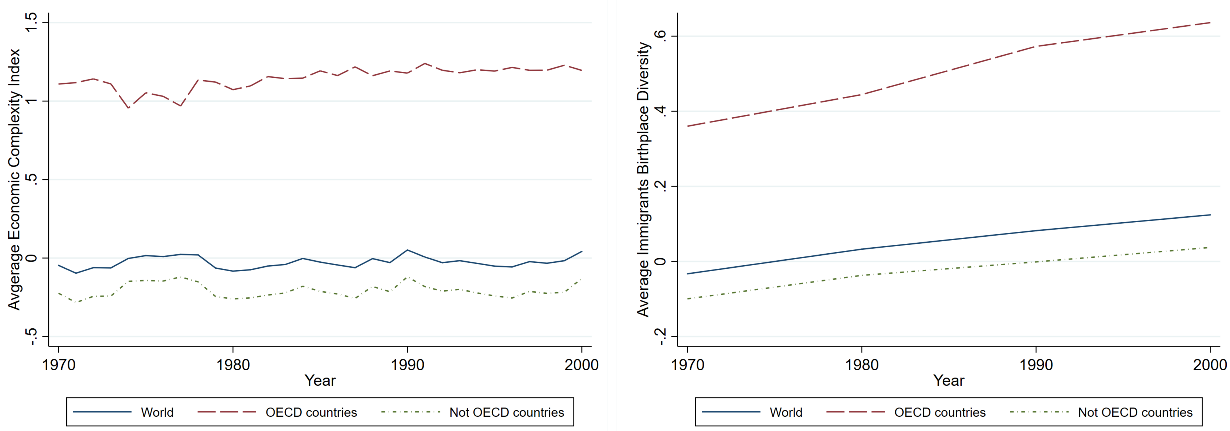
First, we suspect that the measure of economic complexity is prone to be influenced by the economic benefits in terms of skills and knowledge complementarities driven by immigration, in turn affecting other previously studied outcomes such as economic growth. Since countries' economic complexity increases with the accumulation of a country's stock of capabilities or know-how (as interpreted by Hidalgo and Hausmann (2009) and others), then we can explain that birthplace diversity should be positively related with countries economic complexity *if* new skills available in a more diverse society complement the available ones.

Second, the nature of the ECI allow us to decomposed between export diversification and product uniqueness, allowing us to explore mechanisms through which diversity influences countries' complexity.

Third, even though the ECI has been shown to be an important determinant of income and growth, the evidence so far is thin when it comes to determining which factors influence the economic complexity of countries.

Figure 3 plots the evolution of the ECI and immigrants birthplace diversity standardized measures over the period 1970-2000 for the whole World, developed and developing countries. Figure 3(a) shows an overall slight increase of the global average level of economic complexity, although there remain important differences between developed (OECD) and developing (non-OECD) economies. In terms of magnitude, the average level of ECI among developed economies is around ten time larger than the average level of developing economies; while the average yearly growth is larger for developing than for developed economies (1% against 0.4%, respectively). In turn, Figure 3(b) plots average birthplace diversity per group of countries. Birthplace diversity follows a positive trend over time, similar to economic complexity, characterized by huge discrepancy between developed and developing economies.

Figure 3: Evolution of Economic Complexity and Migrants Birthplace Diversity



(a) Economic Complexity

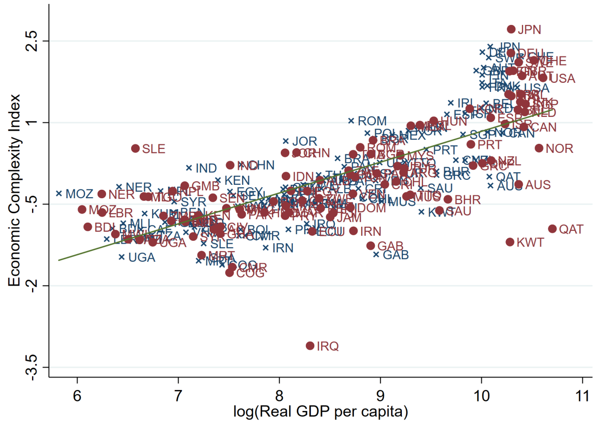
(b) Immigrants Birthplace Diversity

Note: authors' calculations using data from Hausmann *et al.* (2014) and Özden *et al.* (2011). Figure (a) plots the country average level of the standardized Economic Complexity, while Figure (b) plots the country average level of the standardized Immigrants Birthplace Diversity. Both figures show the World level, OECD countries and Not OECD countries.

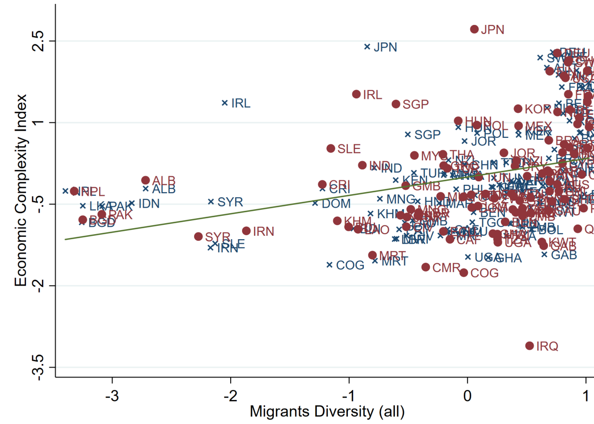
Figure 4(a) plots the country-year standardized level of ECI over logarithm of GDP per capita, showing a positive and clear relation between the level of complexity and growth across countries, consistently with Hidalgo and Hausmann (2009) and Hausmann *et al.* (2014), among others. Figure 4(b) shows a similar although less prominent positive relation between economic complexity and migrants birthplace diversity. This relation remains similar and stable even when we differentiate between highly educated and less educated migrants (Figures 4(c) and 4(d), respectively).

The descriptive evidence, therefore, points toward a clear positive relation between economic complexity and birthplace diversity. However, a question arise: What is the relation between economic complexity,

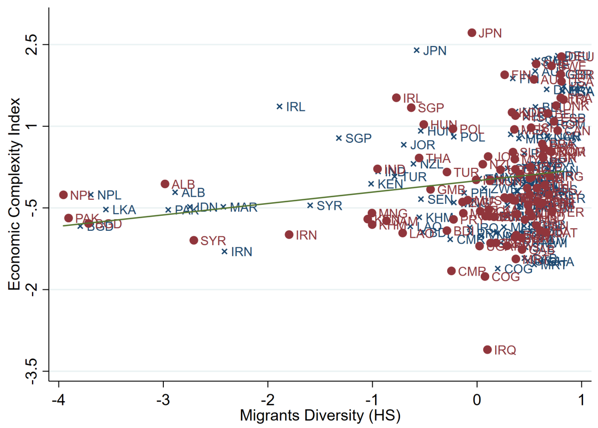
Figure 4: Economic Complexity and Migrants Birthplace Diversity



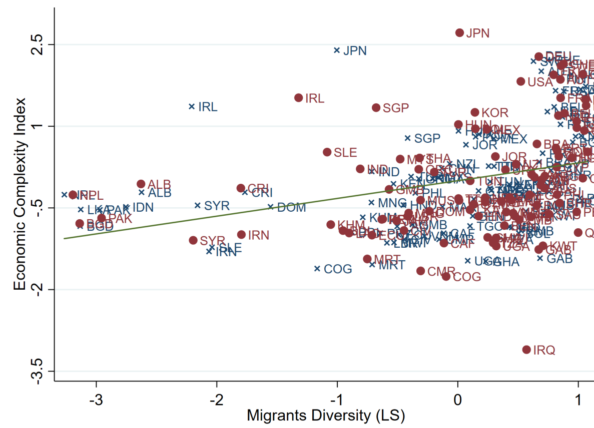
(a) GDP per Capita



(b) All Migrants



(c) HS Migrants



(d) LS Migrants

Note: authors' calculations using data from Hausmann *et al.* (2014) and ADOP (2015). Figures (a), (b), (c) and (d) plot the country level of the standardized Economic complexity index (y-axis) on the country level of the logarithm of real GDP per capita (fig (a)) and on the standardized migrants birthplace diversity index by skill group. All the correlations are statistically significant. Blue cross shows the country-observation in the 1990, while red dot shows the country-observation in the 2000.

diversity and economic growth? Knowing from the literature that birthplace diversity positively affects countries' income, as an additional motivational exercise, we replicate in Table 1 columns (1), (3) and (5) the main analysis of Alesina *et al.* (2016), with our sample of countries. Using their same specification and controls, we consistently estimate a strong and positive relation between diversity and GDP per capita across countries (the latter in log terms).[8] In columns (2), (4) and (6) we include as additional control the ECI of the country. Interestingly the associated coefficient is always positive and significant. Note that when ECI is controlled for there is a significant reduction of the partial correlation of diversity and GDP per capita

[8]The set of controls includes both time variant and time-invariant controls. The set of time-variant controls includes measures of quality of institutions from Polity IV, population size, human capital and trade openness. Time-invariant controls are measures of absolute latitude, average temperature, landlocked dummies, measures of ethnic, genetic and linguistic diversity and measures of local diseases.

of around 20%. This suggests that ECI is, in part, one of the channels through which birthplace diversity affects income as shown in Alesina *et al.* (2016).

In the next section we dig deeper into the relationship between birthplace diversity and economic complexity.

Table 1: Replication of Alesina *et al.* (2016) controlling for Economic Complexity

Dep. Variable:	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS
$ln(GDP)$	*All Skill*		*High Skill*		*Low Skill*	
Div^{Mig}	0.177***	0.141**	0.133***	0.100*	0.167***	0.135**
	(0.052)	(0.055)	(0.051)	(0.054)	(0.051)	(0.053)
Mig	0.227***	0.214***	0.257***	0.232***	0.221***	0.206***
	(0.067)	(0.062)	(0.064)	(0.066)	(0.068)	(0.063)
ECI		0.196**		0.205**		0.197**
		(0.081)		(0.085)		(0.081)
Observations	200	200	200	200	200	200
Countries	100	100	100	100	100	100
Adj. R-Square	0.90	0.90	0.89	0.89	0.89	0.90
Controls	✓	✓	✓	✓	✓	✓
Regional FE	✓	✓	✓	✓	✓	✓

Note: authors' calculations using data from Hausmann *et al.* (2014) and Alesina *et al.* (2016). Standard errors are clustered at country level. * p<0.1, ** p<0.05, *** p<0.01. The dependent variable is the logarithm of real GDP per capita. Each regression includes the full set of controls of Table 3 of Alesina *et al.* (2016).

3 Empirical Strategy

We follow Alesina *et al.* (2016) as benchmark model for our analysis. We first replace their dependent variable with the $ECI_{d,t}$ as defined above. Our baseline specification is:

$$ECI_{d,t} = \alpha + \beta Div^{Mig}_{d,t,s} + \gamma Mig_{d,t,s} + \delta ln(GDPpc)_{d,t} + \Theta X_{d,t} + \Lambda Z_d + \eta_t + \epsilon_{d,t}. \tag{12}$$

Where d indexes a country and t a year. The main variable of interest is $Div^{Mig}_{d,t,s}$, which represents a skill-specific birthplace diversity index among migrants. In particular $s \in \{All, HS, LS\}$, which represent all migrants (All), high-skilled (HS), and low-skilled (LS).

As controls we include a number of variables. First, $Mig_{d,t,s}$, which measures the skill-specific share of migrants in destination country d at year t and skill group s. We also include as an additional control the logarithm of real GDP per capita in the country under consideration (d), which aims to control for the level of development of the countries. We know that, indeed, there is a strong correlation between ECI and income per capita, and thus, by controlling for it, we are estimating whether birthplace diversity can explain higher economic complexity regardless of income per capita. The vector $X_{d,t}$ contains a set of country-time variant characteristics. This exahustive list of controls aims to reduce concerns that our estimates are driven by omitted variable bias from a number of observables. The controls include: (i) quality of institutions measured using the Polity-2 score from the Polity IV database (Marshall *et al.*, 2013), (ii) population size from the UN database (United Nations, 2013) (iii) country-level aggregated years of schooling from Barro and Lee (2013), (iv) real trade openness from the PWT 8.0, (v) trade diversity of imports and exports (measured following Feenstra *et al.*, 2005), and (vi) the weighted average of GDP per capita from all origin countries of migrants to country d using migration shares as weights. The vector Z_d contains a set of country time invariant characteristics, such as: (i) landlocked dummies from Head *et al.*, 2010, (ii) absolute latitude and share of population within 100km of an ice-free coast (iii) average temperature and precipitation, (iv) continent fixed effects, (v) measures of ethnic, linguistic and genetic diversity of the local population from Alesina *e*t al., 2003 and Ashraf and Galor, 2013, and (vi) measures of diseases like malaria, tuberculosis incidence and yellow fever from World Bank, 2013. η_t represents year fixed-effects.

Estimating equation (12) using OLS allow us to obtain an estimate of the partial correlation between economic complexity and migration diversity, captured by the estimate of β. However, in spite of the inclusion of the full battery of time-variant and invariant controls for destination countries specified above, it is likely that unobserved factors captured by $\epsilon_{d,t}$ may be correlated both with immigrants' diversity and economic complexity. In that case our estimated OLS coefficient would be a biased estimate of the causal effect of diversity on economic complexity. To mitigate such potential threat we firstly minimize the potential unobserved factors driven by immigration. To clarify whether the effects of immigration diversity are driven by a diverse set of skills and competences and not by the fact that immigrants are coming from more economically complex societies, we follow the literature on epidemiological effects (Spilimbergo (2009), Docquier *et al.* (2016) and Valette (2018)) by including as additional control an origin-specific effect $\overline{ECI}^{w}_{d,t,s}$, which is the weighted average of ECI at origin. Using as weights the shares of migrants by skill and origin, $\overline{ECI}^{w}_{d,t,s}$ is a measure of exposure to economic complexity driven by international migration of destination country d at year t. Namely, we compute it as follows:

$$\overline{ECI}^{w}_{d,t,s} = \sum_{o} \Big(\frac{MIG_{d,o,t,s}}{\sum_o MIG_{d,o,t,s}} \overline{ECI}^{1990-2000}_{o} \Big) \tag{13}$$

where $\overline{ECI}^{1990-2000}_{o}$ is the average ECI at origin country o over the period 1990-2000. If the effect of immigration is mainly driven by the level of economic complexity brought by immigrants, the estimated coefficient related to birthplace diversity should become statistically not significant once that term is controlled for.

Another potential omitted factor related to immigration and diversity is that immigration can have a persistent/long run effect. If that is the case, the partial correlation β estimated in equation (12) can include also the effects of previously arrived immigrants. To precisely estimate the effect on economic complexity of recently arrived immigrants, we use Özden *et al.* (2011)'s data and compute a measure of immigrants birthplace diversity and share in 1960. Since data on the skill composition of immigrants are not available, we compute it over the whole immigrant population. Those controls will capture whether the effect of immigration are driven by previous cohorts of immigrants or by third/second immigration diversity.

Although we minimize the potential bias driven by omitted factors related to immigration, there could still be omitted factors related to destination countries that will bias our estimates. To deal with this issue, we perform two different analysis. First, we include in our cross-country sample country fixed-effects, accounting for time-invariant country specific unobserved heterogeneity. Being aware of the short time dimension of our data set (T=2), the time-invariant factor can capture the majority of the variation associated to the dependent variable, providing imprecise estimates of the coefficients associated to our main variable of interest. Second, to overcome this issue driven by the reduced time dimension, we merge and harmonize the Database on Immigrants in OECD Countries (DIOC) and the ADOP (2015) data to have a short panel of 22 destination countries covering three time periods (1990, 2000 and 2010).[9] Due to data limitations and the fact that DIOC focuses on developed economies, we cannot cover the same sample as in our main cross-country analysis. However having such panel of countries is quite unique, since for each of them we have the bilateral education specific migration stocks from 194 origin countries,[10] allowing us to build our education-specific birthplace diversity indexes for the relevant immigrant populations. Over this small panel we then estimate the following equation:

$$ECI_{d,t} = \alpha + \beta_p Div^{Mig}_{d,t,s} + \gamma_p Mig_{d,t,s} + \delta_p ln(GDPpc)_{d,t} + \zeta_p HC_{d,t} + \eta_d + \eta_t + \epsilon_{d,t}. \tag{14}$$

The estimated coefficient β_p captures the partial correlation of diversity among migrants on economic complexity after controlling for all time-invariant characteristics of destination countries through the inclusion of country fixed-effects η_d. Moreover we also include in the specification a standardized index of human capital, $HC_{d,t}$, based on the average years of schooling attained from Barro and Lee (2013) and taken from the Penn World Table (Feenstra *et al.* (2015)). Including such index allows us to compare the effects of diversity and human capital on economic complexity.[11] Equation (14) is empirically demanding, given the small number of countries and the short time dimension of the panel (T=3). However finding significant estimates even in such short panel analysis should minimize concerns related to omitted variable bias.

3.1 Identification Strategy

Since immigration is not a random process, another possible threat to our empirical strategy is the presence of a potential reverse causation bias. Countries that are more economically complex are also, on average, richer. Due to better economic conditions, those countries could attract a larger inflow of immigrants coming from a wider range of origin countries. If that is the case, the estimated coefficient β in equation (12) will not give us

[9] The harmonization of those data sets is explained in the Appendix A

[10] The only exception is Germany, since from DIOC we have information only for 51 origin countries.

[11] Docquier *et al.* (2019) show that human capital has a bigger role than birthplace diversity in explaining the economic growth of US States over the period 1990-2010.

any relevant information about the causal relation between diversity and economic complexity. To deal with reverse causation, we follow the literature and use two different two-stage least square estimators to predict skill-origin specific bilateral stocks of migrants. Our first IV strategy is based on Alesina *et al.* (2016), who proposed a gravity model to predict the bilateral stocks of immigrants. To minimize the possible violation of the exclusion restriction, the gravity model proposed is quite parsimonious. The model of bilateral migration is defined as:

$$\begin{aligned} MIG_{d,o,t,s} =& \alpha + \gamma_1\ Pop_{d,1960} + \gamma_2\ Dist_{d,o,t} + \gamma_3\ Border_{d,o} + \gamma_4\ Lang\ Off_{d,o} \\ &+ \gamma_5\ Lang\ ethn_{d,o} + \gamma_6\ Colony_{d,o} + \gamma_7\ Time\ Zone_{d,o} + \zeta_{o,t} + \eta_t + \epsilon_{d,o,t,s} \end{aligned} \tag{15}$$

where $MIG_{d,o,t,s}$ is the stock of immigrants from country of origin o in destination country d with education s at year t. The set of controls includes the population in destination country d in 1960, the bilateral weighted distance, the presence of a common border, dummies for common official and ethnic minority languages, previous colonial ties and time zone differences.[12] We also include year fixed-effects to capture common trends across countries (η_t) and origin-year specific fixed effects, to capture origin country specific trends ($\zeta_{o,t}$). Due to the high number of zeros given by empty bilateral corridors, we estimate equation (15) using a Pseudo-Poisson Maximum Likelihood estimator (PPML) (see Santos Silva and Tenreyro (2006)). We then use the predicted coefficients associated to the estimated gravity equation (15) to predict the skill-specific bilateral stocks of migrants ($\widetilde{MIG}^{Grav}_{d,o,t,s}$) and, with them, we compute the predicted skill-specific birthplace diversity indexes ($\widetilde{Div}^{Grav}_{d,t,s}$) and immigration shares ($\widetilde{Mig}^{Grav}_{d,t,s}$). Since the estimated bilateral stocks are less driven by destination countries economic factors, we then use the resulting predicted diversity and share of immigration as IVs.

The second IV approach is based on a shift-share methodology (Card (2001), Ottaviano and Peri (2006), Docquier *et al.* (2019)). The intuition of this approach is to use past migration settlement patterns as predictor of subsequent migration flows due to network effects. Those predicted flows should be uncorrelated (or at least less correlated) with current levels of economic complexity and development. The variation in aggregate flows of immigrants across origin countries, which is mainly driven by origin-specific push factors, are allocated to our sample of destination countries according to the early distribution of immigrants from the same country of origin. Hence such shift-share instruments produce variation in immigration across destination countries over time due to the interaction between previously established immigrants' settlements and current emigration flows. Following Moriconi *et al.* (2018), we construct skill-specific bilateral stocks taking into account the aggregate variation of immigrant flows by skill and origin, and apply it to the same distribution of immigrants by origin. Using Özden *et al.* (2011) we compute the initial presence of immigrants from origin country o in

[12] Data and methodology from Head *et al.* (2010) and CEPII. The weighted distance is based on distances between the biggest cities in the countries weighted by the share of the city in the overall country population.

destination country d in 1970 as share of the total immigrants from the same origin country as follow:

$$sh^{mig}_{d,o,1970} = \frac{MIG_{d,o,1970}}{\sum_d^D MIG_{d,o,1970}} \tag{16}$$

where $MIG_{d,o,1970}$ is the stock of migrants from origin country o in destination country d in 1970. Then we compute from the ADOP (2015) data for $t \in \{1990, 2000\}$ the total amount of immigrants coming from country o with education s living in our sample of countries as follow:

$$TOT\ MIG_{o,t,s} = \sum_d^D MIG_{d,o,t,s}. \tag{17}$$

Finally we can compute the predicted bilateral stocks of immigrants in destination country d with education s in year t as follow:

$$\widetilde{MIG}^{SS}_{d,t,s} = TOT\ MIG_{o,t,s} * sh^{mig}_{d,o,1970}. \tag{18}$$

We will use the imputed bilateral stocks from equation (18) to compute predicted skill-specific measures of diversity ($\widetilde{Div}^{SS}_{d,t,s}$) and migration shares ($\widetilde{Mig}^{SS}_{d,t,s}$), which will be our instrumental variables. Since the initial distribution of immigrants is not skill-specific, this avoids the potential threat of destination countries specific capabilities to attract low/high educated immigrants. However, as pointed out for example by Goldsmith-Pinkham *et al.* (2018) or by Jaeger *et al.* (2018)), persistent local conditions that could influence initial immigrants location and economic complexity of destination countries could threaten our identification. To reduce this possible threat, we also perform our shift-share approach using the initial distribution of immigrants in an earlier period (namely, 1960).

4 Main Results

Table 2 presents our baseline results from equation (12). The reported coefficients capture the relation between diversity and economic complexity. From column (1) to (4) we use as specification the same model as Alesina *et al.* (2016), while from column (5) to (8) we include also the logarithm of real GDP per capita. Columns (1) and (5) estimates are related to the overall diversity driven by immigrants, while in the other columns we report education-specific estimates. In column (1) we can appreciate a strong and positive correlation between birthplace diversity and economic complexity. Since the variable of interest and the dependent variable are standardized with mean zero and standard deviation equal to one, we can assess the magnitude of the effects. Given all the other factors equal, a one standard deviation increase in the birthplace diversity index is associated with an increase in the economic complexity index of the destination country by 0.18 standard deviation.

A reasonable prior is that if the main contribution of immigration diversity to countries' economic complexity is through skill complementarities, then estimates that rely on birthplace diversity among high-skilled

Table 2: OLS regression on Economic Complexity Index

Dep. Variable: ECI	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS	(7) OLS	(8) OLS
Div^{Mig}_{All}	0.179***				0.125*			
	(0.054)				(0.064)			
Mig_{All}	0.068				-0.002			
	(0.075)				(0.071)			
Div^{Mig}_{HS}		0.160***		0.077		0.120**		0.087
		(0.056)		(0.089)		(0.059)		(0.076)
Mig_{HS}		0.119		0.113		0.042		0.081
		(0.083)		(0.111)		(0.089)		(0.112)
Div^{Mig}_{LS}			0.166***	0.101			0.115*	0.045
			(0.053)	(0.082)			(0.062)	(0.084)
Mig_{LS}			0.074	-0.001			0.006	-0.049
			(0.076)	(0.090)			(0.071)	(0.086)
$ln(GDPpc)$					0.306**	0.300**	0.306**	0.301**
					(0.128)	(0.119)	(0.127)	(0.126)
$HH\ Index\ Trade$	0.010	0.135	-0.009	0.069	0.151	0.216	0.141	0.204
	(0.355)	(0.384)	(0.354)	(0.372)	(0.330)	(0.361)	(0.331)	(0.359)
Observations	200	200	200	200	200	200	200	200
Countries	100	100	100	100	100	100	100	100
Adj. R-Square	0.74	0.74	0.73	0.73	0.75	0.75	0.75	0.75
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Regional FE	✓	✓	✓	✓	✓	✓	✓	✓

Note: authors' calculations using data from Hausmann *et al.* (2014) and Alesina *et al.* (2016). Standard errors are clustered at country level. * $p<0.1$, ** $p<0.05$, *** $p<0.01$. The dependent variable is a standardized measure of the Economic Complexity Index. Each regression includes the full set of controls of Table 3 of Alesina *et al.* (2016).

migrants should result in a higher point estimate. As Alesina *et al.* (2016) and Docquier *et al.* (2019) point out, diversity among college educated immigrants has a positive and significant effect on countries and US states' economic performance while diversity among low-educated immigrants has a smaller and less robust effect. In that sense, investigating the skill-specific relation between diversity and economic complexity provides a more stringent tests of this prior. We present the skill-specific results from columns (2) to (4), and from columns (6) to (8). The relation between economic complexity and skill-specific diversity appears robust across immigrants education groups, and there is no statistical difference between the point estimates of low and high educated diversity indexes. Including both high and low educated birthplace diversity indexes (col. (4)) produces no significant results due to the high correlation between the two measures. However this will no longer be the case (i.e., the two coefficients become statistically different and actually low-skill diversity becomes insignificant) when we conduct a heterogeneity analysis and control for more characteristics in the next set of regressions.

The coefficient for the relative size of the immigrant population (the share of foreign-born) is also positive but not precisely estimated. The inclusion of GDP per capita as an additional control (col. (5) to (8)) results in a reduction of about 25% the size β. However the estimated coefficients remain positive and statistically significant, in particular for diversity among highly educated immigrants.[13] Those results suggest a strong a positive correlation between diversity and economic complexity, regardless of income.[14]

Since our sample includes both developed and developing economies, we can test for non-linearities in the relation between birthplace diversity and economic complexity. Our prior here is that in an economy with a low level of complexity, where the structure of the economy is such that most workers perform manual tasks and work in very small teams or even individually, there may be not much complementarity to expect to start with. Immigrants may well bring new sets of skills and knowledge but these cannot be combined in production due to the structure of the economy. In addition, other complementary inputs such as physical capital, infrastructures or best managerial practices may be missing as well, preventing poor countries from taking advantage of the opportunities linked to diversity.

Table 3: OLS regression on Economic Complexity Index - ECI Terciles

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Dep. Variable: ECI	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
	First Tercile				*Second Tercile*				*Third Tercile*			
Div_{All}^{Mig}	-0.069				0.165**				-0.201			
	(0.140)				(0.071)				(0.147)			
Mig_{All}	0.075				0.065				0.344			
	(0.055)				(0.064)				(0.405)			
Div_{HS}^{Mig}		-0.183		-0.286		0.141***		0.128**		-0.046		0.314
		(0.163)		(0.194)		(0.047)		(0.058)		(0.131)		(0.266)
Mig_{HS}		0.163		0.091		-0.009		-0.072		0.037		0.163
		(0.109)		(0.117)		(0.063)		(0.085)		(0.210)		(0.222)
Div_{LS}^{Mig}			-0.061	0.098			0.142*	0.024			-0.187	-0.462
			(0.142)	(0.095)			(0.078)	(0.087)			(0.128)	(0.287)
Mig_{LS}			0.071	0.078			0.066	0.090			0.253	0.370
			(0.052)	(0.050)			(0.064)	(0.098)			(0.259)	(0.305)
$ln(GDPpc)$	0.012	0.013	0.015	-0.019	-0.030	0.063	-0.013	0.037	1.079***	1.083***	1.065***	0.877***
	(0.110)	(0.113)	(0.111)	(0.119)	(0.115)	(0.074)	(0.118)	(0.125)	(0.214)	(0.202)	(0.219)	(0.213)
Observations	67	67	67	67	67	67	67	67	66	66	66	66
Countries	42	42	42	42	44	44	44	44	36	36	36	36
Adj. R-Square	0.18	0.25	0.18	0.24	0.31	0.37	0.28	0.34	0.80	0.78	0.80	0.80
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Regional FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Note: authors' calculations using data from Hausmann *et al.* (2014) and Alesina *et al.* (2016). Standard errors are clustered at country level. * p<0.1, ** p<0.05, *** p<0.01. The dependent variable is a standardized measure of the Economic Complexity Index. Each regression includes the full set of controls of Table 3 of Alesina *et al.* (2016) plus the logarithm of GDP. The sample of countries is splitted by ECI terciles: first tercile (col. (1)-(4)), second tercile (col (5)-(8)) and third tercile (col. (9)-(12)).

In an attempt to capture potential heterogeneous effects, we therefore split our sample of countries by terciles of ECI and perform a subsample analysis for each tercile of the distribution. Table 3 presents

[13]Note that the inclusion of GDP per capita could be a "bad control" for our analysis if it is determined simultaneously with our index of diversity (see Angrist and Pischke, 2009). Yet, we include it to reduce possible bias due to the established relation between countries' average level of development and their economic complexity.

[14]Since our benchmark analysis includes measures of import and export diversification, our estimates are capturing more than a relation between birthplace diversity and diversification of export and import baskets. Similarly, our estimates are robust to including as an additional control a measure of net inflow foreign net investment. These results are available upon request.

the results by tercile. Columns (1) to (4) present the estimates for the first tercile, columns (5) to (8) for the second tercile and finally columns (9) to (12) for the upper tercile of the distribution. As in Table 2, we present the estimates both for overall immigration diversity and for education-specific immigrants birthplace diversity. There are two main results we can take from Table 3. First, immigrants birthplace diversity, in particular among highly educated immigrants, is positively and strongly associated with economic complexity of countries at intermediate levels of economic complexity (namely, in the 2^{nd} tercile of the distribution). Birthplace diversity among highly educated immigrants is strongly significant also in the more demanding specification where we control for both skill-specific measure of diversity (col. (8)). The correlation between diversity and economic complexity is not significant in the other terciles of the distribution. Those results suggest that diversity contributes to the accumulation of knowledge and abilities required to build up the economic complexity of developing countries but is less important for developed or underdeveloped countries. Second, real GDP per capita is highly correlated with economic complexity only in the upper tercile of the distribution. Intuitively, this result suggests that birthplace diversity is a factor explaining economic complexity but only when the local population has the *right* amount of skills and competences to be complemented by those brought by the immigrants.

We further test whether the positive and significant relation between diversity among immigrants and economic complexity is robust to using alternative measures of diversity. Table B-2 in Appendix B shows the estimates over the sample of countries belonging to the 2^{nd} ECI tercile after including the number of immigrants' countries of origin represented (col. (2)), and after replacing the diversity indexes with polarization indexes (col. (3)) and Theil indexes and its decomposition (col. (4) to (6)). The estimates confirms the direction of the estimates: lower concentration (as measured by the Theil index , which is consistent with high diversity) is associated with higher economic complexity. Moreover, the significant coefficient associated to the between component of the Theil index (col. (6)) suggests that the positive relation between diversity and economic complexity is mainly driven by an expansion of the set of origin countries, which implies an introduction of new skills and competences in the destination country.

As explained in Section 3, the strong partial correlation found between diversity and complexity could be driven by unobserved factors. Immigrants coming from a wide range of origin countries can bring with them a wide set of skills and competences, but they can also bring with them knowledge driven by the economic complexity of their origin country. To clarify whether the effect of immigration on economic complexity is driven by diversity, following Valette (2018) we include in our specification a measure of exposure to economic complexity driven by international migration ($\overline{ECI}_s^w$). Table 4 presents the estimates for the full sample (col. (1) to (6)) and for the subsample of countries in the 2^{nd} tercile of ECI (col. (7) to (12)). Replacing the skill-specific birthplace diversity index with the skill-origin-specific term does not produce any statistically significant relation between the latter term and economic complexity, as odd columns in Table 4 show. Once we include simultaneously the birthplace diversity index and the origin-specific term, the partial correlation between birthplace diversity and economic complexity remains positive and statistically significant. The correlation is stronger among countries belonging to the 2^{nd} tercile of the ECI distribution. Moreover the origin-specific term is not statistically significant in the full sample, while it is negatively correlated with economic complexity in the second tercile subsample. Those results suggest that the effect of immigration on economic complexity is driven by the variety of complementary skills and competences captured by birthplace diversity, and not by a simple transfer of complexity from origin to destination countries. Moreover, the negative and significant estimates associated to the origin-specific term among countries belonging to the 2^{nd} ECI tercile suggests that the beneficial relation between migration and complexity is not necessarily driven by an increase of skills and competences from complex societies, but from less complex counties. However, such negative relation can be also explained by the fact that those countries are attracting more immigrants from low economically complex countries. Investigating the effect of diversity in a panel of the 51 US states, Docquier *et al.* (2019) find a similar negative relation between the origin-specific term and states economic growth.

Table 4: OLS regression on Economic Complexity Index
Origin Specific Effect

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Dep. Variable: ECI	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
Sample:	*Overall Sample*						2^{nd} *Tercile*					
Education	*All*	*All*	*HS*	*HS*	*LS*	*LS*	*All*	*All*	*HS*	*HS*	*LS*	*LS*
Div_s^{Mig}		0.128**		0.123**		0.117*		0.162***		0.157***		0.148***
		(0.064)		(0.060)		(0.061)		(0.052)		(0.045)		(0.050)
Mig_s	-0.018	-0.005	0.026	0.033	-0.008	0.003	0.022	0.033	-0.026	-0.078	0.019	0.034
	(0.075)	(0.071)	(0.094)	(0.093)	(0.076)	(0.072)	(0.068)	(0.056)	(0.070)	(0.060)	(0.067)	(0.057)
$\overline{ECI}_s^w$	0.007	-0.009	-0.002	0.006	0.028	0.015	-0.116	-0.257***	-0.147	-0.233**	-0.121	-0.256***
	(0.111)	(0.110)	(0.112)	(0.109)	(0.113)	(0.114)	(0.086)	(0.078)	(0.127)	(0.088)	(0.076)	(0.079)
Observations	200	200	200	200	200	200	67	67	67	67	67	67
Countries	100	100	100	100	100	100	44	44	44	44	44	44
Adj. R-Square	0.74	0.75	0.74	0.75	0.74	0.75	0.24	0.44	0.24	0.51	0.25	0.43
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Regional FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Note: authors' calculations using data from Hausmann *et al.* (2014) and Alesina *et al.* (2016). Standard errors are clustered at country level. * $p<0.1$, ** $p<0.05$, *** $p<0.01$. The dependent variable is a standardized measure of the Economic Complexity Index. Each regression includes the full set of controls of Table 3 of Alesina *et al.* (2016), excluding the weighted average of GDP per capita at origin countries using migration shares as weights. The skill-specific origin-specific term, computed as the weighted average of the Economic Complexity at the origin, where the weights are the migration shares by origin country (see equation (13)) is included in all the regressions.

Another potential threat that could bias our estimates is the long-run/persistent effect of immigration. As pointed out by Alesina *et al.* (2016), our measure of birthplace diversity could also capture the effect of immigrants previously arrived in the destination country, and so our estimates would capture a combination of diversity driven by recent and older immigration. We thus include as additional control a measure of lagged diversity, to capture the effect of diversity driven by previous cohorts. We use Özden *et al.* (2011) data and construct as additional controls measures of immigrants diversity and shares in 1960. Table 5 presents the estimates on the full sample (col. (1)-(4)) and on the 2^{nd} tercile (col. (5)-(8)). Including past diversity instead of current diversity in columns (1) and (5) shows that past diversity is not significantly correlated with the current level of economic complexity. The inclusion of past diversity as additional control in columns (2) to (4) affects the precision of the estimates related to current diversity in the full sample. This result suggests that part of the strong correlation between current diversity and economic complexity is driven by previous cohorts of immigrants. Not surprisingly, current and past diversity are strongly correlated (at 0.70). However, the estimates for current immigration diversity over the subsample of countries belonging to the 2^{nd} tercile are positive and significant, in particular among highly educated immigrants (col. (6)-(8)). This result confirms the relevant role of diversity in skills and competences driven by immigration in developing economies, expressed by a strong a positive partial correlation between diversity and economic complexity.

Table 5: OLS regression on Economic Complexity Index
Robustness by previous level of diversity

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Variable:	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
ECI								
Sample:	*Overall Sample*				2^{nd} *Tercile*			
Education:	*All*	*All*	*HS*	*LS*	*All*	*All*	*HS*	*LS*
$Div^{Mig}_{All,1960}$	0.107	0.013	0.044	0.027	0.116	-0.025	0.003	0.020
	(0.079)	(0.098)	(0.088)	(0.099)	(0.073)	(0.089)	(0.078)	(0.090)
$Mig_{All,1960}$	0.024	0.044	0.015	0.034	0.063	0.013	0.043	0.017
	(0.076)	(0.088)	(0.080)	(0.089)	(0.073)	(0.086)	(0.062)	(0.092)
Div^{Mig}_{s}		0.116	0.099	0.097		0.182*	0.142**	0.131
		(0.079)	(0.067)	(0.077)		(0.097)	(0.056)	(0.102)
Mig_s		-0.032	0.034	-0.017		0.060	-0.019	0.059
		(0.073)	(0.090)	(0.073)		(0.076)	(0.067)	(0.078)
$ln(GDPpc)$	0.336***	0.307**	0.296**	0.308**	0.079	-0.032	0.056	-0.014
	(0.116)	(0.127)	(0.121)	(0.127)	(0.087)	(0.118)	(0.097)	(0.123)
Observations	200	200	200	200	67	67	67	67
Countries	100	100	100	100	44	44	44	44
Adj. R-Square	0.75	0.75	0.75	0.75	0.21	0.27	0.34	0.24
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Regional FE	✓	✓	✓	✓	✓	✓	✓	✓

Note: authors' calculations using data from Hausmann *et al.* (2014) and Alesina *et al.* (2016). Standard errors are clustered at country level. * p<0.1, ** p<0.05, *** p<0.01. The dependent variable is a standardized measure of the Economic Complexity Index. Each regression includes the full set of controls of Table 3 of Alesina *et al.* (2016).

5 Robustness

After assessing a stable and positive correlation between immigrants diversity, particularly among highly educated immigrants, and economic complexity, this section investigates the robustness of those results.

We first test the robustness of our estimates to time-invariant unobserved heterogeneity, by including in our sample country fixed-effects. The inclusion of time-invariant factors with a reduced time dimension (T=2) is extremely demanding from an empirical point of view, since the majority of the variation can be captured by country fixed-effects. Being aware of that, Table 6 presents the skill-specific estimates over the whole sample (col. (1) to (4)) and over the sample of countries belonging to the 2^{nd} tercile of the Economic Complexity index. Even though the estimates on the overall sample are not precisely estimated, columns (6) and (8) show that the positive relation between diversity among college educated immigrants and economic complexity is robust to the inclusion of time-invariant country-specific factors. However, given the small sample size and the short time dimension, those estimates should be interpreted cautiously.

Table 6: OLS regression on Economic Complexity Index - Country fixed effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Variable: ECI	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
Sample:	*Overall Sample*				2^{nd} *Tercile*			
Div^{Mig}_{All}	0.105				-0.026			
	(0.084)				(0.195)			
Mig_{All}	-0.348*				-0.039			
	(0.206)				(0.451)			
Div^{Mig}_{HS}		0.022		-0.087		0.198**		0.397***
		(0.052)		(0.091)		(0.081)		(0.138)
Mig_{HS}		-0.017		-0.040		-0.483***		-0.658***
		(0.119)		(0.113)		(0.166)		(0.215)
Div^{Mig}_{LS}			0.106	0.175			-0.024	-0.371*
			(0.077)	(0.126)			(0.200)	(0.192)
Mig_{LS}			-0.212	-0.207			0.067	-0.073
			(0.156)	(0.159)			(0.391)	(0.219)
$ln(GDPpc)$	-0.005	-0.090	-0.054	-0.024	0.007	0.041	-0.021	0.117
	(0.192)	(0.220)	(0.193)	(0.186)	(0.223)	(0.131)	(0.223)	(0.125)
Observations	200	200	200	200	46	46	46	46
Countries	100	100	100	100	23	23	23	23
Adj. R-Square	0.93	0.92	0.92	0.92	0.42	0.60	0.42	0.61
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Country FE	✓	✓	✓	✓	✓	✓	✓	✓

Note: authors' calculations using data from Hausmann *et al.* (2014) and Alesina *et al.* (2016). Standard errors are clustered at country level. * p<0.1, ** p<0.05, *** p<0.01. The dependent variable is a standardized measure of the Economic Complexity Index. Each regression includes the full set of controls of Table 3 of Alesina *et al.* (2016) plus the logarithm of GDP. Moreover, each specification includes country fixed effects.

Table 7 presents the results of our short panel over OECD developed countries. Due to data limitations, our short panel cover 22 destination countries[15] over three periods (1990, 2000 and 2010). However, thanks to the bilateral structure and the skill composition available in the harmonized data set we are able to compute skill specific measures of birthplace diversity for each destination country. Odd columns present the estimates of equation (14) without country specific control, while even columns include as additional control measures of human capital and GDP per capita. Moreover all the estimates include year and country fixed

[15] The 22 destination countries in analysis are: Australia, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Japan, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Spain, Sweden, United Kingdom, United States.

Table 7: Panel Analysis on Economic Complexity Index

Dep. Variable: ECI	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS	(7) OLS	(8) OLS
Div_{All}^{Mig}	0.120**	0.159***						
	(0.045)	(0.043)						
Mig_{All}	-0.353**	-0.300**						
	(0.162)	(0.130)						
Div_{HS}^{Mig}			0.085**	0.151**			0.197	0.286*
			(0.041)	(0.066)			(0.145)	(0.155)
Mig_{HS}			-0.132	-0.033			-0.157*	-0.057
			(0.081)	(0.095)			(0.083)	(0.092)
Div_{LS}^{Mig}					0.052	0.104**	-0.076	-0.102
					(0.048)	(0.049)	(0.117)	(0.127)
Mig_{LS}					-0.097	-0.104	-0.153	-0.145
					(0.082)	(0.095)	(0.095)	(0.101)
HC		0.517**		0.559**		0.563**		0.554**
		(0.201)		(0.246)		(0.214)		(0.236)
$ln(GDPpc)$		-0.008		-0.317		-0.033		-0.323
		(0.307)		(0.545)		(0.414)		(0.449)
Observations	66	66	66	66	66	66	66	66
Countries	22	22	22	22	22	22	22	22
Adj. R-Square	0.95	0.96	0.95	0.96	0.95	0.96	0.95	0.96
Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Country FE	✓	✓	✓	✓	✓	✓	✓	✓

Note: authors' calculations using data from Hausmann *et al.* (2014), ADOP (2015), DIOC and Penn World Table. Standard errors are clustered at country level. * p<0.1, ** p<0.05, *** p<0.01. The dependent variable is a standardized measure of the Economic Complexity Index. Each regression includes country and year fixed effects. HC is the standardized measure of human capital (Penn World Table). $ln(GDP)$ is the logarithm of real GDP per capita. The list of countries in the sample is the following: Australia, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Japan, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Spain, Sweden, United Kingdom, United States.

effects, controlling for common trend and for time-invariant country specific factors. The main results of this robustness are two. First, there is a strong and positive correlation between birthplace diversity and economic complexity also in a panel dimension. The results are stronger among highly educated immigrants. Including country specific controls enhances the effect of diversity on economic complexity. Those results should minimize the possible concerns related to an omitted variable bias. Moreover, the negative and statistically significant coefficient associated to the share of immigrants can be explained by countries characterized by a sizeable growth of economic complexity while almost absent or negative growth in terms of immigration share, due to restrictive migration policies (e.g. Japan and Mexico). Second, human capital is another strong and good predictor of economic complexity. Since the measure of human capital is standardized as birthplace diversity,[16] we can compare the magnitude of the two coefficients. Indeed, the effect of human capital on economic complexity are on average 3 times bigger than the effect of birthplace diversity. Those results are

[16]Both measures are standardized with mean zero and standard deviation equal to one.

in line with Docquier *et al.* (2019), which finds, using US data, that the effect of human capital on GDP growth is up to four times larger than birthplace diversity measured on the college educated immigrants.

Table 8: IV regression on ECI

Dep. Variable: ECI	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A - Overall Sample									
Div^{Mig}_{All}	0.221*	0.116**	0.219*						
	(0.113)	(0.059)	(0.114)						
Mig_{All}	0.003	-0.014	-0.003						
	(0.064)	(0.065)	(0.064)						
Div^{Mig}_{HS}				0.260**	0.110**	0.261**			
				(0.115)	(0.054)	(0.119)			
Mig_{HS}				0.045	-0.064	-0.038			
				(0.086)	(0.135)	(0.135)			
Div^{Mig}_{LS}							0.189*	0.109*	0.195*
							(0.108)	(0.057)	(0.107)
Mig_{LS}							0.008	0.020	0.029
							(0.064)	(0.073)	(0.071)
Observations	200	200	200	200	200	200	200	200	200
Countries	100	100	100	100	100	100	100	100	100
K-P rk Wald F-stat	13.97	425.25	7.07	16.77	29.41	8.79	16.49	126.61	8.34
Adj. R-Square	0.74	0.75	0.74	0.74	0.75	0.73	0.74	0.75	0.74
Panel B - 2^{nd} Tercile									
Div^{Mig}_{All}	0.223***	0.176***	0.228***						
	(0.053)	(0.048)	(0.054)						
Mig_{All}	0.091*	0.096*	0.106*						
	(0.051)	(0.051)	(0.055)						
Div^{Mig}_{HS}				0.166***	0.147***	0.159***			
				(0.047)	(0.036)	(0.050)			
Mig_{HS}				-0.002	-0.066	-0.066			
				(0.046)	(0.058)	(0.056)			
Div^{Mig}_{LS}							0.224***	0.159***	0.235***
							(0.058)	(0.052)	(0.058)
Mig_{LS}							0.101*	0.112**	0.132**
							(0.056)	(0.056)	(0.066)
Observations	67	67	67	67	67	67	67	67	67
Countries	44	44	44	44	44	44	44	44	44
K-P rk Wald F-stat	40.47	517.43	17.69	25.91	15.95	10.73	24.48	626.49	10.89
Adj. R-Square	0.29	0.30	0.29	0.35	0.33	0.32	0.24	0.27	0.23
Instr Div	✓		✓	✓		✓	✓		✓
Instr Mig		✓	✓		✓	✓		✓	✓

Note: authors' calculations using data from Hausmann *et al.* (2014), Alesina *et al.* (2016) Özden *et al.* (2011), and ADOP (2015). Standard errors are clustered at country level. * p<0.1, ** p<0.05, *** p<0.01. This table shows the effect of birthplace diversity on our standardized measure of Economic Complexity,after controlling for the full set of controls of Table 3 of Alesina *et al.* (2016) and for the logarithm of real GDP per capita. Birthplace diversity is instrumented with predicted stocks through shift-share methodology based on the 1970 distribution, while the migration share is instrumented with predicted stocks from gravity model.

In Table 8 we test whether our results hold after instrumenting our main variable of interest. As it is presented in Section 3, we have two different sets of instrumental variables. Descriptive statistics across different standardized instrumental variables and actual birthplace diversity and immigration share are available in

Table B-4, Table B-5 and Table B-6 in Appendix B . Overall there is a strong correlation across different measures of diversity (actual and predicted). However, looking at correlations in Tables B-5 and B-6, we can see that both IV approaches have some issues with the prediction of the stocks of highly educated immigrants. The gravity model is not able to predict properly the composition of the bilateral stocks of highly educated migrants in terms of countries of origin (i.e. diversity index), but it is able to produce a good predictor of the skill-specific share of immigrants. On the other hand, the shift-share approach is able to predict the composition in terms of countries of origin of college-educated immigrants, while it fails to predict the size of immigrants population.[17] To properly identify both endogenous variables we then instrument diversity indexes with IVs based on predicted stocks from the shift-share approach, while we instrument skill-specific shares of immigrants with IVs based on predicted stocks from the gravity model. Table 8 reports the second-stage estimates of our IV approach. Columns (1), (4) and (7) show the estimates when we instrument only the skill-specific diversity index, while columns (2), (5) and (8) present the estimates when the skill-specific share of immigrants is treated as endogenous variable. Finally, columns (3), (6) and (9) shows the estimates when both variables are simultaneously instrumented. Moreover our analysis is done over the whole sample of countries (Panel A) and the subsample of countries belonging to the 2^{nd} tercile of the economic complexity distribution (Panel B). Looking at both panels, the instrumental variables have enough predictive power to explain both endogenous variables when we instrument them separately or simultaneously. The F-stat are always above the critical values specified by Stock and Yogo (2002). The estimates associated to birthplace diversity indexes are always positive and significant on the overall sample (Panel A), confirming the results presented in Table 2. The estimates presented in Panel B are positive and precisely estimated at 1% level. Moreover, all the 2SLS point estimates are larger than the OLS estimates presented in Table 3 when we instrument both endogenous variables, suggesting the correction of an attenuation bias.

Overall we can see that the coefficients associated to birthplace diversity are positive and significant across different regressions, confirming the strong and positive relation between diversity and economic complexity in countries that belongs to the middle of the economic complexity distribution.[18]

6 Discussion

Our analysis shows a clear and strong relation between immigrants birthplace diversity and economic complexity. Our estimates are stronger when focusing on college educated immigrants and among developing countries. Those results are robust to the inclusion of historical diversity, origin-specific effects, country fixed-effects and panel analysis. Combining those robustness with our IV estimates, we mitigate the potential reverse causality and omitted variable bias threats, reducing concerns of our results being driven by endogeneity.

[17]Docquier *et al.* (2019) recognize the same weakness of the shift-share methodology to predict the share of college-graduate immigrants.

[18]Second-stage estimates with alternative shift-share stocks based on the 1960 distribution are available in Table B-7 in Appendix B .

However, some questions remain open. For instance, what are the mechanisms through which immigrants diversity positively affects economic complexity? Empirically determining these mechanisms is an important part of our future agenda, but here we provide some insights.

Simply put, a higher Economic Complexity Index is driven by two components. First, a more diverse export basket in terms of the number of industries that are exported competitively; and second, a smaller average number of countries that export those industries competitively (a low ubiquity of those industries). We extend our analysis by testing whether birthplace diversity of a country contributes both to its export basket diversity and the uniqueness of its products.

Following the literature, we compute different indices to measure the diversification of a country's export basket (see Imbs and Wacziarg (2003), Cadot *et al.* (2011), Bahar and Santos (2018)) using -consistently with the construction of ECI above- 4-digit classification of industries according to SITC. We first construct an index that measures concentration of a country's export basket (e.g., a larger value implies higher concentration, or less diversification), namely the Herfindahl-Hirschman index (HHI here onwards).[19] Assuming there are $i = 1, 2, ..., n$ products and $d = 1, 2, ..., m$ countries, the HHI for any country d in time t is computed using the following formula:

$$HHI_{d,t} = \sum_{i=1}^{n} s_{i,d,t}^{2} \tag{19}$$

where $s_{i,d,t}$ is the share product i in country d's export basket at time t. We complement it with two more measures. First, the total number of 4-digit SITC industries exported with any value higher than zero; as well as with the total number of products exported by a country d at time t with RCA values above one, consistent with the economic complexity literature. Naturally, for these two measures a higher value implies more export diversification (and less concentration). These measures can be mathematically expressed as:

$$products_{d,t} = \sum_{i=1}^{n} \times 1[exp_{i,d,t} > 0] \tag{20}$$

$$productsRCA_{d,t} = \sum_{i=1}^{n} \times 1[rca_{i,d,t} \geq 1] \tag{21}$$

We then also compute a number of indices to measure average ubiquity of countries' exports. Firstly we compute the concentration of each product across global export basket, applying the HHI index formula to a product p, as follows:

[19] We also compute alternative indices of product diversification: the Gini coefficient, the Theil Index with its decomposition in the Theil-within and Theil-between. Changes in the Theil-within can be interpreted as changes in concentration due to more concentration of industries that already existed in the export basket of the country, while the Theil-between can be interpreted as changes in export concentration due to the appearance or disappearance of export lines. Results related to those indices and others are reported at Tables B-8 and B-10 in the Appendix B .

$$HHI_{p,t} = \sum_{j=1}^{m} s_{p,j,t}^2 \tag{22}$$

where j represents a country, and $s_{p,j,t}$ is the share of product p in country j's export basket at time t. A higher value of $HHI_{p,t}$ implies that exports of product p in time t are concentrated on fewer countries. We then compute a country level measure which we name HHI^p, and is constructed as:

$$HHI_{d,t}^p = \sum_{i=1}^{n} s_{i,d,t} \times HHI_{i,t} \tag{23}$$

In other words, $HHI_{d,t}^p$ is a sum of *product-level* concentration indices weighted by the relative size of each product in country's d export basket at time t. A higher value of $HHI_{d,t}^p$ implies that country d exports goods that are exported by fewer countries at time t. Consistently with above, we produce similar measures for number of countries that export the product p based on the following two definitions:

$$countries_{p,t} = \sum_{j=1}^{m} \times 1[exp_{p,j,t} > 0] \tag{24}$$

$$countriesRCA_{p,t} = \sum_{j=1}^{m} \times 1[rca_{p,j,t} \geq 1] \tag{25}$$

Then we use these measures to construct a country-level variable that computes the average export shares weighted by number of countries that export each product exported by country c as follow:

$$products_{d,t}^p = \sum_{i=1}^{n} s_{i,d,t} \times countries_{i,t} \tag{26}$$

$$productsRCA_{d,t}^p = \sum_{i=1}^{n} s_{i,d,t} \times countriesRCA_{i,t} \tag{27}$$

In other words, *lower* levels of $products_{d,t}^p$ and of $productsRCA_{d,t}^p$ imply that country-year pair d and t is exporting products that, on average, are less ubiquitous.

Table 9 shows the correlation across those measures of diversification, export uniqueness and our measures of birthplace diversity, GDP per capita as well as ECI. The table documents that, naturally, all diversification measures are highly and significantly correlated among each other. But, in addition, the table also documents that economic complexity and immigration diversity are positively correlated with measures of export diversification (or, in fact, negatively correlated with measures of *concentration*). Concerning measures of export uniqueness we find a negative and significant correlation between our weighted measures of competitive products and diversity, suggesting that more diversity is related with less ubiquitous exported goods.

Table 9: Correlations of diversity, diversification, ubiquity, income and complexity

	(1) ECI	(2) $ln(GDPpc)$	(3) Div_{All}^{Mig}	(4) Div_{HS}^{Mig}	(5) Div_{LS}^{Mig}	(6) HHI	(7) $products$	(8) $productsRCA$	(9) HHI^p	(10) $products^p$
ECI	1									
$ln(GDPpc)$	0.720***	1								
Div_{All}^{Mig}	0.338***	0.447***	1							
Div_{HS}^{Mig}	0.209**	0.295***	0.869***	1						
Div_{LS}^{Mig}	0.326***	0.427***	0.993***	0.841***	1					
HHI	-0.557***	-0.308***	-0.109	-0.0477	-0.0872	1				
$products$	0.697***	0.716***	0.342***	0.144*	0.321***	-0.568***	1			
$productsRCA$	0.811***	0.584***	0.305***	0.171*	0.291***	-0.613***	0.754***	1		
HHI^p	0.137	-0.0717	0.00592	0.0660	0.00936	-0.353***	-0.0348	0.201**	1	
$products^p$	0.306***	0.139*	0.00228	-0.0268	-0.0221	-0.339***	0.316***	0.369***	-0.213**	1
$products\ RCA^p$	-0.499***	-0.364***	-0.188**	-0.175*	-0.185**	0.134	-0.290***	-0.458***	-0.402***	0.244***

Note: authors' calculations using data from Hausmann *et al.* (2014), Alesina *et al.* (2016), and ADOP (2015). This table show correlations among variables with their level of significance * p<0.1, ** p<0.05, *** p<0.01.

Table 10 estimates equation (12), but this time uses the export basket diversification and uniqueness measures as dependent variable. The top panel presents the results for the diversification measures, while the bottom panel presents the estimates for the ubiquity measures. Panel A shows across most specifications that birthplace diversity is associated with higher diversification (again, note that point estimates are negative when the dependent variable is a concentration index and positive when using number of products). While not all the point estimates are statistically different from zero based on the conventional confidence intervals, we do find significant coefficients for the number of products with RCA above one (columns 7 to 9).[20] For instance, an increase of one standard deviation in our birthplace diversity index is associated with an increase in the number of industries exported competitively by around 12 (roughly a 10 percent increase, based on the sample mean of about 120, with standard deviation equal to 81).[21] Panel B shows the partial correlations between skill-specific migration diversity and export ubiquity. Overall the estimates are not statistically different from zero, however the direction of the point estimates is toward a reduction of products ubiquity.[22]

All in all, our evidence points to birthplace diversity explaining higher levels of complexity, in part, through the diversification of country's export basket, though the evidence is not very robust across diversification measures.

[20] Removing from the set of controls measures of diversity in import and export does not affect the size and the precision of the estimates significantly.

[21] Table B-10 presents the results for alternative diversification measures. The direction of the results is similar, and we find significant coefficients associated with the Gini index. From the estimates, an increase of one standard deviation of birthplace diversity index is associated with a drop of the Gini coefficient around 0.012 (roughly 1.3 percent based on the sample mean equal to 0.93).

[22] Table B-9 proposes an alternative way to test the effect of diversity on products ubiquity. Using countries averages of Product Complexity Index over the top 3, 5 and 10 products in terms of exports, we test whether the best products export become more unique. The Product Complexity Index –an index that mirrors the ECI– measures how much a given product is complex: higher values implies products that are exported by fewer countries and that those countries are highly diversified. Even though the direction of the estimates indicates higher uniqueness, the estimates are not statistically different from zero.

Table 10: OLS regression on Diversification and Ubiquity Measures

	HHI			$products$			$productsRCA$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<u>Panel A - Diversification</u>									
Div^{Mig}_{All}	-0.007			7.975			11.438**		
	(0.019)			(11.653)			(4.558)		
Mig_{All}	-0.024			-5.090			1.525		
	(0.020)			(14.611)			(4.615)		
Div^{Mig}_{HS}		-0.006			-11.402			12.687***	
		(0.018)			(12.545)			(4.567)	
Mig_{HS}		-0.033			-7.664			-0.544	
		(0.031)			(15.428)			(6.356)	
Div^{Mig}_{LS}			-0.006			8.248			10.354**
			(0.018)			(11.028)			(4.464)
Mig_{LS}			-0.022			-5.235			2.488
			(0.019)			(14.274)			(4.682)
$ln(GDPpc)$	0.040	0.040	0.038	42.394*	52.321**	42.326*	8.637	10.891	8.543
	(0.026)	(0.029)	(0.026)	(21.602)	(20.421)	(21.372)	(7.281)	(7.248)	(7.322)
Observations	200	200	200	200	200	200	200	200	200
Countries	100	100	100	100	100	100	100	100	100
Adj. R-Square	0.43	0.44	0.43	0.79	0.79	0.79	0.79	0.79	0.79

	HHI^p			$products^p$			$productsRCA^p$		
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
<u>Panel B - Ubiquity</u>									
Div^{Mig}_{All}	0.000			0.306			-0.025		
	(0.002)			(1.533)			(0.801)		
Mig_{All}	0.008***			-2.519*			-2.171***		
	(0.002)			(1.358)			(0.693)		
Div^{Mig}_{HS}		0.002			1.567			-0.734	
		(0.002)			(1.418)			(0.757)	
Mig_{HS}		0.008***			-2.017			-2.451**	
		(0.002)			(2.029)			(0.942)	
Div^{Mig}_{LS}			0.000			-0.063			0.020
			(0.002)			(1.512)			(0.798)
Mig_{LS}			0.008***			-2.424*			-2.116***
			(0.001)			(1.341)			(0.663)
$ln(GDPpc)$	-0.014***	-0.013***	-0.013***	1.801	0.737	1.953	-1.398	-1.217	-1.449
	(0.004)	(0.003)	(0.004)	(2.449)	(2.478)	(2.444)	(1.586)	(1.525)	(1.578)
Observations	200	200	200	200	200	200	200	200	200
Countries	100	100	100	100	100	100	100	100	100
Adj. R-Square	0.45	0.43	0.44	0.78	0.78	0.78	0.59	0.59	0.59
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Regional FE	✓	✓	✓	✓	✓	✓	✓	✓	✓

Note: authors' calculations using data from Hausmann *et al.* (2014) and Alesina *et al.* (2016). Standard errors are clustered at country level. * p<0.1, ** p<0.05, *** p<0.01. This table shows the effect of birthplace diversity on different measures of product diversification and ubiquity: Herfindahl-Hirschman Index on products exported (col. (1)-(3)), number of exported products (col. (4)-(6)), number of exported products with RCA>1 (col. (7)-(9)), weighted measures of countries product concentration (col. (10)-(12)), weighted sum of products uniqueness (col. (13)-(15)) and weighted sum of products with RCA bigger than one uniqueness (col. (16)-(18)).

7 Conclusion

We investigate the relationship between immigration diversity and economic complexity with the prior that a more diverse immigration, especially if immigrants are highly-skilled, brings with it new knowledge and skills that can serve to expand the economic complexity of the receiving economy (i.e., its capacity to develop a set of competitive and unique industries).

We address this question in a sample of 100 countries over the 1990-2000 period for which we could collect data on their levels of economic complexity and on the structure (by country of origin and skill level) of their immigration. In line with the literature, we use the Herfindahl index of diversity in our baseline regressions; we also use the Theil index and its decomposition to refine our interpretation of the results in the robustness section.

We find that the birthplace diversity of immigrants is strongly and positively associated with countries' economic complexity. Increasing birthplace diversity by one standard deviation is associated with an increase in economic complexity by 0.18 standard deviations. This holds particularly true among college-educated migrants and for countries at intermediate levels of economic complexity. Moreover, the results hold in a short panel of developed countries, accounting for countries time-invariant unobserved factors, suggesting that a minimum threshold of complexity should be reached to take advantage of the opportunities linked to diversity.

Those results are robust to controlling for past diversity, country-fixed effects, and for origin-specific effects (i.e., for whether immigrants come from countries which are themselves complex or not). They are also robust to instrumenting migration using two different IV strategies, one based on a pseudo-gravity model, and one of the shift-share methodology. When we use the Theil index, we find that the positive relationship between diversity and complexity is mostly driven by the "between" component of the index, suggesting that the extensive margin of immigration (that is, the diversity of origins) matters. This is consistent with our interpretation of the results in terms of skill complementarity. Finally, we show that the results are driven by the diversification component of the economic complexity index rather than by its uniqueness component Considering the full set of results, we conclude that immigration diversity is an important building block of economic complexity.

References

Ager, P., and Brückner, M. (2013). Cultural diversity and economic growth: Evidence from the US during the age of mass migration. *European Economic Review*, 64:76-97.

Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., & Wacziarg, R. (2003). Fractionalization. *Journal of Economic Growth*, 8(2): 155-194.

Alesina, A., & La Ferrara, E. (2005). Ethnic diversity and economic performance. *Journal of economic literature*, 43(3):762-800.

Alesina, A., Harnoss, J. and Rapoport, H. (2016). Birthplace diversity and economic prosperity. *Journal of Economic Growth*, 21:101-138.

Alesina, A., Murard, E., and Rapoport, H. (2019). Immigration and Preferences for Redistribution in Europe National Bureau of Economic Research. WP No. w25562

Angrist, J., and Pischke, J. S. (2009). Mostly harmless econometrics: an empiricists guide. Princeton University Press (392 p.).

Ashraf, Q., & Galor, O. (2013). The 'Out of Africa' hypothesis, human genetic diversity, and comparative economic development. *American Economic Review* 103(1): 1-46.

Artuc, E., Docquier, F., Özden, C., and Parsons, C. (2015). A Global Assessment of Human Capital Mobility: The Role of Non-OECD Destinations. *World Development*, 65:6-26.

Bahar, D., Hauptmann, A., Özgüzel, C., and Rapoport, H. (2018). Let their Knowledge Flow: The Effect of Returning Refugees on Export Performance in the Former Yugoslavia. CESifo Working Paper. No. 7371

Bahar, D., and Rapoport, H. (2018). Migration, knowledge diffusion and the comparative advantage of nations. *The Economic Journal*, 128(612):F273-F305.

Bahar, D., and Santos, M. A. (2018). One more resource curse: Dutch disease and export concentration. *Journal of Development Economics*, 132:102-114.

Bahar, D., Choudhury, P., & Rapoport, H. (2019). Migrant Inventors and the Technological Advantage of Nations. HBS Working Paper 19-119

Balassa, B. (1965). Trade liberalisation and "revealed" comparative advantage. *The Manchester School*, 33(2): 99-123.

Barro, R., and Lee, J. (2013). A new data set of educational attainment in the world, 1950-2010. *Journal of Development Economics*, 104:184-198.

Böheim, R., Horvath, T., and Mayr, K. (2012). Birthplace diversity of the workforce and productivity spill-overs in firms. WIFO Working Papers No. 438.

Borjas, G. J. (2019). Immigration and Economic Growth. National Bureau of Economic Research. WP No. w25836.

Bove, V., & Elia, L. (2016). Migration, diversity, and economic growth. *World Development*, 89:227-239.

Cadot, O., Carrère, C., and Strauss-Kahn, V. (2011). Export diversification: What's behind the hump? *Review of Economics and Statistics*, 93(2):590-605.

Cadot, O., Carrere, C., & Strauss-Kahn, V. (2013). Trade diversification, income, and growth: what do we know?. *Journal of Economic Surveys*, 27(4):790-812.

Card, D. (2001). Immigration Inflows, Native Outflows and the Local Labor Market Impacts of Higher Immigration. *Journal of Labor Economics* 19: 22-64.

Desmet, K., Ortuño-Ortín, I., & Wacziarg, R. (2012). The political economy of linguistic cleavages. *Journal of development Economics*, 97(2):322-338.

Docquier, F., Machado, J., and Sekkat, K. (2015). Efficiency gains from liberalizing labor mobility. *The Scandinavian Journal of Economics*, 117(2):303-346.

Docquier, F., Lodigiani, E., Rapoport, H., and Schiff, M. (2016). Emigration and Democracy. *Journal of Development Economics*, 120: 209-223.

Docquier, F., Turati, R., Valette, J. and Chrysovalantis, V. (2019). Birthplace diversity and Economic Growth: Evidence from the US States in the Post-World War II Period. *Journal of Economic Geography* forthcoming.

Felipe, J., Kumar, U., Abdon, A., & Bacate, M. (2012). Product complexity and economic development. *Structural Change and Economic Dynamics*, 23(1):36-68.

Feenstra, R. C., Lipsey, R. E., Deng, H., Ma, A. C., & Mo, H. (2005). World trade flows: 1962-2000. National Bureau of Economic Research WP No. w11040.

Feenstra, R. C., Inklaar, R. & Timmer, M. P. (2015). The Next Generation of the Penn World Table. *American Economic Review* 105(10): 3150-3182.

Fulford, S., Petkov, I., and Schiantarelli, F. (2017). Does it matter where you came from? Ancestry composition and economic performance of U.S. counties, 1850-2010. Boston College Working Papers in Economics No. 875.

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2018). Bartik Instruments: What, When, Why, and How. National Bureau of Economic Research. WP No. w24408

Hausmann, R., and Klinger, B. (2007). The structure of the product space and the evolution of comparative advantage. Center for International Development at Harvard University, WP No. 146

Hausmann, R., and Hidalgo, C. A. (2011). The network structure of economic output. *Journal of Economic Growth*, 16(4):309-342.

Hausmann, R., Hidalgo, C. A., Bustos, S., Coscia, M., Simoes, A., and Yildirim, M. A. (2014). The atlas of economic complexity: Mapping paths to prosperity. Mit Press.

Head, K., Mayer, T., & Ries, J. (2010). The erosion of colonial trade linkages after independence. *Journal of International Economics*, 81(1), 1-14.

Hidalgo, C. A., Klinger, B., Barabási, A. L., and Hausmann, R. (2007). The product space conditions the development of nations. *Science*, 317(5837):482-487.

Hidalgo, C. A., and Hausmann, R. (2009). The building blocks of economic complexity. *Proceedings of the national academy of sciences*, 106(26):10570-10575.

Hjort, J. (2014). Ethnic divisions and production in firms. *Quarterly Journal of Economics*, 129(4):1899-1946.

Imbs, J., and Wacziarg, R. (2003). Stages of diversification. *American Economic Review*, 93(1):63-86.

Jaeger, D. A., Ruist, J., and Stuhler, J. (2018). Shift-share instruments and the impact of immigration. National Bureau of Economic Research. WP No. w24285.

Javorcik, B. S., Lo Turco, A., & Maggioni, D. (2018). New and improved: Does FDI boost production complexity in host countries?. *The Economic Journal*, 128(614):2507-2537.

Lyons, E. (2017). Team production in international labor markets: Experimental evidence from the field. *American Economic Journal: Applied Economics*, 9(3): 70-104.

Marshall, M. G., Gurr, T. R., & Jaggers, K. (2013). Polity IV project: Political regime characteristics and transitions, 1800-2012. Center for Systemic Peace.

Montalvo, J. G., & Reynal-Querol, M. (2005). Ethnic polarization, potential conflict, and civil wars. *American economic review*, 95(3): 796-816.

Moriconi, S., Peri, G., and Turati, R. (2018). Skill of the Immigrants and Vote of the Natives: Immigration and Nationalism in European Elections 2007-2016. National Bureau of Economic Research. WP No. w25077.

Ortega, F., and Peri, G. (2014). Openness and income: The roles of trade and migration. *Journal of International Economics*, 92(2):231-251.

Ottaviano, G. I., and Peri, G. (2006). The economic value of cultural diversity: evidence from US cities. *Journal of Economic geography*, 6(1):9-44.

Özden, Ç., Parsons, C. R., Schiff, M., and Walmsley, T. L. (2011). Where on earth is everybody? The evolution of global bilateral migration 1960–2000. *The World Bank Economic Review*, 25(1):12-56.

Ozgen, C., Peters, C., Niebuhr, A., Nijkamp, P. and Poot, J. (2014). Does cultural diversity of migrants employees affect innovation? *International Migration Review*, 48(9):377–416.

Parrotta, P., Pozzoli, D., and Pytlikova, M. (2014). Does labor diversity affect firm productivity? *European Economic Review*, 66:144-179.

Silva, J. S., and Tenreyro, S. (2006). The log of gravity. *The Review of Economics and statistics*, 88(4):641-658.

Spilimbergo, A. (2009). Democracy and foreign education. *American economic review*, 99(1):528-43.

Solow, R. M. (1956). A contribution to the theory of economic growth. *The Quarterly Journal of Economics*, 70(1):65-94.

Stock, J. H., and Yogo, M. (2002). Testing for weak instruments in linear IV regression. National Bureau of Economic Research. Technical Working Paper No. 284

Trax, M., Brunow, S., and Suedekum, J. (2015). Cultural diversity and plant-level productivity. *Regional Science and Urban Economics*, 53:85-96

U.N. Population Division. (2013). World population prospects: The 2013 revision. Retrieved from, http://

Valette, J. (2018). Do migrants transfer productive knowledge back to their origin countries? *The Journal of Development Studies*, 54(9):1637-1656.

World Bank Group (2013). World development indicators 2013. World Bank Publications, 2013.

Appendix A Harmonization of DIOC and ADOP data

To minimize the potential bias due to omitted variable, we combined data of DIOC (Database on Immigrants in OECD counties) (http://www.oecd.org/els/mig/dioc.htm) and ADOP (2015) data (https://perso.uclouvain.be/frederic.docquier/filePDF/ADOP_AgeOfEntry.xlsx), to have a panel of 22 destination countries over the period 1990 to 2010. Since DIOC is available over 2000 and 2010 while ADOP is available over 1990 and 2000, we had to harmonize our data sources. Moreover we take DIOC data as main reference in order to harmonize ADOP 1990 data with the rest of DIOC data. Our harmonization process took the following steps:

1. we compute the bilateral skill specific weight in the year 2000, such that ADOP data should be equal to DIOC, namely: $MIG^{DIOC}_{s,2000} = MIG^{ADOP}_{s,2000} * w_{s,2000}$
2. we re-weight the 1990 ADOP bilateral stocks using the weights for the 2000, namely $MIG^{ADOP_w}_{s,1990} = MIG^{ADOP}_{s,1990} * w_{s,2000}$
3. for bilateral stocks where we could not compute the weight due to missing values or zeroes, we compute the bilateral skill specific growth rate between 2000 and 2010 with DIOC and we assume a constant linear trend also for 1990. If we define $g_{s,00}$ the bilateral skill specific growth rate, then we can compute the missing bilateral skill migration stocks in the 1990 as follow: $MIG^{ADOP_w}_{s,1990} = MIG^{DIOC}_{s,2000}/(1+g_{s,00})$.

After this procedure we have a panel of 25 destination countries with 194 origin countries. However, due to missing information in the 1990 for Luxembourg, Czech Republic and Slovak Republic on our ECI data, our final panel will be on 22 destination countries.

Appendix B Additional Tables

Table B-1: Correlates across different measures of Diversity

Diversity Measures	(1) Div_{HS}^{Mig}	(2) Mig_{HS}	(3) Div_{LS}^{Mig}	(4) Mig_{LS}
$N\#Origin$	0.263***	0.293***	0.264***	0.341***
Pol_{HS}^{Mig}	-0.386***	-0.0845	-0.343***	-0.173*
Pol_{LS}^{Mig}	0.0269	-0.181*	-0.160*	-0.144*
$Theil_{HS}^{Mig}$	-0.901***	-0.118	-0.775***	-0.177*
$Theil_{LS}^{Mig}$	-0.761***	-0.217**	-0.911***	-0.180*
$Theil_{HS}^{W,Mig}$	-0.639***	0.125	-0.567***	0.116
$Theil_{LS}^{W,Mig}$	-0.465***	0.0119	-0.646***	0.0979
$Theil_{HS}^{B,Mig}$	-0.234***	-0.235***	-0.184**	-0.281***
$Theil_{LS}^{B,Mig}$	-0.254***	-0.229**	-0.208**	-0.285***

Note: Correlations * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table B-2: OLS regression on Economic Complexity Index - Different Diversity Indexes

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Variable: *ECI*	OLS	OLS	OLS	OLS	OLS	OLS
Sample:	2^{nd} *Tercile*					
Diversity Measures	Div^{Mig}	Div^{Mig}	Pol^{Mig}	$Theil^{Mig}$	$Theil^{W,Mig}$	$Theil^{B,Mig}$
HS Index	0.128**	0.127**	-0.042	-0.142*	-0.107	-0.283**
	(0.058)	(0.059)	(0.054)	(0.072)	(0.074)	(0.123)
LS Index	0.024	0.025	0.092	0.055	0.085	0.256*
	(0.087)	(0.088)	(0.054)	(0.077)	(0.079)	(0.135)
Mig_{HS}	-0.072	-0.072	-0.048	-0.064	-0.073	-0.056
	(0.085)	(0.085)	(0.071)	(0.094)	(0.084)	(0.097)
Mig_{LS}	0.090	0.088	0.105	0.086	0.112	0.078
	(0.098)	(0.099)	(0.094)	(0.110)	(0.111)	(0.115)
N#Origin		0.000				
		(0.001)				
Observations	67	67	67	67	67	67
Countries	44	44	44	44	44	44
Adj. R-Square	0.34	0.37	0.21	0.24	0.15	0.19
Controls	✓	✓	✓	✓	✓	✓
Regional FE	✓	✓	✓	✓	✓	✓

Note: authors' calculations using data from Hausmann *et al.* (2014) and Alesina *et al.* (2016). Standard errors are clustered at country level. * p<0.1, ** p<0.05, *** p<0.01. The dependent variable is a standardized measure of the Economic Complexity Index. Each regression includes the full set of controls of Table 3 of Alesina *et al.* (2016) plus the logarithm of GDP. The sample of of countries is composed by countries belonging to the 2^{nd} tercile of the ECI distribution. Each column shows different approaches to measure skill-specific birthplace diversity among immigrants: fractionalization index (col. 1), fractionalization index plus the number of migrants' countries of origin (col. 2), polarization index (col. 3), theil index (col. 4) and its decomposition: the within component (col. 5) and between component (col. 6).

Table B-3: PPML regression on bilateral migration stocks like AHR

	(1) PPML	(2) PPML	(3) PPML
Dep var.	MIG_{All}	MIG_{LS}	MIG_{HS}
Pop destination 1960	0.389***	0.374***	0.443***
	(0.133)	(0.126)	(0.153)
Bilateral weighted distance	-3.367***	-3.484***	-3.414***
	(0.838)	(0.818)	(1.105)
Colonial relationship	1.607***	1.578***	1.908***
	(0.292)	(0.321)	(0.301)
Common ethnic language	1.854***	1.688***	2.388**
	(0.678)	(0.571)	(0.958)
Common official language	-0.193	-0.177	-0.319
	(0.585)	(0.555)	(0.729)
Common border	1.827***	1.934***	1.008***
	(0.237)	(0.245)	(0.173)
Horizontal Time difference	0.300***	0.229***	0.454***
	(0.113)	(0.084)	(0.164)
Observations	71060	71060	71060
Countries	190	190	190
R-Square	0.42	0.49	0.19
Year FE	✓	✓	✓
Origin*Year FE	✓	✓	✓

Note: authors' calculations on ADOP (2015) and Alesina *et al.* (2016) data.

Table B-4: Descriptive Statistics Migration

Variable	(1) Obs	(2) Mean	(3) Std. Dev.	(4) Min	(5) Max
Div_{All}^{Mig}	200	0	1	-3.396	1.026
Div_{HS}^{Mig}	200	0	1	-3.955	0.827
Div_{LS}^{Mig}	200	0	1	-3.266	1.077
$Div_{All}^{Mig,Grav}$	200	0	1	-4.034	0.943
$Div_{HS}^{Mig,Grav}$	200	0	1	-5.968	0.664
$Div_{LS}^{Mig,Grav}$	200	0	1	-3.744	0.976
$Div_{All}^{Mig,SS}$	200	0	1	-3.225	1.118
$Div_{HS}^{Mig,SS}$	200	0	1	-3.283	1.044
$Div_{LS}^{Mig,SS}$	200	0	1	-3.156	1.163
Mig_{All}	200	0	1	-0.594	4.712
Mig_{HS}	200	0	1	-0.625	5.551
Mig_{LS}	200	0	1	-0.587	4.910
Mig_{All}^{Grav}	200	0	1	-0.594	4.850
Mig_{HS}^{Grav}	200	0	1	-0.638	5.440
Mig_{LS}^{Grav}	200	0	1	-0.578	4.831
Mig_{All}^{SS}	200	0	1	-0.689	6.0108
Mig_{HS}^{SS}	200	0	1	-0.293	10.299
Mig_{LS}^{SS}	200	0	1	-0.633	6.343

Note: authors' calculations on Özden *et al.* (2011), Alesina *et al.* (2016) and ADOP data.

Table B-5: Correlations across birthplace diversity indices

	(1) Div_{All}^{Mig}	(2) Div_{HS}^{Mig}	(3) Div_{LS}^{Mig}	(4) $Div_{All}^{Mig,Grav}$	(5) $Div_{HS}^{Mig,Grav}$	(6) $Div_{LS}^{Mig,Grav}$	(7) $Div_{All}^{Mig,SS}$	(8) $Div_{HS}^{Mig,SS}$
Div_{HS}^{Mig}	0.869***							
Div_{LS}^{Mig}	0.993***	0.841***						
$Div_{All}^{Mig,Grav}$	0.337***	0.347***	0.348***					
$Div_{HS}^{Mig,Grav}$	0.108	0.174*	0.108	0.708***				
$Div_{LS}^{Mig,Grav}$	0.358***	0.350***	0.372***	0.991***	0.639***			
$Div_{All}^{Mig,SS}$	0.685***	0.584***	0.700***	0.224**	0.0622	0.239***		
$Div_{HS}^{Mig,SS}$	0.637***	0.637***	0.645***	0.244***	0.0722	0.258***	0.927***	
$Div_{LS}^{Mig,SS}$	0.683***	0.559***	0.700***	0.209**	0.0495	0.226**	0.994***	0.890***

Note: authors' calculations on Alesina *et al.* (2016), ADOP (2015) and Özden *et al.* (2011) data. This table show correlations among variables with their level of significance * p<0.1, ** p<0.05, *** p<0.01.

Table B-6: Correlations across shares of migrants

	(1) Mig_{All}	(2) Mig_{HS}	(3) Mig_{LS}	(4) Mig_{All}^{Grav}	(5) Mig_{HS}^{Grav}	(6) Mig_{LS}^{Grav}	(7) Mig_{All}^{SS}	(8) Mig_{HS}^{SS}
Mig_{HS}	0.869***							
Mig_{LS}	0.996***	0.836***						
Mig_{All}^{Grav}	0.980***	0.871***	0.972***					
Mig_{HS}^{Grav}	0.794***	0.834***	0.771***	0.824***				
Mig_{LS}^{Grav}	0.963***	0.851***	0.959***	0.987***	0.774***			
Mig_{All}^{SS}	0.811***	0.663***	0.816***	0.798***	0.722***	0.796***		
Mig_{HS}^{SS}	0.0208	0.0362	0.0218	0.0112	0.235***	0.00274	0.223**	
Mig_{LS}^{SS}	0.810***	0.651***	0.818***	0.808***	0.665***	0.840***	0.963***	0.137

Note: authors' calculations on Alesina *et al.* (2016), ADOP (2015) and Özden *et al.* (2011) data. This table show correlations among variables with their level of significance * p<0.1, ** p<0.05, *** p<0.01.

Table B-7: Alternative IV Estimates - Shift-share based on 1960 distribution

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dep. Variable: ECI									
Panel A - Overall Sample									
Div_{All}^{Mig}	0.296***	0.116**	0.294***						
	(0.091)	(0.059)	(0.090)						
Mig_{All}	0.011	-0.014	0.006						
	(0.067)	(0.065)	(0.066)						
Div_{HS}^{Mig}				0.361***	0.110**	0.355***			
				(0.119)	(0.054)	(0.122)			
Mig_{HS}				0.052	-0.064	-0.022			
				(0.091)	(0.135)	(0.141)			
Div_{LS}^{Mig}							0.279***	0.109*	0.286***
							(0.086)	(0.057)	(0.085)
Mig_{LS}							0.018	0.020	0.038
							(0.066)	(0.073)	(0.072)
Observations	200	200	200	200	200	200	200	200	200
Countries	100	100	100	100	100	100	100	100	100
K-P rk Wald F-stat	22.75	425.25	11.57	17.08	29.41	9.08	27.65	126.61	14.05
Adj. R-Square	0.73	0.75	0.73	0.71	0.75	0.71	0.73	0.75	0.73
Panel B - 2^{nd} Tercile									
Div_{All}^{Mig}	0.108	0.176***	0.117*						
	(0.067)	(0.048)	(0.068)						
Mig_{All}	0.067	0.096*	0.085						
	(0.049)	(0.051)	(0.052)						
Div_{HS}^{Mig}				0.074	0.147***	0.038			
				(0.070)	(0.036)	(0.082)			
Mig_{HS}				0.009	-0.066	-0.060			
				(0.045)	(0.058)	(0.068)			
Div_{LS}^{Mig}							0.106	0.159***	0.123*
							(0.066)	(0.052)	(0.070)
Mig_{LS}							0.072	0.112**	0.102*
							(0.050)	(0.056)	(0.058)
Observations	67	67	67	67	67	67	67	67	67
Countries	44	44	44	44	44	44	44	44	44
K-P rk Wald F-stat	20.88	517.43	9.53	8.44	15.95	2.85	20.30	626.49	8.22
Adj. R-Square	0.28	0.30	0.28	0.30	0.33	0.20	0.26	0.27	0.26
Instr *Div*	✓		✓	✓		✓	✓		✓
Instr *Mig*		✓	✓		✓	✓		✓	✓

Note: authors' calculations using data from Hausmann *et al.* (2014), Alesina *et al.* (2016) Özden *et al.* (2011), and ADOP (2015) data. Standard errors are clustered at country level. * p<0.1, ** p<0.05, *** p<0.01. This table shows the effect of birthplace diversity on our standardized measure of Economic Complexity,after controlling for the full set of controls of Table 3 of Alesina *et al.* (2016) and for the logarithm of real GDP per capita. Birthplace diversity is instrumented with predicted stocks through shift-share methodology based on the 1960 distribution, while the migration share is instrumented with predicted stocks from gravity model.

Table B-8: Correlations of diversity and alternative diversification measures

	(1) ECI	(2) $ln(GDPpc)$	(3) Div_{All}^{Mig}	(4) Div_{HS}^{Mig}	(5) Div_{LS}^{Mig}	(6) Gini on Export	(7) MaxMin	(8) Log Var	(9) Theil Index	(10) Theil Index (w)
ECI	1									
$ln(GDPpc)$	0.720***	1								
Div_{All}^{Mig}	0.338***	0.447***	1							
Div_{HS}^{Mig}	0.209**	0.295***	0.869***	1						
Div_{LS}^{Mig}	0.326***	0.427***	0.993***	0.841***	1					
Gini on Export	-0.844***	-0.636***	-0.343***	-0.218**	-0.331***	1				
MaxMin	-0.588***	-0.352***	-0.129	-0.0480	-0.107	0.629***	1			
Log Var	-0.724***	-0.485***	-0.216**	-0.101	-0.197**	0.836***	0.923***	1		
Theil Index	-0.785***	-0.545***	-0.248***	-0.119	-0.227**	0.880***	0.896***	0.976***	1	
Theil Index (w)	-0.697***	-0.360***	-0.156*	-0.105	-0.141*	0.812***	0.873***	0.938***	0.925***	1
Theil Index (b)	-0.623***	-0.660***	-0.315***	-0.0987	-0.293***	0.641***	0.569***	0.643***	0.723***	0.407***

Note: authors' calculations using data from Hausmann *et al.* (2014) and ADOP (2015). This table show correlations among variables with their level of significance * p<0.1, ** p<0.05, *** p<0.01.

Table B-9: OLS regression on Countries Top Products Ubiquity

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dep. Variable: PCI	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
Products		*Top 3*			*Top 5*			*Top 10*	
Div_{All}^{Mig}	0.034			0.040			0.071*		
	(0.055)			(0.046)			(0.042)		
Mig_{All}	-0.073			-0.014			0.027		
	(0.055)			(0.055)			(0.043)		
Div_{HS}^{Mig}		0.035			0.027			0.063	
		(0.055)			(0.047)			(0.039)	
Mig_{HS}		-0.084			0.004			0.041	
		(0.070)			(0.062)			(0.049)	
Div_{LS}^{Mig}			0.027			0.035			0.068
			(0.052)			(0.044)			(0.042)
Mig_{LS}			-0.064			-0.011			0.027
			(0.055)			(0.054)			(0.042)
$ln(GDPpc)$	0.118	0.118	0.115	0.152	0.152	0.153	0.151	0.160*	0.153
	(0.133)	(0.129)	(0.132)	(0.126)	(0.121)	(0.125)	(0.101)	(0.095)	(0.101)
Observations	200	200	200	200	200	200	200	200	200
Countries	100	100	100	100	100	100	100	100	100
Adj. R-Square	0.49	0.49	0.49	0.58	0.58	0.58	0.69	0.69	0.69
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Regional FE	✓	✓	✓	✓	✓	✓	✓	✓	✓

Note: authors' calculations using data from Hausmann *et al.* (2014) and Alesina *et al.* (2016). Standard errors are clustered at country level. * p<0.1, ** p<0.05, *** p<0.01. The dependent variable is the country average Product Complexity index over the top 3, top 5 and top 10 products in terms of competitive export. Each regression includes the full set of controls of Table 3 of Alesina *et al.* (2016).

Table B-10: OLS regression on Diversification and Ubiquity Measures

	Maxmin			*Log Var*			*Gini Index on Export*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Div_{All}^{Mig}	-0.014			-0.129			-0.010***		
	(0.020)			(0.080)			(0.003)		
Mig_{All}	-0.021			-0.104			0.000		
	(0.020)			(0.077)			(0.003)		
Div_{HS}^{Mig}		-0.014			-0.121			-0.012***	
		(0.019)			(0.075)			(0.003)	
Mig_{HS}		-0.029			-0.125			0.000	
		(0.032)			(0.121)			(0.004)	
Div_{LS}^{Mig}			-0.011			-0.111			-0.010***
			(0.019)			(0.078)			(0.003)
Mig_{LS}			-0.020			-0.099			-0.000
			(0.019)			(0.073)			(0.003)
$ln(GDPpc)$	0.039	0.038	0.036	0.071	0.049	0.059	-0.008*	-0.009*	-0.009*
	(0.030)	(0.032)	(0.029)	(0.120)	(0.128)	(0.119)	(0.005)	(0.005)	(0.005)
Observations	200	200	200	200	200	200	198	198	198
Countries	100	100	100	100	100	100	100	100	100
Adj. R-Square	0.48	0.48	0.48	0.62	0.62	0.61	0.83	0.83	0.83
	Theil Index			*Theil Index Within*			*Theil Index Between*		
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Div_{All}^{Mig}	-0.130			-0.103			-0.026		
	(0.089)			(0.077)			(0.030)		
Mig_{All}	-0.079			-0.071			-0.002		
	(0.084)			(0.078)			(0.039)		
Div_{HS}^{Mig}		-0.106			-0.149**			0.045	
		(0.085)			(0.068)			(0.029)	
Mig_{HS}		-0.110			-0.107			-0.001	
		(0.136)			(0.115)			(0.040)	
Div_{LS}^{Mig}			-0.115			-0.088			-0.026
			(0.086)			(0.075)			(0.028)
Mig_{LS}			-0.075			-0.067			-0.001
			(0.081)			(0.076)			(0.038)
$ln(GDPpc)$	0.048	0.028	0.038	0.079	0.097	0.068	-0.037	-0.071	-0.038
	(0.133)	(0.143)	(0.132)	(0.113)	(0.115)	(0.112)	(0.059)	(0.053)	(0.058)
Observations	198	198	198	198	198	198	200	200	200
Countries	100	100	100	100	100	100	100	100	100
Adj. R-Square	0.70	0.70	0.70	0.57	0.58	0.57	0.70	0.70	0.70
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Regional FE	✓	✓	✓	✓	✓	✓	✓	✓	✓

Note: authors' calculations using data from Hausmann *et al.* (2014) and Alesina *et al.* (2016). Standard errors are clustered at country level. * p<0.1, ** p<0.05, *** p<0.01. This table shows the effect of birthplace diversity on different measures of product diversification: Max min on export (col. (1)-(3)), Log Var on Export (col. (4)-(6)), Gini index on Export (col. (7)-(9)), Theil Index (col. (10)-(12)), Theil Index within (col. (13)-(15)) and Theil Index between (col. (16)-(18)).