I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14772

Migration and Cultural Change

Hillel Rapoport
Sulin Sardoschau
Arthur Silve

OCTOBER 2021

I Z A Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14772

Migration and Cultural Change

Hillel Rapoport
Paris School of Economics, University Paris 1 Panthéon-Sorbonne, CEPII, LISER and IZA

Sulin Sardoschau
Humboldt University and IZA

Arthur Silve
Université Laval

OCTOBER 2021

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Migration and Cultural Change*

We propose a novel perspective on migration and cultural change by asking both theoretically and empirically – and from a global viewpoint – whether migration is a source of cultural convergence or divergence between home and host countries. Our theoretical model derives distinctive testable predictions as to the sign and direction ofconvergence for various compositional and cultural diffusion mechanisms. We use the World Value Survey for 1981-2014 to build time-varying measures of cultural similarity for a large number of country pairs and exploit within country-pair variation over time. Our results support migration-based cultural convergence, with cultural remittances as its main driver. In other words and in contrast to the populist narrative, we find that while immigrants do act as vectors of cultural diffusion, this is mostly to export the host country culture back home.

JEL Classification: F22, 015, Z10
Keywords: migration, cultural change, globalization

Corresponding author:
Hillel Rapoport
Paris School of Economics
48 Boulevard Jourdan
75014 Paris
France
E-mail: Hillel.Rapoport@psemail.eu

* We thank Alberto Alesina, Simone Bertoli, George Borjas, Gustavo De Santis, Frederic Docquier, William Easterly, Jeffry Frieden, Oded Galor, Libertad Gonzalez, Thierry Mayer, Florian Mayneris, Marion Mercier, Nathan Nunn, Mathias Thoenig, Thierry Verdier, Romain Wacziarg, David Weil, and participants at various conferences including the AFD-World Bank Migration and Development Conference, the OECD-CEPII conference on Immigration in OECD countries, World Bank Globalization: Contents and Discontents, NEUDC 2018, CEA / CDESG conference 2018, 4th Workshop on Migration, Health, and Well-Being, AFSE 2018, EEA 2018, Migration Worskshop at the Barcelona summer forum 2019, SIOE 2019, CESifo Migration Workshop 2020, 1st EBRD and Kings College Immigration Workshop 2020, and seminar audiences at Laval, Lausanne, Erasmus University, Louvain, Luxembourg, IDB, Humboldt University, Paris 1 Panthéon-Sorbonne, UQAM, PSE, Göttingen, Lugano, AMSE, and Brown.

1 Introduction

Globalization is not just about trade and financial flows, it is also about culture (Norris & Inglehart, 2009).[1] For some, it creates a new hybrid world culture (Pieterse et al., 2015), gives rise to a global village (McLuhan & Fiore, 1968), or to an Americanization of the world (Ritzer, 2012). Others, such as Huntington (1993) or Inglehart & Norris (2003) believe instead (with important nuances) that it will make cultural differences worldwide more salient and lead to cultural polarization.

Migration is a key mediating factor for globalization. Given that values and attitudes are embedded in people and transported with them, migration has a strong potential to affect culture. Previous literature has looked at this question from various angles, trying to identify cultural diffusion from natives to immigrants ("assimilation" see e.g. Abramitzky et al. 2014, 2020; Fouka et al. 2020), from immigrants to natives ("dissemination" see e.g., Miho et al. 2019; Giuliano & Tabellini 2020) or from immigrants to home communities ("cultural remittances", see e.g., Barsbai et al. 2017).[2] This paper is first to look at these mechanisms jointly, in a global context, and within a unified theoretical and empirical framework.

We first develop a theoretical model of migration-based cultural change that investigates how migration affects cultural proximity between home and host countries. The model integrates the main mechanisms of migration-based cultural change put forth in previous literature into a unified theoretical framework, including compositional changes in the host and home societies as well as several cultural diffusion mechanisms. It delivers distinctive testable predictions as to the sign and direction of convergence (i.e., do host countries converge toward the culture of home countries, or the opposite?) for each of these mechanisms.

We then empirically test the predictions of our model for a large number of country-pairs between the years 1981 and 2014. We use the World Value Survey to develop three different time-varying bilateral cultural distance measures and exploit within country pair variation over time. Specifically, we look at the effect of lagged migration (five years prior) on cultural similarity between origin and

[1] Similar to the rest of the economics literature, we use the term "culture" in a very broad sense, as a set of attitudes, preferences, beliefs and values that govern individual behavior and determine aggregate social, political and economic outcomes. For an overview of the economics of culture, see Alesina & Giuliano (2015).

[2] We review this literature in more detail in Section 2.

destination countries, including country-time fixed effects, country-pair fixed effects and several time-varying bilateral controls. We find a significantly positive correlation between migration and cultural similarity. To probe the robustness of our results and mitigate endogeneity concerns, we include lagged cultural similarity on the right hand side to account for factors that influence cultural dynamics and migration simultaneously, use different time-lags for migration (which also allows us to discriminate between static and dynamic mechanisms), and employ an instrumental variable strategy (see Section 5).

In addition, our theoretical model is crucial for uncovering the mechanisms that drive our baseline results. While case studies and natural experiments at the local level are ideal empirical settings to identify specific mechanisms in isolation, the theoretical model married with the cross-country empirical analysis can tell us something about the net-effect of various competing mechanisms. Not only does it outline the sources of cultural convergence (compositional changes at destination, diffusion, and cultural remittances) it also describes the conditions under which one of these mechanisms is more likely to be at play. Specifically, the model makes predictions about how the cultural selection of the migrant pool – as determined by underlying economic and cultural incentives to emigrate – will favor one mechanism over the other.

Our evidence supports cultural remittances as the main driver behind cultural convergence. Immigrants disseminate the host country's culture to their home country, thereby pulling culturally the home country towards the host country. This is at odds with the (populist) narrative that immigrants disseminate their cultural norms at destination, putting Western culture under threat, or contributing to a cultural "great replacement"; if anything, this suggests instead that immigration can promote the diffusion of the host culture abroad.

The paper belongs to the recent empirical literature in cultural economics, to which it contributes through its focus on the time-varying dimensions of culture (rather than on its persistence) and on "relative culture" (i.e., on cultural proximity to others, rather than on culture as an isolated set of preferences and beliefs). This literature emphasizes the deep-rooted determinants of culture (Ashraf & Galor, 2013; Alesina & Giuliano, 2015; Giuliano & Nunn, 2020; Galor & Savitskiy, 2018; Bazzi et al., 2020) and its implications for comparative economic development (Braudel, 1987; Landes, 1998; Guiso et al., 2006; Alesina & Fuchs-Schündeln, 2007; Aghion et al., 2010; Spolaore & Wacziarg, 2009a, 2013; Desmet et al., 2017). Most related to our paper, various contributions have analyzed the recent cultural dynamics within and across countries (Alesina et al., 2017; Desmet & Wacziarg, 2018; Bertrand & Kamenica, 2018; Falk et al., 2018).[3] So far, trade is the only dimension of globalization that has been analyzed from this perspective. Olivier et al. (2008) show in a theoretical model that goods market integration can lead countries to diverge culturally through a mechanism of "cultural specialization." Conversely, Maystre et al. (2014) show both theoretically and empirically that in a world with two cultural goods (one global, one local), opening up to trade will lead to both countries consuming more of the global good and less of the local good; they conclude that trade liberalization can also lead to cultural convergence. Our analysis acknowledges the potential role of trade flows, for which it finds no robust results.

The rest of this paper is organized as follows: Section 2 reviews the literature on the various channels of migration-based cultural change. Section 3 introduces a theoretical model of migration-driven cultural change. In Section 4, we document the data sources and elaborate on the meaning and statistical measurement of bilateral cultural similarity. Section 5 presents our baseline specification and results. Section 6 investigates mechanisms while Section 7 concludes.

2 Channels of migration-based cultural change

This section presents the main channels linking migration and cultural change documented so far. We organize them according to their order of introduction in the theoretical model, starting with compositional/static channels before turning to diffusion/dynamic mechanisms.

We start with the basic observation that migrants are not a random sample of their home country population; rather, they are self-selected along a number of dimensions such as age, gender, education, wealth or ethnic background. For example, it is well known that migrants tend to disproportionately come from the younger and more educated segments of the adult population. A less salient and yet fundamental dimension of self-selection into migration is the cultural dimension, i.e. *Cultural Selection*. In the political realm and following Hirschman (1970, 1978), who finds that migrants are typically positively selected in terms

[3] For example, Falk et al. (2018) run their own survey on cultural values worldwide, emphasizing trust, reciprocity, risk aversion and altruism as main cultural dimensions. Their cross-sectional results show that within-country differences in culture are as large if not larger than between-country differences.

of their support for democracy, this is called the *exit effect*. A well-known historical example of such exit is that of the 48ers, these German political refugees who fled their home country for the US in the aftermath of the failed revolutions of 1848. The political consequences of their arrival are explored in Dippel & Heblich (2020), and those of their departure in Barsbai & Rapoport (2020). Examples of cultural selection on other dimensions of culture include moral values (Casari et al., 2018), fertility (Livi-Bacci, 2012), risk attitudes (Jaeger et al., 2010), religiosity (Docquier et al., 2020a), gender discrimination (Ruyssen & Salomone, 2018), or the individualism/collectivism divide (Knudsen, 2019). To the extent that emigrants opt for destinations that are culturally closer to them, emigration will tend to deepen cultural differences across countries; as such, cultural self-selection can be seen as a source of cultural divergence.

Second, migration also creates compositional changes in the destination country. As Collier (2013, p. 67) puts it, "Migrants bring not only the human capital generated in their own societies; they also bring the moral codes of their own societies." If immigrants are not a perfect cultural match to the host population but are somewhat representative of the cultural-mix of their home country population, we expect origin and destination countries to become culturally closer by mere *Cultural Mixing*. As such, cultural mixing can be seen as a source of cultural convergence. This is all the more relevant in the absence of assimilation, that is, if immigrants remain loyal to the home country culture. And indeed, there is evidence of intertemporal and intergenerational persistence of cultural traits such as the importance of family ties, women's labor force participation, fertility rates, preferences for redistribution, and more (Giuliano, 2007; Fernandez & Fogli, 2009; Luttmer & Singhal, 2011; Giavazzi et al., 2019). Additionally, Bazzi et al. (2019) show that the relative size of the immigrant v. native groups is an important determinant in maintaining or not ethnic attachments over time.

Third, immigrants may dilute the host country's culture through mere mixing, as we have seen, but also by disseminating their own preferences, norms, and values to the host population. We will call this channel *Cultural Dissemination*. If powerful, such dissemination can lead to cultural convergence between the host and home countries. In spite of its prominence in the public debate, there is little well-identified empirical evidence of such dissemination. In sociology and cultural theory, the concept of cultural dissemination is incorporated within the idea of "hybrid culture," which stipulates that culture is non-static and that both natives and non-natives constantly interact and renegotiate the cultural order (Anthias, 2001; Papastergiadis, 2018). Evidence of cultural dissemination in the economics literature is scarce. A well-known example (but, arguably, limited in scope) is the adoption by New York-based UN diplomats from Western democracies of a relaxed attitude toward parking violations and fines while diplomats coming from countries with high levels of corruption hold on to their low standards (Fisman & Miguel, 2007). Giuliano & Tabellini (2020) show that the historic presence of European immigrants in the United States led to a more liberal political ideology and to stronger preferences for redistribution among US born individuals today. The authors hypothesize that this effect was driven by cultural diffusion from immigrants to natives, showing that the effect is strongest for US counties with higher inter-group contact (as proxied by inter-marriages and residential integration). Miho et al. (2019) exploit Stalin's ethnic deportations during WWII as a natural experiment. The pattern of deportation 'exogenously' exposed certain local populations of Siberia to deportees of either German or Chechen ethnic background. They find that such exposure and contact significantly altered those populations' attitudes, which diverged with respect to gender equality or women's labor force participation. Similarly, Schmitz & Weinhardt (2019) show that West Germans exposed to East German migrants after reunification adopted some of the latter's attitudes, again with respect to gender roles.

Fourth, we consider the *Cultural Assimilation* of migrants, that is, their adoption of the host country culture. Assimilation has been observed in various contexts, especially in the United States in the late nineteenth and early twentieth centuries (Abramitzky et al., 2020), during the 'Age of Mass Migration.' Assimilation has often been inferred from the adoption of American-sounding names or from inter-marriage (Abramitzky et al., 2014; Fouka et al., 2020; Biavaschi et al., 2017; Saavedra, 2018), as well as from immigrants' adoption of local gender norms such as female labor force participation (Blau et al., 2011). Similar assimilation patterns have been observed for Muslims immigrants in the US and in Germany (Norris & Inglehart, 2012). Assimilation would mitigate the effects of both cultural mixing and cultural dissemination, as the diffusion of culture now takes place from natives to immigrants (Giavazzi et al., 2019). In this sense, assimilation can be seen as a source of divergence.

Finally, migrants can transfer host-country cultural values and norms back home. Migration has been widely described as a transformative experience for the migrants, an experience along which they are exposed to new cultural, social and institutional norms. Once absorbed, these norms can be transferred through family, social and community networks. To describe this process, sociologists have coined the term "social remittances" (Levitt, 1998; Levitt & Lamba-Nieves, 2011), which we relabel *Cultural Remittances* for this paper. Social remittances have been documented in the realm of political preferences (Spilimbergo, 2009; Batista & Vicente, 2011; Chauvet & Mercier, 2014; Barsbai et al., 2017; Tuccio et al., 2019), fertility (Fargues, 2007; Beine et al., 2013; Bertoli & Marchetta, 2015; Daudin et al., 2019) and beyond. For example, Barsbai et al. (2017) show in the context of Moldova that the emigration wave that started in the aftermath of the Russian crisis of 1998 strongly affected electoral outcomes and political preferences in that country during the following decade, in opposite directions depending on whether people migrated to Russia or to Western democracies. In an experimental setting Batista & Vicente (2011) show that emigrants promote better institutions at home, particularly if they themselves live in (or have returned from) countries with better governance. Most recently, Tian et al. (2020) use mobile phone data to uncover transmission of social distancing practices from Mexican immigrants in the US to their social contacts in Mexico. Note that emigrants typically remain part of the cultural narrative in the home country and are often portrayed as role models (Kandel & Massey, 2002). In any event, if cultural remittances yet another mechanism of cultural convergence, it differs from 'mixing' or 'dissemination' in that it suggests that it is the home country that is being pulled toward the host country culture.

3 A model of migration-based cultural change

Migration contributes to cultural change both in the origin and in the destination countries. We envisage a number of mechanisms, which we can broadly regroup in a compositional model of migration (migrants carry values and norms from the origin to the destination country) and various diffusion effects (interpersonal diffusion of values and norms). We start with the compositional effect of migration. We assume that there are two components to the decision to migrate: a cultural and an economic component. If the economic motive dominates, migrants reflect the cultural mix of their origin country, which brings countries closer together. Conversely, if cultural homophily dominates, **cultural**

self-selection into migration pulls countries apart.

In a second step, we introduce a dynamic component to cultural formation. The process of intergenerational transmission of cultural norms and values features diffusion across group boundaries. We use a framework inspired from Bisin & Verdier (2000) to consider diffusion of values from natives to migrants (**assimilation**), from migrants to natives (**dissemination**, for want of a better term), and from migrants to their origin community (**cultural remittances**).

We provide additional perspectives on the model and its assumptions in the appendix. In particular, in what follows, asterisks should be understood as pointing to a deeper discussion in Appendix A.

3.1 A compositional model of migration and cultural change

We consider migration from home country A to host country B.* The relative size of country B is n. Individuals in both countries can be characterized by their cultural type i or j. The share of type-i individuals in each country is given by q^A and q^B. Type-i individuals are more frequent in country B, and scarcer in country A: $q^A < q^B$.

Migration has two effects on an individual's utility: a change in economic opportunities, and a change in the cultural environment. We assume that individuals from country A who contemplate migrating have heterogeneous expectations of economic gain (net of costs) of migration. Let g be the typical (lifetime) economic gain of an individual when migrating. At the start of the stage game, g is distributed in the population according to a cumulative distribution function $\mathcal{G}$ with support on $\mathbb{R}$. Assuming quasi-linear preferences,* the pool of type-i country A nationals who wish to migrate is composed of anyone such that

$$\beta g + (1 - \beta)(f(q^B) - f(q^A)) \geq 0, \quad (1)$$

with f the function by which cultural preferences are translated into utility units. We assume that individuals are *homophilic*, so that their utility increases in the share of same-type individuals in the country where they live: f is an increasing function. An individual would only emigrate to a foreign country where there are fewer people of the same cultural type if the economic gains can compensate him or her for it. And vice versa, individuals may be willing to incur an economic loss if their cultural benefits are high enough. β characterizes the relative weight of the economic motive in the migration decision, and $1-\beta$ the weight of the cultural motive. For simplicity, we assume that cultural types and economic gains are uncorrelated.*

We model the migration process as follows. At the start of the period, individuals discover their net economic gain from migrating g. Condition (1) then defines the pool of potential migrants. In the pool of migrants, the share of type-i individuals is given by π. From this pool, one individual is randomly selected to migrate. Each individual who migrates changes the cultural composition of both countries. This updates the pool of potential migrants dynamically. Within each period, the cultural composition of the two country evolves according to dynamics that can be simply written:

$$\begin{cases} \dot{q}^A = q^A - \pi \\ n\dot{q}^B = \pi - q^B. \end{cases} \tag{2}$$

The game reaches its equilibrium when the pool of potential migrants is empty.

The cultural composition of the pool of migrants comes from Condition (1). To simplify the notations, let us introduce $\mathcal{G}_i \equiv \mathcal{G}((1-\beta)(f(q^A)-f(q^B))/\beta)$ the fraction of type-i individuals not interested in moving, and similarly $\mathcal{G}_j \equiv \mathcal{G}((1-\beta)(f(1-q^A)-f(1-q^B))/\beta)$. $1-\mathcal{G}_i$ (resp. $1-\mathcal{G}_j$) is the fraction of type-i (resp. type-j) individuals who wish to migrate. The equilibrium is characterized by $\mathcal{G}_i = \mathcal{G}_j = 1$.

At each successive draw of a new migrant, the probability that her type is i is

$$\pi = \frac{q^A(1-\mathcal{G}_i)}{q^A(1-\mathcal{G}_i)+(1-q^A)(1-\mathcal{G}_j)}. \tag{3}$$

Notice that if $\beta = 0$, $\mathcal{G}_i = 0$, $\mathcal{G}_j = 1$, and $\pi = 1$. As scarce-type individuals leave, the incentive to migrate becomes stronger for remaining scarce-type individuals and weaker for abundant-type individuals, so that the pool of potential migrants remains the same. In equilibrium, country A keeps only other-type individuals, and the share of same-type individuals has increased in country B. The two countries diverge mechanically by *cultural self-selection*. Conversely, if $\beta = 1$, $\mathcal{G}_i = \mathcal{G}_j$, and $\pi = q^A$. The pool of potential migrants is a culturally representative sample of the home country population. q^B decreases as migrants start arriving. The two countries converge mechanically by *cultural mixing*. Finally, if $\beta \in]0,1[$, $\mathcal{G}_i < \mathcal{G}_j$ and $\pi > q^A$. The scarce cultural type is over-represented among migrants.

This compositional model of migration (COM) is heavily tilted towards predicting that migration results in cultural divergence between countries, but it does leave some place for doubt. Thanks to Eq. (2), if we have even a little homophily (ie. $\beta < 1$), $\dot{q}^A < 0$: the scarce cultural type becomes progressively scarcer in the home country (thus, the home country moves away from the host country). In the host country, π may be larger or smaller than q^B. If cultural homophily is strong enough, $\pi > q^B$, and the relatively abundant type becomes progressively more abundant (thus, the host country also moves away from the home country). In that case, unambiguously, the model predicts that the two countries will diverge. Conversely, if cultural homophily is weaker, $\pi < q^B$, the relatively abundant types become scarcer. If cultural homophily is weak enough, we may even find cultural convergence between countries A and B. However, COM does predict, without ambiguity, that migration results in cultural divergence between the home population / country and the *native population* in the host country:

Prediction COM1: Migration can both lead to cultural convergence (through cultural mixing) or divergence (if cultural self-selection is powerful enough) between home and host *countries*. Migration should lead to cultural divergence between home and host *populations*.

COM may be combined with standard economic intuitions to make further predictions. Consider, for instance, a uniform increase in the economic gain of migration by a fixed amount Δg relative to baseline. Again, to fix ideas, we assume $q^A < q^B$. A type-i individual wishes to emigrate iff $\beta(g + \Delta g) \geq (1 - \beta)(f(q^A) - f(q^B))$. The fraction of individuals not interested in emigrating is now given by $\mathcal{G}_i(q^A, q^B, \Delta g) = \mathcal{G}\left(-\Delta g + (1-\beta)/\beta\left(f(q^A) - f(q^B)\right)\right)$ among type-i individuals and $\mathcal{G}_j(q^A, q^B, \Delta g) = \mathcal{G}\left(-\Delta g + (1-\beta)/\beta\left(f(1-q^A) - f(1-q^B)\right)\right)$ among type-j ones. For a log-concave distribution function $\mathcal{G}$, the ratio $(1 - \mathcal{G}_j)/(1 - \mathcal{G}_i)$ increases with Δg, and π decreases. There are more would-be migrants of both types but the cultural selection effect becomes relatively less important. Therefore, the cultural mixing effect dominates.

Prediction COM2: Uniformly higher economic gains from migration should result in stronger convergence, or weaker divergence.

Now, consider the magnitude of the convergence for culturally distant or near countries. If two countries are farther apart, corresponding to a lower (negative) value of $f(q^A) - f(q^B)$ and a higher (positive) value of $f(1-q^A) - f(1-q^B)$, $1 - \mathcal{G}_i$ is larger and $1 - \mathcal{G}_j$ smaller. The cultural selection effect is strengthened and we would expect stronger divergence or weaker cultural convergence (π increases). This paints the picture of cultural clusters of countries that converge inside the clusters and possibly grow farther apart between clusters.

Prediction COM3: Cultural convergence should be stronger for relatively similar countries. Large cultural divides between countries should widen even further.

The compositional model offers the potential for various extensions and an even larger set of predictions.* For the purpose of this analysis, we stick to the most simple set-up of the model, where we focus on unidirectional migration, and neglect – for now – any spillovers between locals and immigrants in the host society. Overall, Prediction COM1 captures the core implication of the model. Predictions COM2, and COM3 rely on comparative statics that essentially come down to the effect of cultural selection on the magnitude of convergence or divergence. Cultural selection increases with initial cultural distance, and decreases with a uniform (positive) shock to the economics gains of migration. In COM, cultural selection should increase the magnitude of divergence, or reduce the magnitude of convergence.

3.2 Cultural diffusion

The model of cultural transmission proposed by Bisin & Verdier (2000) illustrates how different cultural types can coexist in equilibrium. Bisin and Verdier suggest that an individual would rather have a same-type offspring: this is another manifestation of their homophily. To that effect, they may invest in the socialization of their offspring. A larger effort increases the chances of a successful socialization. If the effort fails, the offspring picks a role model at random in the population. In equilibrium, the effort decreases with the frequency of your own type. This yields a structural expression of the cultural equilibrium q^*.

With probability τ_i, a type-i individual successfully socializes her offspring as a type-i. With probability $1 - \tau_i$, her offspring chooses a role model from the relevant population. In that population, the proportion of type-i individuals is χ. With socialization costs $H(\tau_i)$, the program of the type-i individual is

$$\max_{\tau_i} \left(\tau_i + (1-\tau_i)\chi\right) V_{ii} + (1-\tau_i)(1-\chi) V_{ij} - H(\tau_i) \tag{4}$$

where V_{ii} is the benefit for the individual of her offspring being of the same type, and V_{ij} of the other type. Under homophily, $V_{ii} > V_{ij}$, and we introduce the notation $\Delta V_i = V_{ii} - V_{ij} > 0$. We also assume quadratic costs $H(\tau) = \tau^2/2$. The problem is adequately concave: a type-i individual provides effort $\tau_i = (1-\chi)\Delta V_i$, and a type-$j$ $\tau_j = \chi \Delta V_j$ (for the problem to be well-defined, we assume that $\Delta V_i < 1/(1-\chi)$ and $\Delta V_j < 1/\chi$). In the population under consideration, with a cultural mix generically denoted q, the cultural equilibrium q^* is reached when the flow of type-i offspring socialized as type-j is equal to the flow of type-j offspring socialized as type-i, ie.

$$g(q,\chi) \equiv \frac{q(1-\chi)(1-\tau_i(\chi))}{(1-q)\chi(1-\tau_j(\chi))} = 1 \tag{5}$$

where χ is a function of q and of the cultural mix of any other group with influence over the socialization of the offspring. Bisin & Verdier (2000) propose a model of within-country cultural change. We extend their model to accommodate changing population boundaries, and in particular, migration in and out of a country. With two countries, we need to make assumptions on who influences whom. Contact seems to be a natural condition to pick a role model. It is likely to derive from living in the same country and from sharing a common nationality. Role models are picked among neighbors, and as shown in the introduction, there is ample evidence that emigrants continue to play a role in cultural change at home. As a result, there are three principal mechanisms of cultural diffusion that we wish to consider: migrants disseminating norms and values to natives (DSM), migrants assimilating into native culture,* and cultural remittances from migrants to their home community (REM). Intergenerational cultural formation is intrinsically dynamic in nature.*

3.2.1 A model of cultural dissemination (DSM)

First, migrants may disseminate norms and values to the native population of the host country B. Keeping the notations from the compositional model, let us write q^B as the share of type-i individuals in the native population, and π the share in the migrant population. A native offspring picks a role model from within the native population with probability η^B, and from the migrants with probability $1-\eta^B$. Overall, the offspring chooses a type-i role model with probability $\chi(q^B, \pi, \eta^B) \equiv (1-\eta^B)\pi + \eta^B q^B$. The cultural equilibrium q^{B*} is characterized by the equation $g(q^{B*}, \chi(q^{B*}, \pi, \eta^B)) = 1$. With a slight shift of notation, and without ambiguity, we rewrite this condition as $g(q^{B*}, \pi, \eta^B) = 1$.

To determine the comparative statics of q^{B*}, we determine the sign of the partial derivatives of g in the appendix.* We show that $\partial g/\partial q > 0$, $\partial g/\partial \pi < 0$, and that $\partial g/\partial \eta$ has the same sign as $\pi - q$. Since g is continuously differentiable, we can apply the implicit function theorem. q^{B*} can be written as a function of π and η^B, and:

- $\partial q^{B*}/\partial \pi > 0$. If migrants reflect their home culture faithfully, their departure does not change the cultural composition at home. However, their arrival pulls on the native culture at destination, leading to cultural convergence between the two countries. Conversely, cultural self-selection into migration* means a shrinking presence of the minority type at home (as in COM). If the migrants are so selected that the minority is overrepresented, even relative to the host population ($\pi > q^B$), it draws the native culture at destination further away. This suggests cultural divergence. Overall, this pattern reminds us of the predictions of COM.

- $\partial q^{B*}/\partial \eta^B$ has the same sign as $q^{B*} - \pi$. If migrants are assumed to be inspirational to natives, the culture of the host society is drawn towards the culture of the migrant group. If we interpret η^B as an inverse proxy for the magnitude of immigration into B, the larger the flow of immigrants, the stronger the effect on the destination culture.

Prediction DSM1: Migration may lead to cultural convergence between home and host populations (cultural mixing with dissemination) or divergence (if cultural self-selection is powerful enough).

We combine DSM with standard economic intuitions to make further predictions. Since the thrust of DSM is that cultural selection is associated with divergence, its predictions are hard to distinguish from the predictions of COM. Contrary to COM, DSM does not preclude cultural convergence between host and home populations. Other predictions, however, are common to the two mechanisms.

Prediction DSM2: Uniformly higher economic gains from migration should result in stronger convergence, or weaker divergence.

Prediction DSM3: Cultural convergence should be stronger for relatively similar countries. Large cultural divides between countries should widen even further.

In the logic of the model, the home population converges back eventually to the same equilibrium cultural mix (except in the extreme case where all type-i individuals have emigrated). This feature distinguishes COM and DSM: if the effect of migration was purely compositional, the home country would be left with permanently fewer scarce types after emigration. Unfortunately, we only know the relative position of countries, and the comparative statics of DSM match those of COM. In our empirical analysis, we will generally analyze their predictions jointly.

3.2.2 A model of cultural remittances (REM)

Second, we consider how migrants may 'remit' values and norms back to their home community. An offspring in the home community picks a role model from the home community with probability η^A, and from the emigrant population with probability $1 - \eta^A$. Overall, the offspring chooses a type-i role model with probability $\chi^A \equiv (1 - \eta^A)\pi + \eta^A q^A$, and a type-$j$ role model with probability $1 - \chi^A$. The cultural equilibrium q^{A*} is characterized by the equation $g(q^{A*}, \pi, \eta^A) = 1$. One more time, we can write q^{A*} as a function of π and η^A, and:

- $\partial q^{A*}/\partial \pi > 0$. The norms and values of of migrants affect cultural formation at home. In line with the intuition of COM, we expect the scarcer cultural type to be over-represented among migrants, and to remain so ($q^A < \pi$) whether they assimilate or not. According to this new mechanism, the share of the relatively scarce type-i individuals would increase in the home country A. This pattern is opposed to the predictions of COM and DSM.

- $\partial q^{A*}/\partial \eta^A < 0$. If migrants are assumed to be inspirational to those who stayed, they pull the cultural mix at home closer to the cultural mix of the migrant group. The more inspirational they are, the stronger the effect in the home country. If we interpret η^A as an inverse proxy for the magnitude of emigration, the larger the flow of migrants, the stronger the effect on the home culture.

In line with COM, $q^A < \chi^A < \pi$: cultural formation at home is drawn toward the host country culture. In contrast with COM and DSM, REM predicts convergence between the home and host populations, even once we exclude the migrants from the cultural mix at destination. This is a first prediction of REM that distinguishes it starkly from COM and DSM.

Prediction REM1: Migration should lead to cultural convergence between home and host *populations*.

Notice that this prediction is unconditional: we always expect convergence if transmission is indeed the dominant mechanism. Identically, we may combine REM with standard economic intuitions to make further predictions. Contrary to COM and DSM, the thrust of REM is that cultural selection helps cultural convergence. Here, stronger cultural homophily increases cultural self-selection into migration, π, which results in stronger diffusion of the cultural norms and values from the host country to the home country. We also expect this effect to increase with time. As a result, it yields predictions opposed to the ones mentioned above.

Prediction REM2: Uniformly higher economic gains from migration should result in stronger divergence, or weaker convergence.

In REM, cultural selection into migration acts as a magnifying force of convergence between countries, instead of divergence, as was suggested by COM or DSM. In contrast with the image of cultural clusters that they paint, REM suggests that convergence is stronger between countries further apart. This paints a different picture that COM and DSM: not that of cultural clusters of countries, but of universal cultural convergence.

Prediction REM3: Cultural convergence should be stronger for dissimilar countries. Relatively similar countries also converge but at a lower rate.

As hinted upon previously, we do not consider these mechanisms as separate models but rather as a system of incentives and dynamics that unravel at the same time, although not necessarily in the same time frame. While the compositional model hints at more short term effects of migration, the transmission mechanisms may reflect how the cultural technology of migration materializes in the longer term. In the very short run, selection into migration may push the home country away from the cultural mix at destination, but that effect is soon overrun by cultural remittances from successful migrants. The empirical analysis will serve as a way to inspect which of the mechanisms dominates. Since the predictions of REM vs. COM and DSM are diametrically opposed, we have the possibility to discriminate between them through our empirical analysis.

4 Data and measurement

This section describes the main data sources and focuses on the measurement of bilateral cultural similarity. In particular, we outline the reasoning behind the selection of statistical distance (or similarity) measures and explain how we apply them to the cultural space.

4.1 Main data sources

Cultural data. The World Value Survey [WVS] consists of nationally representative surveys among 400,000 respondents in 6 waves between 1981 and 2014. It includes questions on political beliefs, family values, religiosity, attitudes, and other dimensions of culture in a repeated cross-section of almost 100 countries. Additionally, we draw from the European Social Survey [ESS], which is also a cross-national representative survey on attitudes, beliefs and behavior patterns of diverse populations conducted every two years since 2001 in more than thirty countries of the European Union and some of its neighbors. Some questions being identical in the WVS and the ESS, we can combine the databases in later years. For instance, the question on generalized trust is available in both WVS and ESS: we can increase the number of country pairs for which we have bilateral cultural similarity indexes from about 6,700 to over 7,800.

Migration data. Migration data comes from the joint OECD and World Bank's Extended Bilateral Migration Database (Özden et al., 2011), which covers bilateral migrant stocks for each decade between 1960 and 2010.[4] Since we do not have data on migration flows, changes in migration stocks over time have to be interpreted as net migration. If from one year to the next there is the same amount of migrants returning to their home country and new migrants entering the destination country, we would not be able to observe this in the data. Therefore, the change in the migration stock will likely underestimate the back and forth migration between two countries and the cultural diffusion associated

[4] https://finances.worldbank.org/Other/Bilateral-Migration-In-2010/hc8y-24bu for the 2010 data.

with it. In addition, we use data from Brücker et al. (2013) [thereafter, IAB] who collected data on immigration to 20 OECD countries by gender, country of origin and educational level, for the years 1980-2010 in 5 years intervals. The authors distinguish between three levels of skill in their data: lower secondary, primary and no schooling (low skilled), high-school leaving certificate or equivalent (medium skilled) and higher than high-school leaving certificate or equivalent (high skilled).

In order to match the bilateral migration data with the WVS waves, we interpolate bilateral migration in five-year increments, assuming a linear growth rate. As the WVS are carried out over the course of 3 to 5 years for each wave, we use the stock of bilateral migrants before the roll out of the next WVS wave, creating a lag of up to five years.

Other data. Our main time-varying bilateral control variables include bilateral trade and bilateral GDP per capita differences. The United Nations ComTrade Database provides yearly bilateral trade flows around the globe, which we average over the periods corresponding to the World Value Survey waves. We benefit from the efforts of the Center for International Data (CID), an organization within the Department of Economics at UC Davis that collects, enhances, creates, and disseminates international economic data, from which we draw harmonized yearly bilateral trade matrices for thousands of country pairs. Data on GDP per capita are taken from the World Bank and used to calculate a bilateral measure of economic distance (GDP per capita difference). Summary statistics for all bilateral measures are presented in Table 1. Our variables of interest, i.e. migration, trade and GDP gaps enter the regression in logs but are presented as volumes in Table 1.[5] We also use data on time-invariant proxies for bilateral cultural distance to assess the plausibility of our time-varying cultural similarity index. We take the linguistic proximity and religious proximity (weighted relevant sub-populations of countries) from Fearon (2003). We also take the fixation index that measures differences in the genetic structure of populations from Cavalli-Sforza et al. (1994). Geographic distance comes from Spolaore & Wacziarg (2009b).

Finally, we use additional data sources such as the World Bank bilateral remittances matrix for 2010, the World Travel Organization "Compendium of

[5]The list of countries that we use in our main empirical specification are presented in Tables C4 and C5 in the appendix.

Table 1: Summary statistics

Source	Independent Variables	Obs.	mean	sd	min	max
World Bank	migrant stock	9,654	2.76e4	2.65e5	0	1.30e7
IAB	low skilled migrant stock	1,982	1.76e4	1.86e5	0	5.29e6
	medium skilled migrant stock	1,982	1.50e4	9.51e4	0	2.63e6
	high skilled migrant stock	1,982	1.93e4	7.95e4	0	1.32e6
UNComtrade & Feenstra	bilateral trade volume (billion USD)	8,430	1.54	9.38	0.00	339
World Bank	GDP per capita gap (thousand USD)	9,680	13.3	13.5	0.00	65.0
Cultural Similarity Index	(standardized)					
	Euclidean	10,646	0	1	-4.43	2.03
	Herfindahl	10,646	0	1	-3.61	5.28
	Canberra	10,646	0	1	-3.76	2.90

Tourism Statistics" from 1995 to 2010 and data from the International Telecommunication Union, from 1960 to today. Due to many missing values in earlier periods, we do not include these data in the set of time-varying controls (some of these would be bad controls, as they directly relate to migration) but we aggregate observations across years and run sub-sample analyses in order to explore heterogeneous effects at the bilateral level as plausibility checks.

4.2 Measuring cultural similarity

In this section, we compare the properties of a set of measures in multidimensional spaces and we discuss how these measures can be adapted for our purpose (i.e., measuring cultural similarity), highlighting their strengths and limitations.

Culture as a multidimensional space

The appropriate choice and careful interpretation of various measures is central to the empirical analysis of cultural convergence. In this section, we introduce different measures that we will use to compare two multidimensional distributions of cultural values (whose means we will call P in country A and Q in country B). With regards to statistical inference, different measures highlight certain aspects of the underlying distributions and let us draw different conclusions about their properties (see Table C2 in the appendix for the complete list of measures discussed in this section).

For the purpose of our analysis, we will use three measures: the Euclidean distance, the Canberra distance, and the Inner Product (or Herfindhal Index). The most well-known group of distance measures are derivatives of the Minkowski norms.[6] In a $d-$dimensional space, the $p-$norm is written as $M_p = (\sum_{i=1}^{d} |P_i - Q_i|^p)^{1/p}$. The Canberra, the Euclidean, and the Chebyshev distances correspond respectively to M_0, M_2, and M_∞. With M_2, we get one of the most commonly used measures: the Euclidean norm, also known as the Pythagorean metric or geometric distance, that weighs equally the various dimensions of culture. For intuitions on edge cases with M_0 for the Canberra norm and M_∞ for the Chebyshev, we can consider two almost identical countries that differ significantly along one cultural dimension. These countries will be characterized as far apart by the Chebyshev distance, but very similar by the Canberra norm. Conversely, two countries that differ a bit according to every dimension will be characterized as further apart by the Canberra norm than by the Chebyshev distance. Another family of statistical similarity measures is the so-called "Inner Product." It is written as $D_I = \sum_{i=1}^{d} P_i * Q_i$ (also known as the Herfdindahl concentration index) and captures the overlap or the number of matches between two distributions P and Q. In the context of cultural values, this measure would capture the idea that two people, one in country A and the other in country B, picked at random would give the same answer to a question in the WVS.

We follow the literature (Desmet et al., 2011; Spolaore & Wacziarg, 2016; Bell et al., 2009; Desmet & Wacziarg, 2018; Bertrand & Kamenica, 2018; Alesina et al., 2017; Maystre et al., 2014) in picking two of the most commonly used measures in cultural economics: the Euclidean distance and the Herfindahl index. We complement these two measures with the Canberra index as it will dampen the impact of outliers on the index of cultural similarity of two populations.

Measures of cultural similarity

There is a vast number of cultural dimensions along which countries can be differentiated from each other, including family values, generalized trust, religiosity, or political or economic ideologies. Maystre et al. (2014), Desmet et al. (2017), and De Santis et al. (2016), for instance, include a rich set of questions to map culture within and across countries. As we are concerned with the dynamic process behind the formation of attitudes, norms, and values, we rely on a cul-

[6] A thorough discussion of the different distance / entropy / divergence measures can be found in Cha (2007)

tural similarity measure that is consistent over time. We only include questions that were asked in the same way across all WVS waves in all the countries or our data set to avoid that changes in the composition of questions are driving our results. Naturally, this requirement will limit the scope of questions that we can include in our cultural similarity measure.

We summarize the set of questions used for our index in Table C1 in the appendix. The cultural dimensions that are covered consistently across WVS waves include: values transmitted to children, priorities in life, generalized trust, views on gender equality, and the sense of control over ones live. We have no prior on which values are most important for the migration decision and which values are more or less easy to assimilate to, diffuse or transmit. We would argue, however, that all of the cultural dimensions that we use for this analysis, such as the importance of religion or women's ability to acquire education, are quite fundamental in their implications for everyday life.

To illustrate and clarify the construction of the cultural proximity measure, we use the following example: the Euclidean distance along the cultural dimension of *important child qualities*. Respondents of the WVS can choose to pick 5 out of 11 possible character traits that they would like to pass on to their children (see Table C1 in the appendix for a complete list of character traits) which yields a set of 11 binary responses (0 or 1) to each characteristic listed. For two randomly picked individuals, the response matrix would look like this:

$$\mathbf{I1}: 0\ 0\ 1\ 1\ 0\ 1\ 0\ 1\ 1\ 0\ 0$$
$$\mathbf{I2}: 1\ 0\ 1\ 1\ 0\ 0\ 0\ 1\ 1\ 0\ 1$$

For each wave, we consider the vector of the shares of people who have picked each characteristic in a country in the Euclidean space with eleven dimensions. To characterize the distance between two countries, we consider the (Euclidean, Canberra, Herfindahl) distance between their respective vectors. Within each wave, we normalize the distribution and multiply it by -1 in order to get a measure of cultural similarity rather than of distance. Notice that because of the normalization, dissimilar countries will be characterized by a negative measure of similarity, and similar countries by a positive one. Normalization also corrects for the dimensionality of cultural sub-spaces. We present the distribution of all three measures in Figure C1 in the appendix (summary statistics on the Cultural Similarity Indexes are also presented in Table 1).

We are agnostic regarding the best distance measure but emphasize that the choice of a single statistical distance measure is associated with a choice in statistical inference that needs to be carefully interpreted. Our empirical analysis will use three different distance measures (Euclidean, Herfindahl, Canberra). In addition, several questions have ordinal, rather than binary responses. None of the distances we consider suggests an easy way to treat such answers. In particular, how far apart do we believe people who answered *very important* are from others who answered *important*, vs. people who answered *not very important* from others who answered *not important at all*? To address that issue, we choose to consider people who pick any different answers as equally dissimilar from each other. Any other approach would require equally strong assumptions on the relative distances between answers. In doing so we follow the economic literature in quantitatively measuring cultural distance based on on qualitative information (Desmet et al., 2017; De Santis et al., 2016).

The three measures capture different distributional aspects of cultural similarity between countries. While the correlation between the three measures is – of course – positive and significant, the correlation coefficients are rather small. The Herfindahl measure correlates to 40% with the Euclidean measure and 57% with the Canberra measure. The correlation between the Euclidean and the Canberra measure is higher and lies at about 68%, which is still rather low considering that the Canberra index "only" corrects for outliers and domain scales (we report the distributions of the three measures in Figure C1 and the correlations between the three measures in Table C3, both in the appendix).

Commonly used cultural distance measures in the literature (see Spolaore & Wacziarg, 2009b, for instance) explain only a small part of the survey-based cultural similarity measure constructed for our purposes. We find weak correlations of our three measures of cultural similarity with geographic and genetic distance, as measured in Cavalli-Sforza et al. (1994), and a larger correlation with religious and linguistic proximity, as measured in Fearon (2003) (we report the correlations in Table C3 and we also provide the corresponding scatterplots in Figure C3, both in the appendix).

5 Empirical analysis and main results

We are interested in the effect of bilateral migration on the change in cultural proximity over time. In the first step of our empirical analysis, we show that migration increases cultural proximity between sending and receiving countries. In a second step, we investigate the mechanisms through which migration leads to cultural convergence. Before we present the results, we discuss our baseline specification and some identification issues.

5.1 Empirical specification

In our baseline specification we follow Egger (2000) in including sending country-time and receiving country-time fixed effects [FE] to account for time-specific shocks to sending and receiving countries. We also include bilateral FE to control for time-invariant characteristics of country pairs, accounting for standard gravity controls (contiguity, geographical distance etc.) as well as, for instance, a common language or a historical relationship such as colonization. Hence, our baseline specification tracks changes within country pairs over time:

$$CS_{ijt} = \beta_0 + \beta_1 Mig_{ij,t-\Delta} + \beta_2 X'_{ij,t-\Delta} + \theta_{ij} + \theta_{it} + \theta_{jt} + \varepsilon_{ijt}$$

CS_{ijt} is the bilateral cultural similarity between countries i and j in period t. Our main coefficient of interest is β_1. Both migration $Mig_{ij,t-\Delta}$ and the vector of time-varying controls (including bilateral trade flows and GDP per capita differences) $X'_{ij,t-\Delta}$ are lagged. As explained, when the WVS wave starts in the middle of the decade, we use bilateral data from the previous period for which we have an observation. For instance, if the WVS wave starts in 1994, we take data from 1990 with a $\Delta = 1$ lag; for $\Delta = 2$ we use data from 1985, etc.

As mentioned in the previous section, migration is measured in stocks whereas trade is measured in flows. Consequently, in the specification with the full set of FE the variation for migration comes from the change in the bilateral stock of migrants (or net migration) and variation for trade comes from a change in the flow of goods between countries. A 1% increase in the bilateral stock of migrants will affect cultural proximity by β_1. Similarly, a 1% increase in the flow of goods will impact cultural proximity by β_2 with the dependent variable standardized to a mean of zero and the standard deviation set to one.

The dependent variable, cultural similarity, is by construction symmetric. The main explanatory variable, migration, is an asymmetric measure that captures the number of migrants from j living in i. Consequently, in our data set each country pair is going to appear twice, where i and j switch roles as receiving or sending country. In other words, cultural similarity between Mexico and the US is going to be the same for a specific year but we will observe Mexican migrants in the US and American migrants in Mexico separately.[7]

We show the raw conditional correlation between migration and cultural similarity first, and then include FE in our regression. For the baseline regressions we use the aggregate measures of cultural similarity, later differentiating between various cultural dimensions. In a next step, we construct the largest balanced panel that we can build from the data (24 countries in 3 waves) to rule out that the overall effect is driven by sample selection.[8] The coefficient β_1 is the net overall effect of all five mechanisms.

5.2 Empirical strategy

Our empirical strategy aims at estimating the effect of migration on cultural similarity between countries. The main identification challenge we face is the endogeneity of both the size and cultural composition of migration. Indeed, more cultural proximity translates into a larger and less culturally-selected pool of migrants; at the same time, bilateral migration (its size as well as its cultural composition) and cultural proximity could be driven by third factors that we may only partially observe and control for.

We address endogeneity as follows. First, our empirical specification includes both country-time (for origin and destination) and bilateral FE. Country-time FE account for the overall cultural, demographic and socio-economic envi-

[7]Our specification with the full set of FE would not allow us to lump together immigrants and emigrants. If we were to look at the effect of all Americans in Mexico and all Mexicans in the US on the cultural proximity between the US and Mexico, we could no longer include origin and destination FE (which country would be destination and which would be origin?). When we run the regression with cumulative migrants stocks in both countries and only include bilateral FE, the results are very similar to those of Column 3 of Table 2.

[8]The imputation methodology for the bilateral migration data set changed between 2000 and 2010, leaving us with more missing values and zeros for the 2010 migration data set (see Özden et al. (2011) for an overview on the methodology). We therefore repeat the analysis, dropping all country pairs for which we observe zeros in 2010 from the analysis. For the whole observation period, this reduces our sample by about 300 observations. The results hold with this modified sample. Results are presented in Table C10 in the appendix.

ronment at destination and origin (thereby controlling for deep-rooted determinants of culture such as absolute geography or genetic make-up of the population(Ashraf & Galor, 2013; Alesina & Giuliano, 2015; Giuliano & Nunn, 2020; Galor & Savitskiy, 2018) as well as for time-specific and country-specific trends of any sort (e.g., migration policies, rise of populism, economic cycles, natural disasters, technological changes, etc.) that can impact both migration and cultural proximity.[9] The inclusion of bilateral FE, on the other hand, implies that we only exploit within-country pair variation over time for identification. Indeed, bilateral FE capture fundamental time-invariant cultural similarities between sending and receiving countries such as ethnic, linguistic or religious proximity, past common (e.g., colonial) history, etc. This should attenuate endogeneity concerns because, together with the overall cultural shifts in destination and origin countries accounted for, these fundamentals of cultural similarity are typically also the ones that drive individual migration decisions (Mayda, 2010; Belot & Ederveen, 2012). Moreover, we also include time-varying controls (trade and income per capita differences) that are potentially important confounders in the migration and cultural proximity relationship.

In addition, we exploit the time-structure of our data in two ways. First, in our baseline specification we lag migration by 5 years. This implies that for reverse causality to explain our results, migrants would have to be able to anticipate cultural changes both at origin and at destination over several years before they move. While this is unlikely, one could still argue that migrants can observe overall trends in cultural convergence and extrapolate them to the future. In order to overcome this concern, we include contemporaneous cultural similarity at the time of migration in a robustness check (migration and cultural similarity in $t-1$ determining cultural similarity in t) and the results remain almost unchanged (see Table C11 in the appendix). In sum, the rich set of FE and time-varying controls, paired with the lagged time structure and the many robustness checks we perform go a long way toward addressing endogeneity.

One may still be concerned about unobserved bilateral time-varying factors driving our results. Typically, cross-country studies of the impact of migration

[9]Origin- and destination-time FE serve another important purpose: they address concerns about multilateral resistance in the context of migration and cultural similarity. In other words, convergence to one specific destination country could also impact the distance to all other countries. The FE structure "anchors" the origin and destination country in the cultural space, thereby identifying the impact of migration on the convergence between two countries holding all other distances constant.

(e.g., on income per capita, trade, etc.) have addressed the endogeneity of the *size* (and, occasionally, the skill-composition) of migration using various instrumental variable [IV] strategies. Most commonly used instruments have been exogenous (e.g., geographic) bilateral gravity variables, interactions of push (e.g., weather shocks) and pull factors, as well as 'shift-share' – or 'Bartik' – instruments based on preexisting migration networks at destination (Felbermayr et al., 2010; Andersen & Dalgaard, 2011; Bellini et al., 2013; Ortega & Peri, 2014; Alesina et al., 2016; Docquier et al., 2016; Beine & Parsons, 2017; Bahar & Rapoport, 2018; Burchardi et al., 2019; Docquier et al., 2020b; Tabellini, 2020).[10]

However, there are a few reasons – beyond the sheer difficulty in finding sensible instruments – why an IV approach may not be appropriate in our context. First, we cannot observe the cultural composition of migration and, therefore, cannot instrument for it. Second and relatedly, while we do observe the number of migrants and could in principle try to instrument for the size of migration, IV estimations only produce a Local Average Treatment Effect (LATE): this is a concern for us because the cultural composition of migration for compliers would clearly be determined by the choice of instruments. In fact, if we were to predict bilateral migration stocks as part of an IV strategy, each candidate set of instruments would produce a different set of compliers with a differently culturally-selected migrant pool. And since we do not observe cultural selection, we could only make vague inferences as to how we expect certain IVs to influence cultural selection into migration for compliers.[11] An IV would therefore not provide a more transparent and reliable estimate for the net effect of migration on cultural similarity; rather, it would amount to performing a sub-sample analysis for country pairs with more or less culturally selected migrants, which is what we do explicitly – with guidance from our theoretical model – in the mechanisms section below. Finally, recall that an implication from our theoretical model is that since cultural similarity decreases cultural selection, any mechanism that relies on a less selected migrant pool (e.g., dissemination or cultural mixing) will be favored by reverse causality; the opposite holds for any mechanism that relies on a more selected migrant pool (e.g., cultural remittances). In other words, if the evidence pointed to cultural

[10] On the limits of shift-share instruments, see Jaeger et al. (2018); Goldsmith-Pinkham et al. (2020).

[11] See Appendix B for an illustration using an IV gravity framework following Feyrer (2019).

remittances as dominant mechanism behind our results (which will indeed be the case), our point-estimate would capture a lower bound effect.

5.3 Convergence or divergence?

In this section we ask whether migration generates cultural convergence or divergence. We proceed with a reduced form analysis, using various sets of fixed effects, different sample compositions, and three distance measures. It is only in Section 6 that, guided by theory, we will exploit heterogeneous effects resulting from different balances between economic and cultural incentives for migration (and, hence, different likely degrees of cultural selection) across country-pairs to gauge the relevance of various candidate mechanisms in driving our results.

Table 2 presents the main results for the full sample.[12] The results are presented for the Euclidean, Herfindahl and Canberra measures of cultural similarity aggregated over all cultural dimensions, including values parents want to transfer to their children, priorities in life, attitudes toward gender equality, generalized trust, and freedom of choice. Column 1 presents the raw correlation between lagged bilateral migration and cultural similarity. We then successively introduce different types of FE. In Column 2 we control for time-varying characteristics of host and home countries. This is particularly important to control for unilateral changes and cultural shifts in the respective countries that are not related to migration. This set of FE also 'anchors' the countries position in the overall cultural space, as it captures the aggregate of their respective distances to all other countries. In Column 3, we introduce bilateral FE, and the full set of FE in Column 4. Finally, in Column 5, we introduce time-varying bilateral control, including bilateral trade flows and the GDP per capita difference between the origin and destination country. We take Column 5 as our preferred specification.

We find a consistent positive association between changes in bilateral migration at time $t-1$ and changes in cultural similarity at time t. This is true unconditionally, and robust to including various sets of FE, to using various measures of cultural proximity, and to controlling for potentially important confounders, namely, bilateral trade and income differences. When we control for time-varying characteristics, the variation explained by our model increases substantially, as expected. Bilateral FE reduce the magnitude of the correlation and the power of the estimation, capturing (and ridding us of) time-invariant

[12]We show the results split by theme in Appendix Table C9

Table 2: Baseline - migration and cultural similarity

	(1)	(2)	(3)	(4)	(5)
Euclidean					
Migration	0.054***	0.083***	0.083***	0.022**	0.029**
	(0.004)	(0.004)	(0.010)	(0.009)	(0.012)
R^2	0.03	0.52	0.91	0.97	0.97
Herfindahl					
Migration	0.030***	0.066***	-0.048***	0.017**	0.023**
	(0.004)	(0.003)	(0.013)	(0.007)	(0.010)
R^2	0.01	0.70	0.86	0.98	0.98
Canberra					
Migration	0.052***	0.084***	0.056***	0.017	0.006
	(0.004)	(0.004)	(0.013)	(0.011)	(0.013)
R^2	0.03	0.60	0.86	0.95	0.96
Dest.-year FE		X		X	X
Orig.-year FE		X		X	X
Bilateral FE			X	X	X
Controls					X
Obs.	7,486	7,486	7,486	7,486	5,875†

*** $p<0.01$, ** $p<0.05$, † 2,400 observations lost due to no information for some country pairs on bilateral trade and GDP per capita. Standard errors in parentheses. Controls: Trade and GDP Gap. Constant is not reported. Migration is the log of the migrant stock at time $t-\Delta$ and Trade is the bilateral trade flow at time $t-\Delta$, GDP Gap is bilateral per capita difference at $t-\Delta$. In our baseline regression Δ represents a five year lag.

determinants of both migration and cultural shifts.[13] With the full set of FE, the magnitude of the correlation is reduced, but the precision of our estimation is improved. Finally, the inclusion of controls does not change our results substantially. In Appendix Table C7, we repeat the analysis for a balanced panel to rule out that compositional changes in our sample drive the results.[14]

[13]We attribute the negative sign for the Herfindahl index in Column 3 to the fact that we do not account for overall cultural shifts at origin and destination countries at time t. The Herfindahl index measures the overlap between two populations (the likelihood that two people picked at random respond in the same way) and is therefore more sensitive to overall shifts in the respective cultural compositions of the countries at hand than the other measures we use.

[14]The magnitude of the effect is much larger than the one we find for the unbalanced panel, suggesting that compositional effects introduce a downward bias, underestimating the role of migration in cultural convergence. Additionally, the magnitude of the coefficient for the Canberra

In our preferred specifications we include trade (in addition to income differences) as control, situating our results in the literature on trade and cultural convergence. Maystre et al. (2014) find a positive association of trade on cultural similarity.[15] In our most demanding specification in Column 5, the effect of bilateral trade on cultural convergence disappears. Maystre et al. (2014) only use data from the 2nd and 4th wave of the WVS, selecting a set of 30 questions and building an index of fractionalization, akin to the Herfindahl Index. They do not include country-time FE and differentiate between different types of trade, which may explain the differing results. Exploiting a larger sample, including all waves of the WVS, using several distance measures as well as adding income differences and migration as additional controls, the coefficient of trade vanishes.

How should we interpret the magnitude of the effect? Focusing on the specification with the full set of FE and on the Euclidean distance, we estimate that a 10% increase in migration is associated with an increase in cultural similarity of 0.2, i.e. one fifth of a standard deviation in the following period. An increase in migration from Italy to the United States by 10% would make the United States and Italy as similar as the United States and the United Kingdom. Similarly, in order for Mexico and the United States to become as similar as the United States and Canada, the United States would have to experience an eight-fold increase in migration from Mexico. This masks – of course – substantial heterogeneity across country pairs. In fact, we show in Appendix Table C8 that our effect is mainly driven by South-North migration.[16]

6 Mechanisms

The results above show that migration is associated with cultural convergence between countries. This is consistent with the cultural mixing effect at destination (COM) with or without dissemination (DSM) and with cultural remittances (REM). In this section, we discriminate between these mechanisms based on the

measure is more than six times higher and is now statistically significant at the 5% level. Overall, the standard errors between the unbalanced and balanced panel remain relatively stable while the magnitude of the effect increases, delivering higher statistical significance overall.

[15]In Table C6 in the appendix, we include trade in the estimation from the beginning successively introducing different fixed effects.

[16]This is taken from the Euclidean distance measured in 2005, where the CS between the US and Italy is 1.2, between the US and the UK 1.4, between the US and Canada 1.76, and between the US and Mexico 0.05

distinctive predictions of the theoretical model.

6.1 Static or dynamic convergence?

We start with a general look at the timing of convergence relative to migration, as a way to discriminate between static and dynamic drivers. Compositional effects (self-selection and mixing) should be immediate, while diffusion effects (dissemination, assimilation, and cultural remittances) should magnify over time (up to some point). We then examine COM1 vs. REM1 directly, excluding migrants from the population of the destination country.

6.1.1 Timing of migration

In Table 3, we report the results of the main specification with different lags for the migration variables. In the baseline regression, we use a five-year lag for migration to let both compositional and diffusion mechanisms unfold.[17] In Columns 1 to 3 of Table 3, we consider a ten-year lag, and in Columns 4 to 6, contemporaneous migration (i.e., migration measured at the same point in time t as our dependent variable). In all of the specifications, we adjust bilateral control variables to match the corresponding lag. We find that with a longer lag, the effect of migration on cultural similarity remains positive and strongly significant. With contemporaneous migration, the coefficient for migration remains positive but loses statistical significance.[18] Overall, we conclude that the effect of migration magnifies as time passes. These results suggest that convergence is driven by a dynamic rather than compositional mechanism, i.e. DSM1 or REM1 rather than COM1. That is not to say that we can exclude compositional effects, rather that they are dominated by another mechanism, especially after a few years.

6.1.2 Excluding immigrants

As seen from the baseline regression, migration is associated with bilateral cultural convergence. COM suggests that the driving force behind cultural

[17]Note that bilateral migration is interpolated between decades so that for instantaneous migration WVS waves may not exactly correspond to the year of interpolation.

[18]We visualize these results in Figure C2 in the appendix, showing that the effects becomes larger as we increase the lag of migration. While the coefficients for the 10-year lag are slightly larger, they are statistically indistinguishable from one another which may be due to the fact that the 10-year lag is measured with more error (smaller sample size).

Table 3: Timing of migration and cultural similarity

	10-year lag			*no lag*		
	(1)	(2)	(3)	(4)	(5)	(6)
	Euclidean	**Herfindahl**	**Canberra**	**Euclidean**	**Herfindahl**	**Canberra**
Migration	0.024*	0.027**	0.040***	0.011	0.006	0.016
	(0.013)	(0.011)	(0.014)	(0.011)	(0.009)	(0.012)
Controls / FE	All	All	All	All	All	All
R^2	0.97	0.98	0.96	0.97	0.98	0.95
Obs.	5,873	5,873	5,873	6,022	6,022	6,022

*** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses. Controls: Trade and GDP Gap. FE: Destination-year, origin-year, bilateral.

convergence should be should be cultural mixing (i.e., more immigrants in the destination country mechanically making home and host countries more similar). Table 3 showed first evidence that compositional effects are likely to not play a major role. Excluding migrants from the host population, however, should allow us to move further from the cultural mixing effect. COM, reduced to self-selection into migration, predicts divergence (COM1). REM1 predicts convergence, while DSM1 remains indeterminate.

In Table 4, we report the results of our main specification on the two waves of the WVS for which we have information about the birthplace of respondents. For 10 countries in wave 2 and 46 countries in wave 3, we can infer the migratory status of respondents. About 5.5% of respondents are foreign-born. We replicate our analysis for this subset of countries excluding the foreign-born from the construction of the aggregate Euclidean distance (the results hold for the Herfindahl and Canberra measures). As in Table 2, we successively introduce fixed effects and time varying controls. Excluding the foreign-born from the analysis does not alter our results. The use of information on respondents' birthplace cuts our sample substantially. The magnitude of our estimates remains virtually identical, but with only 9 countries covered in both waves with information on the country of origin of the respondent, the results lose significance when adding country-pair fixed FEs (and the R^2 is close to one). To the extent that we can interpret this result, it supports REM1 against COM1.

Table 4: Excluding foreign-born respondents

	(1)	(2)	(3)	(4)	(5)
Euclidean cultural similarity with migrants					
Migration	0.064***	0.074***	-0.020	0.032	0.007
	(0.009)	(0.007)	(0.074)	(0.084)	(0.148)
R^2	0.04	0.63	0.99	1.00	1.00
Euclidean cultural similarity without migrants					
Migration	0.064***	0.074***	-0.012	0.031	0.003
	(0.008)	(0.007)	(0.073)	(0.084)	(0.147)
Dest.-year FE		X		X	X
Orig.-year FE		X		X	X
Bilateral FE			X	X	X
Controls					X
R^2	0.04	0.63	0.99	1.00	1.00
Obs.	1,475	1,475	1,475	1,475	838†

*** p<0.01, † 600 observations lost due to no information for some country pairs on controls. Standard errors in parentheses. Controls: Trade and GDP gap. The sample is reduced to only the countries and WVS waves (second and third) for which information on respondent's country of birth was available (losing about 80% of our observations).

6.2 Economic vs. cultural gains from migration

In the previous section, we established that compositional effects are not a likely driver of migration-based cultural convergence. Here we look more closely at predictions 2 and 3 (specifically DSM2 and DSM3 versus REM2 and REM3) which are the mirror image one of the other and address the economic versus cultural gains from migration. The combination of cultural and economic gains from migration gives us insight into the cultural selection of the migrant pool.

In our theoretical model, we have shown that relatively larger economic gains from migration are associated with a less culturally selected migrant pool. Cultural distance between sending and receiving countries is associated with a more selected migrant pool (as cultural gains from migration increase with cultural distance). DSM2/DSM3 and REM2/REM3 yield opposite predictions as to how the cultural selection of the migrant pool will affect cultural convergence. This allows us to distinguish between the two mechanisms empirically. The dissemination mechanism predicts that less culturally selected migrants will bring more convergence; the cultural remittance mechanism predicts instead that more cultural selection will boost convergence.

To disentangle the two effects, we start by exploiting information on skills at the individual level and look at the effect of a change in the skill pattern of bilateral migration on bilateral cultural similarity. At the country-level, we identify country pairs that are at the same time economically distant *and* culturally similar (i.e., where we expect the migrant pool to be less culturally selected). Since culturally dissimilar countries may also offer the largest economic gains from migration, we cannot look at these two elements separately. Rather we use the combination of the two to unambiguously test the theoretical predictions.

6.2.1 Cultural selection at the individual level: the role of skills

Skilled labor is scarce in poor countries and abundant in rich countries. In international migration economics, this basic observation translates into an expectation of higher economic incentives to migrate for low-skilled migrants than for high-skilled migrants (Borjas, 1987; Chiquiar & Hanson, 2005; McKenzie & Rapoport, 2010): we expect less cultural sorting at lower skill levels. According to COM and DSM, low-skill migration would boost cultural convergence, while REM predicts the opposite (and conversely for highly skilled migrants).

In Table 5, we report the results of our main specification where we have decomposed migrant stocks by skill level. To obtain this decomposition, we use the IAB data set, where migrants are defined as highly-skilled if they have a college degree or above, and low-skill otherwise. The IAB data set covers 20 OECD destinations, which yields a limited sample of about 1,700 observations of immigration to OECD countries from both non-OECD and other OECD countries. In all columns, we use the full set of FE and time-varying controls.

We find a consistently negative relationship between low-skill migration and cultural similarity (significant for the Euclidean and Herfindahl index) and a consistently positive relationship between high-skill migration and cultural similarity. While we can only make limited inference from these results, we take them as suggestive evidence favoring REM2 over COM2/DSM2.

6.2.2 Cultural selection at the country-pair level

In Table 6, we report the results of our main specification, using the balanced panel of countries constructed for Table C7, and distinguishing country pairs

Table 5: Economic vs. cultural gains from migration - skilled migration and cultural similarity

	(1) **Euclidean**	(2) **Herfindahl**	(3) **Canberra**
low skilled	-0.085**	-0.065*	-0.069
	(0.043)	(0.037)	(0.065)
high skilled	0.082*	0.058^{+}	0.032
	(0.044)	(0.038)	(0.067)
Controls / FE	All	All	All
R^2	0.99	0.99	0.96
Obs.	1,717	1,717	1,717

+ $p < 0.15$; * $p < 0.1$; ** $p < 0.05$. Standard errors in parentheses. The IAB data set distinguishes between skilled (high-school degree and above) and unskilled labor.

Table 6: Economic vs. cultural gains from migration - culturally close and economically distant

	(1) **Euclidean**	(2) **Herfindahl**	(3) **Canberra**
Migration	0.068***	0.043***	0.058***
	(0.017)	(0.013)	(0.020)
Migration**CSED*	-0.041^{+}	-0.014	-0.129***
	(0.028)	(0.021)	(0.041)
Controls / FE	All	All	All
R^2	0.93	0.96	0.91
Obs.	1,359	1,359	1,359

+ $p < 0.15$; *** $p < 0.01$. Standard errors in parentheses. *CSED* is a bilateral dummy variable that indicates above median culturally similar and above median economically distant (less selected migrant pool). Panel sample of countries.

that are economically distant and culturally similar. We compute the median economic distance (as difference in GDP per capita) as well as the median cultural similarity of all three measures for all country pairs in that sample. We use the panel data set to make sure that the cut-off calculated at the beginning of the observation period (1995) holds for all country pairs consistently afterwards. We create a dummy variable called *CSED* to characterize country pairs that are both culturally similar (above the median) and economically distant (above the median). A dummy equal to one captures country pairs for which we expect *less* cultural selection into migration.

We find that convergence dampens when migration is less culturally selected. In line with the results of the previous section and Table 5, cultural convergence is stronger when there is cultural self-selection into migration, in support for predictions REM2/3, against predictions COM2/3 and DSM2/3.

6.3 Plausibility checks

In the previous sections, we used the predictions of our theoretical model to discriminate between possible mechanisms that could yield the same reduced form result of migration-based cultural convergence. We first focused on the compositional versus dynamic components and then turned to the mechanisms behind dynamic convergence. Our analysis suggests that cultural remittances are the dominant mechanism that explains why migrations drives cultural convergence. We now move away from the stricter model predictions and run a series of plausibility checks testing for the cultural remittance mechanism. Specifically, we use additional data sources to investigate whether country pairs that have higher intensity of interactions also converge culturally at a faster rate. We use financial remittances, as well as bilateral international travel and telecommunications data to proxy for emigrants' ties to their home community.

Since most of these data are only available for later years (for instance, a comprehensive global bilateral remittances matrix is only available for 2010), we cannot include them as controls (we would also run into the risk of overfitting the model). Instead, we run sub-sample analyses for country pairs that have above and below median interactions with their home countries. We run three different types of sub-sample analysis: above and below median bilateral remittances, above and below median bilateral travel (number of non-resident visitors from the migration destination country), as well as above and below median bilateral telecommunication.[19] For all of these measures we find that the coefficient for migration is larger and significantly positive for countries that interact more with their home communities (we report these results in Tables

[19]The bilateral telecommunication variables are generated by interacting unilateral variables such as sending and receiving country access to international voice traffic, number of fixed lines, and number of cellular or broadband subscriptions.

C12, C13, C14 in the appendix).

The cultural remittances mechanism also suggests that a wider diversity of destinations for emigration would dilute the effect of migration on convergence.[20] Conversely, countries that have concentrated emigration flows to a limited set of destination countries (such as Mexico or Albania for instance, who mainly emigrate to the United States and Germany respectively) should experience a stronger convergence effect than countries with a very diverse set of destination countries. We calculate an emigration concentration index[21] and – as above – run an above and below median sub-sample analysis. In Table C15 in the appendix, we show that migration is positive and significant for countries that have concentrated destination countries. We take all of the above plausibility checks as further evidence for the cultural remittances channel.

7 Conclusion

Migrants are agents of cultural change. They affect the cultural dynamics of the societies they join as immigrants and of the societies they leave as emigrants. Immigrants change the culture of host societies through both mixing and dissemination (or "infusion," in the words of Jefferson[22]) of the values and norms they absorbed during their childhood. When those values and norms are seen as inferior, or as a source of erosion of a local culture valued for itself, then immigration creates cultural tensions. Such tensions have resurfaced very prominently in many Western democracies facing increased immigration and have been associated with the recent rise of populism in some of them (Norris & Inglehart, 2019; Guriev & Papaioannou, 2020; Rodrik, 2020).

[20]The cultural mixing (COM) and dissemination (DSM) effects suggest that a diversity of immigration origins would dilute the effect of migration on convergence. We do not find support for this in the data, as we report in Table C16 in the appendix.

[21]We compute the share of emigration to each destination over total emigration for each origin and take the sum of squares of these shares (this corresponds to the Herfindahl-Hirschman Index [HHI] of concentration), and we split the sample along the median HHI value, which lies at 0.57. We present the density plot for this index in Figure C4 in the appendix.

[22]Thomas Jefferson, *Notes on Virginia*, 1785: "Every species of government has its specific principles. Ours perhaps are more peculiar than those of any other in the universe. (...) To these nothing can be more opposed than the maxims of absolute monarchies. Yet from such we are to expect the greatest number of emigrants. They will bring with them the principles of the government they leave, imbibed in their early youth; (...) These principles, with their language, they will transmit to their children. In proportion to their numbers, they will share with us the legislation. They will infuse into it their spirit, warp and bias its directions, and render a heterogeneous, incoherent and distracted mass."

This paper starts by acknowledging the variety of factors that determine migration-based cultural change: mixing and dissemination, but also other factors that have received less attention. In particular, the fact that cultural attraction to the destination plays a significant role in the decision to migrate (i.e., cultural selection and sorting), or the fact that migrants, more often than not, tend to export the host country culture back home (i.e., cultural remittances), make the global dynamics of migration-based cultural more complex. Making sense of this complexity requires a strong theoretical framework to structure ideas and formulate distinctive testable predictions allowing to discriminate among competing theories. When we turn to the empirical analysis, we first establish that migration generates cultural convergence between home and host countries. Given that this could result from many channels, we then examine various candidate mechanisms under the guidance of our theoretical model, first disqualifying purely compositional effects and then pointing to cultural remittances as the dominant dynamic force driving global cultural convergence. In other words and in contrast to the populist narrative, we find that while immigrants do act as vectors of cultural diffusion, this is mostly to export the host country culture back home.

Our work can be extended in multiple directions. First, our analysis focuses on the bilateral dimension of cultural convergence. However, the finding of an average bilateral convergence is compatible with different global patterns of cultural change. For instance, we can imagine that countries jointly converge to a global hybrid culture, to a benchmark culture (e.g., 'Americanization') or to two or more cultural poles (by which we would observe global cultural polarization *à la* Huntington). These global patterns may be co-determined by international migration flows, inviting more theoretical and empirical explorations of this question. Second, we identify the average effect of immigration and emigration on cultural proximity between countries, but we are silent on their effects on within-country cultural heterogeneity (Desmet & Wacziarg, 2018; Bertrand & Kamenica, 2018). Arguably, immigration can bring about more (e.g., through hybridization) or less (e.g., through polarization arising from a native cultural backlash against immigration) within-country cultural homogeneity. Together with the effects of migration on global cultural change, this is clearly a fundamental question that deserves to be explored further.

References

Abramitzky, R., Boustan, L. P., & Eriksson, K. (2014). A nation of immigrants: Assimilation and economic outcomes in the age of mass migration. *Journal of Political Economy*, 122(3), 467–506.

Abramitzky, R., Boustan, L. P., & Eriksson, K. (2020). Do immigrants assimilate more slowly today than in the past? *American Economic Review: Insights*, 2(1), 125–41.

Aghion, P., Algan, Y., Cahuc, P., & Shleifer, A. (2010). Regulation and distrust. *The Quarterly Journal of Economics*, 125(3), 1015–49.

Alesina, A. & Fuchs-Schündeln, N. (2007). Good-bye Lenin (or not?): The effect of communism on people's preferences. *American Economic Review*, 97(4), 1507–28.

Alesina, A. & Giuliano, P. (2015). Culture and institutions. *Journal of Economic Literature*, 53(4), 898–944.

Alesina, A., Harnoss, J., & Rapoport, H. (2016). Birthplace diversity and economic prosperity. *Journal of Economic Growth*, 21(2), 101–38.

Alesina, A., Tabellini, G., & Trebbi, F. (2017). Is Europe an optimal political area? *Brookings Papers on Economic Activity*, 48(1), 169–234.

Andersen, T. B. & Dalgaard, C.-J. (2011). Flows of people, flows of ideas, and the inequality of nations. *Journal of Economic Growth*, 16(1), 1–32.

Anthias, F. (2001). New hybridities, old concepts: the limits of'culture'. *Ethnic and racial studies*, 24(4), 619–41.

Ashraf, Q. & Galor, O. (2013). Genetic diversity and the origins of cultural fragmentation. *American Economic Review*, 103(3), 528–33.

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

Barsbai, T. & Rapoport, H. (2020). Exit and voice – Germany, 1848-1933. Mimeo at the Paris School of Economics.

Barsbai, T., Rapoport, H., Steinmayr, A., & Trebesch, C. (2017). The effect of labor migration on the diffusion of democracy: evidence from a former Soviet Republic. *American Economic Journal: Applied Economics*, 9(3), 36–69.

Batista, C. & Vicente, P. C. (2011). Do migrants improve governance at home? evidence from a voting experiment. *The World Bank Economic Review*, 25(1), 77–104.

Bazzi, S., Fiszbein, M., & Gebresilasse, M. (2020). Frontier culture: The roots and persistence of "rugged individualism" in the united states. *Econometrica*, 88(6), 2329–2368.

Bazzi, S., Gaduh, A., Rothenberg, A. D., & Wong, M. (2019). Unity in diversity? How intergroup contact can foster nation building. *American Economic Review*, 109(11), 3978–4025.

Beine, M., Docquier, F., & Schiff, M. (2013). International migration, transfer of norms and home country fertility. *Canadian Journal of Economics/Revue canadienne d'économique*, 46(4), 1406–30.

Beine, M. & Parsons, C. R. (2017). Climatic factors as determinants of international migration: Redux. *CESifo Economic Studies*, 63(4), 386–402.

Bell, A. V., Richerson, P. J., & McElreath, R. (2009). Culture rather than genes provides greater scope for the evolution of large-scale human prosociality. *PNAS*, 106(42), 17671–4.
Bellini, E., Ottaviano, G. I., Pinelli, D., & Prarolo, G. (2013). Cultural diversity and economic performance: evidence from European regions. In R. Crescenzi & M. Percoco (Eds.), *Geography, institutions and regional economic performance* (pp. 121–41). Springer.
Belot, M. & Ederveen, S. (2012). Cultural barriers in migration between OECD countries. *Journal of Population Economics*, 25(3), 1077–105.
Bertoli, S. & Marchetta, F. (2015). Bringing it all back home–return migration and fertility choices. *World Development*, 65, 27–40.
Bertrand, M. & Kamenica, E. (2018). *Coming apart? Cultural distances in the United States over time*. NBER Working Paper 24771, National Bureau of Economic Research.
Biavaschi, C., Giulietti, C., & Siddique, Z. (2017). The economic payoff of name americanization. *Journal of Labor Economics*, 35(4), 1089–116.
Bisin, A. & Verdier, T. (2000). "Beyond the melting pot": Cultural transmission, marriage, and the evolution of ethnic and religious traits. *The Quarterly Journal of Economics*, 115(3), 955–88.
Blau, F. D., Kahn, L. M., & Papps, K. L. (2011). Gender, source country characteristics, and labor market assimilation among immigrants. *The Review of Economics and Statistics*, 93(1), 43–58.
Borjas, G. J. (1987). Self-selection and the earnings of immigrants. *The American Economic Review*, 77(4), 531–53.
Braudel, F. (1987). Grammaire des civilisations. *Paris, Arthaud-Flammarion*.
Brücker, H., Capuano, S., & Marfouk, A. (2013). *Education, gender and international migration: Insights from a panel-dataset 1980-2010*. mimeo, Institut für Arbeitsmarkt und Berufsforschung.
Burchardi, K. B., Chaney, T., & Hassan, T. A. (2019). Migrants, ancestors, and foreign investments. *The Review of Economic Studies*, 86(4), 1448–86.
Casari, M., De Paola, M., Ichino, A., Marandola, G., Michaeli, M., & Scoppa, V. (2018). *Civicness drain*. Discussion Paper 11955, IZA Institute for Labor Economics.
Cavalli-Sforza, L. L., Cavalli-Sforza, L., Menozzi, P., & Piazza, A. (1994). *The history and geography of human genes*. Princeton university press.
Cha, S.-H. (2007). Comprehensive survey on distance/similarity measures between probability density functions. *International Journal of Mathematical Models and Methods in Applied Sciences*, 1(4), 300–7.
Chauvet, L. & Mercier, M. (2014). Do return migrants transfer political norms to their origin country? evidence from Mali. *Journal of Comparative Economics*, 42(3), 630–51.
Chiquiar, D. & Hanson, G. H. (2005). International migration, self-selection, and the distribution of wages: Evidence from Mexico and the United States. *Journal of political Economy*, 113(2), 239–81.
Collier, P. (2013). *Exodus: How migration is changing our world*. Oxford University Press.

Daudin, G., Franck, R., & Rapoport, H. (2019). Can internal migration foster the convergence in regional fertility rates? evidence from 19th century France. *The Economic Journal*, 129(620), 1618–92.
De Santis, G., Maltagliati, M., & Salvini, S. (2016). A measure of the cultural distance between countries. *Social Indicators Research*, 126(3), 1065–87.
Desmet, K., Le Breton, M., nu Ortín, I. O., & Weber, S. (2011). The stability and breakup of nations: a quantitative analysis. *Journal of Economic Growth*, 16(3), 183–216.
Desmet, K., Ortuño-Ortín, I., & Wacziarg, R. (2017). Culture, ethnicity, and diversity. *American Economic Review*, 107(9), 2479–513.
Desmet, K. & Wacziarg, R. (2018). *The cultural divide*. NBER Working Paper 24630, National Bureau of Economic Research.
Dippel, C. & Heblich, S. (2020). Leadership and social movements: The *Forty-Eighters* in the Civil War. *American Economic Review*. Forthcoming.
Docquier, F., Lodogiani, E., Rapoport, H., & Schiff, M. (2016). Emigration and democracy. *Journal of Development Economics*, 120, 209–223.
Docquier, F., Tansel, A., & Turati, R. (2020a). Do emigrants self-select along cultural traits?: Evidence from the MENA countries. *International Migration Review*, 54(2), 388–422.
Docquier, F., Turati, R., Valette, J., & Vasilakis, C. (2020b). Birthplace diversity and economic growth: Evidence from the US states in the post-World War II period. *Journal of Economic Geography*, 20(2), 321–54.
Egger, P. (2000). A note on the proper econometric specification of the gravity equation. *Economics Letters*, 66(1), 25–31.
Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., & Sunde, U. (2018). Global evidence on economic preferences. *The Quarterly Journal of Economics*, 133(4), 1645–92.
Fargues, P. (2007). The demographic benefit of international migration: Hypothesis and application to Middle Eastern and North African contexts. In Ç. Özden & M. Schiff (Eds.), *International Migration, Economic Development & Policy* chapter 5, (pp. 161–82). The World Bank and Palgrave Macmillan.
Fearon, J. D. (2003). Ethnic and cultural diversity by country. *Journal of economic growth*, 8(2), 195–222.
Felbermayr, G. J., Hiller, S., & Sala, D. (2010). Does immigration boost per capita income? *Economics Letters*, 107(2), 177–9.
Fernandez, R. & Fogli, A. (2009). Culture: An empirical investigation of beliefs, work, and fertility. *American Economic Journal: Macroeconomics*, 1(1), 146–77.
Feyrer, J. (2019). Trade and income—exploiting time series in geography. *American Economic Journal: Applied Economics*, 11(4), 1–35.
Fisman, R. & Miguel, E. (2007). Corruption, norms, and legal enforcement: Evidence from diplomatic parking tickets. *Journal of Political Economy*, 115(6), 1020–48.
Fouka, V., Mazumder, S., & Tabellini, M. (2020). *From Immigrants to Americans: Race and Assimilation during the Great Migration*. CEPR Discussion Paper 14396, Center for Economic Policy Research.
Galor, O. & Savitskiy, V. (2018). *Climatic Roots of Loss Aversion*. Working Papers 2018-1, Brown University, Department of Economics.

Giavazzi, F., Petkov, I., & Schiantarelli, F. (2019). Culture: Persistence and evolution. *Journal of Economic Growth*, 24(2), 117–54.

Giuliano, P. (2007). Living arrangements in Western Europe: Does cultural origin matter? *Journal of the European Economic Association*, 5(5), 927–52.

Giuliano, P. & Nunn, N. (2020). Understanding cultural persistence and change. *Review of Economic Studies*. forthcoming.

Giuliano, P. & Tabellini, M. (2020). *The Seeds of Ideology: Historical Immigration and Political Preferences in the United States*. NBER Working Paper 27238, National Bureau of Economic Research.

Goldsmith-Pinkham, P., Sorkin, I., & Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8), 2586–624.

Guiso, L., Sapienza, P., & Zingales, L. (2006). Does culture affect economic outcomes? *Journal of Economic Perspectives*, 20(2), 23–48.

Guriev, S. & Papaioannou, E. (2020). The political economy of populism. *Journal of Economic Literature*. Forthcoming.

Hirschman, A. O. (1970). *Exit, voice, and loyalty: Responses to decline in firms, organizations, and states*, volume 25. Harvard university press.

Hirschman, A. O. (1978). Exit, voice, and the state. *World Politics*, 31(1), 90–107.

Huntington, S. P. (1993). The clash of civilizations? *Foreign Affairs*, 72(3), 22–49.

Inglehart, R. & Norris, P. (2003). The true clash of civilizations. *Foreign policy*, 135, 62–70.

Jaeger, D. A., Dohmen, T., Falk, A., Huffman, D., Sunde, U., & Bonin, H. (2010). Direct evidence on risk attitudes and migration. *The Review of Economics and Statistics*, 92(3), 684–9.

Jaeger, D. A., Ruist, J., & Stuhler, J. (2018). *Shift-share instruments and the impact of immigration*. NBER Working Paper 24285, National Bureau of Economic Research.

Jefferson, T. (1998 [1785]). *Notes on the State of Virginia*. Penguin Classics.

Kandel, W. & Massey, D. S. (2002). The culture of Mexican migration: A theoretical and empirical analysis. *Social Forces*, 80(3), 981–1004.

Knudsen, A. S. B. (2019). *Those Who Stayed: Individualism, Self-Selection and Cultural Change during the Age of Mass Migration*. Discussion Paper 19-01, University of Copenhagen, Department of Economics.

Landes, D. (1998). *The Wealth and Poverty of Nations: Why Some Are So Rich and Some Are So Poor*. New York: Norton.

Levitt, P. (1998). Social remittances: Migration driven local-level forms of cultural diffusion. *International migration review*, 32(4), 926–48.

Levitt, P. & Lamba-Nieves, D. (2011). Social remittances revisited. *Journal of Ethnic and Migration Studies*, 37(1), 1–22.

Livi-Bacci, M. (2012). *A short history of migration*. John Wiley & Sons.

Luttmer, E. F. & Singhal, M. (2011). Culture, context, and the taste for redistribution. *American Economic Journal: Economic Policy*, 3(1), 157–79.

Mayda, A. M. (2010). International migration: A panel data analysis of the determinants of bilateral flows. *Journal of Population Economics*, 23(4), 1249–74.

Maystre, N., Olivier, J., Thoenig, M., & Verdier, T. (2014). Product-based cultural change: Is the village global? *Journal of International Economics*, 92(2), 212–30.
McKenzie, D. & Rapoport, H. (2010). Self-selection patterns in Mexico-US migration: the role of migration networks. *The Review of Economics and Statistics*, 92(4), 811–21.
McLuhan, M. & Fiore, Q. (1997[1968]). *War and peace in the global village*. Wired Books, Incorporated.
Miho, A., Jarotschkin, A., & Zhuravskaya, E. (2019). *Diffusion of Gender Norms: Evidence from Stalin's Ethnic Deportations*. CEPR Discussion Paper 13865, Center for Economic Policy Research.
Norris, P. & Inglehart, R. (2009). *Cosmopolitan communications: Cultural diversity in a globalized world*. Cambridge University Press.
Norris, P. & Inglehart, R. F. (2012). Muslim integration into Western cultures: Between origins and destinations. *Political Studies*, 60(2), 228–51.
Norris, P. & Inglehart, R. F. (2019). *Cultural Backlash: Trump, Brexit, and Authoritarian Populism*. Cambridge University Press.
Olivier, J., Thoenig, M., & Verdier, T. (2008). Globalization and the dynamics of cultural identity. *Journal of International Economics*, 76(2), 356–70.
Ortega, F. & Peri, G. (2014). Openness and income: The roles of trade and migration. *Journal of international Economics*, 92(2), 231–51.
Özden, Ç., Parsons, C. R., Schiff, M., & 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.
Papastergiadis, N. (2018). *The turbulence of migration: Globalization, deterritorialization and hybridity*. John Wiley & Sons.
Pieterse, J. N. et al. (2015). *Globalization and culture: Global mélange*. Rowman & Littlefield.
Ritzer, G. (2012). Grobalization. In G. Ritzer (Ed.), *The Wiley-Blackwell Encyclopedia of Globalization*. Wiley Online Library.
Rodrik, D. (2020). *Why does Globalization Fuel Populism? Economics, Culture, and the Rise of Right-wing Populism*. NBER Working Paper 27526, National Bureau of Economic Research.
Ruyssen, I. & Salomone, S. (2018). Female migration: A way out of discrimination? *Journal of Development Economics*, 130, 224–41.
Saavedra, M. H. (2018). Kenji or Kenneth? Pearl Harbor and Japanese-American assimilation. Available at SSRN: https://ssrn.com/abstract=3218512.
Schmitz, S. & Weinhardt, F. (2019). *Immigration and the Evolution of Local Cultural Norms*. Discussion Paper 12509, IZA Institute for Labor Economics.
Spilimbergo, A. (2009). Democracy and foreign education. *American Economic Review*, 99(1), 528–43.
Spolaore, E. & Wacziarg, R. (2009a). The diffusion of development. *The Quarterly Journal of Economics*, 124(2), 469–529.
Spolaore, E. & Wacziarg, R. (2009b). The diffusion of development. *The Quarterly journal of economics*, 124(2), 469–529.
Spolaore, E. & Wacziarg, R. (2013). How deep are the roots of economic development? *Journal of Economic Literature*, 51(2), 325–69.

Spolaore, E. & Wacziarg, R. (2016). Ancestry, language and culture. In V. Ginsburgh & S. Weber (Eds.), *The Palgrave Handbook of Economics and Language* (pp. 174–211). Palgrave-Macmillan.

Tabellini, M. (2020). Gifts of the immigrants, woes of the natives: Lessons from the age of mass migration. *The Review of Economic Studies*, 87(1), 454–486.

Tian, Y., Caballero, M. E., & Kovak, B. K. (2020). *Social learning along international migrant networks*. Research Paper 2020/12, University of Nottingham, GEP.

Tuccio, M., Wahba, J., & Hamdouch, B. (2019). International migration as a driver of political and social change: evidence from Morocco. *Journal of Population Economics*, 32(4), 1171–203.

Online Appendix

A Cultural diffusion across group boundaries

One-way migration. In the logic of the model, the distribution of economic gains associated with migration and differences in the cultural compositions of the two countries drive bilateral cultural selection into migration between countries A and B. The empirical section of this paper focuses on the comparative static predictions of the model, which we get at with shifters of the relative economic and cultural gains of migration. In our analysis, what determines an increase in the cultural selection into migration from country A to country B also determines an increase in the cultural selection from country B to country A. If we predict for instance that more cultural selection into migration from A to B means that countries A and B should diverge (cf. the COM mechanism), then we would also expect more cultural selection into migration from B to A; in addition, such selection should also translate (still in the COM model) into more cultural divergence. Overall, extending the model to two-way migration would not qualitatively change our predictions.

Quasi-linear preferences. We use quasi-linear preferences for their analytical convenience and clarity of interpretation: in particular, such preferences make an economic gain and cultural preferences commensurable. This assumption is not crucial to the model's interpretation: in particular, the absence of a wealth effect is irrelevant here.

Economic gains and cultural types. We simplify the discussion by assuming that expected economic gain and cultural types are not correlated. If they were correlated, we would have to envisage a situation where the majority-type has a greater overall incentive to migrate, despite a cultural reluctance to do so. Even then, our comparative statics would remain unaffected. The distinctive Predictions 2 and 3 of the various mechanisms would remain the same. Predictions 1 would be reversed. The paradox is interesting, but it is not supported by the data.

Further predictions of the compositional model. We have based COM on the description of migrants drawn in succession before the process reaches its equilibrium. Each migrant contributes to changing the cultural composition at home and at destination. As the minority types select into migration, other minority types at home might join the pool of potential migrants, while majority types may drop out of it. To the extent that we can derive a dynamic argument from this mechanism, cultural selection into migration now would increase cultural selection into migration in the future (until the pool of potential migrants is possibly depleted): the cultural selection effect becomes more powerful as time grows. The simultaneous development of network effects, however, may also affect the gain of migration, and limits the pertinence of pushing too far the logic of the mechanism in that direction. There is another parameter of the model that we have not discussed. Eq. (2) suggests that social mixing is more powerful for a smaller destination country (small n). This looks intuitive: it reflects the relative importance of the immigrants in the cultural mix at destination. Unfortunately, this does not transform into a clear testable prediction on the effect on cultural distance. When cultural homophily dominates (β low enough for $\pi > q^B$), divergence will be stronger in smaller destination countries. When the economic incentive dominates ($\pi < q^B$), divergence is stifled, or convergence emphasized, for smaller destination countries. Strictly speaking, we need to distinguish between three situations: destination size dampens convergence when $\dot{q}^B < \dot{q}^A < 0$, emphasizes divergence when $\dot{q}^A < \dot{q}^B < 0$, and dampens divergence when $\dot{q}^A < 0 < \dot{q}^B$. In the data, we observe that migration is associated with cultural convergence. Per COM, this means that the most frequent situation would the first one. Both the first and the second situations imply a negative relationship between the size of the destination country and the effect of migration on cultural convergence. How to check this in the data? In a regression of cultural distance on migration (with country-pair fixed effects), a negative coefficient means that migration brings countries closer together. In that case, if we add migration×destination country size as a regressor, COM implies that its coefficient must be positive. Even if we admit this interpretation of the model, we still need to take into account a second difficulty: n may well be correlated with crucial parameters of the model, such as the economic gain of migration, the size of the migrant community in the destination country, etc.

Assimilation. It may seem like we would characterize the full cultural equilibrium in the host country, with both assimilation and dissemination, by the joint conditions $g(q^{B*}, \pi^*, \eta^B) = 1$ and $g(\pi^*, q^{B*}, \eta^m) = 1$. We do not believe that our model is well-suited for that purpose. Indeed, initial differences across countries point to differences in homophily, as measured by ΔV_i and ΔV_j, between migrants and natives. Characterizing the full cultural equilibrium would also require a careful discussion of the intergenerational transmission of homophily. This goes beyond our purpose here, although we discuss what a similar model would tell us about assimilation below.

Time frame of cultural formation. Taken at face value, the Bisin & Verdier model of cultural formation occurs in the time frame of successive generations. For a number of theoretical and empirical reasons (in particular, that generations overlap), we are agnostic as to the exact time frame in which we should expect cultural formation to occur. Nevertheless, the dynamic nature of these three mechanisms matters empirically.

Partial derivatives of g. In the main text, we have defined:

$$g(q, \pi, \eta) = \frac{q}{1-q} \frac{1-(1-\eta)\pi - \eta q}{(1-\eta)\pi + \eta q} \frac{1-(1-(1-\eta)\pi - \eta q)\Delta V_i}{1-((1-\eta)\pi + \eta q)\Delta V_j},$$

where q is interpreted as the cultural mix of the group under consideration, π the out-group that influences cultural formation in the in-group, and η the strength of within-group cultural socialization. $1-\eta$ characterizes the influence of the out-group, and we have sometimes taken the relative size of the out-group as a proxy for $1-\eta$. We introduce the notation $\partial_x g$ to denote the partial derivative of g with respect to variable x. We get

$$\partial_q g/g = \frac{1}{q(1-q)} - \frac{\eta}{\chi(1-\chi)} + \eta \frac{\Delta V_i + \Delta V_j - \Delta V_i \Delta V_j}{(1-\tau_i)(1-\tau_j)}$$

For a well-defined economic problem, $\Delta V_{i,j} < \min\{1/\chi, 1/(1-\chi)\}$, which cannot be larger than 2. For $\Delta V_{i,j} \in [0, 2]$, the third term is nonnegative. We consider the first two terms together. Their sign is the same as $(\chi(1-\chi) - \eta q(1-q)/(1-\eta)$, which we can write as a second-order polynomial of π as $-(1-\eta)\pi^2 + (1-2q\eta)\pi + q^2\eta$. To show that $\partial_q g$ is positive, it is enough to show that this polynomial is positive. Its discriminant $1 - 4\eta q(1-q)$ is nonnegative, and therefore the polynomial has two roots. Between these two roots, the polynomial is positive. The product of the roots is negative, so one is negative and the other one positive. The expression of the positive root taken at $\eta = 0$ is 1, and increases with η, meaning that the positive root is larger than 1. As a conclusion, for any $\pi \in [0, 1]$, the polynomial, and $\partial_q g$, are positive.

We proceed in the same way to sign $\partial_\eta g$:

$$\frac{\partial_\eta g}{(\pi - q)g} = \frac{1}{\chi(1-\chi)} - \frac{\Delta V_i(1-\tau_j) + \Delta V_j(1-\tau_i)}{(1-\tau_i)(1-\tau_j)}$$

This expression has the same sign as a second-order polynomial in χ. Proceeding in the same way, we can show that for any $\Delta V_{i,j} \in (0, 2]$, this polynomial is negative. The same reasoning holds for $\partial_\pi g$.

A discussion of assimilation. Immigrants may adopt norms and values from natives. An immigrant offspring picks a role model from within the migrant community with probabilty η^m, and from the native population with probability $1 - \eta^m$. Overall, the offspring chooses a type-i role model with probability $\chi(\pi, q^B, \eta^m) = (1-\eta^m)q^B + \eta^m\pi$. The cultural equilibrium π^* is characterized by the equation $g(\pi^*, q^B, \eta^m) = 1$. As in the DSM model, we can write π^* as a function of q^B and η^m. First, $\partial\pi^*/\partial q^B > 0$. The intuition closely resembles that of DSM. Native culture contributes to cultural formation among migrants. Assimilation may thus increase the representation of the minority type among migrants (for $q^B > \pi$) or decrease it, when migrants are highly selected. Second, $\partial\pi^*/\partial\eta^m$ has the same sign as $\pi^* - q^B$. If natives are assumed to be inspirational to migrants, they pull the cultural mix among migrants closer to destination: migrants are assimilated into native culture.

In the interesting and empirically relevant case where $q^B > \pi$, assimilation predicts that cultural selection is a factor of divergence between countries A and B: the cultural mix of the migrants moves closer to the cultural mix at destination. In the first order, assimilation attenuates cultural mixing at destination (i.e., less convergence / more divergence). Qualitatively, assimilation also provides an additional justification for the over-representation of the scarce cultural type in the home country among emigrants (i.e., $\pi > q^A$).

Cultural self-selection into migration. In the exposition of COM, DSM, and of assimilation above, we have assumed that the decision to migrate was made without taking into consideration the socialization costs in either countries. Since taking them into consideration would reinforce self-selection into migration, and would not affect our comparative statics, we prefer to keep the exposition simpler. The optimal socialization effort of a type-i individual when i's offspring finds a type-i role model with probability χ is $\tau_i^*(\chi)$, as defined by Eq. (4). τ_i^* is a decreasing function. Since the same-type potential role models are more abundant in the destination country and in the migrant population (in other words, $q^B > q^A$ and $\pi > q^A$, the socialization cost of the scarce-type individual would be lower after migration. Conversely, the socialization cost of the abundant-type would be higher. Effectively, this is an indirect effect of homophily. It is less costly to socialize an offspring in an environment that resembles the parent. There are in fact three motivations to migrate: an economic gain, direct homophily, which favors cultural selection, and indirect homophily, whereby you want to socialize your offspring in the right environment. Direct and indirect homophily have complementary effects on the decision to migrate: endogenizing socializing costs in the decision to migrate would reinforce the mechanism exposed here.

B Discussion on identification

As an illustration of the discussion on identification in Section 5.2, we run a 2SLS regression using a dynamic gravity framework to instrument the bilateral migrant stock. We follow Feyrer (2019) in interacting time-invariant bilateral distance measures with time dummies to instrument for the outcome of interest. Feyrer (2019) predicts bilateral trade with sea distance and air distance interacted with a year dummy and includes a set of fixed effects. In this gravity framework the time-dummies can be interpreted as time-varying weights on the two distance measures. In his case, for instance, the weight on air distance before the advent of air travel should be zero and change once technological change allows for this bilateral dimension to play a role. We apply this concept to bilateral migration, taking into account the traditional bilateral time-invariant predictors in migration-gravity frameworks: linguistic and religious proximity, geographic and genetic distance, as well as colonial ties. We interact these variables with time dummies and control for trade and income differences and include destination-time, origin-time as well as bilateral fixed effects. Given the set of controls and fixed effects, we can limit concerns about violations in the exclusion restriction and can produce a reasonably strong first stage (with a Cragg-Donald Wald F statistic of 10.5). We find a larger and highly significant positive effect of migration on cultural similarity.

Table B1: 2SLS Estimation - migration and cultural similarity

	(1) **Euclidean**	(2) **Herfindahl**	(3) **Canberra**
Migration	0.271***	0.436***	0.480***
	(0.079)	(0.082)	(0.104)
Controls / FE	All	All	All
R^2	-0.23	-0.97	-0.70
Obs.	3,141	3,141	3,141
Cragg-Donald Wald F statistic		10.5	
Anderson canon. corr. LM statistic		98.1	
Sargan statistic		18.1	

*** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses. The instrument for migration is the interaction between year t and several bilateral distance measures including: genetic distance, linguistic proximity, religious proximity, colonial ties, and geographic distance.

However, we shy away from taking these results too seriously because of the following. Firstly, it is difficult to assess the degree of cultural selection within the set of compliers. In other words: how does the LATE compare to the ATE? We do not know how overall global changes in climate, technology or migration policy (all captured in the time dummy) would affect the cultural selection of migrants given a certain linguistic proximity, for instance. While the time-invariant factors would likely produce a more culturally selected set of compliers, the effect of global changes on the relevance of certain distance measures is not clear cut. Does technological progress in automated translation make linguistic proximity less relevant or do increasing migration restrictions make language skills more important in gaining access to certain destination countries? This is a-priori not obvious and therefore it is hard to make inferences about the cultural selection of migrants in the pool of compliers. Secondly, we are instrumenting for the size of the migrant stock. We find in this context that the impact is significantly larger in the 2SLS regression. We know from the model that the size would only magnify the scope of the different mechanisms that can lead to convergence. Detecting a larger effect can mean various things: the migrant pool is more selected and therefore we have evidence for the remittance channel; or the migrant pool is less selected and therefore we have evidence on the dissemination channel. Since we cannot assess whether the migrant pool is more or less selected for compliers, the results do not yield any meaningful insights. We therefore dedicate the mechanisms section to testing our model predictions explicitly, which will allow us to discriminate between the proposed mechanisms before further validating our results with various plausibility checks.

Table B2: First Stage - Gravity instrument

	Bilateral Migration
	(1)
Genetic Distance ∗ year	0.000
	(0.000)
Linguistic Proximity ∗ year	-0.001
	(0.001)
Releigious Proximity ∗ year	0.019***
	(0.003)
Colonial Ties ∗ year	-0.038***
	(0.014)
Geographic Distance ∗ year	-0.000**
	(0.000)
Trade	-0.011
	(0.027)
GDP Gap	-0.031
	(0.049)
Controls / FE	All
R^2	0.97
Obs.	3,141
Cragg-Donald Wald F statistic	10.5
Anderson canon. corr. LM statistic	98.1
Sargan statistic	18.1

*** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses.

C Appendix Tables and Figures

Figure C1: Kernel Density of Standardized Cultural Similarity Indexes

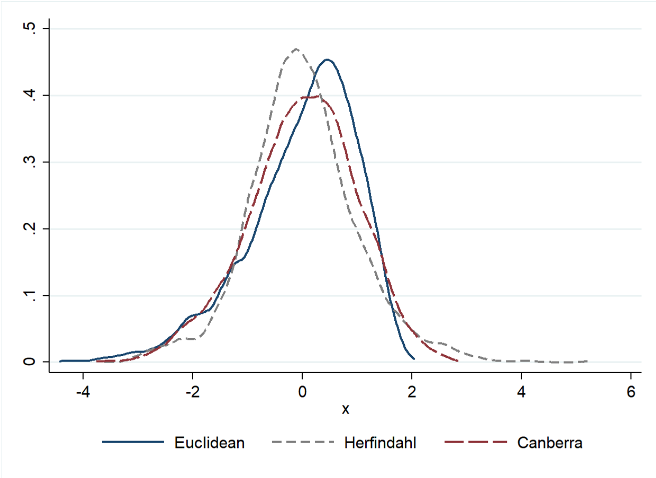
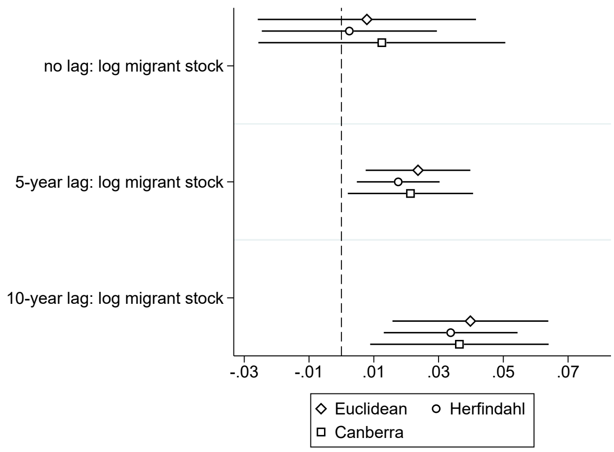
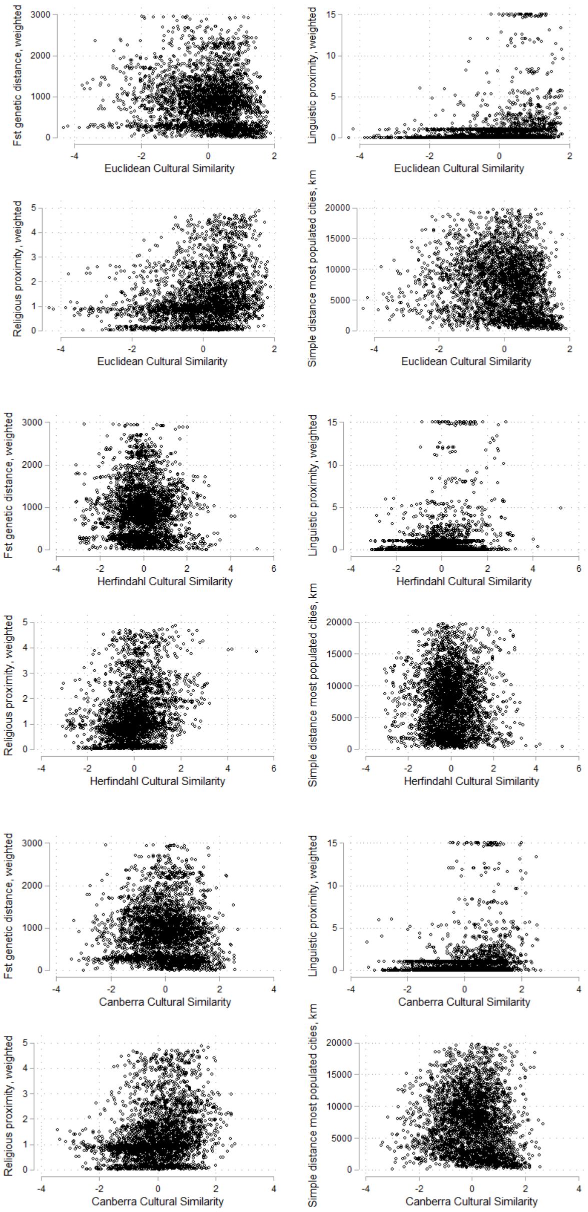


Figure C2: Baseline Regression with different lags of migration



Coefficient plot for baseline regression with full set of fixed effects, instantaneous migration with no lag, migration with five year lag (as in baseline regression) and lagged migration with ten year lag. Dependent variables are the three different cultural similarity measures.

Figure C3: Scatterplot main outcome variables and other distance measures



Coefficient plot for baseline regression with full set of fixed effects, instantaneous migration with no lag, migration with five year lag (as in baseline regression) and lagged migration with ten year lag. Dependent variables are the three different cultural similarity measures.

Figure C4: Kernel Density Plot for the Emigration Concentration Index (HHI)

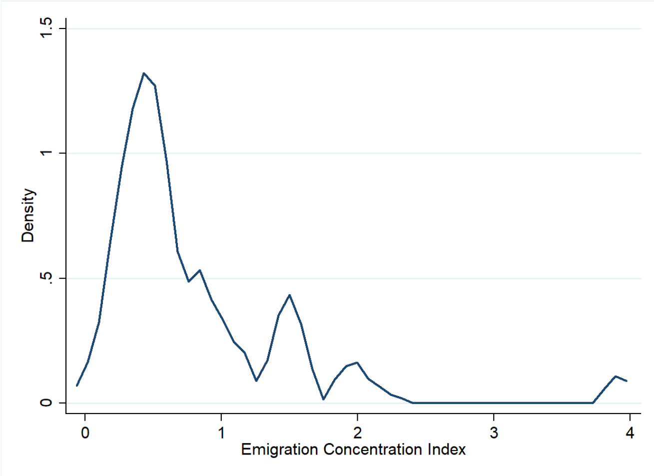


Table C1: Selected WVS Questions along Cultural Dimensions

Dimension	**WVS Question**	**Options**	**Response Scale**
Values to Children	*Here is a list of qualities that children can be encouraged to learn at home. Which, if any, do you consider to be especially important? Please choose up to five!*	Independence Hard work Feeling responsibility Imagination Tolerance Thrift Determination Religious faith Unselfishness Obedience Self-expression	binary
Priorities in Life	*For each of the following, indicate how important it is in your life.*	Family Friends Leisure Time Politics Work Religion	Very important Rather important Not very important Not at all important
Generalized Trust	*Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?*	Most ppl can be trusted Need to be very careful	binary
Gender Equality	*Do you agree with the following statement?*	When jobs are scarce, men should have more right to a job than women (i)Being a housewife is just as fulfilling as working for pay (ii) On the whole, men make better political leaders than women do (iii)A university education is more important for a boy than for a girl	Agree Neither Disagree Strongly Agree Agree Disagree Strongly Disagree
Control over Life	*How much freedom of choice and control you feel you have over the way your life turns out*	No Choice at all A great deal of Choice	Scale 1 to 10

Table C2: Selected Statistical Distance Measures

Minkowski	$D_M = \sqrt[p]{\sum_{i=1}^{d} \lvert P_i - Q_i \rvert^p}$
Euclidean	$D_E = \sqrt{\sum_{i=1}^{d} (P_i - Q_i)^2}$
Canberra	$D_{Ca} = \sum_{i=1}^{d} \frac{\lvert P_i - Q_i \rvert}{P_i + Q_i}$
Chebyshev	$D_{Ch} = \max_i \lvert P_i - Q_i \rvert$
Inner Product	$D_I = \sum_{i=1}^{d} P_i * Q_i$

Table C3: Cross-correlation of various distance measures

		(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1)	Euclidean Cultural Similarity	1.000						
(2)	Herfindahl Cultural Similarity	0.410	1.000					
		(0.000)						
(3)	Canberra Cultural Similarity	0.683	0.570	1.000				
		(0.000)	(0.000)					
(4)	Simple distance between capitals, km	-0.145	-0.013	-0.089	1.000			
		(0.000)	(0.316)	(0.000)				
(5)	Religious proximity, weighted	0.254	0.261	0.269	-0.143	1.000		
		(0.000)	(0.000)	(0.000)	(0.000)			
(6)	Linguistic proximity, weighted	0.213	0.173	0.223	-0.162	0.492	1.000	
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)		
(7)	Fst genetic distance, weighted	-0.081	-0.021	-0.045	0.332	-0.201	-0.231	1.000
		(0.000)	(0.101)	(0.000)	(0.000)	(0.000)	(0.000)	

Table C3 shows the pairwise correlations between various cultural distance measures with significance levels in brackets. Euclidean, Herfindahl and Canberra Cultural Similarity Indexes are averaged over all country pairs and years. We take the linguistic proximity and religious proximity (weighted relevant sub-populations of countries) from Fearon (2003). We also take the fixation index that measures differences in the genetic structure of populations from Cavalli-Sforza et al. (1994). Geographic distance comes from Spolaore & Wacziarg (2009b).

Table C4: List of countries - Part 1

country	1980	1990	2000	2010
Albania	0	0	1	0
Algeria	0	0	1	1
Andorra	0	0	0	0
Argentina	1	1	1	1
Armenia	0	0	0	1
Australia	1	0	0	1
Azerbaijan	0	0	0	1
Bahrain	0	0	0	1
Bangladesh	0	0	1	0
Belarus	0	0	0	1
Brazil	0	1	0	1
Bulgaria	0	0	0	0
Burkina Faso	0	0	0	0
Canada	0	0	1	0
Chile	0	1	1	1
China	0	1	1	1
Colombia	0	0	0	1
Cyprus	0	0	0	1
Czech Republic	0	0	0	0
Dominican Republic	0	0	0	0
Ecuador	0	0	0	1
Egypt, Arab Rep.	0	0	1	1
El Salvador	0	0	0	0
Estonia	0	0	0	1
Ethiopia	0	0	0	0
Finland	1	0	0	0
France	0	0	0	0
Georgia	0	0	0	1
Germany	0	0	0	1
Ghana	0	0	0	1
Guatemala	0	0	0	0
Hungary	0	0	0	0
India	0	1	1	1
Indonesia	0	0	1	0
Iran, Islamic Rep.	0	0	1	0
Iraq	0	0	0	1
Israel	0	0	1	0
Italy	0	0	0	0

Table C4 lists the countries used in our analysis (specifically, the sample that is used in our baseline specification). Year refers to the year of the migration data. Zeros and ones indicates whether all variables (WVS, migration, trade, GDP) are available for that year.

Table C5: List of countries - Part 2

country	1980	1990	2000	2010
Japan	1	1	1	1
Jordan	0	0	1	1
Kazakhstan	0	0	0	1
Korea, Rep.	1	0	0	0
Kuwait	0	0	0	1
Kyrgyz Republic	0	0	1	1
Lebanon	0	0	0	1
Libya	0	0	0	1
Malaysia	0	0	0	1
Mali	0	0	0	0
Mexico	1	1	1	1
Morocco	0	0	1	1
Netherlands	0	0	0	1
New Zealand	0	0	0	1
Nigeria	0	1	1	1
Norway	0	0	0	0
Pakistan	0	0	1	1
Peru	0	0	1	1
Philippines	0	0	1	1
Poland	0	0	0	1
Qatar	0	0	0	1
Romania	0	0	0	1
Russian Federation	0	0	0	1
Rwanda	0	0	0	1
Saudi Arabia	0	0	1	0
Singapore	0	0	1	1
Slovak Republic	0	0	0	0
Slovenia	0	0	0	1
South Africa	1	0	1	1
Spain	0	1	1	1
Sweden	0	0	0	1
Switzerland	0	1	0	0
Thailand	0	0	0	1
Trinidad and Tobago	0	0	0	1
Tunisia	0	0	0	1
Turkey	0	1	1	1
Uganda	0	0	1	0
Ukraine	0	0	0	1
United Kingdom	0	0	0	0
United States	0	0	1	1
Uruguay	0	0	0	1
Uzbekistan	0	0	0	1
Venezuela, RB	0	0	1	0
Vietnam	0	0	1	0
Yemen, Rep.	0	0	0	1
Zambia	0	0	0	0
Zimbabwe	0	0	1	1

Table C5 lists the countries used in our analysis (specifically, the sample that is used in our baseline specification). Year refers to the year of the migration data. Zeros and ones indicates whether all variables (WVS, migration, trade, GDP) are available for that year.

Table C6: Baseline - migration and cultural similarity including controls

	(1)	(2)	(3)	(4)
Euclidean				
Migration	0.066***	0.059***	0.064***	0.029**
	(0.005)	(0.005)	(0.013)	(0.012)
Trade	0.014***	0.032***	0.022*	-0.003
	(0.005)	(0.007)	(0.013)	(0.013)
GDP Gap	-0.235***	-0.270***	0.037	-0.012
	(0.009)	(0.009)	(0.027)	(0.023)
R^2	0.13	0.57	0.92	0.97
Herfindahl				
Migration	0.035***	0.049***	-0.028*	0.023**
	(0.006)	(0.004)	(0.017)	(0.010)
Trade	0.001	0.027***	-0.041**	-0.011
	(0.005)	(0.005)	(0.016)	(0.011)
GDP Gap	-0.109***	-0.220***	0.021	0.003
	(0.009)	(0.007)	(0.033)	(0.019)
R^2	0.03	0.73	0.87	0.98
Canberra				
Migration	0.060***	0.065***	0.042***	0.006
	(0.005)	(0.005)	(0.016)	(0.013)
Trade	0.009*	0.030***	-0.009	-0.015
	(0.005)	(0.006)	(0.016)	(0.015)
GDP Gap	-0.172***	-0.221***	0.027	0.020
	(0.009)	(0.008)	(0.032)	(0.026)
R^2	0.08	0.63	0.87	0.96
Obs.	5,875	5,875	5,875	5,875
Dest.-year FE		X		X
Orig.-year FE		X		X
Bilateral FE			X	X

*** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses. Table 2 shows the main results of this analysis. All three cultural similarity measures are reported, successively introducing the FE. Column 1 of each measure shows results with no FE, Column 2 introduces origin and destination FE, Column 3 shows specification with only bilateral FE, and the Column 4 shows results with country-pair, destination-time and origin-time FE.

Table C7: Balanced panel - migration and cultural similarity

	(1) **Euclidean**	(2) **Herfindahl**	(3) **Canberra**
Migration	0.058***	0.040***	0.038**
	(0.015)	(0.012)	(0.019)
Controls / FE	All	All	All
R^2	0.93	0.96	0.91
Obs.	1,359	1,359	1,359

*** p<0.01, ** p<0.05. Standard errors in parentheses. Controls: Trade and GDP Gap. FE: Destination-year, origin-year, bilateral. Selected country pairs remain in the data set for the 1995, 2005, and 2010 waves of the WVS. In order to rule out concerns about results being potentially driven by changes in the WVS sample over time, we construct a balanced panel. The sample will help reduce the noise due to unit heterogeneity. For instance, there may be endogenous reasons for which countries have not participated in different rounds of the WVS, or there are systematic differences in lags in observations that are correlated with cultural proximity. We can alleviate these concerns by picking the three WVS waves with the highest country coverage (wave 3 with 53 countries, wave 5 with 58, and wave 6 with 59). The largest 3-wave balanced panel we can construct is composed of 24 countries over the third (1995-1998), fifth (2005-2009) and sixth (2010-2014) waves of the WVS. The results of the balanced panel regressions are presented in Table C7. We present our preferred baseline specification with the full set of fixed effects and time-varying controls.

Table C8: Migration corridors - migration and cultural similarity

	North-North			*North-South*			*South-South*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Eucl.	**Herf.**	**Canb.**	**Eucl.**	**Herf.**	**Canb.**	**Eucl.**	**Herf.**	**Canb.**
Migration	-8.9e-3	-0.028	-0.29**	0.046**	0.046***	0.060***	0.030*	0.024*	0.017
	(0.074)	(0.063)	(0.13)	(0.019)	(0.017)	(0.023)	(0.017)	(0.013)	(0.018)
Controls / FE	All	All	All	All	All	All	All	All	All
R^2	0.96	0.99	0.95	0.99	0.98	0.98	0.96	0.98	0.96
Obs.	621	621	621	2670	2670	2670	2584	2584	2584

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors in parentheses. Controls: Trade and GDP Gap. FE: Destination-year, origin-year, bilateral. North is defined as a country that is member of the OECD in 2014. South is the rest of the world. We define the "North" as the group of countries which belonged to the OECD in the year 2014, and the "South" as the rest of the world. While this classification is a-theoretical, the three resulting corridors exhibit characteristics that are informative from the perspective of our model: North-North as well as South-South migration occur between culturally and economically more similar (or less dissimilar) countries than for North-South migration; in addition, these pairs of countries exhibit on average relatively balanced immigrant and emigrant stocks. In contrast, North-South migration occurs between culturally and economically more distant countries with highly imbalanced migrant stocks within pairs. The Euclidean cultural similarity between North-North and South-South in our sample lies at 0.52 and 0.17 respectively, while the Euclidean cultural similarity between North-South is substantially lower and lies at -0.35. Similarly, the average income per capita gap between North-North and South-South lies at 13,600 and 7,100 USD, respectively. The North-South income gap is at 22,700 USD on average.

Table C9: Baseline - migration and cultural similarity by theme

	(1) VtC	(2) Imp	(3) Trst	(4) Gndr	(5) Free
Euclidean					
Migration	0.022*	-0.004	0.027+	0.017	0.034*
	(0.013)	(0.008)	(0.016)	(0.014)	(0.020)
R^2	0.96	0.98	0.95	0.96	0.91
Obs.	5,869	5,639	5,790	5,847	5,874
Herfindahl					
Migration	0.008	-0.003	0.004	0.018	0.026*
	(0.008)	(0.006)	(0.009)	(0.013)	(0.014)
R^2	0.98	0.99	0.99	0.97	0.96
Obs.	5,869	5,639	5,875	5,847	5,874
Canberra					
Migration	0.004	-0.005	-0.026+	0.013	0.019
	(0.013)	(0.011)	(0.016)	(0.015)	(0.016)
R^2	0.95	0.97	0.95	0.96	0.94
Obs.	5,869	5,639	5,790	5,847	5,869
Controls / FE	All	All	All	All	All

* p<0.1, + p< 0.15. Standard errors in parentheses. Preferred specification split by theme. VtC: values transmitted to children. Imp: priorities in life. Trst: generalized trust. Gndr: gender equality. Free: control over life. The full questions are summarized in Table C1.

Table C10: Robustness check - Accounting for methodology change in migration data

	(1) **Euclidean**	(2) **Herfindahl**	(3) **Canberra**
Migration	0.033***	0.024**	0.003
	(0.012)	(0.010)	(0.014)
Controls / FE	All	All	All
R^2	0.97	0.98	0.96
Obs.	5,588	5,588	5,588

*** p<0.01, ** p<0.05. Standard errors in parentheses. Baseline specification for the sub-sample of countries for which we observe non-zero migration in 2010. We lose about 300 observations. This number of observations lost is relatively small since the data set is comprised of countries for which we have information on beliefs and preferences from the World Value Survey and for which bilateral trade and GDP data is available. This already restrict our sample and therefore limits the number of country pairs for which the change in methodology from 2000 to 2010 was relevant.

Table C11: Robustness check - Lagged dependent variable

	(1) **Euclidean**	(2) **Herfindahl**	(3) **Canberra**
Migration	0.046**	0.041**	0.049*
	(0.022)	(0.018)	(0.026)
Lagged cultural similarity	-0.044	-0.187***	-0.092**
	(0.043)	(0.043)	(0.045)
Controls / FE	All	All	All
R^2	0.94	0.97	0.93
Obs.	1,036	1,036	1,036

*** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses. Baseline specification in the panel sub-sample. We include the respective lagged dependent variable (Euclidean, Herfindahl and Canberra Cultural Similarity Index) measured at the same time as our main variable of interest, migration, and control variables.

Table C12: Plausibility checks - intensity of interaction proxied by remittances

	Above median remittances			*Below median remittances*		
	(1)	(2)	(3)	(4)	(5)	(6)
	Euclidean	**Herfindahl**	**Canberra**	**Euclidean**	**Herfindahl**	**Canberra**
Migration	0.044***	0.037***	0.027+	0.016	0.020	0.012
	(0.015)	(0.013)	(0.018)	(0.019)	(0.016)	(0.021)
Controls / FE	All	All	All	All	All	All
R^2	0.97	0.98	0.95	0.98	0.98	0.97
Obs.	2,975	2,975	2,975	2,297	2,297	2,297

*** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses. Baseline specification, with the sample split into country pairs that have remittance flows and those that do not have recorded remittance flows (as the median remittances are zero) to proxy intensity of interaction between diaspora and home community.

Table C13: Plausibility checks - intensity of interaction proxied by travel (non-resident visitors)

	Above median travel			*Below median travel*		
	(1)	(2)	(3)	(4)	(5)	(6)
	Euclidean	**Herfindahl**	**Canberra**	**Euclidean**	**Herfindahl**	**Canberra**
Migration	0.075***	0.048***	0.004	0.005	0.013	0.011
	(0.021)	(0.018)	(0.027)	(0.019)	(0.016)	(0.020)
Controls / FE	All	All	All	All	All	All
R^2	0.97	0.98	0.95	0.97	0.98	0.97
Obs.	1,726	1,726	1,726	1,726	1,726	1,726

*** p<0.01. Standard errors in parentheses. Baseline specification, with the sample split into country pairs that have above median bilateral travel. The data comes from the UNWTO's "Compendium of Tourism Statistics," providing data on non-resident visits between origin and destination between 1995 and 2018. The median is calculated from the average number of bilateral visits between 1995 and 2010. Standard errors in parentheses.

Table C14: Plausibility checks - intensity of interaction proxied by telecommunication

	Above median			*Below median*		
	(1)	(2)	(3)	(4)	(5)	(6)
	Euclidean	**Herfindahl**	**Canberra**	**Euclidean**	**Herfindahl**	**Canberra**
Panel A	*landline subscriptions*			*landline subscriptions*		
Migration	0.048***	0.026**	-0.019	-0.013	0.001	-0.005
	(0.015)	(0.013)	(0.020)	(0.019)	(0.015)	(0.020)
R^2	0.98	0.99	0.96	0.97	0.98	0.97
Obs.	2,934	2,934	2,934	2,933	2,933	2,933
Panel B	*mobile subscriptions*			*mobile subscriptions*		
Migration	0.038***	0.021*	-0.018	0.007	0.031+	0.035
	(0.013)	(0.012)	(0.017)	(0.024)	(0.019)	(0.025)
R^2	0.98	0.98	0.96	0.97	0.98	0.97
Obs.	2,934	2,934	2,934	2,933	2,933	2,933
Panel C	*international calls*			*international calls*		
Migration	0.040+	0.056**	-0.039	0.018	0.019+	0.020
	(0.027)	(0.022)	(0.034)	(0.016)	(0.012)	(0.016)
R^2	0.97	0.98	0.96	0.96	0.97	0.96
Obs.	2,259	2,259	2,259	2,258	2,258	2,258
Controls / FE	All	All	All	All	All	All

*** p<0.01, ** p<0.05, * p<0.1, + p< 0.15. Standard errors in parentheses. Baseline specification, with the sample split along the median for three measures of bilateral telecommunication. The data comes from the UN International Telecommunication Union. These data are not available at the bilateral level so we construct three different bilateral measures. In Panel A and B we multiply the share of people in origin and destination that have a landline or a mobile phone subscription. This indicates the probability that two individuals picked at random at destination and origin could communicate with one another. In Panel C, we use data on the total international voice traffic (fixed and mobile) per person. We multiply this measure of "intensity of international communication" for destination and origin country and take the log.

Table C15: Plausibility checks - diverse and concentrated emigration destinations

	Diverse destinations			*Concentrated destinations*		
	(1)	(2)	(3)	(4)	(5)	(6)
	Euclidean	**Herfindahl**	**Canberra**	**Euclidean**	**Herfindahl**	**Canberra**
Migration	-0.004	-0.009	-0.026	0.050***	0.040***	0.022
	(0.023)	(0.019)	(0.026)	(0.014)	(0.012)	(0.016)
Controls / FE	All	All	All	All	All	All
R^2	0.98	0.99	0.97	0.96	0.97	0.95
Obs.	2,876	2,876	2,876	2,999	2,999	2,999

*** p<0.01. Standard errors in parentheses. Baseline specification, with the sample split along the median value of an emigration concentration index specifically created for this analysis. The index measures for each source country the overall concentration of destination countries. Source countries with a large share of their emigrants residing in only few destination countries will receive a high score (Mexico's score is at 3.88) and countries that have a very diverse set of destination countries receive a low score (the US's score is at 0.60, France's 0.09, the median concentration index lies at 0.53).

Table C16: Plausibility checks - diverse and concentrated origin countries

	Diverse origins			*Concentrated origins*		
	(1)	(2)	(3)	(4)	(5)	(6)
	Euclidean	**Herfindahl**	**Canberra**	**Euclidean**	**Herfindahl**	**Canberra**
Migration	0.026	0.024+	-0.008	0.024+	0.018+	0.008
	(0.019)	(0.016)	(0.023)	(0.016)	(0.013)	(0.017)
Controls / FE	All	All	All	All	All	All
R^2	0.98	0.98	0.96	0.96	0.98	0.96
Obs.	3,215	3,215	3,215	2,660	2,660	2,660

+ p< 0.15. Standard errors in parentheses. Baseline specification, with the sample split along the median value of an immigration concentration index similar to the index in Table C15. The index measures for each source country the overall concentration of origin countries.