International Migration and Cultural Convergence*

Hillel Rapoport† Sulin Sardoschau‡ Arthur Silve§

This draft version: February 25, 2018

Abstract

Does international migration contribute to cultural convergence or divergence between sending and receiving countries? We investigate this question both theoretically and empirically. We first develop a compositional model of international migration and cultural change, where divergence arises from self-selection on cultural traits and convergence arises from social mixing. The model is then adapted to allow for horizontal and vertical cultural transmission following Bisin & Verdier (2000). The model yields a rich set of predictions, which we test empirically using panel data from the World Value Survey and bilateral migration data for the period 1981-2014. We exploit within country-pair variation in cultural proximity over time and find support for the cultural transmission hypothesis. As the model with cultural transmission predicts, migration generates bilateral cultural convergence even if we exclude migrants from the pool of respondents in both countries (hence eliminating social mixing). It is also more likely in the long-run, especially after controlling for economic incentives to migrate and for initial cultural distance, which is consistent with the cultural transmission hypothesis (but not with compositional changes). Interestingly, international migration appears as a stronger and more robust driver of cultural convergence than trade. The results hold for a large set of time-varying cultural distance measures along different statistical and topical dimensions.

Keywords: convergence, migration, culture
JEL classification: F22, Z10

---

*We thank Gustavo De Santis, Frédéric Docquier, Thierry Verdier and participants at the 9th AFD-World Bank Migration and Development Conference, EUI, Florence, June 2016; the 5th OECD-CEPII Conference on Immigration in OECD countries, Paris, December 2016; the OxDev Workshop, Oxford, September 2016; the Cultural Diversity Workshop, University Paris 1 Pantheon-Sorbonne, May 2016; and the Conflict and Migration Workshop, Louvain-la-Neuve, May 2017.

†Paris School of Economics, University Paris 1 Pantheon-Sorbonne, and CEPII. Email: hillel.rapoport@psemail.eu

‡Paris School of Economics and University Paris 1 Pantheon-Sorbonne. Email: sulin.sardoschau@psemail.eu

§Université Laval, Department of Economics. Email: arthur.silve@ecn.ulaval.ca.

1 Introduction

Globalization facilitated and accelerates the intensity of interactions between countries. This holds for economic interactions such as trade in goods and services, capital flows and the movement of people, as well as for cultural exchanges. Cultural and economic interactions often reinforce one another: firms such as IBM open up offices and invest in different countries, thereby introducing a new corporate culture into other labor markets (see Hofstede, 2001, for a survey of IBM employees in 30 different countries); American firms export goods in fashion, entertainment, or technology to many countries in the world, thus spreading the "American Way of Life;" migrant communities bring their own cultural goods and traditions to their host communities (Fernandez, 2011). At the same time, cultural distance may be a strong impediment to international exchange (Guiso et al., 2006).

We use the term culture in a very broad sense, as a set of attitudes, preferences, beliefs and values that govern individual behavior and thereby determine aggregate social, political and economic outcomes(see Alesina & Giuliano, 2015, for an overview of the economics of culture). Our level of trust in others, the determinants of status in a society, our degree of patience, of individualism, our beliefs about the correct trade- offs between efficiency and equity, the "proper" roles for women (or the proper size of a family) are all examples of beliefs/preferences that have differed across societies and over time, with strong implications for comparative economic development (Braudel, 1987; Landes, 1998; Alesina & Fuchs-Schündeln, 2007; Aghion et al., 2010; Spolaore & Wacziarg, 2009, 2013). Economists tend to assume that systematic differences in preferences across social groups are exogenous.[1]This paper relaxes this assumption by shedding light on the determinants of cultural formation itself. It follows recent research devoted to the two-way relationshipbetween culture and trade. In one of the earlier explorations of the effects of cultural ties on bilateral trade, Rauch & Trindade (2002) show that ethnic Chinese networks have a significant impact on the trade in differentiated goods through the creation of business and social ties. Similarly, Felbermayr & Toubal (2010) show that cultural proximity is an important determinant of bilateral trade volumes, using a time-varying measure based on bilateral scores from the Eurovision Song Contest (a very popular pan-European television show). Additionally, there is a stream of literature that shows that trade in itself can have an impact on the cultural proximity of two countries, particularly through trade in differentiated goods with a high "cultural content" such as movies, clothing and other life style articles (Disdier et al., 2010). With different assumptions, two closely related articles yield opposite outcomes. Olivier et al. (2008) show that goods market integration can lead countries to diverge culturally. They assume that there are two goods that each embody a "cultural type" and can be

[1]Guiso et al. (2006) define culture as "those customary beliefs and values that ethnic, religious, and social groups transmit fairly unchanged from generation to generation".

purchased and consumed at market price. When two countries decide to engage in trade, the distributions of cultures (as captured by the share of people consuming each cultural good) become more dissimilar across countries due to a change in the relative price of the cultural goods. The model even predicts that some cultures that existed in autarky may disappear when a country opens to international trade. Conversely, Maystre et al. (2014) examine what happens when the production side is better described by monopolistic competition. An idiosyncratic good is available only in the home country and a global good can be traded and consumed in both countries. In this setup, they show that opening up to trade and the resulting market size effect will lead to both countries consuming more of the global good and less of the idiosyncratic good. They conclude that trade liberalization can also lead to cultural convergence. Each of these two papers provides important intuitions. Considered jointly, they also illustrate how sensitive the convergence vs. divergence result is to the form of the interaction.

This sensitivity extends beyond trade and culture. In this paper, we explore the link between migration and culture. We find the same theoretical ambivalence on the causal effect of migration on culture. Let us start with a few basic observations. First, migrants are not a random sample of their home country population; rather, they are self-selected on a number of dimensions (e.g., age, gender, education, wealth, ethnic background), including culture (Docquier et al., 2017). If migrants leave their home country to live at a destination that is culturally closer to them, they will leave behind a more dissimilar population. We call this effect the *Cultural Selection-Effect*.[2] Hence, the "exit" or "selection" effect, which has to do with the non-randomness of migration, implies a compositional change for the remaining population, e.g. divergence through **cultural sorting**; it is likely to be minor except where relatively small numbers can make a difference, as is the case, arguably, in domains such as scientific or political innovation.

Second, with the compositional change at the home population also comes a change in cultural composition in the destination country. If immigrants are not a perfect cultural match to the host population, but culturally similar to the home population, we expect origin and destination countries to converge culturally by mere **social mixing**. This approach assumes that culture is embedded in people and thus carried with them, introducing their values into the host society and consequently making home and host populations more similar. This is in line with the epidemiological approach to culture, which shows that even if individuals face the same socioeconomic and institutional setting, migrants carry a part of their culture (values and norms of their places of birth or their parents' places of birth) with them into these environments and thus differ culturally from other natives and non-immigrant descendants (Fernandez, 2011). Social mixing is part of what we call this the *Diffusion-Effect*, namely the part of cultural diffusion that takes place in the host communities. The idea of a cultural melting pot is intuitively

[2]Hirschman (1981) introduced a similar *exit-effect* to refer to self-selection along political attitudes. In particular, he finds that migrants are typically positively selected in terms of their support for democracy.

compelling and a phenomenon very visible in big Western cities. However, there is little evidence that immigrant communities are sizable enough to have a significant impact on the average cultural stance in the host population. Additionally, cultural assimilation of (even large) immigrant communities counterbalances the longer-run effects of social mixing. Depending on the balance between selection effect and social mixing, international migration may lead to cultural convergence or divergence.

Third, even abroad, the diaspora plays a role in the cultural formation in their home country. Migrants contribute to the diffusion of values as they carry with them home-country knowledge and social norms and *remit* back new values, preferences or beliefs to their home communities. For instance, in Morocco, Egypt, and Turkey, Fargues (2006) shows birth rates are affected differently when a member of a household emigrated to a high fertility (Gulf) versus low fertility (West) destination country. Exposure to new norms shapes the preferences of migrants and are then passed on to friends and family back home through so called **social remittances**. Additionally, emigrants typically remain part of the cultural narrative and are often even portrayed as success stories and role models. For instance, Kandel & Massey (2002) show that young Mexicans consider migration as a "marker in the transition to manhood." Similarly, the 2013 Country Migration Report of the Philippine government mentions that "a culture of migration has settled in, particularly in some regions, where the aspirations of youngsters are molded by the examples of migrants". European institutions have also long recognized the role of migrants in shaping aspirations, norms, and values in their home countries and has established the Entrepreneurship 2020 Action Plan for migrants, based on the idea that migrants serve as role models and facilitate outreach to specific groups. More anecdotally, the French have an expression to describe the migrant, glorified by successes overseas, and becoming a role model: l'Oncle d'Amérique. We call this effect the *Cultural Diffusion-Effect*, namely the part of cultural diffusion that takes place in the home community.

We suggest a unified framework to disentangle the two effects. In this framework, migration can either lead to cultural convergence or divergence. The decision to migrate is motivated by both cultural homophily and a universal quest to improve one's material situation. The model first examines the compositional effect of migration. If cultural self-selection is the primary motive, we predict that migration should lead to cultural divergence. Conversely, if economic gain dominates, then migration leads to social mixing and cultural convergence. Second, we consider cultural dynamics, using the framework first suggested in Bisin & Verdier (2000). They suggested that culture is transmitted either vertically (through the parents) or horizontally (through a role model). If we assume that migrants continue to be role models for the population which remained at home, then culture diffuses across borders. If that is the case, then we predict that migration should lead to cultural convergence.

We bring the various predictions of the model to the data. Our empirical analysis focuses on the dynamic aspect of culture. If culture is not deterministic but an outcome of socio-economic conditions, then naturally our norms, attitudes and preferences must adjust over time as these conditions change. Previous works on cultural proximity developed static measures to capture how culturally close or distant two social groups (De Santis et al., 2016; Spolaore & Wacziarg, 2014), but such distances change over time. Neglecting the dynamic component of culture can substantially hinder our understanding of its determinants. To address this issue, we use the World Value Survey, which covers issues of beliefs, family values, religiosity, in a cross-section of almost 100 countries over six waves. Our results are robust to using various time-varying bilateral distance measures. We also construct two balanced panels, the first of 24 countries for which we have three waves of data, and the second of 12 countries for which we have four. Our results are also robust to these alternative specifications, thereby addressing possible issues of increasing coverage and changes in the number and nature of questions asked. The choice of a coherent and comparable set of questions and measures is a significant contribution to the young literature on cultural proximity.

Lastly, we extend and deepen the analysis on the determinants of cultural proximity to migration and allow for two competing effects to unravel in the data. The empirical analysis confirms the main predictions of our model. For over six thousand country pairs, cultural diffusion dominates between 1970 and 2010: we observe cultural convergence. This results is robust to all our statistical measures and the inclusion of a large set of fixed effects (exploiting the within country-pair variation over time) and survives the horse race with other potential time-varying determinants of culture such as bilateral trade or income differences.

The results can be summarized as follows. As the model with cultural transmission predicts, migration generates bilateral cultural convergence even if we exclude migrants from the pool of respondents in both countries (hence eliminating social mixing). It is also more likely in the long-run, especially after controlling for economic incentives to migrate and for initial cultural distance, which is consistent with the cultural transmission hypothesis (but not with compositional changes). Interestingly, international migration appears as a stronger and more robust driver of cultural convergence than trade. The results hold for a large set of time-varying cultural distance measures along different statistical and topical dimensions.

The paper is organized as follows: section 2 introduces a theory of migration-driven cultural change. In section 3, we document the data sources and elaborate on the meaning and statistical measurement of bilateral cultural similarity. Section 4 shows the results from our empirical analysis, comparative statics, and some robustness checks. Section 5 concludes.

2 A theory of migration-based cultural change

We start with the compositional effect of migration. We assume that there are two components to the migration decision: a cultural and an economic component. Depending on which motive dominates, we show that there is cultural convergence or divergence. In a second step, we introduce the dynamic component of cultural transmission thanks to a Bisin & Verdier (2000) setting, where migrants continue to play a role in the cultural equilibrium of their home community. With this setting we encompass all of the three basic observations we made about the role of migrants in global cultural formation. The compositional model will give insights into **cultural selection** as well as one dimension of cultural diffusion, namely the effect of **social mixing**. The transmission model will highlight the second dimension of the diffusion effect, that is cultural convergence through **social remittances**.

2.1 A compositional model of migration and cultural change

We consider two countries A and B. We assume that only the nationals of country A (the home country) are considering migrating to B (the host country). For simplicity, we only consider one-directional migration but all results can be generalized to migration in both directions (A to B and B to A). 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_i^A and q_i^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 net 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 (CDF) $\mathcal{G}$ with support on $\mathbb{R}$. Assuming quasi-linear preferences,[3] the pool of type-i country A nationals who wish to migrate is composed of anyone such that

$$\beta g \geq (1-\beta)(f(q_i^A) - f(q_i^B)), \tag{1}$$

with f the function by which cultural preferences are translated in 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. β characterizes the relative weight of the economic motive in the migration decision, and $1-\beta$ the weight of the cultural motive. When $\beta = 0$, only homophily matters in the

[3]We use quasi-linear preferences for their analytical convenience and clarity of interpretation. This assumption is not crucial to the model's interpretation: in particular, the absence of a wealth effect is irrelevant here.

decision to migrate: potential migrants are attracted by a higher share of same-type individuals in the destination country. In that case, only one type of individuals is likely to move. Conversely, if $\beta = 1$, only the expectation of economic gains counts in the decision to migrate. Potential migrants compare the cost of migration to the prospect of economic gains, but do not consider the cultural change involved. As a result, the pool of potential migrants is a culturally representative sample of the home country population (assuming that expected economic gain and cultural type are not correlated, an assumption that we discuss later on).

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. 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. To simplify the notations, let us introduce $\mathcal{G}_i \equiv \mathcal{G}((1-\beta)(f(q_i^A) - f(q_i^B))/\beta)$ the fraction of type-i individuals not interested in moving, and similarly $\mathcal{G}_j \equiv \mathcal{G}((1-\beta)(f(1-q_i^A) - f(1-q_i^B))/\beta)$. $1 - \mathcal{G}_i$ (resp. $1 - \mathcal{G}_j$) is the fraction of type-i (resp. type-j) individuals who wish to migrate. To further simplify notations, let us also take i as the generic cultural type. For instance, we will write $q \equiv q_i$. As a result, $q_j = 1 - q$. At each successive draw of a new migrant, the probability that her type is i is

$$\pi \equiv \pi_i(q^A, q^B) = \frac{q^A(1-\mathcal{G}_i)}{q^A(1-\mathcal{G}_i) + (1-q^A)(1-\mathcal{G}_j)} \tag{2}$$

With each period, the cultural composition of the two country evolves according to dynamics which can be simply written:

$$\begin{cases} \dot{q}^A = q^A - \pi \\ n\dot{q}^B = \pi - q^B. \end{cases} \tag{3}$$

The game reaches its equilibrium when the pool of potential migrants is empty, that is, when $\mathcal{G}_i = \mathcal{G}_j = 1$. The system thus displays intuitive features. Let us first examine separately the two motives for migration: cultural homophily and the prospect of economic gain. To simplify the discussion, we may assume that type-i individuals are more frequent in country B, and scarcer in country A: $q^A < q^B$ (from now on, we maintain this assumption).

If $\beta = 0$, individuals only rely on cultural homophily when deciding to migrate. In this case, the pool of potential migrants is composed of all individuals of the relatively scarce type i, who wish to move to country B, where their type is relatively abundant, and no abundant-type individual. Then $\mathcal{G}_i = 0$, $\mathcal{G}_j = 1$, and $\pi = 1$. As scarce-type individuals leave, the incentive to migrate becomes stronger for each remaining scarce-type individual, 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. In that case, the cultural selection effect dominates, which leads mechanically to cultural divergence between the two countries.

If $\beta = 1$, individuals only consider their prospect of economic gain. In that case, $\mathcal{G}_i = \mathcal{G}_j$, 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 social mixing effect dominates, which leads mechanically to cultural convergence between the two countries.

Finally, in the more interesting case, $\beta \in]0, 1[$: individuals take both motives into consideration. $\mathcal{G}_i < \mathcal{G}_j$ and finally, $\pi > q^A$: with cultural homophily, the scarce cultural type is over-represented in the migrant population. Thanks to Eq. (3), it is straightforward that $\dot{q}^A < 0$: the scarce cultural type becomes progressively scarcer in the home country. In the host country, we must consider two cases, as π may be larger or smaller than q^B. If $\pi > q^B$, that is if the cultural homophily largely dominates the decision to migrate, then the relatively abundant type becomes progressively more abundant. As type-i become scarcer in country A and more abundant in country B, the pool of type-i migrants expands (up to a point which may or may not encompass the whole type-i population), while the pool of type-j migrants shrinks. In equilibrium, the two countries have diverged culturally. Conversely, if $\pi < q^B$, the prospects of economic gains are prominent (for high values of β). The share of type-i individuals decreases in both countries. This suggests the possibility of cultural convergence between countries A and B for high enough values of β, when the economic motive is the principal driver of migration. Formally, we define divergence as $(q^A - q^B)(\dot{q}^A - \dot{q}^B) > 0$, and convergence otherwise.

We have made the assumption that expected economic gain and cultural type were uncorrelated across individuals, a useful assumption to understand the mechanism at hand. In particular, this ensures that if $\beta = 1$, the composition of the pool of potential migrants reflects the cultural composition of the origin country. If we do not make the assumption, $\beta = 1$ still ensures that the economic motive is the only motive for migration, but the prospect of economic gains is unevenly distributed between the two types. Actually, in that case, it is sufficient to compare $\mathcal{G}_i$ to $\mathcal{G}_j$ to characterize the cultural evolution. $\mathcal{G}_i < \mathcal{G}_j$ means that there are more type-i than type-j potential migrants. This would mimic the cultural selection effect, and stifle the social mixing effect. Conversely, $\mathcal{G}_i > \mathcal{G}_j$ would strengthen the social mixing effect and attenuate the cultural selection effect. In the remainder of the paper, we maintain the assumption that wealth and cultural types are uncorrelated.

This compositional model (CM) brings together social mixing, arguably the most intuitive mechanism that links migration to cultural composition, and cultural selection, another rather intuitive argument. The model suggests a set of clear empirical predictions, and would help interpret the data. For instance, empirical support for CM with cultural convergence would be interpreted to imply the domination of the economic motive of migration.

We make a practical distinction between the host *population*, which does not include migrants and the host *country*, which comprises both the native and the migrant population. As its name indicates, the compositional model offers a mechanistic view of migration, one in which preferences are not affected through social interactions or active policy. The compositional model does not incorporate assimilation of migrants in the host population, e.g. migrants core beliefs are sticky and not impacted by the host population. And reversely, the host population does not adopt the preferences of the migrant population. In general, there are neither cultural spillovers from migrants to natives or from natives to migrants. We will relax these assumptions in section 2.2. and allow for mutual influence between the host and the migrant populations, as well as influence of the migrants on the cultural mix in their home country. But for now, let us examine the predictions of this mechanistic model.

If economic incentives are not the only driver of migration (i.e. $\beta < 1$), we predict that the scarce type becomes scarcer at home. We thus obtain a first prediction:

CM prediction 1

- Migration can both lead to cultural convergence (through social mixing) or divergence (if social mixing is not powerful enough) between home and host *countries*.
- Migration should lead to cultural divergence between home and host *populations*

More precisely, in the generic case where $q^A < q^B$, convergence corresponds to $\dot{q}^B < \dot{q}^A < 0$. Let us introduce a few notations: A and B are the two countries before migration, and A' and B' after migration. $D(x, y)$ is a measure of cultural distance between countries x and y. Recall that in this model we focus on one-way migration flows from A to B. With these notations, convergence would be accounted for by $D(A', B') < D(A, B)$. That is, we have convergence if the cultural distance after migration is smaller than the cultural distance before migration. As mentioned in the above prediction, migration in the compositional model can lead to both convergence and divergence between home and host countries. However, once we exclude migrants, we rule out social mixing and therefore must have cultural divergence if the compositional model holds.

CM 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 the 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 who wish to emigrate is now given by $\mathcal{G}_i(q^A, q^B, \Delta g) = \mathcal{G}\left(-\Delta g + \frac{1-\beta}{\beta}\left(f(q_i^A) - f(q_i^B)\right)\right)$ among type-i individuals and

$\mathcal{G}_i(q^A, q^B, \Delta g) = \mathcal{G}\left(-\Delta g + \frac{1-\beta}{\beta}\left(f(1-q_i^A) - f(1-q_i^B)\right)\right)$ among type-j ones. For distribution functions $\mathcal{G}$ without fat tails, the ratio $(1-\mathcal{G}_j)/(1-\mathcal{G}_i)$ decreases with Δg, and so does π.There are more candidates to migration in both subpopulations but the cultural selection effect becomes less important relative to the economic motive. Therefore, the effect of social mixing becomes more important

CM prediction 2

Uniformly more economic gains from migration should result in stronger convergence, or less divergence.

Now, consider what happens for countries that are initially closer or further apart culturally speaking. For country pairs that are further apart, that is, for lower (negative) values of $f(q^A) - f(q^B)$ and, correspondingly, larger (positive) values of $f(1-q^A) - f(1-q^B)$, we have a large $1-\mathcal{G}_i$, a smaller $1-\mathcal{G}_j$, and a high π, possibly close to 1. The cultural selection effect is strengthened and we would expect more divergence or less cultural convergence. In sum, initial cultural proximity favors cultural convergence, but initial cultural distance deters it, to the point where we might observe divergence.

CM prediction 3

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[4] However, for the purpose of this analysis, we stick to the most

[4]There is another parameter of the model that we have not discussed yet. Eq. (3) 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 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 CM, 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, CM implies that its coefficient

simple set-up of the model, where we focus on unilateral migration, and neglect - for now - any spill overs between locals and immigrants in the host society. Overall, CM prediction 1 captures the core implication of the model. Prediction 2 makes propositions about the comparative statics behind the economic versus cultural gains from migration. Lastly, prediction 3 hints at the initial conditions (culturally close or distant countries) that drive the dynamics of the model.

2.2 Intergenerational Cultural Transmission with Migrants as Role Models

The model of cultural transmission suggested by Bisin & Verdier (2000) illustrates how different cultural types can coexist in equilibrium. Individuals socialize their offspring with a probability of success directly linked to the costly socialization effort they provide. Bisin and Verdier suggest that a type-i individual would rather have a type-i offspring: this is another manifestation of their homophily. To that effect, they may invest in the socialization of their offspring, at a cost. 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^*.

Extending that framework to two countries entails making assumptions on who can be taken as a role model, and where. Contact seems to be a natural condition to choose a role model. Contact is most likely to derive from living in the same country, but also from a common nationality. Role models are picked among neighbors, but we have shown in the introduction that there is ample evidence that emigrants continue to play a role in cultural transmission at home as well (we discuss this assumption later).

Keeping the notations from the compositional model, let us write $q \equiv q^A$ the share of type-i individuals in home country A, and π is the share in the migrant population. In line with the intuition of the compositional model, we assume that $\pi > q$. η is the probability that a role model will be picked from the emigrant population. With probability $1 - \eta$, the role model is picked from the home population.

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 migrant or home populations. In that case, the offspring chooses a type-i role model with probability $\chi \equiv (1-\eta)q+\eta\pi$, and a type-$j$ role model with probability $1-\chi$. Notice that $q < \chi < \pi$. With socialization costs $H(\tau_i)$, the program of the type-i individual is

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

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.

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 $H(\tau) = \frac{1}{2}\tau^2$ and, to avoid corner solutions, we assume $\Delta V_i, \Delta V_j < 1$. The problem is adequately concave: a type-i individual provides effort $\tau_i = (1 - \chi)\Delta V_i$, and a type-$j$ $\tau_j = \chi \Delta V_j$. The cultural equilibrium in country A is reached when $q(1 - \tau_i)(1 - \chi) = (1 - q)(1 - \tau_j)\chi$. We get the following characterization of the cultural equilibrium q^*:

$$\frac{q^*}{1 - q^*} = \frac{(1 - \eta)q^* + \eta\pi}{1 - (1 - \eta)q^* - \eta\pi} \frac{1 - ((1 - \eta)q^* + \eta\pi)\Delta V_j}{1 - (1 - (1 - \eta)q^* - \eta\pi)\Delta V_i} \tag{5}$$

This expression defines implicitly q^* as a function of η and π. According to the implicit function theorem, $\frac{\partial q^*}{\partial \eta} > 0$ (computations in Appendix). This shows that if migrants are assumed to be inspirational to those who stayed, then, contrary to what the compositional model predicts, the share of the relatively scarce type-i individuals may actually increase in the home country A. The more inspirational they are, the stronger the effect in the home country.

In CM, convergence required the economic motive to dominate strongly. This transmission model (TM) greatly expands the range of parameters where we expect convergence. Furthermore, we can also show that $\frac{\partial q^*}{\partial \pi} > 0$. This counterbalances the predictions of the simple compositional model. In the compositional model, the stronger the economic motive, the more likely we are to observe cultural convergence. Conversely, here, stronger cultural homophily increases π, which results in a larger diffusion effect in the home country in the medium term.

In the exposition of TM we have implicitly assumed that the decision to migrate was made without taking into consideration the alternative socialization costs in both countries. 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 your homophily. If you want to socialize your offspring to resemble you, it is less costly to do so in an environment that resembles you. There are now three motivations to migrate: an economic gain, direct homophily, which favors cultural selection, and indirect homophily, whereby you want to socialize your offspring in an environment that resembles you. 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. At no cost to the generality of the model, we can therefore dispense with it in the name of simplicity.

Convergence is a more likely scenario according to TM than according to CM, and the stronger the selection effect, the stronger the push towards convergence. In contrast with CM, TM considers that the home, host, and migrant populations can evolve. TM predicts convergence between the home and the host populations, even when excluding the migrants. Such an empirical observation could not be explained by CM. It implies a cultural diffusion mechanism.

TM prediction 1

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. We believe that there is a kernel of truth in both CM and TM. We expect that they unravel at different speeds in the data. While the compositional model hints at immediate effects of emigration, e.g. the mechanics of migration, the transmission may reflect how the cultural technology of migration materializes in the longer term. Contrary to CM, TM does not provide such a clear time frame for the effect of migration. Migration from A to B between times t and $t+1$ might have an impact on the cultural mix of country A before $t+1$, and the effect may persist after $t+1$.

We have made the parsimonious assumption that emigrants still influence the cultural dynamics at home. If anything else matters, we might anticipate that a migrant offspring is likely to be influenced by a role model from the host population (more than from home). Whether π is smaller or larger than q^B, the share of type-i individuals in the migrant population remains larger than at home. We would also expect that a host offspring is socialized within the host community, with a very small probability of choosing a role model in the migrant population. If these intuitions are correct, they would only serve to reinforce the mechanism we have outlined. We have not considered whether migrants assimilate the structural determinants of the cultural equilibrium of the host country (ΔV_i^B and ΔV_j^B) or keep their own, and thus, how immigration affects the cultural equilibrium. In any case, since $\pi > q$, it is enough that $\eta > 0$ for our result to hold. We have also not considered how the structural determinants of the cultural preferences in the host population may be affected by immigration. It is possible that the host population adopt new cultural habits brought by immigrants, thus bringing the two countries closer together, though we can also imagine a reaction against these new habits. Though our model is agnostic towards the sign and magnitude of this structural effect, we believe it to be small and slow relative to the cultural diffusion-effect.

Identically, we may combine TM with standard economic intuitions to make further predictions. Since the thrust of TM is that cultural selection helps convergence, it yields opposed predictions relative to CM:

TM prediction 2

Uniformly more economic gains from migration should result in stronger divergence, or weaker convergence.

In TM, cultural selection into migration acts as a magnifying force of convergence between countries, instead of divergence, as was suggested by CM. CM paints the picture of clusters of countries, with countries within clusters made closer by migration, and clusters growing apart. Conversely, TM would not favor the emergence of clusters. TM suggests that convergence is stronger between countries further apart.

TM prediction 3

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 CM and TM as separate models but rather as a system of incentives and dynamics that unravel at the same time. The empirical analysis will serve as a way to inspect which of the two mechanisms dominates. Since the predictions of the models are diametrically opposed (see Table below), we have the possibility to discriminate between CM and TM through our empirical analysis.

Table 1: Comparison of Empirical Predictions

Mechanism / Implications	**Compositional Model**	**Transmission Model**
Home & host *populations*	Cultural divergence	Cultural convergence
$\upuparrows$ Economic gains	More cultural convergence	Less cultural convergence
$\upuparrows$ Initial cultural distance	Less cultural convergence	More cultural convergence

3 Cultural Distance Measures

3.1 Data

One of the main sources for cultural values in a society is the World Value Survey (WVS), which consists of nationally representative surveys among 400,000 respondents in 6 waves between 1981 and 2014. The WVS includes questions on beliefs, values and political attitudes in a repeated cross-section of almost 100 countries. Thanks to the large topical coverage of the survey questions, we are able to distinguish between various cultural features that range from political beliefs, family values, religiosity, and other dimensions of various cultures (we elaborate on the different dimensions of culture in more detail in section 4.1). Additionally, we draw from questions of 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 so that we can increase the number country pairs for which we have bilateral cultural similarity indexes from about 6,700 to over 7,800.

Migration data comes from the joint OECD and World Bank's Extended Bilateral Migration Database, which covers migration flows for each decade between 1960 and 2010. In addition, we make use of data from Brücker et al. (2013) (IAB) who collected data on migration into 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). If we expect differential effects of migration on cultural similarity depending on the educational level of migrants, we will be able to tease out potential counterbalancing effects with the help of the more detailed (but also more restricted) IAB data set. The United Nation ComTrade Database provides yearly bilateral trade flows around the globe, which we average over the time periods corresponding to the WVS waves.

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.

To rule out classical concerns about unbalanced panel data sets and to facilitate interpretation of our coefficients, we construct a balanced panel which consists of 24 countries for which we have both migration, trade, and WVS data in the years 1995, 2005 and 2010. Additionally, we construct another panel that covers 12 countries for which we have corresponding data for 4 consecutive waves in 1995, 2000, 2005, and

2010. In addition to the analysis of the unbalanced panel with the full set of countries, this will allow us to estimate more precisely how an increase in bilateral migration is associated with a change in the bilateral cultural distance.

3.2 Cultural Dimensions

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. (2015), and De Santis et al. (2016) include a vast set of questions with highly imbalanced coverages in terms of WVS waves or set of countries. This paper is concerned with the dynamic process behind the formation of attitudes, norms, and values over time with a particular focus on migration as a main driver in the cultural approximation between countries. This makes a consistent measure over time more important than in other studies. Naturally, this requirement will limit the scope of questions that we can include in our cultural distance measure. For our purpose, we include the cultural dimensions of the World Value Survey that are available in at least *all* of the 4 waves between 1994 and 2014 (most of them cover all 6 waves). In doing so, we avoid compositional effects that can either come from the selection of questions or countries that are missing in a given wave.

Most important for our purpose, the third wave of the World Value Survey asked what the most important value is that individuals would like to pass on to their children, choosing between thrift, obedience, determination, or religious faith. Figure 1 shows that in both Germany and Japan, an overwhelming majority of the population wants to pass on thrift and determination to their children, while a large share of spaniards want their children to be obedient. This is a reminder against simple geographical categorization of cultural preferences. The same question has been extended to account for multiple responses in wave 3 and all subsequent waves, where respondents can choose five out of eleven possibilities (good manners, independence, hard work, feeling of responsibility, imagination, tolerance and respect for other people, thrift saving money and things, determination perseverance, religious faith, unselfishness, obedience). This set of questions is most appropriate to test our model of intergenerational cultural transmission.

In Figure 2, we analyze the cultural proximity of countries using the values that individuals would like to pass on to their offspring. We conduct a principal component analysis, which identifies two dimensions (linear combinations of the average shares of respondents choosing one of the eleven options in the most recent WVS waves of 2010 to 2014) to discriminate countries. Unsurprisingly, Northern European countries, such as Sweden, the Netherlands, and Germany, seem to share common educational values. So do several Middle Eastern countries: Kuwait, Iraq, Qatar, Libya, Palestine, and Egypt. The apparent proximity between Taiwan, Japan, and Northern European countries may come across as more of a surprise, and sheds new light on the perception of cultural distance across countries.

The WVS captures other dimensions of culture. One question concerns the importance of family, friends, leisure, work, politics or religion in the life of respondents (for each item, they can choose between very important, rather important, not very important, not at all important). We use this question in addition to the "values to children" question to highlight the difference of cultural values in the intensive margin. In Figure 3, we replicate the principal component analysis on this question. A simple comparison between Figure 2 and 3 reveals that broadly similar groupings of countries, whether we consider the values passed on to children or priorities in life.

The model is founded upon the idea of the selection effect in migration. We have no reason to expect that all cultural dimensions are equivalent in that regard. Our analysis includes other dimensions of cultural preferences captured in the WVS (a comprehensive summary of the WVS questions used is given in Table 3). First, while attitudes about religion and family are persistent, trust can change quite dramatically (Putnam, 2001). The WVS asks: *Do you think people can generally be trusted or you cannot be too careful?*, a question which has been widely used to characterize social capital (Knack & Keefer, 1997; Alesina & La Ferrara, 2002; Portes, 2014). Trust differs qualitatively from personal beliefs, such as religiosity for instance, in that it is highly dependent on reciprocity and the environment. Finally, our analysis will include views on gender equality and the feeling of control over one's life.

In a first step, we will create an aggregate measure of cultural proximity for all questions that meet our coverage criterion and then split this aggregate variable into thematic clusters along the lines that we have described above to shed light on the sources of cultural convergence and possible heterogeneity in cultural dimensions.

3.3 Statistical dissimilarity

The appropriate choice and careful interpretation of statistical distance measures is central to the empirical analysis of cultural convergence. In this section, we introduce different examples of distance and entropy measures that we will use to compare national distributions of cultural values (which we will call P and Q in the following). With regards to statistical inference, different statistical measures highlight certain aspects of the underlying distributions and let us draw different conclusions about their properties.

The most well-known group of distance measures are derivatives of the Minkowski norms, which is written as $M_p = \sqrt[p]{\sum_{i=1}^{d} |P_i - Q_i|^p}$. The Euclidean, the Chebyshev, and the Canberra distances correspond respectively to M_2, M_∞ and M_0. How to interpret these three distances, when applied to cultural differences? A thorough discussion of the different distance / entropy / divergence measures can be found in Cha (2007). For now, it is sufficient to have a few intuitions in mind. Two almost identical countries that differ significantly along one cultural dimension will be characterized as far apart by the Chebyshev distance, but very similar by the Canberra norm. Conversely, two countries which differ a bit according to every dimension will be characterized as further apart by the Canberra norm than by the Chebyshev distance. The Euclidean norm is the usual Pythagorean metric. In a statistical analysis, the Chebyshev distance increases the weight of outliers, while the Canberra distance decreases it. Joint consideration of the three measures alleviates possible concerns as to the interpretation of a cultural distance.

We also consider statistical measures of *overlap*, also known as *inner product*. These measures of overlap give an idea about the number of matches in two distributions. In the context of cultural values, this measure would capture the idea that two people in country A and county B would give the same answer to a question in the World Value Survey. In Figure 4 and 5, we present the relationship between the Euclidean Distance and the Herfindahl Index, which is one form of the overlap measures. Both measures are standardized to normal ($\mu = 0$ and $\sigma = 1$) and use questions on priorities in life and desired qualities of character for children (we explained the exact content of the WVS questions in more detail above). Pearson's correlation is highly significant at the 1% level but only lies at 0.45 and 0.20 for priorities and child qualities respectively. The two measures capture two different aspects of cultural distance: the inverse of the Euclidean distance tells us how close on average the responses to a question in the WVS was, whereas the Herfindahl index tells us how probable it is that the same response to a question was given. Although these two interpretations may overlap in some cases, the analysis of very similar and dissimilar country pairs reveals some differences. For example, whereas the Euclidean measure ranks the United States and South Africa as the most similar in choice of child qualities and China and Argentina as the most dissimilar, the Herfindahl Index identifies Germany and Switzerland as the most similar country pair and Great Britain and Poland as the most dissimilar one.

We illustrate the construction of the cultural proximity measure with the help of the following example: the Euclidean distance along the cultural dimension of *Values desired to inherit to Children*. 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 2 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:

$$\mathbf{A}: 0\,0\,1\,1\,0\,1\,0\,1\,1\,0\,0$$
$$\mathbf{B}: 1\,0\,1\,1\,0\,0\,0\,1\,1\,0\,1$$

We calculate the share of people that have chosen the first characteristic for each country and wave and do the same for all of the possible answers (this yields the shares of people by characteristic, country and wave). We then calculate the Euclidean distance by taking the squared difference of the share of people in country i that have picked the first character trait and the share of people in country j. We repeat and sum this for all

11 dimensions. We then take the square root of this term, standardize it to normal for each wave, and multiply it by -1 in order to turn the distance measure to a proximity measure. The normalization of the measures has two advantages. First, it gives an easy way to interpret the sign of the measure, that is more similar countries have a positive distance measure, more dissimilar ones have a negative sign. Second, it makes them more comparable across cultural dimensions (note that the number of questions available for each cultural dimension varies largely), as these measures are sensitive to the number of questions included. This problem can be remedied by taking the overall standard deviation and mean for each wave for all countries into account.

In this paper, we are agnostic towards the choice of the best distance measure. However, we do 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 make use of three different distance measures (the Euclidean, the Herfindahl Index, and the Canberra relative distance) to highlight various forms of cultural convergence or divergence. 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., 2015; De Santis et al., 2016).

Figure 6 documents the distribution of the Cultural Similarity Index (measured as the Euclidean distance) for 21 countries over the course of 20 years. In these countries the WVS question on "Priorities in Life" was asked in all waves between 1995 and 2010, which ensures that our comparison is consistent over time. A first glance at the distribution of the CSI reveals a tightening over time. Negative values represent culturally distant country pairs, positive values represent culturally close countries (the maximum value is 1). The distribution of the CSI tightens for values larger zero, indicating that more countries become more similar and that there does not seem to be a polarization over time (e.g. a lot of very similar and many very dissimilar country pairs). All of this is a first hint towards a convergence of values across countries, which we will test more systematically in our empirical analysis.

4 Empirical Analysis

We are interested in the effect of bilateral migration on the change in cultural proximity over time. In our most demanding specification we include fixed effects for each country pair, home-time and destination-time fixed effects, which allows us to track changes within country pairs over time, irrespective of non-time varying bilateral determinants as well as time specific shocks (economic, political, environmental etc.) to sending and receiving countries. The equation writes as follows:

$$CS_{ijt} = \beta_0 + \beta_1 Mig_{ij,t-\Delta} + \beta_2 Trade_{ij,t-\Delta} + \theta_{ij} + \theta_{it} + \theta_{jt} + \varepsilon_{ijt}$$

CS is the bilateral cultural similarity between countries i and j over time. Our main coefficient of interest is β_1. Both migration $Mig_{ij,t-\Delta}$ and trade $Trade_{ij,t-\Delta}$ are lagged. As explained above, when the WVS wave starts in the middle of the decade, we use bilateral migration and trade data from the previous point in time for which we have an observation. For instance, if the WVS wave starts at 1994, we take data from 1990 with a $\Delta = 1$ lag; for $\Delta = 2$ we use data from 1985 etc. We follow Egger (2000) in including sending country-time and receiving country-time fixed effects, as well as bilateral fixed effects to control for non-time-varying characteristics of country pairs, accounting for above-mentioned traditional gravity controls (contiguity, geographical distance etc.).

We show the raw conditional correlation between migration and cultural similarity first, then including fixed effects in our OLS regression. For the OLS regressions we use the aggregate measure of cultural similarity, later differentiating between various cultural dimensions. In a next step, we construct two balanced panels of different time frames and country coverage (40 countries in 2 waves; 21 countries in 4 waves) to rule out that the overall effect is driven by changes in the sample of countries (a compositional effect, though not the one we describe in CM). The coefficient β_1 is an aggregate effect of both selection and diffusion (melting pot and transmission), that is the sign of the coefficient will tell us which of the two effects dominates.

4.1 Baseline Results

The baseline results serve as a first look at the net effect of migration on cultural proximity (including the effects of cultural selection, social mixing and cultural transmission). The reduced form analysis first establishes whether we observe overall convergence or divergence. The subsequent sub-sample and heterogeneity analysis will try to get at the mechanisms. If we observe cultural divergence, our model implications tell us that this must come from cultural selection. If we observe convergence, this can either stem from social mixing (Composition Model) or cultural transmission/social remittances (Transmission Model).

4.1.1 OLS Results

Table 4 presents the results of the OLS regression. Overall, we find that migration in the previous period (e.g. 5 years prior) is associated with cultural convergence, controlling for the full set of fixed effects. The results are presented for the Euclidean, the Herfindahl, and the Canberra measure of cultural similarity aggregated over all cultural dimensions, including values desired to inherit to children, priorities in life, ideas about gender equality, generalized trust, and freedom of choice. Columns 1,4, and 7 present the raw correlation between bilateral migration and cultural similarity exhibiting a positive and strongly significant result at the 1% significance level, indicating that culturally similar countries also experience high levels of bilateral migration. We introduce country pair fixed effects. These reduce the magnitude of the correlation, as well as the power of the estimation, as expected if they capture (and rid us of) time-invariant determinants of both migrations and cultural shifts. Adding destination-time and origin-time fixed effects reduces the magnitude of the correlation further, but it also improves the precision of the estimation, as expected if country-specific shocks had introduced a systematic bias in our initial specification, because they affected either migration flows or cultural shifts, but not the both of them at the same time. Nevertheless, the result holds when we introduce these fixed effects. They do not rid us of all possible concerns with this interpretation, but we would argue that this is reasonable evidence that a change in bilateral migration in $t-1$ leads to an increase in cultural similarity in t. The positive effect of migration on cultural similarity is overall consistent for the three different statistical measures of cultural similarity.

When comparing the effect of trade and migration, our results stand in contrast to previous findings on the effect of trade on cultural similarity (Maystre et al., 2014): we find no systematically positive effect of trade on cultural proximity for any of the statistical distance measures. However, Maystre et al. (2014) only use data from the 2nd and 4th wave of the World Value Survey, selecting a set of 30 questions and building an index of fractionalization, similar to the *Inner Product*. The authors do not include country-time fixed effects and differentiate between different types of trade, which may explain the differing results. Depending on the measure and inclusion of fixed effects, the coefficient of trade switches sign and loses statistical significance, particularly when we exploit within country-pair variation over time.

Figures 7 to 9 illustrate the outcomes of Table 4, differentiating between the topical dimensions that have been subsumed under the aggregate measure. The markers represent the point estimates for all specifications (without fixed effects, with bilateral fixed effects, and with bilateral and origin/destination-time fixed effects) on migration and trade respectively. The most demanding specification (marked in green) consistently provides positive coefficients, though not always statistically significant, hinting to country pairs becoming more culturally similar when they experience bilateral migration. This result is particularly clear for *values that are desirable to pass on to children*, the cultural dimension that best fits the idea of intergenerational cultural transmission.

4.1.2 Balanced Panel

In order to rule out concerns about results being potentially driven by changes in the WVS sample over time, we construct two balanced panel data sets. The balanced panel regression reduces the noise introduced through 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 through the construction of a balanced panel. In our constructed balanced panel we pick the three WVS waves with the highest country coverage (wave 3 with 53 countries, wave 5 with 58, and wave 6 with 59) such that we can track 24 countries for the period between 1994 and 2014. 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 largest 4-wave balanced panel we can construct is composed of 12 countries.

The results of the balanced panel regressions are presented in Table 5. 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.[5] In Table 6, we present the panel regression results for the disaggregated measures. Splitting the measure into its thematic subcomponents gets rid of a significant share of the variation explained through the model despite the inclusion of a large set of fixed effects. In fact, the R-squared drops from over 90% to an average 60% and even down to 35% in some cases. This shows that the aggregation of the different thematic dimensions of the cultural similarity index has an explanatory value in itself by showing that cultural convergence cannot easily be reduced to a single dimension of culture. All of the panel specifications find positive and significant effects of migration of the convergence of attitudes with regards to gender attitudes and the feeling of having choice and control over ones life. In the majority of cases, the sign of the migration coefficient is positive and consistent with the dominance of diffusion that we have found for the aggregate measure.

Our reduced form analysis exploits within country-pair variation over time and distinguishes between different measures of cultural proximity. We consistently find that migration contributes to cultural convergence between countries. As we have already mentioned, this result is consistent with both the Compositional and the Transmission Model. In the following section, we will try to identify the dominating mechanism, that

[5]In both unbalanced and balanced panel regressions the coefficient on the Herfindahl measure of Cultural Similarity turns negative when controlling for bilateral fixed effects indicating that the *Measure of Overlap* is affected differently by origin and destination time-specific shocks than our Minkowski type distance measures (Euclidean and Canberra). This uncovers some of the conceptual differences between the statistical measures and illustrates how important the inclusion of a complete set of measures and fixed effects is in order to draw conclusions about the empirical relationship of our two variables of interest. We are somewhat confident in our prediction of cultural convergence, as the inclusion of all fixed effects leaves us with consistently strong and positive coefficients for bilateral migration.

is to show whether convergence comes from social mixing or cultural transmission.

4.2 Social Mixing vs. Social Remittances

Testing CM and TM prediction 1

Now that we have established in the baseline regression that migration is associated with bilateral cultural convergence, we will now try to discriminate between the two potential mechanisms that could drive convergence in our model: social mixing and the transmission of values through social remittances. Since the data does not allow us to test these mechanisms directly, we will rely on a comparative analysis, which we base on implications that we derive from the model. We will first look at the timing of migration and its implications for CM versus TM. Then we will attempt to rule out cultural convergence through social mixing, by excluding migrants from the WVS sample. And lastly, we will try to proxy intensity of interactions between sending and receiving countries with bilateral remittances. We are aware that these measures are imperfect but we hope that in combination, all of the results will lead us to become more confident towards one mechanism versus another.

4.2.1 Timing of Migration

Let us first start with the timing between the instance where we observe migration and the subsequent bilateral cultural proximity. We expect that cultural convergence from social mixing should be instantaneous to migration, whereas the transmission mechanism should be magnified over time. The top half of Table 7 replicates the baseline results for which we lagged migration by 5 years. The bottom half reports the results when lagging migration and trade by 5 additional years increasing the overall lag to 10 years. We have fewer observations, which increases the standard deviation accordingly. However, the magnitude of the effect that we measure is also more than twice larger (for the most demanding specification with all fixed effects), and overall more statistically significant. These results are consistent with the compositional model having immediate effects, and the transmission effect to progressively gain traction, when the migrant has had time to learn the language, become more familiar with the local culture and eventually transmits these values to the home community.

4.2.2 Excluding Immigrants

As we have established in the baseline regression, we find that migration is associated with bilateral cultural convergence. Our compositional model suggests that the driving force is social mixing, e.g. more immigrants in the destination country, mechanically make home and host countries more similar. We have also established in CM prediction 1 that home and host *populations* should *diverge*, irrespective of whether we find overall cultural convergence between countries. Consequently, if we exclude migrants when calculating the cultural distance between countries, we should be able to partially rule out the effect of social mixing. If we still find convergence, this could suggest that the transmission mechanism is dominating.

To discriminate between predictions CM1 and TM1, we focus on the two waves of the World Value Survey 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 born in a different country. We replicate our analysis for this subset of countries excluding the foreign-born from the construction of the aggregate Euclidean distance measure and report the results in Table 8 (the results hold for the Herfindahl and Canberra measure of cultural distance). Excluding the foreign-born from the analysis does not alter our results: we still observe convergence. This supports TM prediction 1, against CM prediction 1. The size and significance of the coefficient are virtually identical. 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 pairs fixed effects. We take these results as further evidence for an effect of international migration on cultural convergence that goes beyond a simple re-allocation of cultural types.

4.2.3 Remittances

The model of cultural transmission with migrants relies on the basic assumption that migrants remain a vital part of the cultural formation in their home communities. This implies that there should be some link leading back to their countries of origin and there remains a certain level of interaction. Since we cannot directly observe interactions between emigrants and their home society (especially difficult for changes over time), we propose bilateral remittances as a proxy for intensity of interactions. Naturally, financial transfers are not equivalent to the transfer of norms or values (so called social remittances). However, there are various reasons to consider remittances as a potential proxy for the intensity of cultural transmission. As discussed in the model, cultural transmission is driven by emigrants remaining a part of the intergenerational cultural transmission. This means that the offspring can pick an emigrant as a role model. Within families, individuals may take after the relative whose remittances help support the family. Their influence can even extend beyond the family, whose affluence might advertise a cultural type in neighboring families. We also interpret remittances as a sign that migrants still have strong family or business ties in their home community. Therefore, remittances are an interesting proxy for the cultural interaction between the diaspora and their home community.

We draw from the World Bank Data Set for bilateral remittances. This data set has one important drawback as remittances on a disaggregate bilateral level have only been available since 2010. This will not allow for a dynamic analysis of remittances for the complete observation period in our analysis. However, we are able to split the sample into country pairs that have recorded remittance flows and country pairs that do not. The split sample is based on a static analysis of bilateral remittances in 2010 and consequently does not adequately represent the actual split in the 1980s. As the size of remittances has increased quite substantially over the last decades,[6] we will be overestimating the number of country pairs with remittances in previous decades, attributing a higher level of interaction between diaspora and home community to some country pairs. Detecting a stronger effect of migration on cultural proximity for this subset of countries will consequently be a conservative (lower bound) estimate, as we overestimate the level of interaction between country pairs in the subset of remittance country pairs.

About 48% of countries have immigrants that send money back to their home countries in 2010. Figure 10 illustrates the relationship between the size of the immigrant community and size of remittances flows in 2010. There is a strong correlation between the size of the diaspora and the volume of remittances that flow back to their home countries. There are still some outliers. Some country pairs have over proportionally high levels of remittances (above the gray line) and country pairs that remit less money than their total migrant stocks would indicate (below the gray line). This can be illustrated at the case of Mexico and the United States (upper right of the graph) with both high levels of migration and remittances but the size of remittances is still comparatively high. Table 14 confirms our intuition that countries with stronger social ties between diaspora and home community tend to converge culturally. Columns 1,3, and 5 show a positive and significant effect of migration on cultural convergence in countries where there are remittances. We control for the full set of fixed effects, with no impact on the magnitude of the effect. Meanwhile, in countries without remittances, the effect is not significant. This does not necessarily mean that migration does not have an effect on cultural convergence in those countries but that cultural divergence through selection may play a significant role and even outweigh convergence through transmission.

4.3 Economic Gains from Migration

Testing CM and TM prediction 2

In this section we will look more closely at the CM and TM predictions 2. We posit in the theoretical section that an increase in economic gains from migration will lead to more cultural convergence in the compositional model, since the migrant pool is less culturally selected, and will lead to less cultural convergence in the transmission model for the same reason. Since we do not have individual level data on the economic returns from migration, we will try to rely on (imperfect) proxies to approach this question. We

[6]The World Bank's Migration and Development Brief (No. 26) has estimated the increase in remittances by a 20-fold since 1990 reaching USD 432 Billion in 2015 and this is expected to rise in the future

will start by including the bilateral income gap to account for the fact that increases in economic distance between countries will also have an impact on the economic gains from migration. In a second step, we will look at skilled versus unskilled migration with the intention to relate skills to economic and cultural benefits from migration.

4.3.1 Bilateral Income Differences

Let us briefly reiterate how we can think about the relationship between bilateral income differences and economic gains from migration. According to CM prediction 2, we expect that an increase of the economic gain from migration would favor the social mixing over cultural selection effect. That is when economic gains become more attractive, the cultural composition of the pool of emigrants will be closer to the overall home population. Looking at TM prediction 2, we expect that a strong selection in the emigrant pool would accelerate the transmission of values from destination to home and thus more convergence once economic incentives are controlled for. In order to discriminate between those two channels, we would like to hold economic gains from migration constant and see how the coefficient for migration changes. Since the data does not allow us to do this on an individual level, we try to proxy aggregate economic gains from migration with GDP gap.

This has several flaws. First, the GDP gap is only a net of cost economic gain from migrating, e.g. with a higher GDP gap, we also expect migration costs to be higher. Second, the overall GDP gap is not sector specific, that is, we do not look at sectors where most immigrants from the sending countries are employed and then look at economic fluctuations there to proxy economic gains from migration (we also expect that sorting into sectors are endogeneous to the expected economic gains in that sector). However note that we do not make an argument about how the volume of bilateral migration changes once economic factor are controlled for. Instead, we suggest that once we control for economic distance, the cultural composition of the migrant pool changes (the migrant pool becomes more culturally selected) and thus its effect on cultural proximity. Also note that we are interested in how the coefficient of migration changes in the medium and long-run, including the full set of fixed effects and GDP gap. This means that we exploit the bilateral changes in economic distance on cultural distance, which in itself is interesting.

We first use GDP gap as a control in our main specification, and we present the results in Table 9. The income gap is negatively correlated with cultural distance in all specifications: countries tend to converge both economically and culturally, or to diverge along both dimensions. The interesting and relevant comparison is the size of the coefficient for migration in this tables versus in our baseline results (Table 7). Controlling for the income gap does not change our qualitative results on the effect of migration on cultural distance. Comparing columns 3, 6, and 9 of the first panel of Table 9 with Table 7, we see that in the short term, controlling for GDP gap reduces the effect of migration but this effect is quite small. However, in the medium run (lagging migration by 10 years), we see a significant increase in the coefficient for migration. If we believe that we can capture parts of the economics gains from migration by controlling for the GDP gap between countries, then we expect the migrant pool to be more culturally selected. A culturally more selected pool of migrants indicates *less* cultural convergence in CM and *more* cultural convergence in TM. We find that in the short-term the coefficient is slightly smaller, which is in line with the instantaneous effect of cultural selection in the compositional model. In the longer-run the coefficient for migration becomes larger, which confirms the transmission model. In both cases however, cultural convergence is still the dominating force.

4.3.2 Lower and High Skilled Migration

Let us now turn to the economic versus cultural gains from migration, looking at differences in skill levels. We expect that low skilled migrants emphasize economic gains from migration more than high skilled migrants. This would imply that the pool of low-skilled migrants is less culturally selected than the pool of skilled migrants. Connecting this to our theoretical model, CM would predict *more* cultural convergence through the emigration of the lower skilled and TM would predict *less* cultural convergence. The opposite holds true for high skilled migrants.

We are able decompose the migrant pool into low and high skill, using the IAB dataset (we define high skilled migrants as those migrants who attended college or received any degree above and lower skilled as any education below a college degree). In terms of data coverage we are limited by two restrictions: coverage through the IAB data set and coverage of both sending and receiving country in the WVS. Those restrictions leave us with about 1,800 observations for immigration to OECD countries from both non-OECD and other OECD countries but rather few longitudinal observations by country-pairs.

Table 10 presents the results of the OLS regression by skill level. We include a new specification that includes country-pair, origin and destination fixed effects as well as year fixed effects, allowing for some variation in country-time specific factors (columns 2, 5, and 8). Particularly the country-pair fixed effects will absorb a large share of the underlying migration costs, as they already encompass geographic distances, common language, past colonial ties or bilateral visa agreements that preceded our observation period. The analysis of migration by skill level allows us to get to heterogeneous migration incentives within countries and also track these changes over time. The previous specification, with country pairs, destination-year, and origin-year fixed effects (columns 3, 6, and 9) is too demanding for the data available and we loose some power. Although the sign of the effect remains consistent throughout all the specifications, it is not statistically significant here. Our results from the fixed effects regression demonstrate a negative relationship between cultural convergence and low-skilled migration, and positive with high-skilled migration. This is again indicative evidence for the transmission mechanism. The pool of high skilled migrants is more culturally selected, which emphasizes the transmission of values from host to home society and therefore increases cultural convergence.

4.4 Initial Cultural Distance

Testing CM and TM prediction 3

In a two-country setting where both countries have a cultural equilibrium with different shares of the two cultural types, migrants of the scarce cultural type at the home country have a cultural incentive to emigrate on top of the potential economic gains from migration. A larger cultural distance between the two countries should emphasize the cultural incentive to emigrate. The compositional model associates this with the cultural selection effect, which might lead to stronger divergence as suggested in CM prediction 3. Conversely, cultural transmission should emphasize the magnitude of convergence in equilibrium as proposed in TM prediction 3.

To discriminate between these contradictory predictions, we use subsample analyses where we look at different types of country pairs. First, we will split the sample based on our similarity indexes at the beginning of our observation period to see whether those country pairs classify as culturally similar or not. In a second step, we divide country pairs geographically, into the global South and North (differentiating between OECD and non OECD countries) and look at migration and cultural convergence within these sub-samples.

4.4.1 Similar Country-Pairs

We split the balanced panel sample along different threshold for initial cultural distance. In this case, we look at the distribution of our Cultural Similarity Indexes at the beginning of our panel sample period in 1995 and split the sample 2 groups according to different thresholds. Countries above and below the median level of cultural similarity in 1995 (50%), all country pairs that were among the 75% most dissimilar (or grouping the 25% most similar), and finally grouping the 90% most dissimilar country pairs (or 10% most similar) together. We present the results of the threshold regression in Table 11. The Table reports the coefficients for migration for the respective sub-samples. As proposed in CM and TM predictions 3, we see more or less cultural convergence depending on initial cultural distance. CM tells us that very similar country pairs should converge more. TM proposes that country pairs that are culturally further apart should converge at a lower rate. If we find that more similar countries converge more slowly, this would confirm the transmission model.

Looking at columns 2,4, and 6 we see that for a sub-sample of increasingly similar country pairs, the coefficient for migration first becomes larger (from column 2 to column

4) and then decreases and loses significance from column 4 to 5. This suggests that there may be some non-linearity in the relationship between speed of convergence and initial cultural distance. Nevertheless, we do find that convergence seems to stem from culturally dissimilar countries, which is in line with the predictions of the transmission model. It also plays into the natural intuition that much of the cultural convergence may be caused by culturally very different countries becoming more similar, rather than very similar countries becoming even more similar.

In order to present the results of Table 11 more intuitively, we create a dummy variable called *Similar* that takes the value 1 if a country pair belongs to the top 10% of most similar countries in 1995 in Panel A. Table 12 reports the results of the regression for all three measures, including the demanding set of fixed effects. Migration has the expected positive and significant coefficient, confirming that bilateral migration accelerates cultural convergence, but it does so at a lower intensity for country pairs that were already similar in the first place, as suggested by the negative sign of the interaction term. As mentioned in the previous section, this is confirming TM prediction 3. In the case of the Euclidean measure, the selection effect even outweighs the diffusion effect for culturally similar country pairs (adding the coefficient of migration and the interaction term leaves a negative overall coefficient of migration for culturally similar countries) and is almost canceled out for the Canberra measure.

4.4.2 South-North Migration

We continue our sub-sample analysis and focus now on the division of our sample according to what is typically classified as North-North, South-North, and South-South migration. We define as North countries all countries that were member states of the OECD at the end of our observation period (2014). The large majority of observations for bilateral migration in our data are South-North and South-South (together about N=6,200) with a minority of observations for North-North (N=700). We can interpret the migration between South and North in various ways. First, South-North country pairs typically have a clear sending and a clear receiving country, which corresponds more closely to our simple model set-up. This means that in our model country A corresponds to the sending non-OECD country and country B is the receiving OECD country. This is a more direct test of the transmission mechanism since convergence through social mixing in both countries (emigrants from A in B and emigrants from B in A) is reduced. Second, we expect that the diaspora from a developing country living in a developed country may potentially keep closer social ties to their host societies or conversely that emigrants from developing countries are more pertinent in remaining in the intergenerational cultural transmission of the home society (e.g. emigrants are more likely to act as role models). Third, we expect that the cultural distance between OECD countries and non-OECD countries may be larger than within OECD countries.

Table 13 reports the results of the sub-sample analysis for each of the cultural distance measures. All of the results come from the specification with the full set of fixed effects. Interestingly, our observed bilateral convergence comes from South-North migration. The coefficient for this sub-sample is three times the size of our baseline results (Table 4, columns 3,6, and 9). For the North-North sample, we find a negative coefficient, which is not significant for the Euclidean measure but becomes significant for the Herfindahl and Canberra measures. We find no effect of migration on cultural convergence between Southern countries. The large and positive effect of migration in the South-North sample confirms our expectations. With i) a clear sending and a clear receiving countries, ii) migrants being more likely to be picked as role models and iii) large initial cultural distance, we observe stronger cultural convergence. We consider this as further evidence for the transmission mechanism.

5 Conclusion

Does international migration contribute to cultural convergence or divergence between sending and receiving countries? We investigate this question both theoretically and empirically. We first develop a compositional model of international migration and cultural change, where divergence arises from self-selection on cultural traits and convergence arises from social mixing. The model is then adapted to allow for horizontal and vertical cultural transmission following Bisin & Verdier (2000). The model yields a rich set of predictions, which we test empirically using panel data from the World Value Survey and bilateral migration data for the period 1981-2014. We exploit within country-pair variation in cultural proximity over time and find support for the cultural transmission hypothesis. As the model with cultural transmission predicts, migration generates bilateral cultural convergence even if we exclude migrants from the pool of respondents in both countries (hence eliminating social mixing). It is also more likely in the long-run, especially after controlling for economic incentives to migrate and for initial cultural distance, which is consistent with the cultural transmission hypothesis (but not with compositional changes). Interestingly, international migration appears as a stronger and more robust driver of cultural convergence than trade. The results hold for a large set of time-varying cultural distance measures along different statistical and topical dimensions.

A potential path for future research is to dive deeper into the direction of convergence, that is whether sending countries move towards receiving countries or vice versa. One of the implications of the compositional model versus the transmission model is that if we observe convergence, this convergence must come from the host towards the home country (social mixing). In the transmission model, we expect that the home country moves towards the host country (social remittances). More micro-level analyses could contribute to our knowledge about the direction of convergence and consequently shed light on the potential creation of a globally homogeneous culture.

References

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. & La Ferrara, E. (2002). Who trusts others? *Journal of Public Economics*, 85(2), 207–34.

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.

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.

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.

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., Ortuño-Ortín, I., & Wacziarg, R. (2015). *Culture, ethnicity and diversity*. NBER working paper 20989, National Bureau of Economic Research.

Disdier, A.-C., Tai, S. H., Fontagné, L., & Mayer, T. (2010). Bilateral trade of cultural goods. *Review of World Economics*, 145(4), 575–95.

Docquier, F., Tansel, A., & Turati, R. (2017). *Do emigrants self-select along cultural traits? Evidence from the MENA countries*. GLO discussion paper 146, Global Labor Organization.

Egger, P. (2000). A note on the proper econometric specification of the gravity equation. *Economics Letters*, 66(1), 25–31.

Fargues, P. (2006). *The demographic benefit of international migration: Hypothesis and application to Middle Eastern and North African contexts*. World Bank Policy Research Working Paper 4050, The World Bank.

Felbermayr, G. J. & Toubal, F. (2010). Cultural proximity and trade. *European Economic Review*, 54(2), 279–93.

Fernandez, R. (2011). Does culture matter? In J. Benhabib, M. O. Jackson, & A. Bisin (Eds.), *Social Economics*, volume 1. North-Holland.

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

Hirschman, A. O. (1981). Exit, voice, and the state. *Essays in Trespassing. Economics to Politics and Beyond*.

Hofstede, G. H. (2001). *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations*. Sage.

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

Knack, S. & Keefer, P. (1997). Does social capital have an economic payoff? a cross-country investigation. *The Quarterly Journal of Economics*, 112(4), 1251–88.

Landes, D. (1998). The wealth and poverty of nations: Why some are so rich and some are so poor.

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.

Olivier, J., Thoenig, M., & Verdier, T. (2008). Globalization and the dynamics of cultural identity. *Journal of International Economics*, 76(2), 356–70.

Portes, A. (2014). Downsides of social capital. *Proceedings of the National Academy of Sciences*, 111(52), 18407–8.

Putnam, R. D. (2001). *Bowling alone: The collapse and revival of American community*. Simon and Schuster.

Rauch, J. E. & Trindade, V. (2002). Ethnic Chinese networks in international trade. *Review of Economics and Statistics*, 84(1), 116–30.

Spolaore, E. & Wacziarg, R. (2009). 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. (2014). *Fertility and Modernity*. Discussion Papers type 779, Department of Economics, Tufts University.

Appendix A Cultural equilibrium with migrants as role models

Define

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

It is easy to obtain

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

Since $\chi > q$, the two first terms add up to a negative amount, and the third term is negative, so that $\frac{\partial g}{\partial q} < 0$. We can also write

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

To find the sign of this expression, let us write that $\frac{\partial g}{\partial \eta} > 0$ iff $\chi(1-\chi)\left(\Delta V_j(1-\tau_i)+\Delta V_i(1-\tau_j)\right) < (1-\tau_i)(1-\tau_j)$. Replacing with the expressions of the main text, this is equivalent to

$$-(\Delta V_i + \Delta V_j - 2\Delta V_i \Delta V_j)\chi^2 + 2\Delta V_j(1-\Delta V_i)\chi - 1 + \Delta V_i < 0$$

For this to be true for all χ it is enough that the discriminant of this expression be negative, which it is for any $\Delta V_i, \Delta V_j < 1$. Then

$$\frac{dq^*}{d\eta} = -\frac{\frac{\partial g}{\partial \eta}}{\frac{\partial g}{\partial q}} > 0.$$

The same reasoning holds for $\frac{dq^*}{d\pi}$.

Appendix B Figures

Figure 1: Distribution of Responses to Values to Children Question in Wave 3 of the WVS for Germany, Japan, and Spain

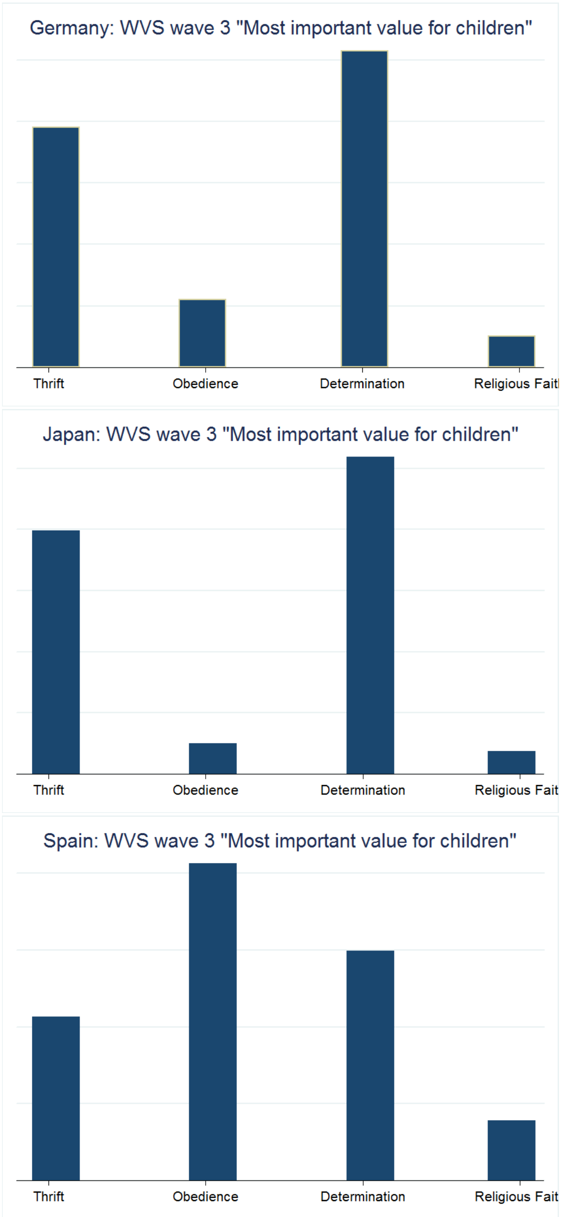


Figure 2: Principal Component Analysis of WVS question on values desired to inherit to children, including independence, hard work, feeling of responsibility, imagination, tolerance and respect for others, thrift saving money and things, determination/perseverance, religious faith, unselfishness, and obedience

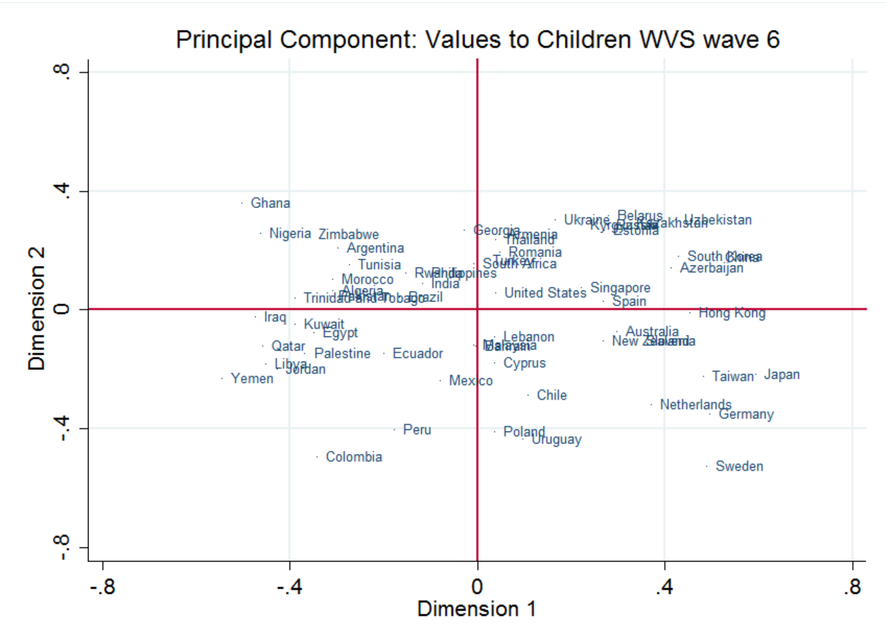


Figure 3: Principal Component Analysis of WVS question on priorities in life,including family, friends, leisure time, politics, work, and religion

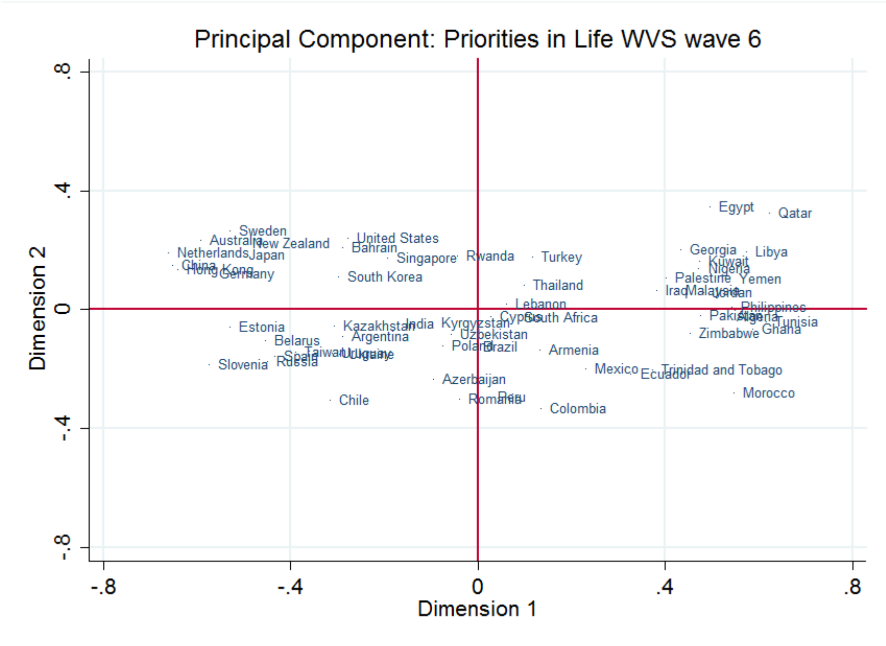


Figure 4: Scatterplot of standardized values for the Herfindahl Index and weighted Euclidean Distance

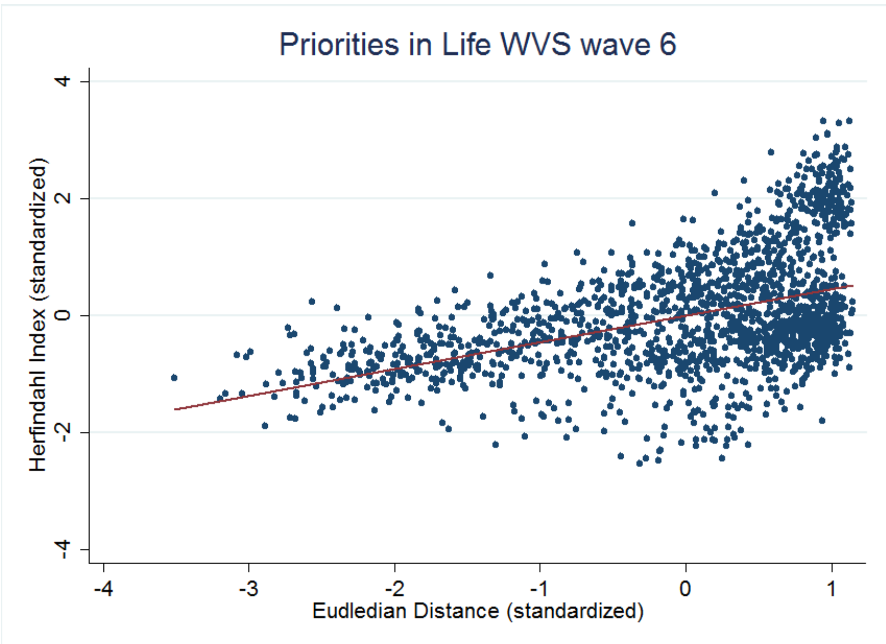


Figure 5: Scatterplot of standardized values for the Herfindahl Index and weighted Euclidean Distance

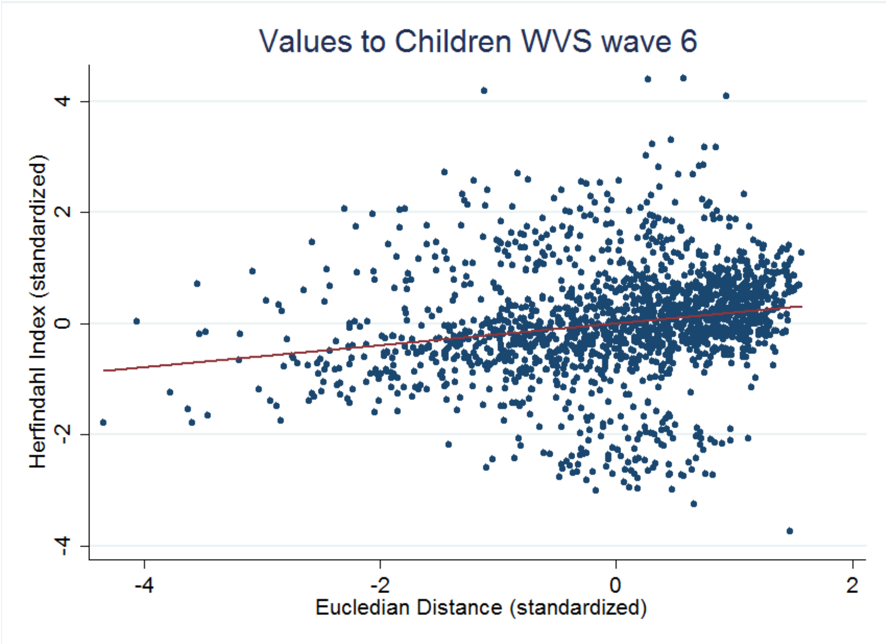


Figure 6: Kernel Density of WVS question on priorities in life over time for 21 countries with observations between 1995 and 2010, showing that the distribution of cultural norms becomes more compact

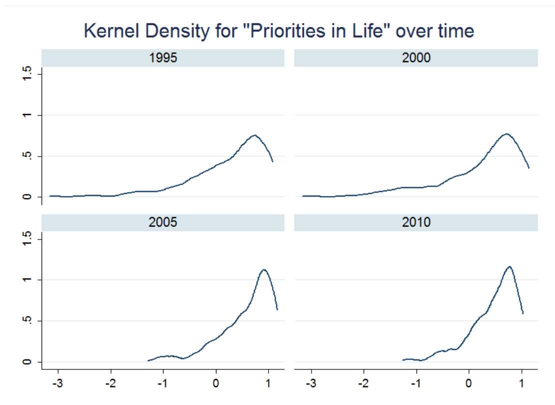


Figure 7: OLS Regression Outcomes for three Specifications with -1*Euclidean Distance

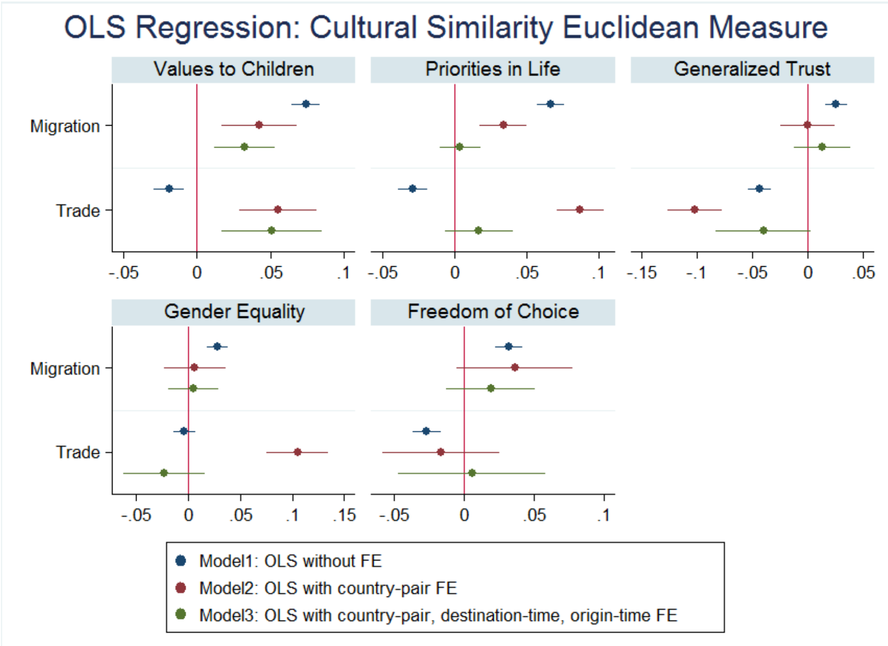


Figure 8: OLS Regression Outcomes for three Specifications with Herfindahl Index

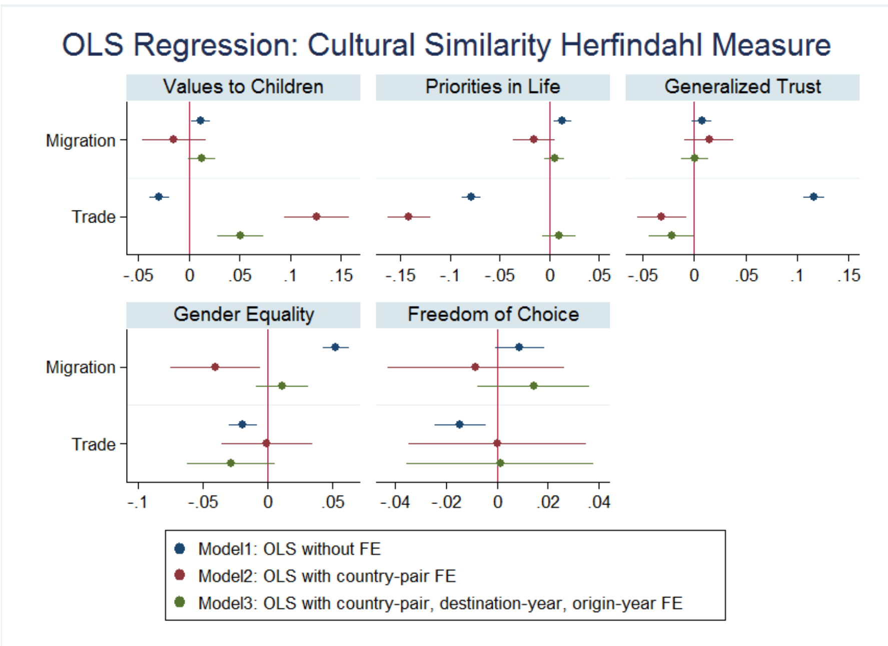


Figure 9: OLS Regression Outcomes for three Specifications with -1*Canberra Index

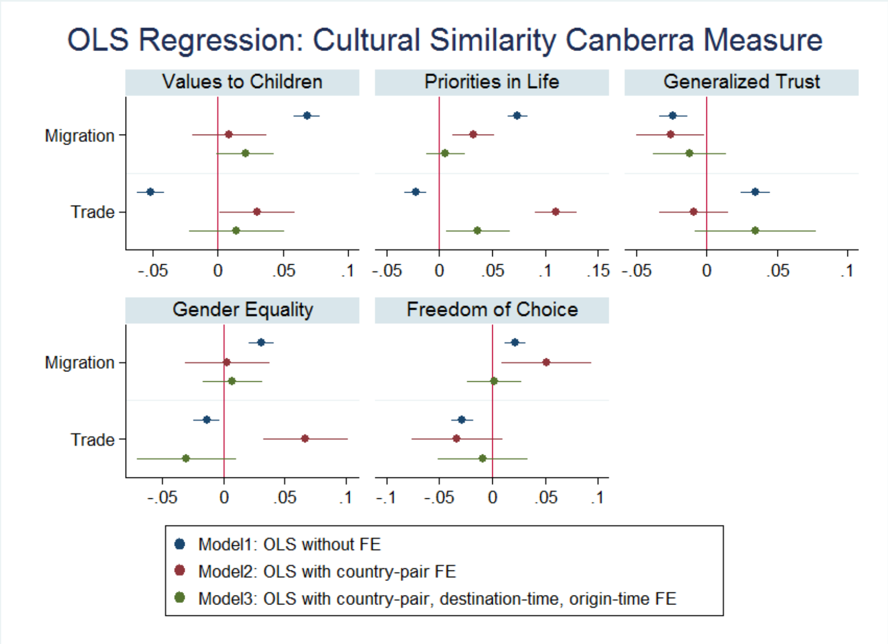
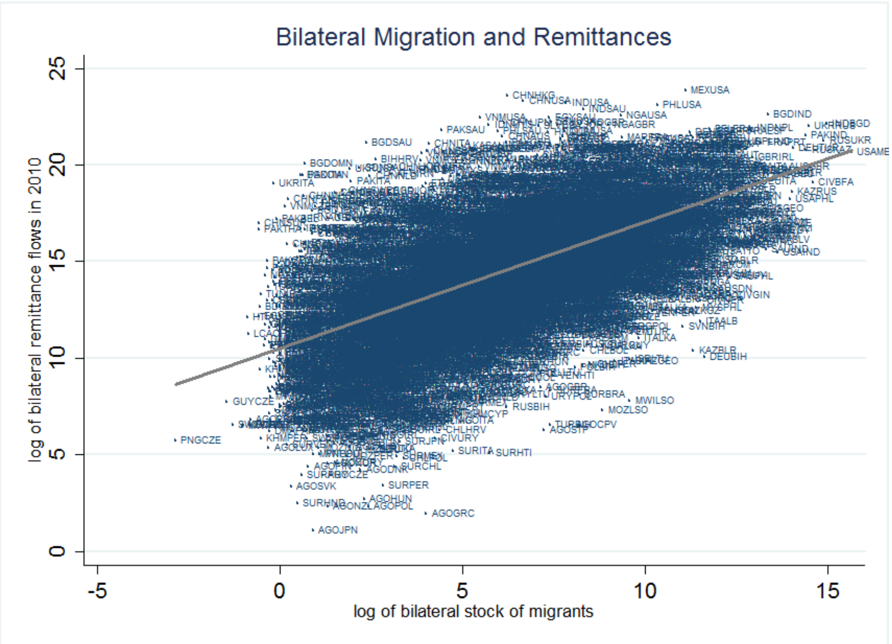


Figure 10: Scatterplot: Log of Remittances and Log of Stock of Migrants in 2010



Appendix C Tables

Table 2: 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 3: Selected World Value Survey 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 4: OLS regression of aggregate Cultural Similarity Measures

	Euclidean			**Herfindahl**			**Canberra**		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Migration	0.0730***	0.0503***	0.0237**	0.0370***	-0.0290*	0.0175**	0.0689***	0.0359**	0.0213*
	(0.00489)	(0.0118)	(0.00982)	(0.00499)	(0.0149)	(0.00776)	(0.00490)	(0.0150)	(0.0118)
Trade	-0.0374***	0.0855***	0.00263	-0.00997*	-0.0339**	8.27e-05	-0.0334***	0.0515***	0.00192
	(0.00518)	(0.0119)	(0.0164)	(0.00528)	(0.0150)	(0.0130)	(0.00519)	(0.0151)	(0.0196)
Constant	0.0651	-2.078***	-0.189	-0.117	0.909***	-0.147	0.0554	-1.278***	-0.111
	(0.0798)	(0.205)	(0.333)	(0.0814)	(0.257)	(0.263)	(0.0800)	(0.259)	(0.398)
Observations	6,983	6,983	6,983	6,983	6,983	6,983	6,983	6,983	6,983
R-squared	0.032	0.910	0.966	0.010	0.859	0.979	0.029	0.855	0.952
Fixed Effects									
Country-Pair	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Dest.-Year	No	No	Yes	No	No	Yes	No	No	Yes
Origin-Year	No	No	Yes	No	No	Yes	No	No	Yes
Model	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS

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

Table 5: Panel Regression Aggregate Cultural Similarity and all Statistical Measures

	Euclidean			Herfindahl			Canberra		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: 24 countries in 3 Waves									
Migration	0.137***	0.164***	0.161***	0.142***	-0.175***	0.119***	0.218***	0.0881	0.136***
	(0.0302)	(0.0542)	(0.0432)	(0.0282)	(0.0588)	(0.0343)	(0.0347)	(0.0537)	(0.0455)
Trade	-0.0763**	0.0626	-0.0723	-0.219***	-0.130*	-0.0482	-0.227***	0.0129	-0.133
	(0.0377)	(0.0640)	(0.115)	(0.0351)	(0.0695)	(0.0912)	(0.0432)	(0.0635)	(0.121)
Constant	0.966	-2.191*	1.038	2.830***	4.145***	-0.840	2.951***	-0.741	1.980
	(0.641)	(1.165)	(2.484)	(0.598)	(1.266)	(1.972)	(0.736)	(1.156)	(2.615)
Observations	1,259	1,259	1,259	1,259	1,259	1,259	1,259	1,259	1,259
R-squared	0.017	0.794	0.928	0.032	0.724	0.949	0.033	0.848	0.941
Panel B: 12 countries in 4 Waves									
Migration	0.166***	0.0884*	0.0298	0.139***	0.0544	0.0455	0.159***	0.199***	0.143***
	(0.0201)	(0.0527)	(0.0612)	(0.0208)	(0.0501)	(0.0513)	(0.0198)	(0.0468)	(0.0508)
Trade	-0.246***	-0.0796	-0.273**	-0.0638**	-0.0160	-0.209**	-0.277***	-0.132***	0.00815
	(0.0290)	(0.0506)	(0.119)	(0.0300)	(0.0482)	(0.100)	(0.0286)	(0.0450)	(0.0991)
Constant	3.629***	0.820	6.019**	-0.0257	-0.177	4.210*	4.399***	0.811	-1.629
	(0.522)	(0.877)	(2.865)	(0.539)	(0.835)	(2.401)	(0.515)	(0.779)	(2.379)
Observations	454	454	454	454	454	454	454	454	454
R-squared	0.158	0.742	0.904	0.106	0.768	0.933	0.179	0.796	0.934
Fixed Effects									
Country-Pair	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Dest.-Year	No	No	Yes	No	No	Yes	No	No	Yes
Origin-Year	No	No	Yes	No	No	Yes	No	No	Yes
Model	Panel	Panel	Panel	Panel	Panel	Panel	Panel	Panel	Panel

Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1, Panel Regression includes 24 countries for which we have WVS observations in waves 3,5, and 6. Countries in Panel A include: ARG AUS CHL CHN COL DEU ESP GEO IND JPN KOR MEX NZL PER POL ROM RUS SVN SWE TUR UKR URY USA ZAF. Countries in Panel B include: ARG CHL CHN ESP IND JPN KOR MEX PER TUR USA ZAF

Table 6: Panel Regression with 24 countries Disaggregate Cultural Similarity and all Statistical Measures

	VtC	PiL	Trust	Gender	Freedom
	(1)	(2)	(3)	(4)	(5)
Euclidean					
Migration	0.0286	0.0120	0.0284	0.0433*	0.0626***
	(0.0196)	(0.0151)	(0.0192)	(0.0222)	(0.0239)
Trade	0.0232	0.00792	0.0156	-0.120**	0.00397
	(0.0522)	(0.0406)	(0.0513)	(0.0590)	(0.0636)
Constant	-0.600	-0.645	-0.608	2.062*	-0.631
	(1.101)	(0.863)	(1.079)	(1.245)	(1.342)
Observations	1,259	1,221	1,219	1,259	1,259
R-squared	0.659	0.672	0.359	0.640	0.725
Herfindahl					
Migration	0.0196*	0.00820	0.00259	0.0324**	0.0399**
	(0.0119)	(0.0100)	(0.00913)	(0.0146)	(0.0161)
Trade	0.00642	0.00637	-0.00367	-0.0883**	0.00524
	(0.0315)	(0.0269)	(0.0243)	(0.0389)	(0.0429)
Constant	-0.686	-0.152	0.0662	1.708**	-0.551
	(0.665)	(0.570)	(0.512)	(0.821)	(0.905)
Observations	1,259	1,221	1,259	1,259	1,259
R-squared	0.893	0.837	0.761	0.838	0.806
Canberra					
Migration	-0.00465	0.0192	-0.0284	0.0495**	0.0452**
	(0.0196)	(0.0170)	(0.0192)	(0.0192)	(0.0200)
Trade	-0.0476	0.0659	-0.0156	-0.0630	-0.0500
	(0.0521)	(0.0457)	(0.0513)	(0.0511)	(0.0527)
Constant	1.140	-1.808*	0.608	0.864	0.900
	(1.099)	(0.970)	(1.079)	(1.077)	(1.087)
Observations	1,259	1,221	1,219	1,259	1,257
R-squared	0.440	0.463	0.359	0.673	0.802

Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1, Panel Regression (with country-pair FE, destination-time, origin-time FE) includes 24 countries for which we have WVS observations in waves 3,5, and 6. These countries include: ARG AUS CHL CHN COL DEU ESP GEO IND JPN KOR MEX NZL PER POL ROM RUS SVN SWE TUR UKR URY USA ZAF

Table 7: OLS regression of aggregate Cultural Similarity Measures

	Euclidean			Herfindahl			Canberra		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Lag: Δ = 1									
Migration	0.0730***	0.0503***	0.0237**	0.0370***	-0.0290*	0.0175**	0.0689***	0.0359**	0.0213*
	(0.00489)	(0.0118)	(0.00982)	(0.00499)	(0.0149)	(0.00776)	(0.00490)	(0.0150)	(0.0118)
Trade	-0.0374***	0.0855***	0.00263	-0.00997*	-0.0339**	8.27e-05	-0.0334***	0.0515***	0.00192
	(0.00518)	(0.0119)	(0.0164)	(0.00528)	(0.0150)	(0.0130)	(0.00519)	(0.0151)	(0.0196)
Constant	0.0651	-2.078***	-0.189	-0.117	0.909***	-0.147	0.0554	-1.278***	-0.111
	(0.0798)	(0.205)	(0.333)	(0.0814)	(0.257)	(0.263)	(0.0800)	(0.259)	(0.398)
Observations	6,983	6,983	6,983	6,983	6,983	6,983	6,983	6,983	6,983
R-squared	0.032	0.910	0.966	0.010	0.859	0.979	0.029	0.855	0.952
Lag: Δ = 2									
Migration	0.0582***	0.0553**	0.0554**	0.0339***	-0.0589**	0.0408**	0.0648***	0.101***	0.0690***
	(0.00873)	(0.0258)	(0.0233)	(0.00829)	(0.0257)	(0.0191)	(0.00889)	(0.0293)	(0.0253)
Trade	-0.0407***	0.0461*	0.0108	-0.00886	-0.150***	-0.0260	-0.0481***	-0.0447	-0.0784*
	(0.00997)	(0.0257)	(0.0407)	(0.00946)	(0.0256)	(0.0334)	(0.0101)	(0.0292)	(0.0442)
Constant	0.369**	-1.312***	-0.609	-0.170	3.413***	-0.0415	0.393**	0.00742	0.947
	(0.162)	(0.453)	(0.864)	(0.154)	(0.450)	(0.710)	(0.165)	(0.514)	(0.939)
Observations	1,840	1,840	1,840	1,840	1,840	1,840	1,840	1,840	1,840
R-squared	0.024	0.941	0.978	0.011	0.934	0.983	0.028	0.927	0.975
Fixed Effects									
Country-Pair	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Dest.-Year	No	No	Yes	No	No	Yes	No	No	Yes
Origin-Year	No	No	Yes	No	No	Yes	No	No	Yes
Model	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS

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

Table 8: OLS Regression With and Without Foreign-Born

	Aggregate Euclidean Distance Measure							
	With Foreign-Born				**Without Foreign-Born**			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Migration	0.0896***	0.0353***	0.0364***	0.0344	0.0873***	0.0344***	0.0353***	0.0335
	(0.0113)	(0.0105)	(0.0104)	(0.0855)	(0.0112)	(0.0106)	(0.0105)	(0.0854)
Trade	-0.0572***	0.0998***	0.101***	-0.0286	-0.0534***	0.101***	0.102***	-0.0348
	(0.0125)	(0.0163)	(0.0161)	(0.117)	(0.0124)	(0.0163)	(0.0161)	(0.117)
Constant	0.262	-1.052***	-1.067***	0.198	0.211	-1.042***	-1.058***	0.319
	(0.187)	(0.303)	(0.298)	(2.202)	(0.187)	(0.303)	(0.299)	(2.200)
Observations	1,308	1,308	1,308	1,308	1,308	1,308	1,308	1,308
R-squared	0.046	0.638	0.656	0.995	0.044	0.635	0.653	0.995
Origin	No	Yes	No	No	No	Yes	No	No
Dest.	No	Yes	No	No	No	Yes	No	No
Dest.-Year	No	No	Yes	Yes	No	No	Yes	Yes
Origin-Year	No	No	Yes	Yes	No	No	Yes	Yes
Country-Pair	No	No	No	Yes	No	No	No	Yes

Analysis includes only countries and WVS waves for which information on respondent's country of birth was available. Countries for wave 2 include: BRA CHL CHN CZE IND JPN MEX NGA SVK TUR Countries for wave 3 include: ARG ARM AUS AZE BGR BIH BLR CHE CHL CHN CZE DEU DOM ESP EST FIN GEO HRV HUN IND JPN LTU LVA MDA MEX MKD MNE NGA NOR NZL PAK PER PHL PRI ROM RUS SRB SVK SVN SWE TUR TWN UKR URY USA VEN ZAF

Table 9: OLS regression of aggregate Cultural Similarity Measures including differences in income levels

	Euclidean			**Herfindahl**			**Canberra**		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Lag: Δ = 1									
Migration	0.0759***	0.0485***	0.0217**	0.0385***	-0.0290*	0.0167**	0.0710***	0.0357**	0.0193
	(0.00466)	(0.0119)	(0.00980)	(0.00494)	(0.0149)	(0.00775)	(0.00478)	(0.0150)	(0.0118)
Trade	-0.0109**	0.0809***	-0.00477	0.00243	-0.0409**	-0.00557	-0.0138***	0.0502***	-0.00426
	(0.00503)	(0.0131)	(0.0164)	(0.00534)	(0.0164)	(0.0130)	(0.00516)	(0.0166)	(0.0197)
GDP Gap	-0.204***	0.0178	-0.0438***	-0.0976***	0.0185	-0.0305***	-0.152***	0.00219	-0.0370**
	(0.00761)	(0.0173)	(0.0133)	(0.00808)	(0.0217)	(0.0105)	(0.00781)	(0.0219)	(0.0159)
Constant	1.368***	-2.136***	0.379	0.509***	0.876***	0.254	1.023***	-1.270***	0.370
	(0.0903)	(0.211)	(0.363)	(0.0958)	(0.266)	(0.287)	(0.0926)	(0.268)	(0.435)
Observations	6,977	6,977	6,977	6,977	6,977	6,977	6,977	6,977	6,977
R-squared	0.123	0.910	0.967	0.030	0.860	0.980	0.079	0.855	0.952
Lag: Δ = 2									
Migration	0.0893***	0.0875**	0.0717**	0.0910***	-0.0246	0.0516*	0.132***	0.159***	0.104***
	(0.0157)	(0.0431)	(0.0362)	(0.0163)	(0.0365)	(0.0267)	(0.0165)	(0.0407)	(0.0348)
Trade	-0.0405*	-0.0264	-0.171	-0.0992***	-0.0268	-0.121	-0.150***	-0.116**	0.104
	(0.0210)	(0.0526)	(0.127)	(0.0219)	(0.0446)	(0.0937)	(0.0221)	(0.0496)	(0.122)
GDP Gap	-0.219***	-0.125*	-0.158***	0.00320	-0.163***	-0.101**	-0.161***	-0.123*	-0.117**
	(0.0290)	(0.0723)	(0.0557)	(0.0303)	(0.0613)	(0.0411)	(0.0306)	(0.0683)	(0.0535)
Constant	2.053***	0.906	4.627	1.155***	2.284***	2.924	3.368***	2.032**	-2.179
	(0.417)	(0.997)	(2.902)	(0.434)	(0.845)	(2.142)	(0.439)	(0.941)	(2.786)
Observations	648	648	648	648	648	648	648	648	648
R-squared	0.110	0.846	0.953	0.050	0.891	0.975	0.134	0.879	0.962
Fixed Effects									
Country-Pair	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Dest.-Year	No	No	Yes	No	No	Yes	No	No	Yes
Origin-Year	No	No	Yes	No	No	Yes	No	No	Yes
Model	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS

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

Table 10: OLS Regression Cultural Similarity by Skill Level

	Euclidean			**Herfindahl**			**Canberra**		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Lower Skilled	0.0214	-0.296***	-0.0355	-0.223***	-0.676***	-0.0300	-0.0513*	-0.439***	-0.00135
	(0.0294)	(0.0478)	(0.0379)	(0.0308)	(0.0694)	(0.0306)	(0.0295)	(0.0681)	(0.0577)
High Skilled	0.0677**	0.270***	0.0233	0.193***	0.513***	0.0150	0.0678**	0.340***	-0.0268
	(0.0281)	(0.0510)	(0.0388)	(0.0294)	(0.0740)	(0.0313)	(0.0281)	(0.0726)	(0.0591)
Constant	-1.603***	-0.544**	0.0913	-1.237***	1.565***	0.317**	-0.516***	0.729**	0.495*
	(0.190)	(0.237)	(0.195)	(0.199)	(0.344)	(0.158)	(0.190)	(0.337)	(0.298)
Observations	1,764	1,827	1,827	1,764	1,827	1,827	1,764	1,827	1,827
R-squared	0.083	0.927	0.984	0.061	0.855	0.990	0.012	0.840	0.959
Dest.-Year	No	No	Yes	No	No	Yes	No	No	Yes
Origin-Year	No	No	Yes	No	No	Yes	No	No	Yes
Country-Pair	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Dest.	No	Yes	No	No	Yes	No	No	Yes	No
Origin	No	Yes	No	No	Yes	No	No	Yes	No
Time	No	Yes	No	No	Yes	No	No	Yes	No

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

Table 11: Threshold Estimation for Initial Cultural Distance

	50% most similar		**25% most similar**		**10% most similar**	
	below	above	below	above	below	above
	(1)	(2)	(3)	(4)	(5)	(6)
Euclidean						
	0.114*	0.141**	0.152***	0.223**	0.163***	-0.117
	(0.0639)	(0.0616)	(0.0509)	(0.0933)	(0.0460)	(0.196)
Herfindahl						
	0.0800*	0.179***	0.108***	-0.0422	0.112***	0.0204
	(0.0477)	(0.0574)	(0.0377)	(0.106)	(0.0349)	(0.326)
Canberra						
	0.0280	0.182**	0.111**	0.0169	0.149***	-0.0285
	(0.0616)	(0.0720)	(0.0490)	(0.220)	(0.0470)	(0.391)

Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1, Threshold estimation with split samples: left columns present coefficient of bilateral migration for the subsample below the respective % threshold and vice versa for the right columns. Panel Regression (with country-pair FE, destination-time, origin-time FE) includes 24 countries for which we have WVS observations in waves 3,5, and 6. These countries include: ARG AUS CHL CHN COL DEU ESP GEO IND JPN KOR MEX NZL PER POL ROM RUS SVN SWE TUR UKR URY USA ZAF

Table 12: Panel Regression by Initial Cultural Similarity

	Euclidean (1)	Herfindahl (2)	Canberra (3)
Migration	0.0688***	0.0414***	0.0499**
	(0.0176)	(0.0129)	(0.0196)
Migration**Similar*	-0.0991*	-0.0213	-0.440***
	(0.0533)	(0.0541)	(0.0825)
Trade	-0.0284	-0.0385	-0.0566
	(0.0464)	(0.0342)	(0.0519)
Constant	0.120	0.590	1.342
	(0.977)	(0.718)	(1.100)
Observations	1,259	1,259	1,259
R-squared	0.648	0.850	0.643
Country-Pair FE	Yes	Yes	Yes
Dest.-Year	Yes	Yes	Yes
Origin-Year	Yes	Yes	Yes
Model	Panel	Panel	Panel

Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1, Similar is a dummy variable defined as belonging to the 10% of most similar country pairs in 1995. Panel Regression (with country-pair FE, destination-time, origin-time FE) includes 24 countries for which we have WVS observations in waves 3,5, and 6.

Table 13: OLS Regression by Countries' Development Status

	North-North			**South-North**			**South South**		
	Euclidean	Herfindahl	Canberra	Euclidean	Herfindahl	Canberra	Euclidean	Herfindahl	Canberra
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Migration	-0.00334	-0.0564*	-0.209***	0.0541***	0.0367***	0.0616***	0.0111	0.0121	0.0129
	(0.0384)	(0.0316)	(0.0595)	(0.0151)	(0.0124)	(0.0192)	(0.0146)	(0.0112)	(0.0162)
Trade	0.0484	0.00482	-0.0660	-0.0120	-0.0199	-0.0394	-0.0392*	-0.0220	0.00725
	(0.0971)	(0.0798)	(0.150)	(0.0256)	(0.0210)	(0.0326)	(0.0228)	(0.0175)	(0.0252)
Constant	-0.360	1.412	4.312	-0.540	-0.145	0.0573	0.909**	0.324	-0.0141
	(2.159)	(1.774)	(3.345)	(0.521)	(0.428)	(0.662)	(0.448)	(0.344)	(0.496)
Observations	702	702	702	3,108	3,108	3,108	3,173	3,173	3,173
R-squared	0.953	0.991	0.944	0.983	0.982	0.966	0.958	0.981	0.960

Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1, OLS regression with full set of fixed effects (bilateral FE, destination-year FE, origin-year FE). Sample split into 3 categories: North-North for bilateral migration between OECD countries (members in 2014), South-South for bilateral migration between non-OECD countries, and South-North for bilateral migration between OECD and non-OECD.

Table 14: OLS Regression Cultural Similarity Split Sample by Remittance Flows

	Euclidean		Herfindahl		Canberra	
	REM	no REM	REM	no REM	REM	no REM
	(1)	(2)	(3)	(4)	(5)	(6)
Migration	0.0278**	-0.00463	0.0229**	-0.00586	0.0312**	-0.00915
	(0.0114)	(0.0253)	(0.00909)	(0.0200)	(0.0137)	(0.0300)
Trade	0.0212	-0.0696	0.00514	-0.0304	-0.0156	0.0477
	(0.0184)	(0.0485)	(0.0147)	(0.0385)	(0.0221)	(0.0575)
Constant	-0.623	1.368	-0.328	0.524	0.139	-0.703
	(0.383)	(0.885)	(0.306)	(0.702)	(0.460)	(1.050)
Observations	5,177	1,178	5,177	1,178	5,177	1,178
R-squared	0.963	0.981	0.976	0.989	0.945	0.978
Country-Pair	Yes	Yes	Yes	Yes	Yes	Yes
Dest.-Year	Yes	Yes	Yes	Yes	Yes	Yes
Origin-Year	Yes	Yes	Yes	Yes	Yes	Yes
Model	OLS	OLS	OLS	OLS	OLS	OLS

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