International Migration and Cultural Convergence *

Hillel Rapoport[†]

Sulin Sardoschau[‡]

Arthur Silve[§]

This draft version: February 28, 2017

Abstract

Does international migration contribute to cultural convergence or divergence? We investigate this question both theoretically and empirically. At a theory level, we 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. We then extend this model to allow for horizontal and vertical cultural transmission following Bisin and Verdier (2001). 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. Indeed, as predicted by the extended model with cultural transmission, bilateral convergence appears even if we exclude migrants from the pool of respondents in both countries (hence eliminating social mixing), and more likely in the long-run, especially after controlling for economic incentives to migrate and for countries that are initially culturally distant. Interestingly, international migration appears as a stronger and more robust driver of cultural convergence than trade. The results are robust to using a large set of time-varying cultural distance measures along different statistical and topical dimensions which we constructed for this study, and to introducing bilateral fixed effects.

*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; and the Cultural Diversity Workshop, University Paris 1 Pantheon-Sorbonne, May 2016

[†]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

The era of globalization with its technological advancements has facilitated and accelerated the intensity of interactions between countries worldwide. This includes the trade in goods and services, the flow of capital, and the movement of people. These cross-country interactions are not purely economical but they also initiate a cultural exchange between countries. There are various examples illustrating how economic and cultural exchange foster and 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 IGM employees in 30 different countries); the United States exporting goods in fashion, entertainment, or technology to many countries in the world spreading the "American Lifestyle"; and migrant communities that bring services and traditions to their host communities, such as their respective cuisines and national holidays. In fact, every economic interaction also entails a cultural component and culture plays a major role in the creation of institutions, the economic integration into world markets, and the wealth of nations more generally. The importance of culture for economic performance has long been recognized by historians and economic historians (Braudel, 1987; Landes, 1998; Greif, 1989, 1994). And yet, it is only relatively recently that culture has become a core research question in economics (Guiso et al., 2006; Fernandez, 2011), especially in the fields of comparative political economy and development (Alesina and Fuchs-Schündeln, 2007; Aghion et al., 2010; Spolaore and Wacziarg, 2009, 2013).

There is no single definition of culture that the economic literature has agreed upon and there are empirical, theoretical, and psychological approaches to its definition (see Alesina and Giuliano (2015) for an overview of definitions). In our context, by culture we mean a set of attitudes, preferences, beliefs and values that govern individual behavior and thereby determine aggregate social, political and economic outcomes. 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. As noted above, a growing body of literature in economics has recognized the importance of culture for economic interactions. However, economists have long assumed 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. In particular, among the many potential determinants of culture, we focus on international migration and investigate its role in shaping preferences and beliefs, disentangling the different mechanisms at play and exploring the link between the movement of people and changes in attitudes.

Migration is a key component in the analysis of economic outcomes, reinforcing and strengthening many economic dimensions of globalization such as trade, investments or technology diffusion[2]. While a lot of research has been devoted to the economic impact of migration in a context of globalization, very little research is dedicated to the cultural sway of migration. This is particularly surprising because we think of values and attitudes to be embedded in people and therefore carried with and spread through them.

[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". This definition mirrors the idea among economists that culture is "fairly unchanged" and thus empirically and theoretically often treated as static and exogeneous

[2]see (Docquier and Rapoport, 2012) for an overview of this literature, and (Bahar and Rapoport, 2016) on migration and dynamic comparative advantage)

In contrast to migration, some research efforts has been devoted to the two-way causal effect between culture and trade. In one of the earlier explorations of the effects of cultural ties on bilateral trade, Rauch and 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 and 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). Theoretical approaches to trade-based cultural convergence have been made by (Olivier, Thoenig, and Verdier, 2008) and (Maystre, Olivier, Thoenig, and Verdier, 2014), where the cultural transmission mechanism is based on the notion of imperfect altruism as put forth by Bisin and Verdier (2000). In their theoretical model, Olivier, Thoenig, and Verdier (2008) show that goods market integration causes a phenomenon of cultural divergence across countries. The authors assume that there are two goods X and Y that each embody a "cultural type" and can be purchased and consumed (exclusively, individuals either pick X or Y) at a certain market price. When two countries decide to engage in trade, the distributions of cultures (as captured by the share of people consuming X and Y respectively) become more dissimilar across countries due to a change in the relative price of goods X and Y. The model predictions go as far as saying that cultures that existed under autarky ultimately disappear when countries open up to trade. Maystre, Olivier, Thoenig, and Verdier (2014), change the model assumptions on the production side (moving from an endowment economy to a standard monopolistic trade model) and on the consumption side (now both cultural goods X and Y can be consumed in different intensities). There is one idiosyncratic good that is only available in the home country and one global good that can be traded and consumed in both countries. Based on this model setting, the authors show that opening up to trade and the resulting market size effect (Krugman, 1979) will lead to both countries consuming more of the global good and less of the idiosyncratic good. They conclude that trade liberalization leads to cultural convergence.

These two theoretical works illustrate that depending on the assumptions about the production function and the demand function, the model produces two different outcomes, one leading to cultural divergence while the other leads to cultural convergence. This encapsulates the intuition that economic interactions between two countries can both lead to a widening or closing of the "cultural gap", depending on certain conditions. In fact, this ambivalent intuition also applies to migration: 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. 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*[3].

However, we do not consider migration to be a cultural zero sum game since the diaspora still has an impact on 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 from their host to their home communities. For instance, Fargues (2006) shows in the case of Morocco, Egypt, and Turkey that depending on the emigration to a high fertility (Gulf) versus low fertility (West) destination country, birth-rates in their respective

[3]A similar concept has been introduced by Hirschman (1981), which he calls *exit-effect* and refers to self-selection along political attitudes. In particular, he finds that migrants are typically positively selected in terms of support for democracy.

households of their home countries are affected. Exposure to new norms in the respective destination shape the preferences of the migrant and are 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. Kandel and Massey (2002) analyze the culture of migration in Mexico and describe that young Mexicans consider migrants as "markers of [...] manlihood". Further anecdotal evidence confirms that migrants significantly impact their home country's culture. For instance, in their 2013 Country Migration Report the Philippine government constitutes 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. We call the notion of migrants remaining a vital part in the cultural development of their home communities the *Cultural Diffusion-Effect*.

In this paper, we try to shed light on the interaction between bilateral migration and cultural proximity. With our theoretical model, we aim to capture - in a unified model - that migration can either lead to cultural convergence or divergence, identifying the conditions under which we would expect one of the two to dominate. We start with a compositional model of international migration, where there are two costs/benefits associated with the emigration decision: a cultural and an economic component. Depending on the weight between those two components, we can show that there is cultural convergence (through social mixing) or cultural divergence (through cultural selection). In a second step and following Bisin and Verdier (2000), we introduce a dynamic dimension to the model by allowing for intergenerational cultural transmission with migrants remaining a part of the cultural development in their home country. We incorporate the notion of migrants as potential role models into the dynamics of cultural transmission and show that international migration can change the cultural composition in the home community and can lead to cultural convergence through cultural diffusion.

We will complement our unified theoretical approach with a rich empirical analysis, where we will focus, in particular, 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, too. Many of the previous work on cultural proximity has developed static measures to capture how culturally close or distant two social groups are (De Santis et al., 2016; Spolaore et al., 2014). However, these distances change over time and disregarding the dynamic component of culture can substantially limit our understanding of its determinants. This is why with our empirical analysis, we will develop a set of time-varying bilateral cultural distance measures, where we make sure that compositional effects do not influence our outcomes (change in the number or types of questions used in the World Value Survey) and to test the robustness of our results with several statistical distance measures drawn from computer sciences. The rigor in choosing a coherent and comparable set of questions and measures over time 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 and shows for over six thousand country pairs between 1970 and 2010 that cultural diffusion over-weighs and that there is 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 paper is organized as follows: In order to structure ideas behind the two forces that may ultimately lead to bilateral cultural convergence and divergence, in section 2 we present 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 a compositional model of migration and cultural change, where we set aside for a moment the dynamic dimension of cultural change and focus on the purely on how the exit of parts of the population influence the cultural composition in the home community. 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). In this compositional model we assume that there are two components to the migration decision: a cultural and an economic component. Depending on the balance between the two migration motives, we show that there is cultural convergence or divergence.In a second step, we introduce in a Bisin and Verdier (2000) setting the dynamic component of cultural transmission, where we acknowledge the exceptional role of migrants in the cultural formation of the home community.

2.1 A compositional model of migration and cultural change

The economy is composed of 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). The relative size of country B is n. Individuals in both country 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. They display "homophily", meaning that their utility increases in the share of same-type individuals in the country where they live.

We assume that individuals from country A who contemplate migrating have hetereogeneous expectations of economic gain / cost of migration. Let us write the typical gain of an individual g. At the start of the stage game, g is distributed in the population according to a cdf $\mathcal{G}$ with support on $\mathbb{R}$. Assuming quasi-linear preferences, the pool of type-i country A nationals who wish to migrate is composed of anyone such that

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

with f the function by which cultural preferences are translated in utility units. To account for homophily, we assume f to be an increasing function.

When $\beta = 0$, only homophily matters in the decision to migrate: potential migrants see a higher share of same-type individuals in the destination country. In that case, only one type of individuals is likely to move, which progressively emphasizes the cultural motivation to migrate other same-type individuals in the home country. Eventually, all same-type individuals migrate, until the home country is homogeneously composed of other-type

individuals. 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. Cultural types and preferences are irrelevant in the migration decision: the pool of potential migrants reflects the cultural composition of the home country.

We model the migration process as follows. At the start of the period, individuals discover their 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. The game reaches its equilibrium when the pool of potential migrants is empty. 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 $\mathcal{G}_j \equiv \mathcal{G}((1-\beta)(f(1-q_i^A) - f(1-q_i^B))/\beta)$. Let us also take i as the generic cultural type. Notations are simplified accordingly. 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 equilibrium is reached (and the dynamics stop) when $\mathcal{G}_i = \mathcal{G}_j = 1$. The system thus described displays intuitive features. Let us first examine separately the two motives for migration: cultural homophily and the prospect of economic gain.

If $\beta = 0$, individuals only rely on cultural homophily when deciding to migrate. In this case, $\mathcal{G}_i = 1 - \mathcal{G}_j$: the pool of potential migrants is composed of all individuals of the relatively scarce type, who wish to move to country B, where their type is relatively abundant. As some of them 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 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 composition of the pool of potential migrants reflects the cultural composition of country A. If initially $q^A > q^B$, then q^B increases as migrants start arriving, and reciprocally. The diffusion 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. To simplify the discussion, we may assume that type-i individuals are more frequent in country B than in country A, so that $q^A < q^B$, $\mathcal{G}_i < \mathcal{G}_j$ 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 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 include 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. The share of type-i individuals decreases in both countries. This suggests the possibility of cultural convergence between countries A and B when β is high enough. Formally, we define divergence as $(q^A - q^B)(\dot{q}^A - \dot{q}^B) > 0$, and convergence otherwise.

2.2 Intergenerational Cultural Transmission with Migrants as Role Models

The standard model of cultural transmission as put forth by Bisin and Verdier (2000) suggests that 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. In a one-country framework, Bisin and Verdier (2000) 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, according to a costly socialization technology. 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 reduced-form cultural equilibrium q^* of the two cultural types in an economy. In a two country setting with a simple migration exit model (that is emigrants have no impact on the intergenerational cultural transmission in their home country), in equilibrium the host country should converge back to q^{B*}. Identically, the population of the home country which did not migrate should converge back to q_{A*} (unless migration depleted completely country A of one type). Overall, the simplistic model would predict that migration does not alter the cultural equilibrium in both the host and the home country.

Let us now adapt Bisin and Verdier's framework to a two-country model, with a twist. We assume that migrants remain potential role models in their country of origin. For simplicity, we assume for now that migrants and host populations do not blend, and we focus our interest on the home country. We keep notations from the compositional model: $q \equiv q^A$ is 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 migrant 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

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

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, and a type-i individual provides effort $\tau_i = (1 - \chi)\Delta V_i$. An identical reasoning yields that a type-j individual provides effort $\tau_j = \chi \Delta V_j$. 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^*:

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

Thanks to the implicit function theorem, we know that q^* can be viewed as a function of η and π. We show in the Appendix that $\frac{dq^*}{d\eta} > 0$. 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 the compositional model, convergence required the economic motive to dominate strongly. With this new model, we greatly expand the range of parameters where we expect convergence. Furthermore, we can also show that $\frac{dq^*}{d\pi} > 0$. This somewhat 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.

2.3 Empirical Predictions

The model introduces two motivations for migration. Based on these two motivations, the compositional model (CM) is the most intuitive and fundamental mechanism that links migration to cultural composition. That model suggests a set of strong empirical predictions. If confirmed by the data, these predictions would help us interpret cultural convergence. In the compositional model, convergence requires a clear domination of the economic over the cultural motive for migration. The transmission model (TM), also quite intuitive, suggests other empirical predictions, which sometimes contradict the predictions of the CM. The data allows us to discriminate between these two competing mechanisms for their explanatory power, with a different interpretation of convergence. In TM, convergence is a more likely scenario, and it increases with the importance of the cultural motive. To discriminate between CM and TM, let us clarify the predictions associated with each one.

CM does not envisage any change of the cultural composition of the migrant and host populations (excluding migrants), only of the home population and of the host country (including migrants). With a stable host population and a scarce type becoming scarcer at home, we obtain a first prediction:

CM prediction 1: home and host populations should diverge culturally.

CM also suggests that the dynamics that underlie convergence imply that $\dot{q}^B < \dot{q}^A < 0$. 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. With these notations, convergence is equivalent to $D(A', B') < D(A, B)$. This provides a second prediction:

CM prediction 2: Conditional on convergence between countries A and B, then we must have $D(A, B') < D(A', B') < D(A, B) < D(A', B)$.

Also consider the consequences of an economic shock, which reduces uniformly the economic benefit from migration by a fixed amount c relative to the baseline. Again, to fix ideas, we assume $q^A < q^B$. A type-i individual wishes to emigrate iff $\beta(g-c) \geq (1-\beta)(f(q^A) - f(q^B))$. The fraction of individuals who wish to emigrate is now given by $\mathcal{G}_i(q^A, q^B, c) = \mathcal{G}\left(c + \frac{1-\beta}{\beta}\left(f(q_i^A) - f(q_i^B)\right)\right)$ among type-i individuals and $\mathcal{G}_i(q^A, q^B, c) = \mathcal{G}\left(c + \frac{1-\beta}{\beta}\left(f(1-q_i^A) - f(1-q_i^B)\right)\right)$ among type-j ones. Whatever $\beta < 1$, c and $\mathcal{G}$, we still have $\mathcal{G}_i < \mathcal{G}_j$ and $\pi > q^A$.

For reasonable distribution functions $\mathcal{G}$, the ratio $\mathcal{G}_i/\mathcal{G}_j$ increases with c, and so does π. While there are fewer candidates to migration in both subpopulations, the cultural selection effect takes more importance relative to the economic motive.

CM prediction 3: A sufficiently large economic shock that reduces uniformly the economic benefit from migration should lead to cultural divergence between A and B.

Conversely, TM does consider the evolution of the cultural composition of the home, the host and the migrant populations. While we have focused on the role of migrants as role models in their origin country, the conclusions of the model are robust to alternative specifications. If anything, we might anticipate that a migrant offspring is more likely to be influenced by a role model from the host population 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. In any case, we believe that the most relevant channel of cultural transmission is the one described by popular wisdom as the "Uncle in America", the successful migrant who returns home glorified by his overseas success. TM predicts convergence between the home and the host populations, even when excluding the migrants. That form of convergence cannot be a compositional effect. It would correspond to a cultural diffusion mechanism, where the relative economic success of migrants advertises their cultural type in their home country.

TM prediction 1: home and host populations should converge culturally.

Contrary to CM, convergence is the only equilibrium outcome of TM. Our prediction on the cultural distances is therefore unconditional, and becomes

TM prediction 2 (weak): In equilibrium, $D(A', B') < D(A', B) < D(A, B)$.

Taking into account our intuition on the suggested alternative specifications, this prediction can be made stronger:

TM prediction 2' (strong): In equilibrium, $D(A', B') < D(A', B) < D(A, B') < D(A, B)$

CM and TM yield opposite assumptions on the motive for migration which explains convergence. In the compositional model, the prospect for economic gain led to a cultural melting pot, and therefore to convergence. In the transmission model, homophily led to the scarce type gaining in influence in equilibrium. As a result:

TM prediction 3: A sufficiently large economic shock that reduces uniformly the economic benefit from migration should increase cultural convergence between A and B.

We consider the two components of the model, the compositional and transmission mechanisms, to unravel at different speeds in the data. While the compositional model hints at immediate effects of emigration, e.g. the mechanics of migration, we consider the transmission to reflect the cultural technology of migration to materialize after longer periods. We will show this more formally in future versions of this paper but we will already incorporate this intuition into our empirical analysis.

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 World Value Survey, 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. Where WVS and ESS questions are identical we are able to expand the number of observations 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 World Bank's and the OECD's joint Extended Bilateral Migration Database that covers migration flows for each decade between 1960 and 2010. In addition, we make us 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 World Value Survey 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 make our coefficients better inter-

pretable, 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. These cultural traits cannot be subsumed under "Western" values and be geographically or economically divided. For instance, 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. Looking at the response to this question (see Figure 1), Japan and Germany seem to be much more similar than Germany and Spain in terms of educational values [4]. 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). In the context of intergenerational cultural transmission, we consider this set of questions to be the one most closely related to our model. The parent derives direct utility out of the socialization of their child but has to exhibit an effort to do so. This effort is less costly in an environment where the share of people with the same value is higher. In other words, if a parent wants his or her child to have religious faith (thrift, obedience or any other value), he or she will have an easier time transmitting this value to the child in a religious country.

Figure 2 presents a principal component analysis for all of the 11 dimensions in the transmission of values to children for the most recent WVS wave of 2010 to 2014. The graph illustrates which countries are the most similar, along two dimensions that present linear combinations of the average share of respondents choosing one of the eleven options. Expectedly, some Northern European countries (such as Sweden, the Netherlands, and Germany) and several Middle Eastern countries (Kuwait, Iraq, Qatar, Libya, Palestine, Egypt) seem to share similar educational values among each other but the cultural proximity between Taiwan and Japan to the rest of Northern Europe may come as a surprise and helps to shed new light on the perception of cultural distance across countries.

Another dimension of culture is how individuals set priorities and evaluate the importance of different aspects of their lives. One WVS question asks their respondents how important (very important, rather important, not very important, not at all important) family, friends, leisure, work, politics or religion are for their lives. We use this question in addition to the "values to children" question to highlight the difference of cultural values in the intensive margin. The comparison between Figure 2 and 3 shows that most of the countries that are defined to be culturally close for the values to children, also think similarly about priorities in life. However, there are still substantial difference in distance between country pairs. For instance, while Sweden and the United States are closer in ideas about priorities in life,

[4] an overwhelming majority in Japan and Germany want to pass on thrift and obedience, whereas a large share of the Spanish would like their children to be obedient

there seem to be larger differences in which values they would like to transmit to their children.

Contrary to attitudes about religion and family, trust is one of the beliefs that can change quite dramatically over a short period of time (Putnam, 2001). We make use of the WVS on generalized trust asking *Do you think people can generally be trusted or you cannot be too careful?*, which has been widely applied to questions about social capital (Knack and Keefer, 1997; Alesina and La Ferrara, 2002; Portes, 2014). With this we intend to capture changes in norms and beliefs that are highly dependent on the environment and on societal reciprocity. They differ greatly from personal belief, like religiosity for instance, and may ultimately be more prone to the emigration of the "trusting and trustworthy" from the home country (selection effect dominates). Additionally, our analysis will include views on gender equality and the feeling of control over one's life (a comprehensive summary of the WVS questions used is given in Table 2).

Different from Maystre et al. (2014), Desmet et al. (2015) or De Santis et al. (2016) (who include a vast set of questions with highly imbalanced coverages in terms of WVS waves or set of countries), the cultural dimensions used for the analysis of cultural proximity obey to one important criterion: availability of the World Value Survey question 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. This is particularly important in our analysis as compared to cited studies, since we put emphasis on cultural convergence rather than deriving a static measure of cultural proximity. In other words, 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 of utmost importance. Naturally, the requirement to complete coverage will limit the scope of questions that we can include in our cultural distance measure. 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 the underlying probability density functions of the national distribution 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 distance, which is written as $M = \sqrt[p]{\sum_{i=1}^{d} |P_i - Q_i|^p}$. For instance, the Euclidean, the Chebyshev, and the Canberra distance are subsumed under Minkowski's generalization but while the Canberra distance normalizes the individual distances to the absolute sum of observation i in both distributions P and Q (relative versus absolute distance), the Euclidean distance applies a simple Pythagorean metric, following the idea that the shortest distance between two points is a straight line. In practice, this means that the Canberra distance gives much less weight to outliers in the higher values and more weight to observations close to the origin. Conversely, the Chebychev distance does give all its weight to these extreme values, supporting the idea that the accurate measure for the difference between probability density functions is the distance

between the two points which are the furthest away from each other. Applying this concept to cultural differences implies that country pairs that may agree on some cultural dimensions but disagree strongly on others are considered to be more dissimilar than country pairs that slightly disagree on all cultural dimensions.

Conceptual differences become even more apparent across rather than within statistical "families"[5]. The approach of Pythagorean distances is somewhat altered when considering 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). The 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

[5]See Cha (2007) for a comprehensive overview and categorization of different distance/entropy/divergence measures

all countries into account.

In this paper, we are agnostic towards the choice of the best distance measure. However, we do emphasize that a exclusive choice of statistical distance measure is associated with a choice in statistical inference that needs to be carefully interpreted. Therefore, 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 spite of the use and careful interpretation of different statistical measures, a discussion of the underlying assumptions that are common to all of these measures in necessary. In fact, all of the measures do not take into account the relative distance between the response possibilities of each question. Some questions (for instance for priorities in life) allow for different degrees in their answers (ordinal response) such as *very important, rather important, rather not important, and not important at all*. Treating the responses as binary does not allow us to make any claims about the relative distance between people that choose *very important* and *important* as compared to the distance between *very important* and *not important at all*. The underlying assumption here is that people that pick the same response are culturally similar and equally dissimilar from individuals that choose any other answer. Any other approach would require to include scores for each answer, imposing an equally strong assumption on relative distances (is the switch from very important to important the same as switching from important to rather not important?).

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) 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 will first show the raw conditional correlation between migration and cultural similarity, successively 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 compositional effects. The coefficient β_1 is an aggregate effect of both selection and diffusion, that is the sign of the coefficient will tell us which of the two effects dominates. We are not able to entirely

separate the two effects in the data but the predictions of the model will lead our comparative statics analysis. In the section on heterogeneous effects, we will approximate the parameters of our model that will either favor or discourage one of the two effects.

A first glance at the distribution of preferences over time (see Figure 6) shows a tightening of the distribution in the preferences over priorities in life for the same set of countries over a period of 20 years. This is a first hint towards a convergence of values across countries. Table 3 presents the results of the OLS regression 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. This result holds even after introducing fixed effects for country pairs, destination-time and origin-time fixed effects. Introducing the fixed effects changes the interpretation of the coefficient, as we now look at within country pair changes over time: 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. However, 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. Depending on the measure and inclusion of fixed effects, the coefficient of trade switches sign and looses statistical significance when looking at within country pair changes over time.

The bottom half of Table 3 reports the results when lagging migration and trade by 5 additional years increasing the overall lag to 10 years[6]. This accounts for possible migration-based effects on cultural convergence that only materialize over a longer period. While we expect the selection effect to be more dominant right after the migration decision, the true effect of diffusion may come in at a much later stage when the migrant had time to learn the language, make himself more familiar with the local culture and eventually transmit these values to his home community. The results in columns 3,6, and 9 confirm this prior. The coefficients of bilateral migration are two to three times higher (and more significant) than in the regressions with a shorter lag between migration and the cultural outcomes as observed through the World Value Survey. We repeat the analysis for a sub set of countries for which we have observations for four WVS waves, which substantially decreases the number of countries (half the countries fall out of the data set). The coefficients are consistently positive and when significant close to the coefficient size in the large Panel. As we identify off of only 144 country pairs instead of 576 as in Panel A, the power decreases significantly and the coefficients are not significant for the Euclidean and Herfindahl measure once we include all fixed effects.

Figures 7 to 9 illustrate the outcomes of Panel A in Table 3 graphically but 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 with varying statistical significance, indicating that over time country pairs become more similar culturally when also experiencing higher levels of bilateral migration. This result is particularly strong in terms of

[6]The requirement here is that there is an observation for migration two periods prior to the outcome variable available through the WVS, which substantially limits the number of observations. The lack of observations comes with an increase in the standard deviation however this is more than offset by the size of the coefficient such that the results are of higher statistical significance

size and significance for the measure that captures *values that are desirable to inherit to children*, which is the cultural dimension most closely associated with our model of intergenerational cultural transmission. In contrast to migration, the coefficient on trade is less robust to the inclusion of fixed effects. Similar to our results in Table 3, we are not able to make strong claims on the effects of trade on cultural convergence[7].

In order to account for potential compositional effect in the sample of countries in the WVS over time, we construct a balanced panel data set. 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 results are presented in Table 4 and confirm our results from the OLS regression with the same sign and stronger statistical significance. However, the size of the coefficient is considerably bigger which suggests that compositional effects in the full sample downward bias the true overall effect, underestimating the diffusion effect. In both Tables 3 and 4 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.

In Table 5, 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 incorporating different domains of cultural heterogeneity. 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.

4.1 Bilateral Income Differences

In our compositional model we allow for economic and cultural drivers to determine the migration decision. We assume that wealth and cultural preferences are independently distributed, such that if only economic reasons determined the migration decision, we would have a representative cultural sample in the pool of potential emigrants. On the other hand, if there is some weight attributed to cultural drivers of migration (given homophily), there will be some cultural selection in the pool of emigrants thus over representing the scarce cultural type at home.

[7]This slightly different from the findings of Maystre, Olivier, Thoenig, and Verdier (2014), who establish a positive link between cultural proximity, particularly for differentiated goods. The authors 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 results in Table 6 present the relationship between migration and cultural proximity controlling for time varying income differences. The coefficient remains positive and significant[8] and robust in magnitude. The coefficient on the income gap variable is negative and strongly significant even when we control for country-pair as well as home and host country time-specific fixed effects. This indicates that fluctuations in the income differences between countries contributes to the cultural divergence between countries. In other words, if country pairs become more different economically, they will also diverge culturally.

Comparing the coefficients for migration in the top half of Table 3 (without controlling for changes in economic distance) and Table 6 (including the bilateral GDP gap), we observe that those in Table 6 are consistently smaller. This implies that controlling fluctuations in economic distance, which translate into larger or smaller economic incentives for emigration, will reduce the diffusion effect of migration in the short-run. This is in line with our compositional model. Once we hold economic incentives fixed, we would expect the cultural selection effect to become more dominant. However, the coefficient remains positive and significant, meaning that the cultural selection effect does not outweigh the cultural diffusion effect. This is further evidence for the notion of migrants as role models, e.g. our theory of intergenerational cultural transmission including migrants. In fact, when we lag bilateral migration to look at more long-term effects, we find that (contrary to the short-term effects) the coefficient increases from Table 3 to Table 6. Again, this is in line with the dynamic model extension which tells us that larger cultural selection into the pool of emigrants will reinforce the cultural diffusion effect through the role model hypothesis.

4.2 High and Low-Skilled Migration

We dive deeper into the economic and cultural incentives of the emigration by differentiating between types of emigrants. Thus, in this section we will make use of IAB data on bilateral migration by skill level. 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. Therefore, we include an additional 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. 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, will allow us to get to heterogeneous migration incentives within countries and also track these changes over time.

Grogger and Hanson (2011) analyze two dominant features in international labor movements. First, the positive selection of individuals into migration and second, positive sorting of migrants across destinations. The authors develop a theoretical model of income maximization that accounts for both features. Their empirical analysis shows that larger skill-related income difference between source-destination pairs are associated with a more educated pool of migrants. However, this observation only holds for wage differences expressed in levels and not in logs. The authors attribute this result to the fact that log wage differences, which capture cross-country differences in returns to

[8]except for the most demanding specification of the Canberra measure, where the coefficient that was weakly significant and now just misses the 10% significance mark

skill, fail to account for bilateral migration patterns because cross-country differences in returns to skill are dwarfed by cross-country differences in labor productivity. Grogger and Hanson (2011) show that skilled migration is ultimately more economically incentivized as compared to low skilled migration. Consequently, high levels of skilled migrants reflect a more economically incentivized pool of emigrants which would in turn favor cultural convergence.

Table 7 presents the results of the OLS regression by skill level. 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. Consistent with our model predictions, most specifications confirm a negative relationship between low skilled migration on cultural similarity and an overall positive effect of high skilled migration and cultural similarity. Those results loose significance when we control for the most demanding set of fixed effects. As mentioned above, the data restrictions leave us with few observations by country pair over time which limits us in terms of statistical power. However, once we look at the variation within country pairs (controlling for country-pair, destination, origin, and time fixed effects but without time specific shocks to home and host country) in columns 2,5, and 8 we find the expected negative coefficient for low skilled migrants and the positive sign for the more educated migrants.

4.3 Excluding Migrants

Our compositional model shows that economically selected migrants will mechanically lead to cultural through social mixing. Extending this model to intergenerational cultural transmission with migrants as role models predicts that when excluding emigrants from cultural composition of the host population, we would still expect cultural diffusion. The World Value Survey does not provide information on respondents' immigrant status in all waves but for some countries in waves 2 (10 countries) and 3 (46 countries) we have information of the birthplace of the respondent. About 5.5% of respondents are born in a country different from the respondent's host country. In Table 8 we replicate our analysis for this subset of countries including and excluding the foreign-born from the construction of the aggregate Euclidean distance measure[9]. Excluding the foreign-born from the analysis does not alter our results, size and significance of the coefficient are virtually identical. We loose power when including country pair fixed effects as we only have 9 countries that are covered in both waves and for which there is information of the respondent's country of origin. 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.

4.4 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. The most direct measure of intensity of interaction between migrants and their home communities are bilateral remittances. Naturally, financial transfers are not equivalent to the transfer of norms or values (so called social remittances). However, households that send remittances back to their home countries can be important drivers of cultural diffusion as they are different from households that emigrate in their entirety and send no remittances. Sending financial support back to their countries of origin implies that these

[9]The results hold also for the Herfindahl and Canberra measure of cultural distance

migrants still have family members or business ties in their home communities. In both cases there is an incentive for increased interaction between diaspora and 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[10], 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 very strong relationship between the size of the diaspora and the volume of remittances that flow back to their home countries. But 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 9 confirms our theoretical prediction that countries with stronger social ties between diaspora and home community tend to converge culturally. Columns 1,3, and 5 show the expected positive and significant sign controlling for the full set of fixed effects and are of a similar magnitude as in the full sample analysis. Additionally, country pairs that have no recorded remittance flows do not seem to experience convergence and the consistently negative (yet statistically insignificant) sign of the coefficients in column 2,4, and 6 indicate that the selection effect over weighs the cultural diffusion effect.

4.5 Initial Cultural Distance

We now turn to one of the more subtle predictions of the simple and extended versions of the model. 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. We show in the model that the difference in q_i^A and q_i^B will have a positive impact on the cultural incentive to emigrate. Consequently, the compositional model would predict that there is more cultural selection and ultimately cultural divergence. Conversely, the cultural transmission model shows that big differences in the autarky cultural equilibrium will lead to a faster convergence of attitudes.Empirically, this translate into the prediction that country pairs that are culturally very different initially will converge at a faster rate that countries that are relatively similar to start with. In other words, very similar countries will experience a slower cultural convergence and may even diverge culturally over time. Figures 10 and 11 show how the average level of bilateral cultural similarity evolves

[10]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

for the set of countries in Panel A and Panel B. We split the samples in two groups, above and below the median level of cultural similarity at the beginning of the panel in 1995 and look at how the average cultural proximity evolves for those groups over time. The graphs show that the level of cultural proximity grows at a higher rate for countries that were initially below the median level of cultural similarity. Moreover, on average country pairs that were similar in 1995 are less similar in 2014, which is suggestive of our model predictions.

We test this hypothesis further by splitting 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 along 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 10. According to our model, we would expect that the coefficient of migration in the right columns (above threshold estimation) becomes smaller and/or turn negative as we move to the right side of the table, that is when we consider more and more country pairs that were among the most similar initially. The regression results for dissimilar countries (similar countries) in columns 1,3, and 5 (columns 2,4, and 6) show that the coefficient for bilateral migration becomes larger (becomes smaller), that is that the diffusion effect largely dominates (selection effect is more prevalent). The turning point for which the countries are "sufficiently similar" to experience cultural divergence (or at least to see a significant jump of the selection effect) is different for each statistical distance measure and lies at least at 75% (this is where the below threshold coefficient becomes larger than the above threshold coefficient for the Herfindahl and Canberra measure). However, our results suggest that for selection to outweigh diffusion, countries need to be culturally extraordinarily similar in the first place. Considering the composition of countries in our panel data set, there are only a few country pairs with that feature (Australia and New Zealand or Ukraine and Russia, for instance).

In a last step, 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 11 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. 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.

5 Conclusion

Preliminary: Does international migration contribute to cultural convergence or divergence? We investigate this question both theoretically and empirically. At a theory level, we 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. We then extend this model to allow for horizontal and vertical cultural transmission following Bisin and Verdier (2001). 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. Indeed, as predicted by the extended model with cultural transmission, bilateral convergence appears even if we exclude migrants from the pool of respondents in both countries (hence eliminating social mixing), and more likely in the long-run, especially after controlling for economic incentives to migrate and for countries that are initially culturally distant. Interestingly, international migration appears as a stronger and more robust driver of cultural convergence than trade. The results are robust to using a large set of time-varying cultural distance measures along different statistical and topical dimensions which we constructed for this study, and to introducing bilateral fixed effects.

References

Philippe Aghion, Yann Algan, Pierre Cahuc, and Andrei Shleifer. Regulation and distrust. *The Quarterly Journal of Economics*, 125(3):1015–1049, 2010. doi: 10.1162/qjec.2010.125.3.1015. URL http://qje.oxfordjournals.org/content/125/3/1015.abstract.

Alberto Alesina and Nicola Fuchs-Schündeln. Good-bye lenin (or not?): The effect of communism on people's preferences. *The American Economic Review*, 97(4):1507–1528, 2007.

Alberto Alesina and Paola Giuliano. Culture and institutions. *Journal of Economic Literature*, 53(4):898–944, 2015.

Alberto Alesina and Eliana La Ferrara. Who trusts others? *Journal of public economics*, 85(2):207–234, 2002.

Dany Bahar and Hillel Rapoport. Migration, knowledge diffusion and the comparative advantage of nations. *Working Paper, Harvard Kennedy School*, 2016.

Alberto Bisin and Thierry Verdier. " beyond the melting pot": Cultural transmission, marriage, and the evolution of ethnic and religious traits. *Quarterly Journal of Economics*, pages 955–988, 2000.

Alberto Bisin and Thierry Verdier. The economics of cultural transmission and the dynamics of preferences. *Journal of Economic theory*, 97(2):298–319, 2001.

Fernand Braudel. Grammaire des civilisations. *Paris, Arthaud-Flammarion*, 1987.

Herbert Brücker, Stella Capuano, and Abdeslam Marfouk. Education, gender and international migration: insights from a panel-dataset 1980-2010. *Retrieved March*, 20:2014, 2013.

Sung-Hyuk Cha. Comprehensive survey on distance/similarity measures between probability density functions. *City*, 1(2):1, 2007.

Gustavoand De Santis, Mauroand Maltagliati, and Silvana Salvini. A measure of the cultural distance between countries. *Social Indicators Research*, 126(3):1065–1087, 2016. ISSN 1573-0921. doi: 10.1007/s11205-015-0932-7. URL http://dx.doi.org/10.1007/s11205-015-0932-7.

Klaus Desmet, Ignacio Ortuño-Ortín, and Romain Wacziarg. Culture, ethnicity and diversity. Technical report, National Bureau of Economic Research, 2015.

Anne-Célia Disdier, Silvio HT Tai, Lionel Fontagné, and Thierry Mayer. Bilateral trade of cultural goods. *Review of World Economics*, 145(4):575–595, 2010.

Frédéric Docquier and Hillel Rapoport. Globalization, brain drain, and development. *Journal of Economic Literature*, 50(3):681–730, 2012.

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

Philippe Fargues. *The demographic benefit of international migration: Hypothesis and application to Middle Eastern and North African contexts*, volume 4050. World Bank Publications, 2006.

Gabriel J. Felbermayr and Farid Toubal. Cultural proximity and trade. *European Economic Review*, 54 (2):279 – 293, 2010. ISSN 0014-2921. doi: http://dx.doi.org/10.1016/j.euroecorev.2009.06.009. URL http://www.sciencedirect.com/science/article/pii/S0014292109000725.

Raquel Fernandez. *Does Culture Matter?*, volume 1. North-Holland, Amsterdam, NL, 2011.

Avner Greif. Reputation and coalitions in medieval trade: evidence on the maghribi traders. *The journal of economic history*, 49.04(04):857–882, 1989.

Avner Greif. Cultural beliefs and the organization of society: A historical and theoretical reflection on collectivist and individualist societies. *Journal of political economy*, 102.5:912–950, 1994.

Jeffrey Grogger and Gordon H. Hanson. Income maximization and the selection and sorting of international migrants. *Journal of Development Economics*, 95(1):42 – 57, 2011. ISSN 0304-3878. doi: http://dx.doi.org/10.1016/j.jdeveco.2010.06.003. URL http://www.sciencedirect.com/science/article/pii/S0304387810000647. Symposium on Globalization and Brain Drain.

Luigi Guiso, Paola Sapienza, and Luigi Zingales. Does culture affect economic outcomes? *The journal of economic perspectives*, 20(2):23–48, 2006.

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

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

William Kandel and Douglas S Massey. The culture of mexican migration: A theoretical and empirical analysis. *Social forces*, 80(3):981–1004, 2002.

Stephen Knack and Philip Keefer. Does social capital have an economic payoff? a cross-country investigation. *The Quarterly journal of economics*, pages 1251–1288, 1997.

Paul R Krugman. Increasing returns, monopolistic competition, and international trade. *Journal of international Economics*, 9(4):469–479, 1979.

David Landes. S.(1999). the wealth and poverty of nations: Why some are so rich and some are so poor, 1998.

Nicolas Maystre, Jacques Olivier, Mathias Thoenig, and Thierry Verdier. Product-based cultural change: Is the village global? *Journal of International Economics*, 92(2):212–230, 2014.

Jacques Olivier, Mathias Thoenig, and Thierry Verdier. Globalization and the dynamics of cultural identity. *Journal of International Economics*, 76(2):356–370, 2008.

Alejandro Portes. Downsides of social capital. *Proceedings of the National Academy of Sciences*, 111(52):18407–18408, 2014.

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

James E. Rauch and Vitor Trindade. Ethnic Chinese Networks In International Trade. *The Review of Economics and Statistics*, 84(1):116–130, February 2002. URL https://ideas.repec.org/a/tpr/restat/v84y2002i1p116-130.html.

Enrico Spolaore and Romain Wacziarg. The diffusion of development. *The Quarterly Journal of Economics*, 124(2):469–529, 2009. doi: 10.1162/qjec.2009.124.2.469. URL http://qje.oxfordjournals.org/content/124/2/469.abstract.

Enrico Spolaore and Romain Wacziarg. How deep are the roots of economic development? *Journal of Economic Literature*, 51(2):325–369, 2013.

Enrico Spolaore, Romain Wacziarg, et al. Fertility and modernity. *Discussion Papers Series*, 779, 2014.

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

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

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

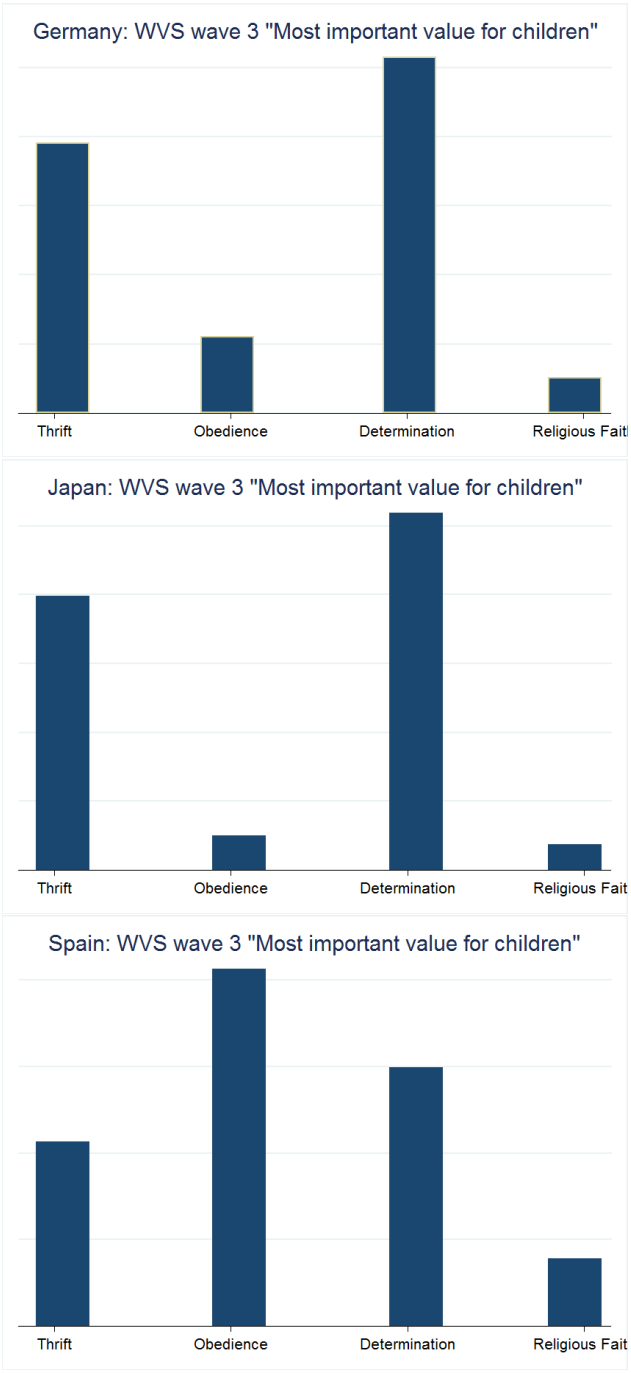


Figure 2: Principal Component Analysis of WVS question on values desired to inherit to children, including independence, hard work, feeling of responsibility, imagination, tolerance and respect for others, thrift saving money and things, determination/perseverance, religious faith, unselfishness, and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

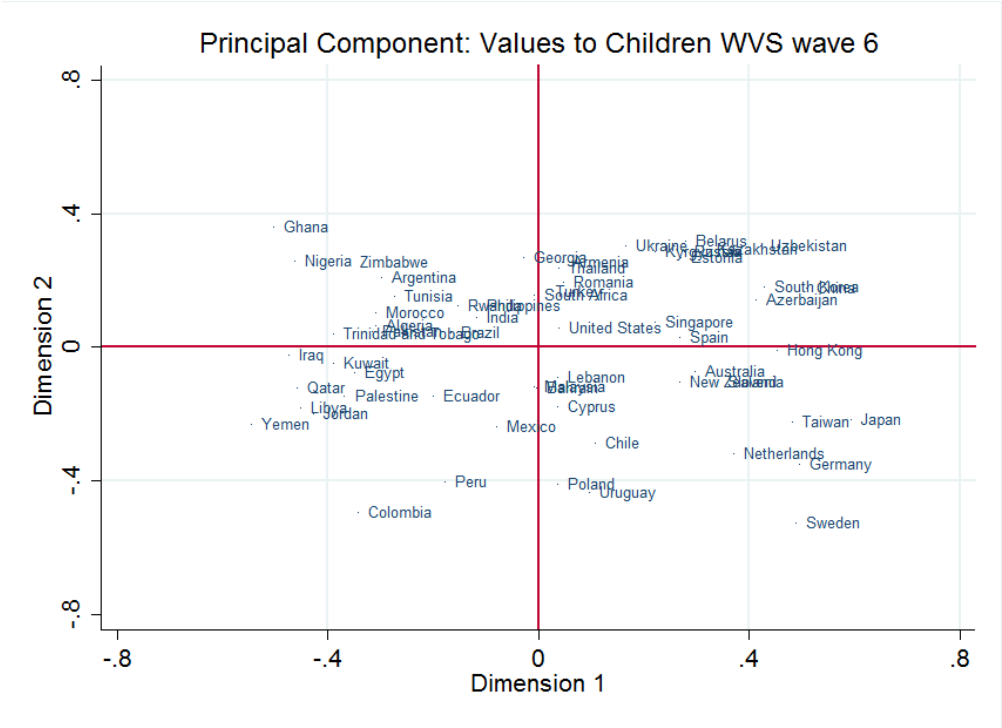


Figure 3: Principal Component Analysis of WVS question on priorities in life,including family, friends, leisure time, politics, work, and religion

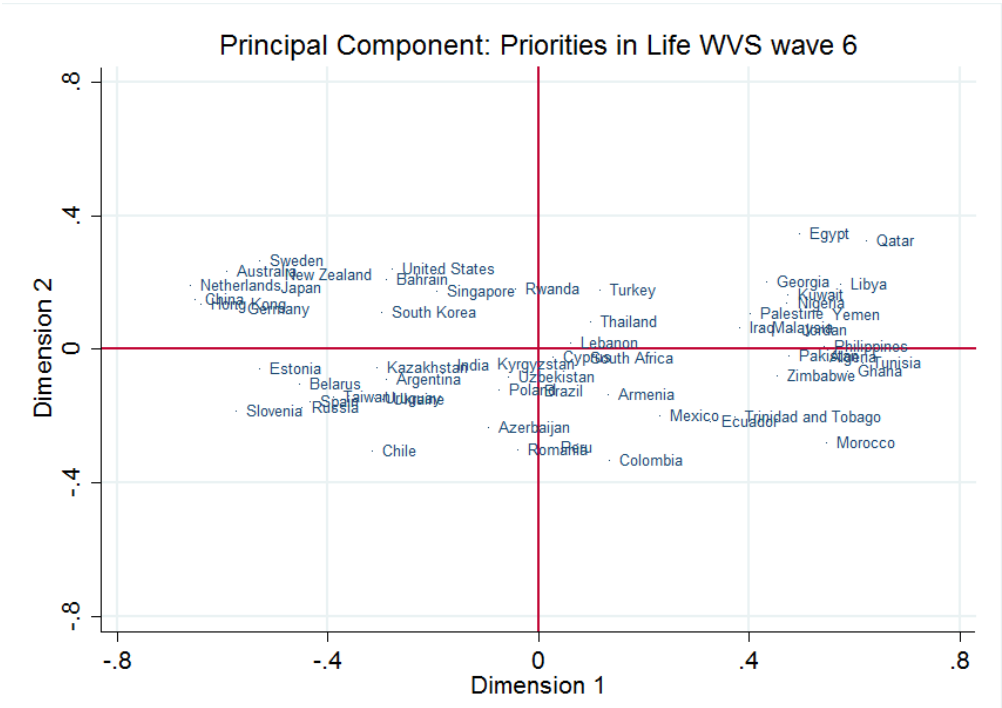


Figure 4: Scatterplot of standardized values for the Herfindahl Index and weighted Euclidean Distance

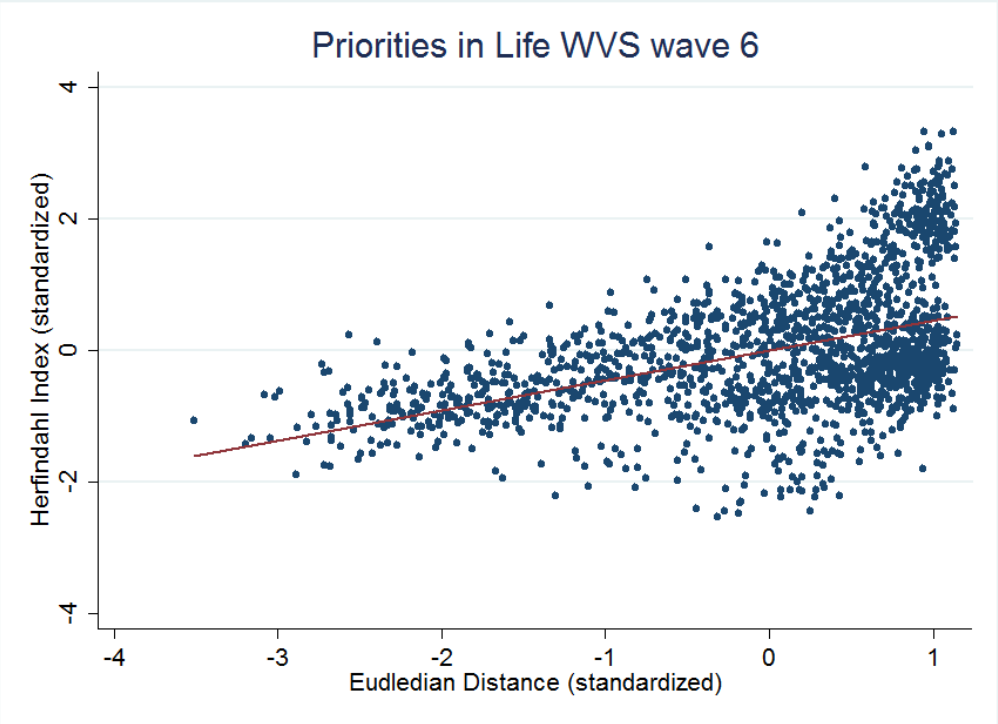


Figure 5: Scatterplot of standardized values for the Herfindahl Index and weighted Euclidean Distance

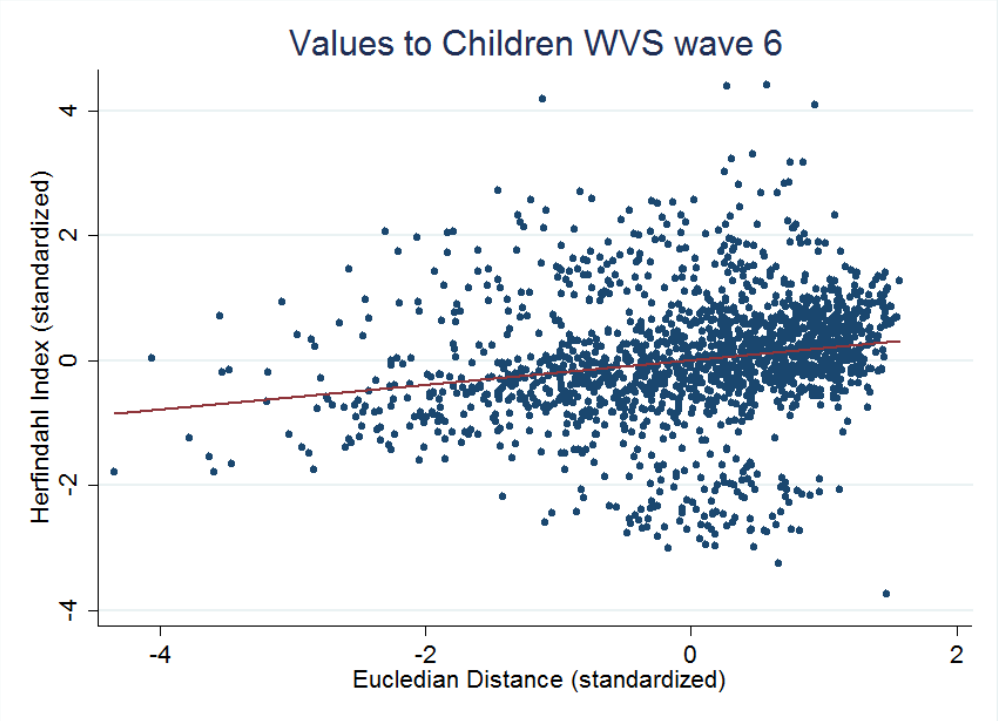


Figure 6: Kernel Density of WVS question on priorities in life over time for 21 countries with observations between 1995 and 2010, showing that the distribution of cultural norms becomes more compact

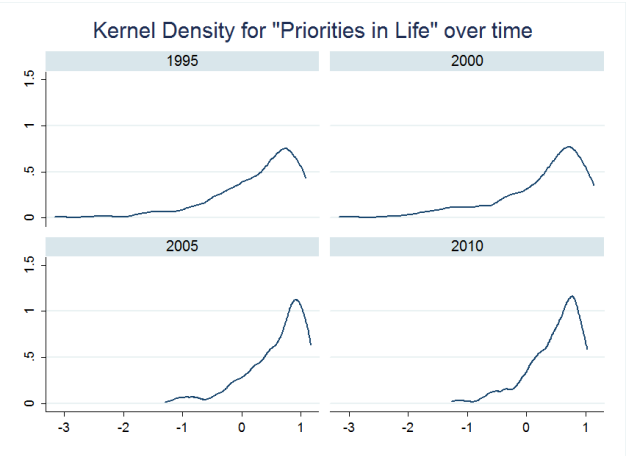


Figure 7: OLS Regression Outcomes for three Specifications with -1*Euclidean Distance

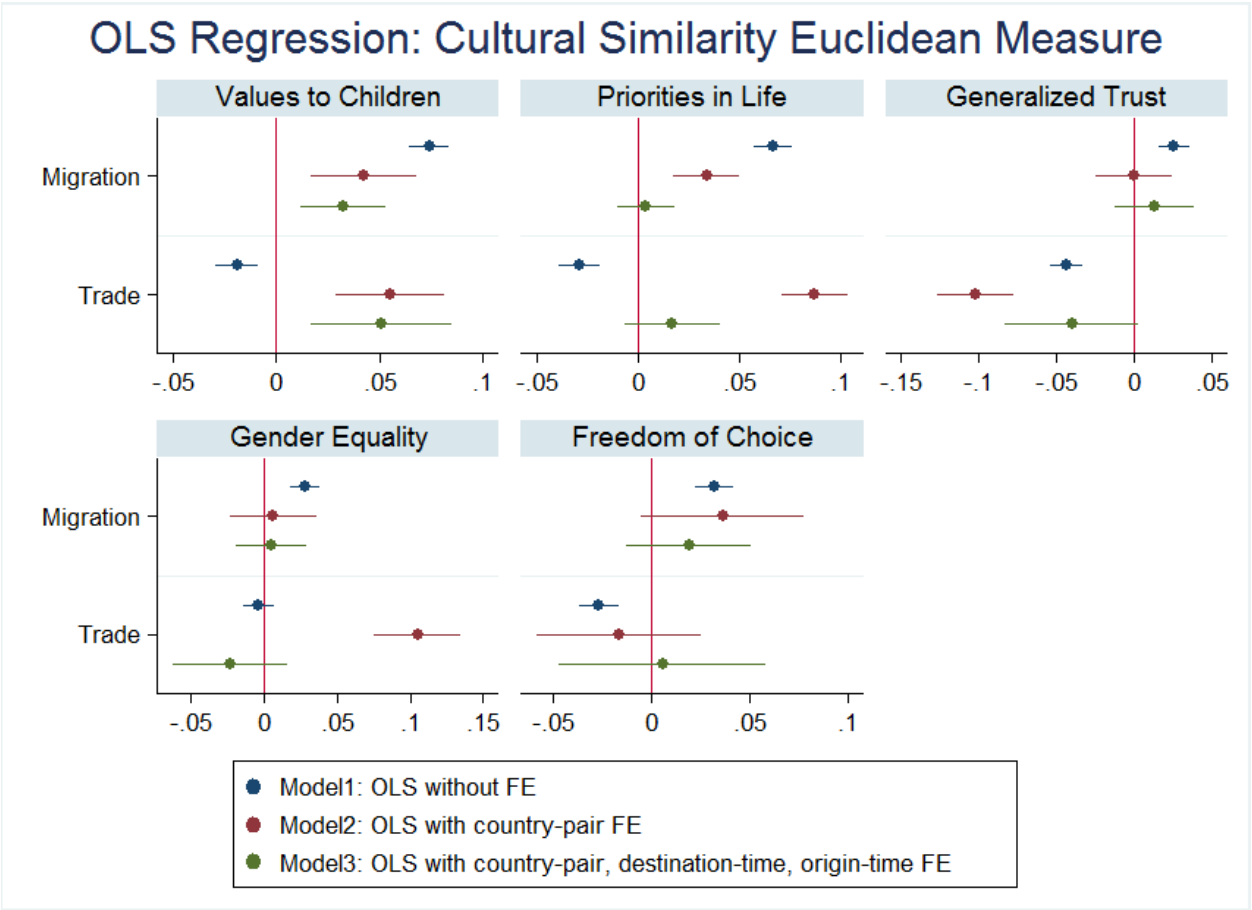


Figure 8: OLS Regression Outcomes for three Specifications with Herfindahl Index

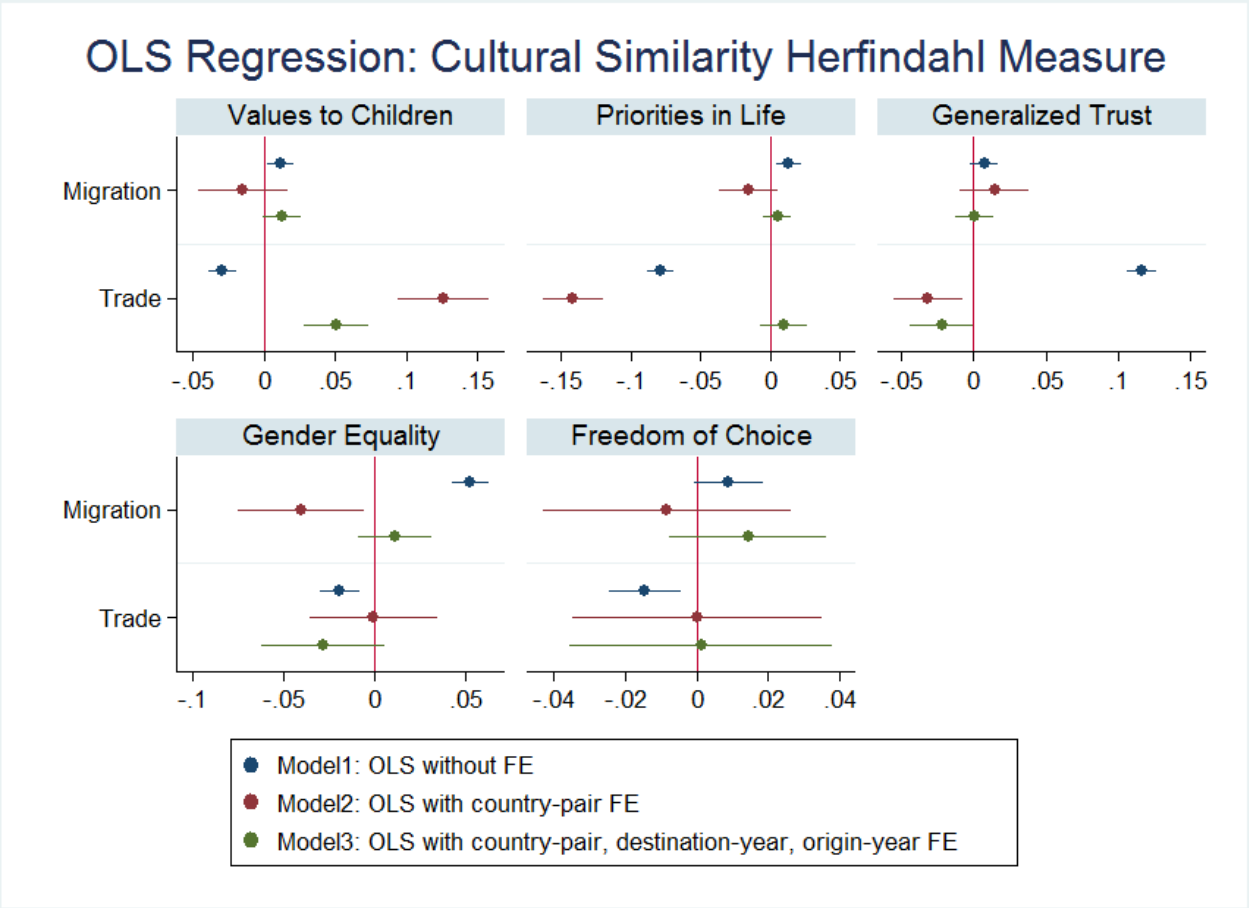


Figure 9: OLS Regression Outcomes for three Specifications with -1*Canberra Index

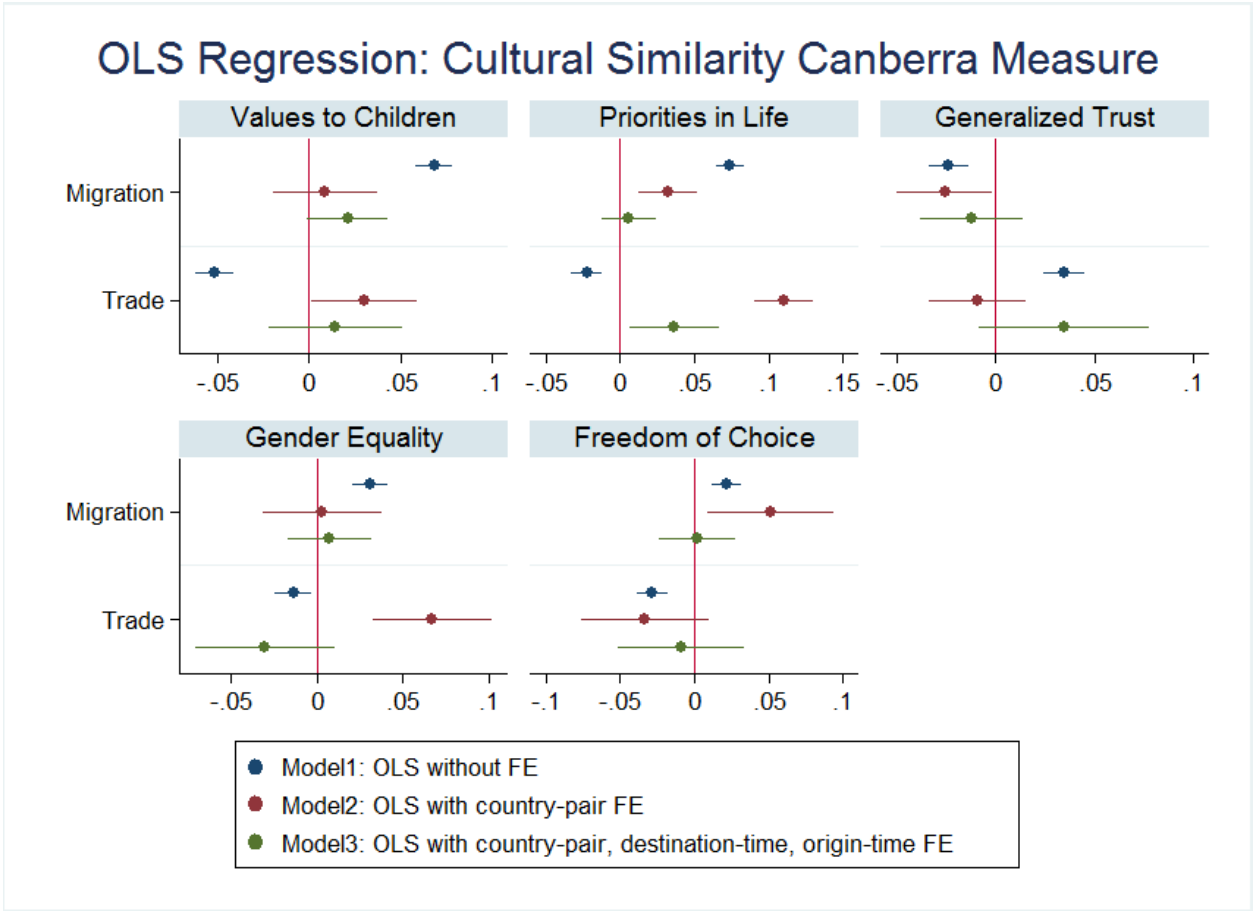


Figure 10: Scatterplot: Log of Remittances and Log of Stock of Migrants in 2010

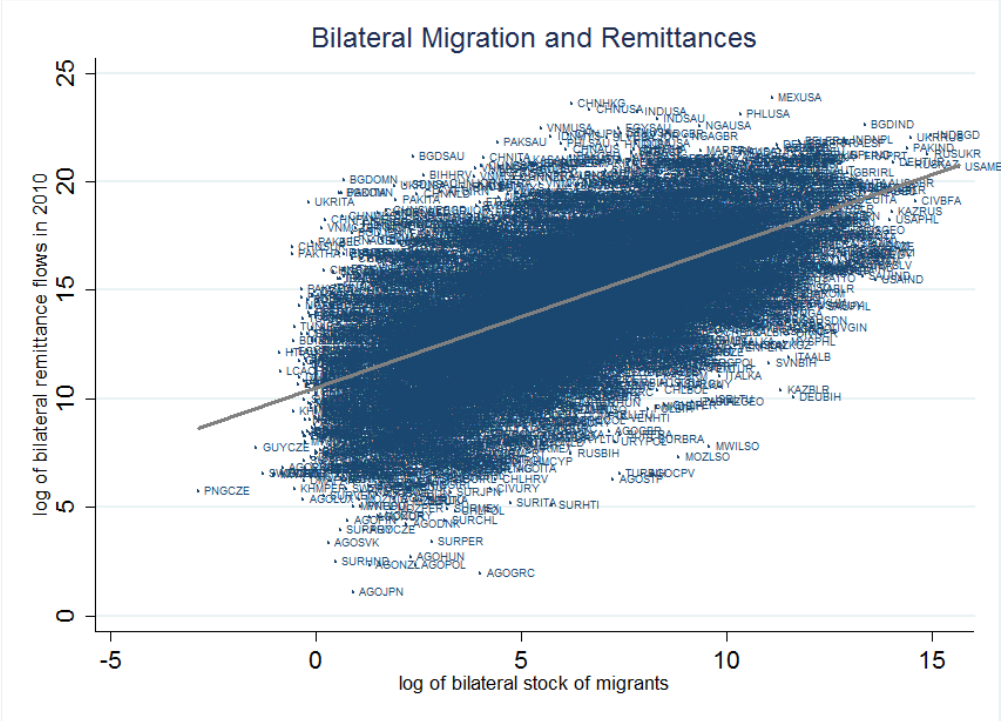


Figure 11: Average Cultural Similarity by year for Panel A (24 countries in 3 waves) split in 2 groups above and below the median Cultural Similarity in the first period

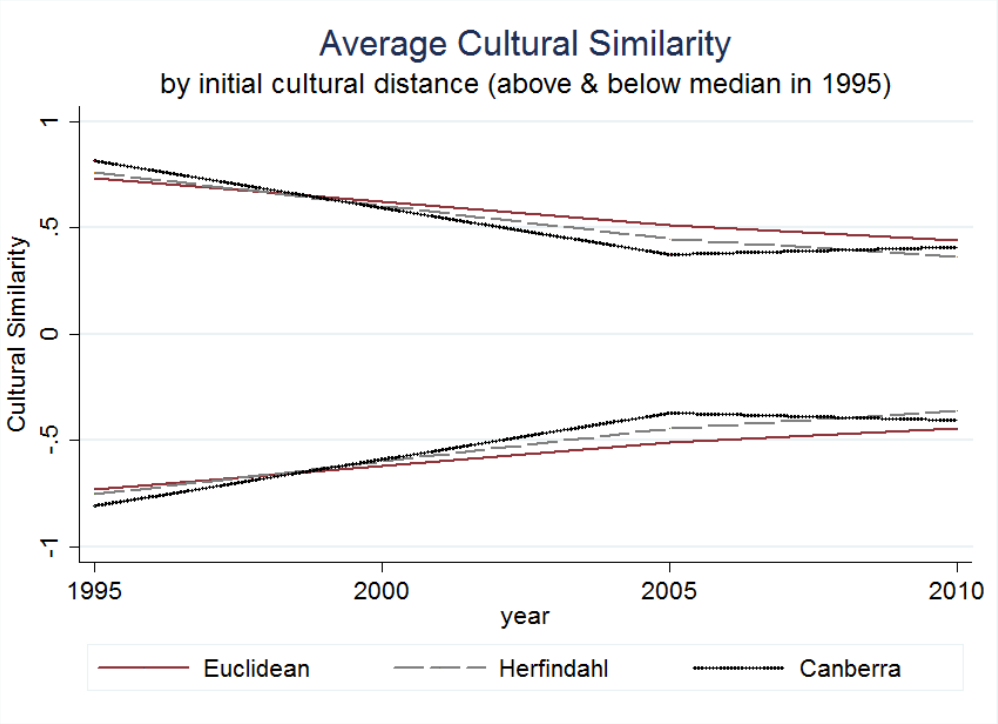
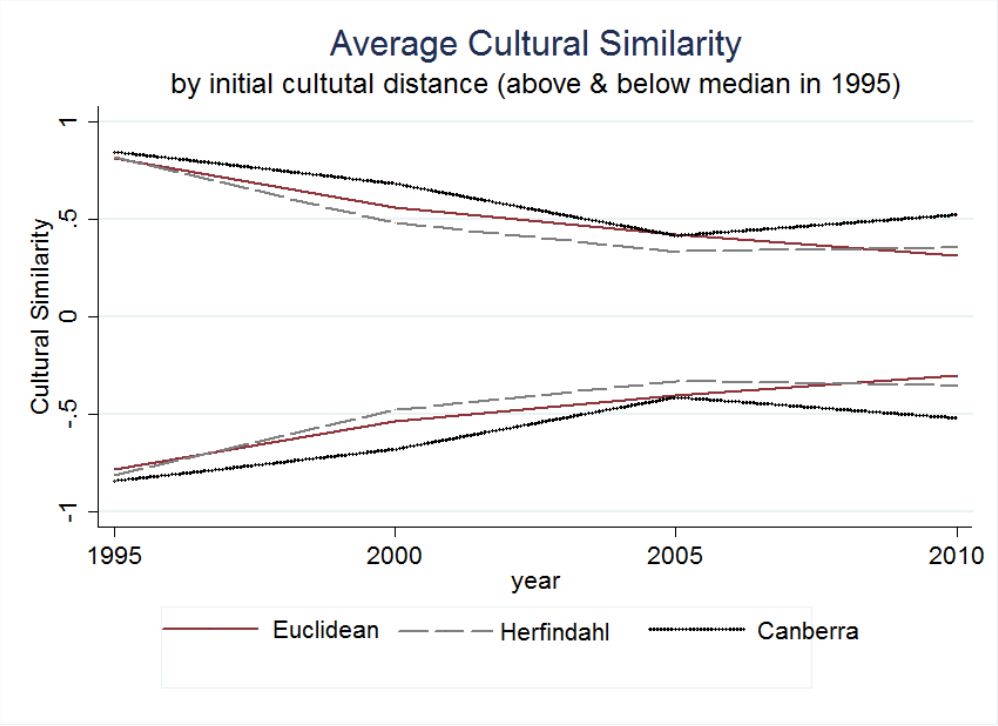


Figure 12: Average Cultural Similarity by year for Panel B (12 countries in 4 waves) split in 2 groups above and below the median Cultural Similarity in the first period



Appendix C Tables

Table 1: 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 2: 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 3: OLS regression of aggregate Cultural Similarity Measures

	Euclidean			**Herfindahl**			**Canberra**		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Lag: $\Delta = 1$									
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: $\Delta = 2$									
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 4: 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 5: 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 6: 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: $\Delta = 1$									
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: $\Delta = 2$									
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 7: 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 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 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,

Table 10: 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 11: Panel Regression by Initial Cultural Similarity

	Euclidean	Herfindahl	Canberra
	(1)	(2)	(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.