Cultural Similarity and Migration*

New evidence from a gravity model of international migration

Tobias Grohmann

European University Institute and GLO.
t.grohmann@rug.nl

This version 26 January 2024.

Abstract

Theory suggests that cultural similarity of countries increases migration flows between them. This paper brings best practices from the trade gravity literature to migration and tests this prediction. Using time-varying and time-invariant similarity variables based on religion, language, genetics, and the World Values Survey, I estimate a theory-consistent gravity model on a panel of international and domestic migration flows (>200 countries, 1990-2019, 5-year intervals). The main results using time-varying similarity variables and implementing a three-way fixed effects structure with origin-year, destination-year, and corridor fixed effects, do not show the hypothesized positive effect of cultural similarity on migration. Instead, religious similarity has a significant negative effect on migration, while WVS-based attitudinal similarities regarding individualism, indulgence, and trust are insignificant. Additional results suggest that cultural selection and sorting can explain these findings, where migrants are attracted by destinations that are culturally similar to their *personal* cultural beliefs rather than the average cultural beliefs of their home country. Results of a two-stage fixed effects (TSFE) procedure and a gravity-specific matching estimator, which both allow theory-consistent estimation of time-invariant similarity variables, such as linguistic and genetic similarity, confirm that the relationship between cultural similarity and migration is more nuanced than previously thought.

Keywords: international migration, culture, gravity model of migration
JEL classification: F22; O15; Z10

* I am grateful to Steven Brakman, Richard Frensch, Hanna Fromell, Olga Griaznova, Johannes Kleinhempel, Tristan Kohl, Vladimir Kozlov, Maite Laméris, Robert Lensink, Milena Nikolova, Panu Poutvaara, Ahmed Skali, Jonathan Wadsworth, Juliette de Wit for helpful comments and suggestions. I also thank participants at the joint IOS/APB/EACES Summer Academy 2023, EPCS 2023, URSI Groningen and the Gender Econ Group in Groningen for insightful discussions. All remaining mistakes are mine.

1 Introduction

Economic theory, which views migration as a human capital investment (Roy 1951; Sjaastad 1962), suggests that migrants choose destinations that are culturally similar to their homes. Cultural similarity refers to the attitudes, beliefs, values, and practices shared between the populations of two countries. The more similar the cultures in the origin and destination country, the lower the migration cost and the higher the returns to migration (e.g., Bodvarsson and Van den Berg 2013). Hence, utility-maximizing migrants increase the returns to their human capital by moving to countries that are culturally similar to their home country, and, all else being equal, cultural similarity increases migration between countries.

Empirical findings generally support this prediction.[1] For instance, as one of the first in the field, Belot and Ederveen (2012) find that several similarity measures based on languages, religions, and attitudinal surveys are positively associated with international migration flows. More recent studies refine and extend their findings, using languages (Adserà and Pytliková 2015; Bredtmann, Nowotny, and Otten 2020), genetic ancestry (Collier and Hoeffler 2018; Krieger, Renner, and Ruhose 2018), cultural trade (Lanati and Venturini 2021), and survey-based cultural attitudes (Caragliu et al. 2013; Wang, De Graaff, and Nijkamp 2016; White and Buehler 2018) as measures of cultural similarity.[2]

A common feature of the above studies is that they estimate gravity models of migration.[3] The gravity model has become an increasingly popular tool to investigate country-level drivers of migration.[4] While the theoretical foundations differ between the two fields,[5] the migration gravity model is closely linked to the gravity model of international trade. In their empirical form, both models represent bilateral flows as a function of origin- and destination-specific

---

[1] These theoretical considerations are usually tested in country-level settings as individual-level data that includes both migration and culture is limited. So far, most individual-level data does not follow migrants across borders, thereby recording migration behavior, but rather records migration intentions such as the Gallop World Poll (e.g., Ruyssen and Salomone 2018).

[2] Many of the cited studies, rather than the effect of cultural similarity on migration, investigate the effect of cultural distance on migration. So, they report negative effects rather than positive effects. Yet, as the measures used express both distance and similarity, findings regarding distance translate easily into findings regarding similarity.

[3] Note that Krieger, Renner, and Ruhose (2018) and Wang, De Graff, and Nijkamp (2016) use different approaches. Among the studies that use gravity models, estimations differ greatly with respect to their theoretical foundations, their empirical specifications, and use of migration data. See Table 8.1 for an overview of the cited studies (Appendix).

[4] Mayda (2010) is probably the first to study country-level determinants of migration using a gravity framework. Following her, the topics studied using this framework include the role of income and migration policies (Ortega and Peri 2013), climate change (Beine and Parsons 2015), migrant networks (Beine, Docquier, and Özden 2011) or visa requirements (Czaika and de Haas 2017).

[5] While migration gravity relies on the discrete choice structure of location decisions, gravity in trade is either derived from a demand-side view using Armington/constant elasticity of substitution (CES) functions or from a supply-side view using a Ricardian structure of supply (e.g. Anderson 2011; Beine, Bertoli, and Fernández-Huertas Moraga 2016; Yotov 2022a).

determinants and bilateral costs.[6] Because of this analogy, both fields face similar empirical challenges when estimating their gravity models. The most important challenges are controlling for multilateral resistance[7] (Anderson and van Wincoop 2003; Bertoli and Fernández-Huertas Moraga 2013), addressing unobserved bilateral heterogeneity (Egger and Nigai 2015), and correctly representing theoretical gravity by including domestic flows in addition to the international flows (Yotov 2022b).

However, while addressing these challenges has become standard in trade (e.g., see Yotov et al. 2016; Larch and Yotov 2023), this does not yet apply to the migration literature. For example, among the above-cited studies, only recent ones account for multilateral resistance (Collier and Hoeffler 2018; Bredtmann, Nowotny, and Otten 2020; Lanati and Venturini 2021). Second, only Lanati and Venturini (2021) account for unobserved bilateral heterogeneity using country-pair fixed effects. Third, no previous study includes domestic migration as a component of the dependent variable.[8] The latter is not only standard in trade (Yotov 2022b) but also required by migration theory as the origin country is part of the choice set of people's location decisions (see Section 4). All this implies that the evidence supporting the hypothesis that cultural similarity increases migration is prone to empirical biases, where omitted variable bias and simultaneity bias prevent causal interpretations, rendering previous results largely correlational.

Against this background, this study aims to bring established estimation techniques for gravity models from trade to the migration literature and investigate whether there is evidence for the hypothesis that cultural similarity increases migration flows. To do so, I estimate a migration gravity model on a panel of migration flows between more than 200 countries between 1990 and 2019 using various measures of cultural similarity. My preferred econometric specification includes domestic migration flows, lagged right-hand-side variables, and a three-way fixed effects structure (3WFE) with origin-year, destination-year, and corridor (asymmetric country pair) fixed effects. To make this FE structure work, I use only time-varying measures of cultural similarity (see below). As is standard in the gravity literature, I obtain results with Poisson pseudo-maximum likelihood estimation (Santos Silva and Tenreyro 2006; Correia, Guimarães, and Zylkin 2020).

[6] Compare, for example, the migration gravity model in Lanati and Venturini (2021) with the trade gravity model in Larch et al. (2018).
[7] Failure to properly control for multilateral resistance has been called the "*gold medal mistake*" in gravity estimations (Baldwin and Taglioni 2006).
[8] An exception from the wider migration literature is the study by Beverelli (2022). In addition, Beine and Parsons (2015) use the number of "stayers" as a denominator in computing bilateral migration rates and thereby also account for domestic flows.

The estimation results are not in line with a positive effect of cultural similarity on migration. My measure of religious similarity (Maoz and Henderson 2013), which I use as a proxy for similarity of cultural beliefs, has a significant, negative effect on migration. In addition, my survey-based, attitudinal measures of cultural similarity – capturing attitudes towards individualism, indulgence, and social trust (Beugelsdijk and Welzel 2018) – do not show statistically significant effects on migration.

A subsequent heterogeneity analysis suggests that these unexpected findings may be explained by cultural selection and sorting of migrants (e.g., Knudsen 2022; Docquier, Tansel, and Turati 2020). More specifically, I find that when religious tolerance in the origin country is low, religious similarity has a negative effect on migration, and when tolerance is high, it has the expected positive effect. This means that migrants from intolerant countries seek destinations that are religiously different from their home countries, and *vice versa*. This is consistent with the idea that, in religiously intolerant countries, religious minorities select into migration and then sort themselves into destinations that are similar to their *own* religious beliefs rather than the average cultural beliefs of their home country. Vice versa, in religiously tolerant countries, emigrants are not selected on religion and sort themselves into destinations similar to their home country's average beliefs. Similar results are obtained concerning individualism. Therefore, cultural similarity may still increase migration, but the effect is not obvious and can, on the aggregate, be masked by cultural selection and sorting.

Finally, the results of a two-stage fixed effects (TSFE) procedure (Egger and Nigai 2015; Honoré and Kesina 2017) and a gravity-specific matching estimator(Baier and Bergstrand 2009), which also allow estimation of time-invariant similarity variables, such as linguistic and genetic similarity (Gurevich et al. 2021; Spolaore and Wacziarg 2018), show that the positive average effect of cultural similarity in earlier literature can be explained in large parts by historical cultural similarity between countries. Contemporary effects of similarity, such as those obtained from the 3WFE, are usually smaller, often insignificant, or even negative. Thus these results, too, contribute to the overall picture that the relationship between cultural similarity and migration is more nuanced than previously thought.

This study makes three main contributions. First, this study shows that the empirical (cross-country) relationship between cultural similarity and migration is more nuanced than previously thought. On the one hand, being the first since Belot and Ederveen (2012) to use a wide range of cultural similarity measures, I provide evidence that the effect depends on which aspect of culture is considered. For instance, while linguistic and genetic similarity increases migration on average, religious similarity decreases migration. In addition, similarity regarding directly measured cultural attitudes towards individualism, indulgence, and trust has no significant effect on migration. Hence, the average effects do not unanimously support the hypothesis that cultural similarity increases migration. On the other hand, I show that cultural selection and sorting of migrants may severely affect the sign and size of average effects. This suggests that cultural similarity with respect to the *individual* beliefs of migrants can increase migration, as predicted by theory. Moreover, because I use advanced gravity techniques, these results are more reliable than the findings of previous studies.

Second, the methodological contribution of this study is to bring established gravity estimation practices from the trade to the migration literature and to present theory-consistent and unbiased migration gravity estimations. I achieve this by, first, implementing estimation approaches that account for multilateral resistance and various kinds of unobserved heterogeneity. These approaches have rarely (3WFE) or never (TSFE, matching) been used in the migration literature before. Furthermore, the 3WFE and matching approaches allow for a causal interpretation of the estimated coefficients (see esp. discussion in Section 7), which is not possible with earlier studies. Second, as a novel input to the migration literature, I include domestic migration flows and their determinants in my gravity estimations. This better represents that people decide between international versus domestic labor markets and so ensures theory-consistent estimation of the discrete location choice-based gravity model of migration.

Finally, this study contributes by using the most recent migration data and measures of cultural similarity available to date. For example, in contrast to previous studies, I focus only on outmigration flows and exclude return and onward migration, whose response to cultural similarity is not subject to the theoretical considerations above. I use appropriately disaggregated bilateral migration flows as the dependent variable (Azose and Raftery 2019; Abel 2019, Version 9). In addition, as the first in the field, I use state-of-the-art measures of time-varying religious similarity (Maoz and Henderson 2013), linguistic similarity (Gurevich et al. 2021), genetic similarity (Spolaore and Wacziarg 2018) as well as cultural attitudes (Beugelsdijk and Welzel 2018).

2 Conceptual and theoretical background

2.1 Cultural Similarity

The members of each society share views about the world and how the world should be. We call such shared views 'culture'. Throughout this paper, I understand culture as the shared set of beliefs among people in a society about (i) which material and immaterial goods are socially desirable (values) and (ii) what is socially desirable behavior (norms).[9] Based on this

[9] This is consistent with standard definitions in the literature. For example, Guiso, Sapienza, and Zingales define culture as "those customary beliefs and values that ethnic, religious, and social groups

definition, *cultural similarity between countries* is the intersection of the sets of cultural beliefs of the two populations living in these countries. Thus, the bigger the overlap of cultural beliefs, the more culturally similar the two countries.

Over the past decades, researchers have collected an ever-growing body of evidence on how culture affects individuals' interactions, how it shapes their economic activities, and, eventually, how it determines the socio-economic outcomes of entire societies and countries (e.g., reviews by Roland 2015; and Alesina and Giuliano 2015). The empirical literature has also studied *cultural similarity* between countries and its implications for economic outcomes. Evidence in the literature suggests, for instance, that technology spreads more rapidly among culturally similar countries (Spolaore and Wacziarg 2009; Bove and Gokmen 2020) and that they trade larger volumes of goods between them (Felbermayr and Toubal 2010; Egger and Lassmann 2015).

2.2 Cultural similarity in migration theory

Economic migration theory rests on the idea that migration is a human capital investment. Migrants maximize their utility by choosing residence locations that yield the highest net returns to their human capital, i.e., the highest net returns to their labor supply (Sjaastad 1962; for a review, see Bodvarsson, Simpson, and Sparber 2015). By migrating, they can relocate to places where they expect higher returns to their skills and effort. While for Sjaastad (1962) the benefits of migration are sufficiently reflected in earnings differentials between origin and destination, more recent literature also considers non-monetary benefits, such as more desirable social, political, religious, or environmental circumstances (e.g., Pedersen, Pytlikova, and Smith 2008; Mayda 2010; Beine and Parsons 2015).

Like any other investment, migration is associated with certain costs. The literature distinguishes between monetary costs (e.g., travel costs, visa requirements, increased expenditure, and language courses) and non-monetary, "psychological" costs (e.g., being separated from family or not being able to exercise certain rights or practices). Against this background, cultural similarity between countries is viewed as a factor that reduces migration costs. Cultural similarity reduces migration costs, for instance, by facilitating a better mutual understanding of behavior and therefore facilitating the transfer of existing human capital and the acquisition of new social capital (see, for example, discussions in Caragliu et al. 2013; Krieger, Renner, and Ruhose 2018). Therefore, the hypothesis follows that cultural similarity between countries increases migration.

transmit fairly unchanged from generation to generation." (2006, 23) Gorodnichenko and Roland define culture as "the set of values and beliefs that people have about how the world (both nature and society) works as well as the norms of behavior derived from that set of values." (2017, 402)

This can also be derived from the benefit point-of-view: If expected earnings at the destination are also a function of individuals' cultural traits, then the cultural similarity between countries makes these earnings more likely. For instance, prospective migrants will not find it attractive to sit too high or too low in the distribution of some important cultural trait at the destination because it will make it difficult for others to work with them. This means that culturally similar destinations are more attractive to migrants compared to less similar destinations, and cultural similarity between countries increases migration.[10]

2.3 Gravity of migration and cultural similarity

The above theoretical considerations are also captured in the RUM-based gravity model of migration. This model uses a discrete choice structure to represent the destination choice of migrants (Appendix 8.2). Individuals pick their country of residence from the set of all destinations, including their origin country, depending on the net benefits they can expect from living in a country. The gravity model then aggregates these individual choices to represent migration flows between countries on the country level. Following Beine, Bertoli, and Fernández-Huertas Moraga (2016), the gravity model of migration is expressed as

$$E[m_{ijt}] = \frac{\phi_{ijt} y_{jt}}{\Omega_{\mathrm{it}}} p_{it} \,, \tag{1}$$

where bilateral migration flows are modeled as origin i's ability to send migrants at a given time, p_{it}, multiplied by the attractiveness of destination j – which depends on the characteristics y_{jt} and accessibility ϕ_{ijt} of destination j – relative to the attractiveness of all other options k, $\Omega_{it} = \sum_{k \neq j} \phi_{ikt} y_{kt}$.

Cultural similarity between countries of an ij-country pair represents one of the factors that increase the attractiveness of destination j. It makes destination j, and therefore also the benefits that can be gained from moving there, more accessible to migrants from origin i. This is in line with the idea that migration is a human capital investment and that migrants seek to maximize returns to their human capital.

Of course, the most culturally similar destination is the country of origin of migrants. Because of the non-zero costs of moving to a place that is culturally different, individuals stay at home

[10] Not only economic theory predicts that migrants select culturally similar destinations. Take, for example, the concept of homophily from sociology and network theory. Homophily is the tendency of people to associate and form ties with similar others, where similarity with others is understood as having common attributes, such as status and values (Lazarsfeld and Merton 1954; McPherson, Smith-Lovin, and Cook 2001; Lawrence and Shah 2020). All economic factors being equal, the tendency to associate with similar others will make culturally similar destinations more attractive to migrants than dissimilar places. Thus, a positive relationship between cultural similarity and bilateral migration flows indicates homophilic migration patterns. As the proverb goes, birds of a feather flock together.

even if they could achieve higher earnings elsewhere.[11] In this situation, international migration occurs when the cultural frictions are low enough, i.e., when an international destination is sufficiently similar to the origin country.

Based on these considerations, I expect to find a positive effect of cultural similarity on the volume of migration flows. This prediction should hold the same for *between* effects as well as for (temporal) *within* effects. That is, not only should cultural similarity positively affect the migration of prospective migrants who choose between different destinations at a time, but it should also positively affect the migration decisions of cohorts of migrants as countries become more similar over time. The methods I apply later will allow me distinguish between within and between effects of cultural similarity on migration.

2.4 Cultural selection and sorting

The above framework assumes implicitly that all individuals in a country hold homogeneous cultural beliefs. However, this ignores the cultural selection and sorting of migrants.[12] More specifically, it fails to account for the potential heterogeneity of cultural beliefs between migrants and stayers and, in turn, for the heterogeneity of cultural motives for migrating or choosing one destination over the other.

Selection and sorting of migrants are well-known issues in migration economics –traditionally regarding other migrant characteristics and, more recently, regarding culture as well. For instance, highly skilled individuals are more likely to migrate, and high-skilled migrants choose destinations with higher returns to skill, while low-skilled migrants chose destinations with lower relative returns to skill (Borjas 1987; Grogger and Hanson 2011). Regarding cultural characteristics e.g., Alesina and Giuliano (2010) show that individuals with strong family ties are less mobile, Knudsen (2022) finds that people with individualistic traits are more likely to emigrate, and Docquier, Tansel, and Turati (2020) show that less religious individuals have stronger migration intentions.

Cultural selection and sorting have implications for the hypothesis that cultural similarity increases migration. For example, a particular religious group is no longer represented or is persecuted by the government in their country of birth. For many such groups, emigration is the only way to find representation and protection. Take the migration of Hindus to India after the independence of Pakistan from the British Empire. While they migrate to a country that is,

[11] In such an aggregate-level framework, the crucial theoretical assumption is that individuals from the same origin are homogeneous regarding their cultural beliefs. This assumption ensures that there is no cultural selection into migration (to specific destinations). See following section.
[12] Selection refers to *who* decides to migrate internationally in the first place while sorting refers to *where* these migrants decide to go depending on the characteristics of destinations (e.g., Grogger and Hanson 2011)

on average, similar to *their* religious beliefs, they do not migrate between two religiously similar countries. Afterall, they are a minority in their place of birth. Similar cases can be imagined for other aspects or dimensions of culture, for example, individualistically-minded people in collectivistic origin countries or speakers of specific native languages in origin countries with other majority languages.

Thus, if cultural selection and sorting are powerful enough, i.e., those who migrate are those who do not fit in at home culturally and seek places that are similar to *their* cultural beliefs, then country-level cultural similarity may even have a negative effect on migration. Initially, I will assume that they are exceptions to the rule and that all individuals in a country hold homogeneous cultural beliefs. However, in an additional analysis I present later, I will reintroduce the idea of cultural selection and sorting of migrants.

3 Previous empirical findings

The following review reports previous empirical findings. The review will show that, while there is general support for the theoretical prediction that cultural similarity increases migration, the evidence is not unambiguous.

The review also introduces measures of cultural similarity that have been used in the literature. There are proxies and direct measures of cultural similarity. Proxy measures approximate the overlap of cultural beliefs between populations by the degree to which these populations share languages, religion, and genetic ancestry. Direct measures are based on the overlap of cultural attitudes between populations, which are collected, for instance, through large surveys such as the World Values Survey. The review is structured according to the measures used in each study.[13] Some authors use cultural distance instead of cultural similarity. Yet, as the measures used express both distance and similarity depending on the sign one uses ($cultsim_{ij} = cultdist_{ij} * (-1)$), findings regarding distance translate easily into findings regarding similarity.

3.1 Linguistic similarity

The most widely used measure of cultural similarity – not only in migration but also in trade – is linguistic similarity. The previous literature uses two types of this proxy: first, common language indicators, which say whether (certain shares of) two populations speak/use the same language (e.g., Gravity Database by CEPII, Mayer and Zignago 2011); and second, linguistic proximity indices that also consider the similarity of languages that are not shared between populations (e.g., Dyen, Kruskal, and Black 1992; Bakker et al. 2009). The former are often

[13] For an overview of studies see Table 8.1 in Appendix 8.1.

used as control variables in studies that focus on objectives other than the effects of cultural similarity. In contrast, the latter are used in studies that focus specifically on the effect of linguistic similarity. Both indices can be based on either spoken or official languages.

Among the studies that focus specifically on linguistic similarity, Belot and Ederveen (2012) find a positive effect on migration flows using the so-called Dyen index as a measure of linguistic similarity, which is based on whether words with the same meaning have a common linguistic derivation (Dyen, Kruskal, and Black 1992)[14]. Their results are robust to controlling for common spoken languages. Adserà and Pytliková (2015) find the same positive effect using three different measures of linguistic similarity: the Dyen index, the so-called Levenshtein index, which relies on the phonetic similarity of words (Bakker et al. 2009), and their proximity measure, constructed from official languages and the distance between language-tree branches from Ethnologue (Lewis 2009; see also Fearon 2003). Their results, too, are robust to controlling for common language and show, additionally, that people are more likely to migrate to countries with a widely spoken language (English) even if their native language is dissimilar. Finally, Bredtmann, Nowotny, and Otten (2020), using the Levenshtein index, not only corroborate the previous positive effect of linguistic similarity on migration but also show that migrant networks and the ability to communicate in the host country language are substitutes as indicated by a strong negative interaction effect between networks and linguistic similarity.

Studies using common languages as a control variable typically find positive effects of linguistic similarity on migration (Bertoli and Fernández-Huertas Moraga 2015; Beine and Parsons 2015; Lucas 2015), although the effects are not always statistically significant (e.g., Lanati and Venturini 2021 when using PPML; Mayda 2010; Ruyssen and Rayp 2014).

3.2 Religious similarity

A further proxy of cultural similarity is based on religion, measuring to which extent religious denominations coincide in two populations(e.g., Maoz and Henderson 2013). Religious beliefs are a reliable indicator of cultural beliefs and strongly influence societal behavior and cultural norms. For example, although religious communities are typically large, adherents share remarkably similar religious beliefs ensured by hierarchical systems of distribution of religious knowledge that (typically) rely on central scriptures, common practices, formal and informal institutions, and a shared history. Such systems create clear distinctions between religious groups whose adherents share homogeneous beliefs and values. This makes religion a good proxy for cultural similarity.

[14] This index focuses on Indo-Germanic languages. An example of the same linguistic derivation would be that English *father*, German *Vater* derive from the Latin word *pater*.

Against this background, it is somewhat surprising that Belot and Ederveen (2012) are the only authors who study the effect of religious similarity on migration flows.[15] They find that religious similarity between countries is associated with increased migration flows (*ibid.*, Table 1 and Table 2). However, this effect is not robust to including migrant networks as a control variable (Table 3). It is unclear whether this is because their initial findings were driven by migration effects on the composition of the destination population or because the sample shrinks because of this inclusion. Hence, there is no unambiguous evidence for a positive effect of religious proximity on migration flows.

3.3 Genetic similarity

Genetic similarity is a further indicator that has been used as a proxy for cultural similarity in economics (e.g., Spolaore and Wacziarg 2009). The rationale for using genetic similarity builds on dual inheritance theory. According to this, culture is a system of inheritance, subject to similar evolutionary processes and forces as genes. These processes affect differentiation between groups and increasing the genetic and cultural distance between groups (Boyd and Richerson 1988; Henrich and McElreath 2003). As both cultural beliefs and genes are inherited from parents to children[16], genes and cultural traits develop together over time; as some groups split from other groups (e.g., through migration, see out of Africa hypothesis), they develop their own genetic and cultural traits and grow further apart concerning their cultural beliefs and genetic heritage. Therefore, the genetic distance between two populations indicates their cultural distance, while genetic similarity indicates cultural similarity.[17]

In the migration literature, Collier and Hoeffler (2018) show that cultural distance proxied by ancestral distance does not seem to affect migration flows between countries, *per se*. However, they report that existing migrant networks at destination have a positive effect on migration between genetically distant populations. Bredtmann et al. (2020), who use genetic distance as a control variable, find a weakly significant, negative effect of genetic distance on migrants' location decisions. Adserà and Pytliková (2015) use genetic distance as a control variable. Their results show that the small and weakly significant negative effect of genetic distance on migration flows turns insignificant when they include existing migrant stocks as a control variable in their preferred specification (table 2, column 8). Finally, Krieger et al. (2018) find that genetic distance has heterogenous effects on educational migrant selection: they report

[15] Docquier et al. (2014) include religious similarity (proximity) as controls in their estimations, but do not report the estimated effects.
[16] Cultural beliefs need not only be inherited from parents. See e.g., Bisin and Verdier (2001) for an account of the transmission of culture, which allows vertical (parent-to-children) and horizontal (peer-to-peer) transmission.
[17] The genetic distance between two population also indicates the time since they shared common ancestors and therefore the time since their cultural believes have diverged (e.g. Spolaore and Wacziarg 2009).

negative educational selection at closer cultural proximity but positive educational selection at higher levels of cultural difference. Overall, the literature suggests a positive effect of genetic similarity between populations on migration flows. Yet, the evidence seems to depend on existing migrant networks at destination.

3.4 Attitudes-based cultural similarity

The measures discussed so far approximate the intersection of the shared cultural beliefs of the populations in two countries. However, these measures are imprecise because populations may share cultural beliefs despite having different languages, religions, or genetic markers. Additionally, proxies merely capture *that there are* similarities or differences in beliefs between populations. They do not speak about the *contents* of the cultural beliefs that populations share or do not share. To remedy this, one can measure cultural attitudes and, thus, cultural similarity between populations directly.

Just as in the wider economic literature, using indices based on cultural dimensions by Hofstede (e.g., 2001) or Inglehart (e.g., 1997) has also enjoyed increasing popularity in the migration literature. Belot and Ederveen (2012) report mixed results regarding cultural similarity among OECD countries, aggregating both Hofstede and Inglehart dimensions of culture into single measures. Only for one of their specifications using Inglehart do they find the expected positive effect on migration. However, they do not find significant effects in any other specification. In contrast, Wang et al. (2016) show that average cultural distance significantly decreases migration between European regions. They obtain their measure of cultural distance by conducting PCA on survey items of the European Social Survey (ESS). White and Buehler (2018), using Hofstede, Globe, and Inglehart dimensions of culture, find that cultural similarity generally increases migration in their sample of OECD countries. While the above results all pertain to composite indices of cultural similarity, the studies by Caragliu et al. (2013) and White and Buehler (2018) indicate that individual cultural dimensions affect migration in different ways. For example, while similarity is associated with higher flows for some dimensions, it is associated with lower migrant flows for others, or not associated at all. Thus, there is ambiguous evidence for a positive effect of cultural similarity on migration using attitudes-based measures.

3.5 Cultural trade

A final proxy for cultural similarity used in the literature is cultural trade. Cultural trade is the volume of bilateral trade in cultural goods (music, arts and crafts, games, etc.) between countries (e.g., Disdier et al. 2010). The cultural trade measure rests on the idea that trade in cultural goods reflects similarity in cultural tastes and, therefore, reflects cultural similarity itself. In migration, only Lanati and Venturini (2021) study cultural trade as a determinant of migration. Their results show a positive effect on international migration flows. This means that the higher the degree of affinity of a country for another country's culture, the larger, on average, the size of migration flows from the former to the latter.

However, in my view, appreciating and importing products from a country does not necessarily imply sharing cultural beliefs. Cultural trade may simply express attraction to others because they are different. Therefore, trade in cultural goods may capture cultural affinity but not necessarily cultural similarity. For these reasons I will not use cultural trade as a proxy in this study.

4 Method

While most of the studies reviewed above estimate gravity equations, only a few estimate gravity models that are consistent with the underlying micro-foundations of migration decisions and/or follow the best empirical practices for gravity estimations in the trade literature (e.g., Yotov et al. 2016; Yotov 2022a). In the following, I describe the main challenges of migration gravity estimations and present the remedies I employ in this study.

4.1 Estimation challenges and proposed remedies

I begin with the challenges posed by theory. First, theory-consistent estimation of the RUM-based gravity model (Equation (1) requires that one includes not only international migration flows m_{ijt} and their determinants but also domestic migration flows m_{iit} and *their* determinants. I use "domestic migration flows" to refer to the number of people who decided not to migrate internationally. This includes those who migrated internally, say, between two cities, and those who did not change residence. The reason to include domestic flows is that, according to the model, the location decision of individuals is over all possible destinations, including their home country: individuals decide about their country of residence based on the relative costs and benefits associated with relocating to that country, which may well be their home country. International migration occurs when the net benefits offered by going to one specific international destination outweigh the net benefits in the origin country (and any other international destination). Hence, omitting those who decided not to emigrate only really captures the relative costs and benefits among international destinations rather than the costs and benefits of international migration compared to staying at home. Hence, estimating a theory-consistent gravity model of international migration requires including domestic m_{iit} 'flows' to the set of observations. To date, empirical applications of the gravity model of international migration largely ignore this property of the theoretical gravity model[18]. The inclusion of domestic flows allows recovered coefficients to be interpretable as the effect of the variable on migrating internationally relative to staying in the home country. Without domestic flows, one only estimates the effect relative to *other* international destinations. This reasoning follows closely the trade literature, where estimating gravity models with domestic flows is a recommended and common practice (Yotov 2012; 2022b).

The second challenge of theory-consistent gravity estimation is to account for so-called multilateral resistance to migration (Bertoli and Fernández-Huertas Moraga 2013), represented by Ω_{it} in Equation (1). In contrast to the trade-gravity model, which has two multilateral resistance terms, one for inward frictions and one for outward frictions, the migration gravity model needs only one. Multilateral resistance to migration captures that decisions to migrate from origin i to destination j depend on alternative destinations and that the attractiveness of alternative destinations can exert a confounding influence on the determinants of bilateral migration (Beine, Bertoli, and Fernández-Huertas Moraga 2016, 502). However, like in trade (Anderson and van Wincoop 2003), multilateral resistance is a theoretical construct, not observed in practice, and must be accounted for through appropriate modelling techniques. The usual migration-specific strategy, which I will also use here, is to include origin-time fixed effects in the estimation, and so absorb Ω_{it} ensuring theory-consistent estimation (e.g., Ortega and Peri 2013).[19]

While the challenges discussed so far are posed by theory, some challenges are more empirical in nature. The first of these empirical challenges is that unobserved bilateral (cost) factors may affect both cultural similarity and international migration decisions. Following the trade literature (Yotov et al. 2016; Larch and Yotov 2023), the standard remedy for (time-invariant) unobserved bilateral heterogeneity is to include asymmetric (i.e., directional) country-pair fixed effects, which I call corridor fixed effects, and to fully exploit the panel structure of the data instead of estimating the gravity model on a series of cross sections. Newly available statistical routines make the use of high-dimensional fixed effect structures (e.g., origin-time, destination-time, corridor) efficient and convenient (e.g., ppmlhdfe by Correia, Guimarães, and Zylkin 2020). An alternative to corridor fixed effects is parameterizing time-invariant migration costs with "standard" gravity variables such as distance, common official language, contiguity, etc. However, the trade literature has shown that while these variables do well in capturing relative trade (and, by analogy, migration) costs, they do not capture the *level* of

[18] An exception is the study by Beine and Parsons (2015), who use emigration rates rather than migration flows (people counts). They compute the emigration rates by using the number of stayers as the denominator when calculating emigration rates.
[19] Note that other approaches have been developed, too (e.g., Bredtmann, Nowotny, and Otten 2020).

these costs sufficiently well (Egger and Nigai 2015; Agnosteva, Anderson, and Yotov 2019). Consider travel costs as a simple example: While relative travel costs to the Netherlands increase when comparing Belgium with Ghana as origin countries, their level may differ between Ghana, Kyrgyzstan, and the United Arab Emirates, which are (roughly) equidistant from the Netherlands. Moreover, the levels of travel costs are likely correlated with cultural differences between countries, that leaving them to be subsumed in the error term would cause omitted variable bias. Hence, the recommendation is to use corridor fixed effects to account for level differences in unobserved bilateral cost factors.

A related empirical challenge is unobserved heterogeneity in the origin and destination country. While many economic and demographic factors could be parameterized by including appropriate control variables on the origin- and destination-country level, e.g., population, GDP, labor market conditions, etc., there are country-level determinants, such as openness to migration or access to international labor markets, that are typically not observed. Therefore, a solution to capturing unobserved country-level heterogeneity, is to use origin-year and destination-year fixed effects. While the use of origin-year FE is already required by theory to account for multilateral resistance, the inclusion of destination-year FE is an empirical requirement.[20]

The above challenges and their remedies give rise to the three-way fixed effects (3WFE) approach, including domestic migration, which I will introduce in the following section. As the 3WFE approach addresses the main challenges of (migration) gravity estimations, it is the main focus of this study.

However, there are further challenges that are particular to the context of this study. For instance, using corridor fixed effects in the 3-way model introduces the subsequent challenge that the fixed effects absorb time-invariant indicators of cultural similarity. In this study this pertains to proxies of cultural similarity based on languages and genetic ancestry for whom there are no time-varying measures. Hence, one cannot recover estimates for these variables of interest using corridor fixed effects. As an alternative, the trade literature has proposed using a two-step procedure, called two-step fixed effects (TSFE), to recover estimates of the time-invariant or slowly moving cultural variables (Honoré and Kesina 2017; Egger and Nigai 2015; Spornberger 2022; Frensch, Fidrmuc, and Rindler 2023). I will offer the two-step procedure as a solution to complement my analysis.

[20] Note a difference between trade and migration gravity. The trade gravity model has two multilateral resistance terms, inward and outward multilateral resistance (e.g., Anderson and van Wincoop 2003), which require importer-year and exporter-year fixed effects (e.g., Yotov et al. 2016; Yotov 2022a). So while both directional time-varying fixed effects are required by theory in trade, in migration, only origin-year FEs are required by theory – destination-year FEs are an empirical requirement.

A further challenge is simultaneity or reverse causality, whereby significant empirical associations occur not only because cultural similarity affects migration but also because migration between two countries causes cultural change in these countries. For example, Rapoport, Sardoschau, and Silve (2021) show that migration is a source of cultural convergence between countries through mechanisms such as compositional changes in the origin and destination countries or cultural diffusion. Suppose that predominantly individualistically-minded individuals emigrate and move to more individualistic destinations, leaving their collectivistic compatriots behind. Then, if cultural selection is powerful enough, origin and destination countries could become more dissimilar. To attenuate concerns about reverse causality I will operate with lagged versions of my time-varying similarity variables.

A final empirical challenge, which is unique to the topic in this study, is the persistence of culture over time. If culture is persistent, then estimated coefficients could conflate long-term and contemporaneous effects of cultural similarity between countries. Yet, not all estimation methods are affected by this. For instance, the 3WFE approach is not affected because the corridor fixed effects absorb historical levels of cultural similarity. It therefore delivers contemporaneous effects of cultural similarity. However, the TSFE model likely conflate long-term effects of historical, cultural similarity from contemporaneous effects (Frensch, Fidrmuc, and Rindler 2023). To address this, I provide an additional specification of the two-stage procedure, which disentangles these two kinds of effects and yields contemporaneous effects.

4.2 Empirical specification and estimation: three-way fixed effects

The main method by which I estimate the gravity model of migration is a 3-way fixed effects approach using origin-year, destination-year, and corridor fixed effects. Following conventions in the trade and migration gravity literature (Beine, Bertoli, and Fernández-Huertas Moraga 2016; Yotov et al. 2016), I rewrite the theoretical gravity model in exponential form such that

$$m_{ijt} = exp\left[\delta_{it} + \delta_{jt} + \delta_{ij} + \beta_1 cultsim_{ijt-1} + \boldsymbol{X}_{ijt-1}\gamma\right] * \eta_{ijt}, \tag{2}$$

where m_{ijt} is the bilateral migrant flow between origin country i and destination country j for each 5-year period t ($T = 6$). $cultsim_{ijt}$ represents cultural similarity between i and j and β denotes the parameter of interest to be estimated.[21] Some of my measures of $cultsim_{ij}$ are time-varying while others are time-invariant, necessitating different estimation approaches as described earlier. The vector $\boldsymbol{X}_{ijt-1}$ represents lags of other time-varying, bilateral determinants of migration, such as migration networks or bilateral agreements. This model

[21] $cultsim_{ijt}$ and all other continuous RHS variables enter as natural logarithms. Therefore, if the values of a variable fall between 0 and 1, I compute $\log(x + 1)$ to make sure that the natural log is well-defined and equal to zero when the original variable is zero.

implements a three-way fixed effects structure of origin-time, δ_{it}, destination-time, δ_{jt} and directional (i.e., asymmetric) migration-corridor fixed effects, δ_{ij}. η_{ijt} is the error term and is in expectation $E\left[\eta_{ijt}|cultsim_{ijt}, \boldsymbol{X}_{ijt}, \delta_{it}, \delta_{jt}, \delta_{ij}\right] = 1$.

Note that m_{ijt} has two components: m_{ijt}^{int}, which represents international bilateral migrant flows, and m_{iit}^{dom}, which represents domestic 'flows'.[22] By including these two components, I make sure to estimate the gravity model consistent with the underlying theory of residence choice. The similarity measures for $cultsim_{ijt}$ and the control variables $\boldsymbol{X}_{ijt}$ also have international and domestic components. Details on the construction of the international and domestic components follow in section 5.

The corridor fixed effects, δ_{ij}, absorb unobserved time-invariant heterogeneity between migration corridors. This includes traditional gravity variables such as geodesic distance, common official language, common border, and past colonial ties as well as whether the corridor is an international (ij) or domestic (ii) corridor. As explained above, the corridor fixed effects absorb (time-invariant) relative migration costs and (time-invariant) levels of bilateral migration costs. This also includes initial levels of cultural similarity between countries. Overall, including corridor fixed effects reduces omitted variables bias and restores cross-sectional independence (Egger and Nigai 2015; Bertoli and Fernández-Huertas Moraga 2015).

The origin-time fixed effects, δ_{it}, control for multilateral resistance to migration to ensure theory consistent estimation in the presence of third-country effects that could confound the influence of cultural similarity on migration between i and j (see Bertoli and Fernández-Huertas Moraga 2013; Ortega and Peri 2013). Additionally, the origin-time FEs also absorb origin-specific determinants of migration – both time-varying and time-invariant, such as financial constraints to migration due to differing levels of economic development (Dao et al. 2018), climatic factors in the origin country (Beine and Parsons 2015), violence and conflict in the origin country, as well as demographic characteristics of the origin country population. The destination-time fixed effects, δ_{jt}, absorb destination-specific observed and unobserved heterogeneity caused by, for example, varying levels of economic performance, labor demand, and immigration policies across destinations and time. They also account for general openness to immigration and accessibility through international labor markets.

The simultaneous inclusion of origin-time and destination-time FE absorbs directional and time-varying differences between countries. For instance, income differences calculated by

[22] As explained above, these domestic 'flows' do not necessarily represent internal movements but represent the number of people who decided not to migrate internationally. To investigate the country-level determinants of international migration, it is not necessary to know, i.e., have data about, whether people changed residence *within* a country.

$gdp_j - gdp_i$ are fully absorbed because they depend, by definition, only on the two GDP terms, which vary over the origin-time and destination-time dimensions. Hence, the combination of δ_{it} and δ_{jt} serves as a control for differences in income and economic opportunities between origin and potential destinations, which are, according to theory, the important drivers of migration location decisions.

To obtain results, I estimate equation (2) with PPML (Santos Silva and Tenreyro 2006) using estimation routines for models with high-dimensional fixed effects structures (Correia, Guimarães, and Zylkin 2020; 2021).

5 Data

5.1 Bilateral migration flows

My dependent variable, bilateral international migrant flows, m_{ijt}, measures the number of people who change their usual residence from one country to another in a given period. These flows are *directional,* capturing migration in either direction between a pair of countries – one flow from A to B and one from B to A. In this study, I use estimated migration flows between all pairs of countries (UN definition) as totals over the six 5-year intervals between 1990 and 2019 (Azose and Raftery 2019; Abel and Cohen 2019).[23] I use these estimates of migration flows, because data on actual, recorded bilateral flows are only available for a small subset of countries. The flow estimates cover migration between all countries. The results in Grohmann and Fromell (2023) show that these flow estimates perform relatively well in gravity estimations and provide good proxies for the underlying flows.

Recent updates of the flow estimates include a disaggregation by migration type,[24] distinguishing between outmigration, transit migration, and return migration. In this paper, I use outmigration, which refers to the migration of i-born individuals from origin i to destination j.[25] The hypothesis about the positive effect of cultural similarity on migration is about *this* specific type of migration: individuals choose destinations that are culturally similar to their *home country* because this reduces costs to migration and increases returns to their skills. While similar considerations may also apply to transit and return migration, I leave detailed investigations into these types of migration for future work.

[23] Because these estimates are based on the UN migrant stock estimates, they carry the same inaccuracies regarding definitional discrepancies. For example, as indicated by the UN documentation (UNDESA 2020), while some stock estimates include refugees, others do not (e.g., many European countries). Moreover, some stock estimates are derived from data on the foreign-born population, while others were derived from data on foreign citizens.
[24] See Abel (2019), Version 9 from October 2022.
[25] Transit migration is about i-born individuals moving from origin $j \neq i$ to destination $k \neq i$. Return migration is about i-born individuals returning from destination j to their country of birth i.

Table 5.1 Summary Statistics

	Mean	sd	min	max	N
Dependent Variable					
migration flows	120,627	7,981,806	0	1,403,834,752	304,245
migration flows (intl.)	1,037	19,246	0	2,846,253	302,894
migration flows (dom.)	26,932,707	116,769,757	604	1,403,834,752	1,351
Attitudes					
coll./ind. similarity	.33	.085	.11	.71	21,651
duty/joy similarity	.37	.027	.25	.54	24,560
distrust/trust similarity	.42	.055	.17	.75	22,743
Proxies					
linguistic similarity	.083	.17	0	1	380,023
religious similarity	.15	.21	0	1.2	167,491
genetic similarity	.96	.023	.89	1	212,359
Control Variables					
migrant networks	118,705	7,875,952	0	1,438,411,392	386,575
migrant networks (intl.)	3,547	67,189	0	12,168,662	384,930
migrant networks (dom.)	27,065,718	117,708,453	734	1,438,411,392	1,645
income distance	19,492	20,401	0	118,582	235,159
EU/EFTA	.012	.11	0	1	386,575
BLA	.0081	.09	0	1	386,575
FTA	.089	.28	0	1	302,718
distance (km)	8,374	4,664	1	19,746	302,718
intl. corridor	1	.065	0	1	386,575
comm. official lang.	.21	.41	0	1	380,023
comm. spoken lang.	.4	.49	0	1	302,718
common border	.012	.11	0	1	302,718
colony ever	.0049	.07	0	1	302,718

As mentioned, m_{ijt} has two components: m_{ijt}^{int}, which represents international bilateral migrant flows, and m_{iit}^{dom}, which represents domestic 'flows'. The international component is taken directly from the data, but the domestic component is missing. So, I compute values for m_{iit}^{dom}, the number of i-country people who have remained in country i in period t, by subtracting the sum of all outmigration flows from country i over a 5-year period from the native-born population in country i at the beginning of each 5-year interval, such that $m_{iit}^{dom} = nativepop_{it} - \sum_j m_{ijt}$.[26] The required population data comes from the UN's Trends in International Migrant Stocks (UNDESA 2020).

5.2 Cultural Similarity

Earlier, I defined *cultural similarity between countries* as the overlap of the cultural beliefs of the two populations living in these countries. The following describes the measures of this overlap I use in this study. The first two, cultural attitudes and religious similarity, are time-varying and therefore suited to be used with the 3WFE approach; the remaining measures, linguistic and genetic similarity, are time-invariant and cannot be used with this approach.

[26] $nativepop_{it}$ is calculated as the difference between the total and foreign-born population (sum of all bilateral migrant stocks) in country i at time t.

However, I will present methods to recover estimates also for the time-invariant measures later in the study.

5.2.1 Attitudes-based cultural similarity

To measure cultural similarity based on cultural attitudes, I use the World Values Survey (WVS) and compute a Herfindahl-style similarity index along three cultural dimensions – Collectivism-Individualism, Duty-Joy, and Distrust-Trust. These dimensions were proposed by Beugelsdijk and Welzel (2018) to form a set of comprehensive yet conceptually and empirically independent cultural dimensions that integrates Hofstede's multidimensional framework of national culture (Hofstede 2001) with Inglehart's theory of cultural change (Inglehart 1997). *Collectivism-Individualism* characterizes the relationship between the individual and the collective, *Duty-Joy* represents the degree of people's restraint or indulgence regarding joyful moments in life, and *Distrust-Trust* reflects people's attitudes and responses towards unstructured situations. Figure 1 shows the evolution of average country scores of these three dimensions since 1990.[27]

For the analysis in this paper, I use the 13 WVS items that Beugelsdijk and Welzel identify for their dimensions (for a list of items, see Appendix 8.4) and compute, for each possible response to these 13 items, the share of respondents who agree with this response. For example, in 2010, 31.2% of respondents in Germany strongly agreed with the statement, 'One of my main goals in life has been to make my parents proud' (item D054), which is used for the individualism-collectivism dimension. 43.1% agreed, 18.9% disagreed, and 6.7% strongly disagreed. I do this per country and period for each item in the list.[28] These shares give the probabilities of the answers a randomly drawn individual from a population would give to the items. I then multiply the shares for each country pair and sum over the products, obtaining a Herfindahl-style index of each item. This represents the overlap of the cultural attitudes of two populations regarding each item. Finally, to get a score for the three dimensions, I take simple averages over the Herfindahl indices of the items of each dimension.[29] The measure is computed as follows

$$attsim_{ijt}^{d} = \frac{1}{B}\sum_{b}\left(\sum_{a} s_{it}^{a_b} * s_{jt}^{a_b}\right), \tag{3}$$

[27] The country scores are calculated using the formulas provided by Beugelsdijk and Welzel (2018).
[28] Appendix 8.4 provides details on the used survey items and on which waves of the WVS correspond with my 5-year periods. I use the Integrated Values Survey (IVS) series (1981-2021), which also includes European Values Survey (EVS) data.
[29] Because of missing data, I compute a few average scores differently. In the coll/ind dimension, when item 'D054' is missing, I take the average over the Herfindahl indices of the remaining four items. In the distrust/trust dimension, when either 'E069_12' or 'E069_17' are missing, I take the average over the Herfindahl indices of the remaining two items.

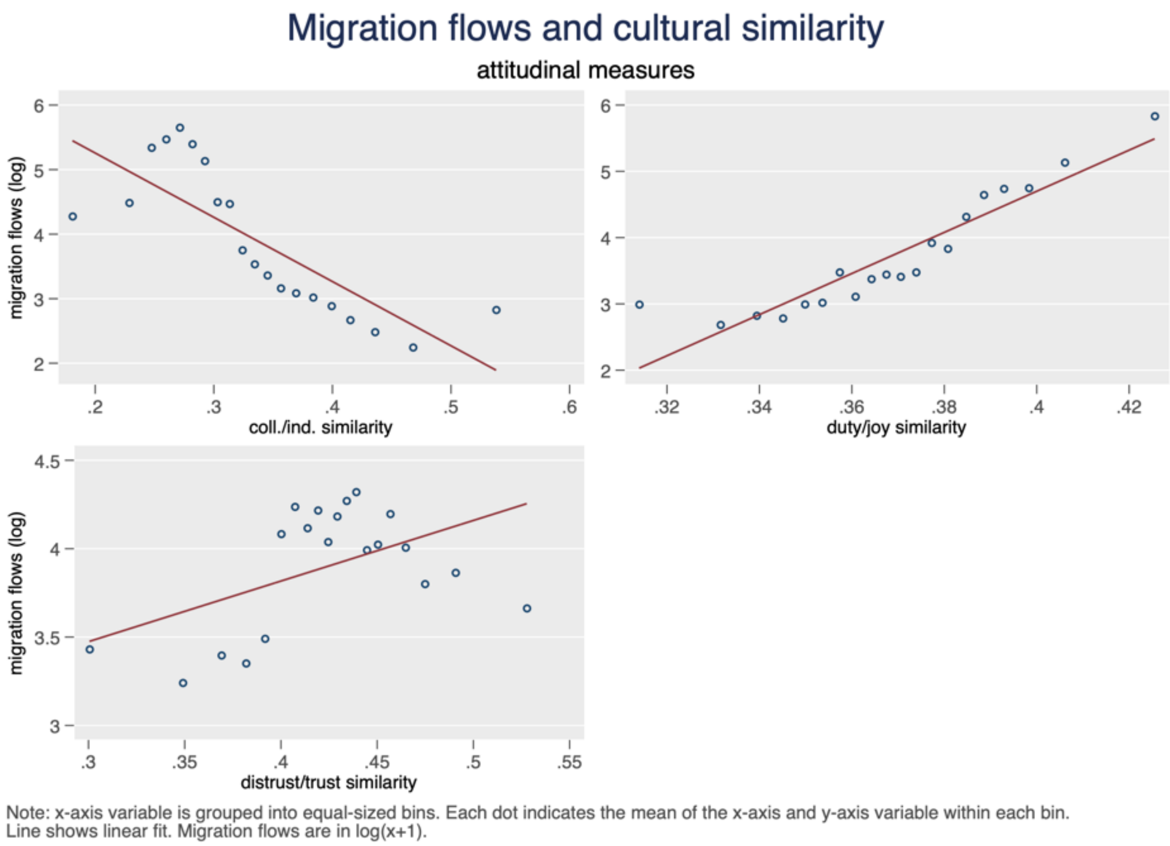

Figure 2: Migration Flows and attitudes-based measures of cultural similarity

where $attsim_{ijt}^{d}$ represents the attitudes-based similarity between country i and country j for dimension d at time t; $s_{it}^{a_b}$ and $s_{jt}^{a_b}$ are the shares of respondents who gave answer option a to item b. B is the number of WVS items used for each dimension: 5 items each for individualism-collectivism and duty-joy, and 3 items for distrust-trust. The computation of the domestic component of these measures follows the same steps, except that $s_{jt}^{a_b} = s_{jt}^{a_b}$. The domestic component is an expression of how homogenous a country's population is with respect to the three dimensions of culture.

The advantage of the attitudinal measures compared to proxy measures of culture is that they directly measure the overlap of cultural beliefs between two populations. Calculating separate dimensions also allows a more nuanced evaluation of the cultural attraction of possible destination countries on migrants compared to using proxies, which capture selected aspects of culture or compared to kitchen-sink style measures that aggregate cultural attitudes over multiple dimensions. Moreover, because the WVS is a *repeated* cross-section, $attsim_{ijt}^{d}$ varies over time, allowing estimation with country-pair fixed effects. I match existing WVS waves with my 5-year periods (see Appendix 8.4) and use 5-year lags in the regressions to preempt concerns about simultaneity.[30] Yet, as WVS data is available for only a relatively few countries (about 100 in my sample), the number of observations is considerably lower than the proxy measures $lingsim$, $relsim$ and $gensim$.

Figure 2 shows binned scatterplots to illustrate the relationship between the three dimensions of culture and migration flows.[31] The scatterplots suggest a nuanced relationship between cultural attitudes and migration flows. Collectivism-individualism similarity, depicted in the top left panel, has a negative relationship with migration flows ($\rho = -0.294$), while duty-joy and distrust-trust similarity show the expected positive relation, although the relationship is weak for distrust-trust similarity ($\rho = 0.271$ and 0.043).[32]

5.2.2 Religious similarity

My first proxy of cultural similarity is religious similarity. As the second proxy of cultural similarity, I use a "Herfindahl-style" measure of the overlap of religious families in the populations of two countries. Rather than focusing only on major religions (e.g., Christianity, Judaism, Islam, Buddhism, nonreligious, etc.) I consider population shares of religious families (e.g., Protestants, Orthodox Jews, Sunni, Shi'a, Mahayana, etc.). To measure religious similarity, I compute, for each country pair, the sum of the products of population shares of religious families using data from the COW World Religion Dataset (Maoz and Henderson 2013):

$$relsim_{ijt} = \sum_{r} s_{it}^{r} * s_{jt}^{r} , \tag{4}$$

Where $relsim_{ijt}$ represents the religious similarity between countries i and j. s_{it}^{r} and s_{jt}^{r} are the population shares of religious families r in countries i and j at time t.[33] The measure represents the religious overlap between the two populations and thus the probability that two randomly drawn individuals from either country belong to the same religious family. Belonging to a religious family means that individuals share a common set of beliefs, rituals, and practices that are determined and maintained by formal and informal institutions, reference to scripture

---

[30] For example, in the period 2010-2014, I use WVS wave 5 (2005-2009) to measure cultural attitudes. This implies, of course, that there may be up to 9 years between when attitudinal similarity was measured and when the corresponding migration occurred.

[31] I use binned scatterplots because regular scatterplots of my data with more than 300,000 observations are too crowded to interpret. In a binned scatterplot, the x-axis variable is grouped into equal-sized bins. Each dot indicates the mean of the x-axis and y-axis variable within each bin, making the data and underlying relationships easier to interpret.

[32] Table of correlations in Appendix 8.3. The correlation coefficients are Spearman rank correlations to better reflect the non-linear relationship between the Poisson-distributed migrant flows and the normally distributed similarity measures.

[33] The Maoz and Henderson (2013) data comes in 5-year intervals between 1945-2010. I use religion shares in 1990 for the period 1990-1994, shares in 1995 for the period 1995-1999, etc. Hence, time t denotes the first year of each 5-year period.

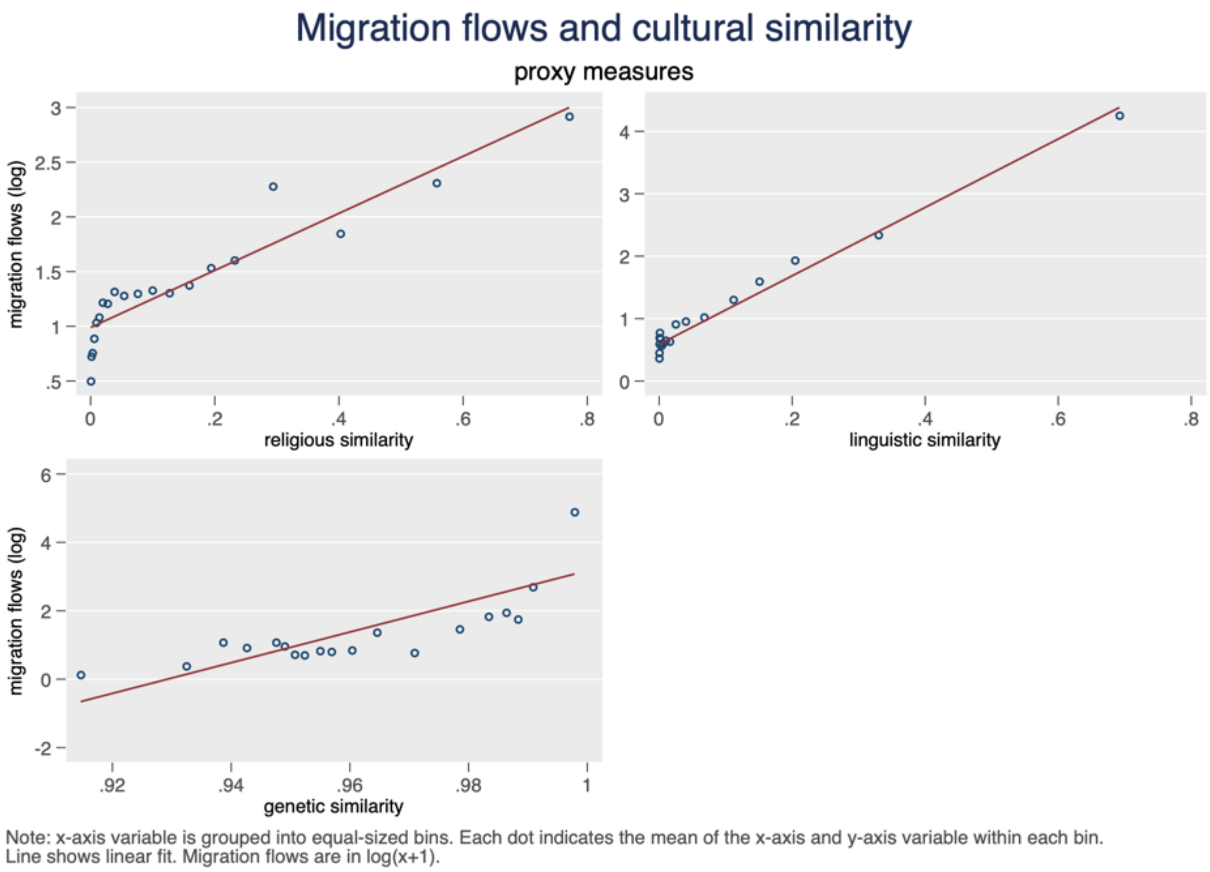

Figure 3 Scatter plot: migration flows and proxy measures of cultural similarity

(if applicable), and a shared history.[34] The measure is, therefore, an indicator for the intersection of religious beliefs and norms of two populations.

It is straightforward to compute the domestic component $relsim_{iit} = \sum_r s_{it}^r * s_{it}^r$ from the Maoz and Henderson (2013) data to complement the observations of domestic migration flows.[35] This represents the likelihood that two randomly drawn individuals from the same country belong to the same religious family. In the case of domestic observations (county $i = j$), it is an indicator of the religious heterogeneity in origin countries.

The Maoz and Henderson (2013) data is used over other resources because it reports religious population shares vary over time. Typical religion data (e.g., CEPII data) is time-invariant (see

[34] See criteria that define religions and religious families in Maoz and Henderson (2013).

[35] For three countries $relsim_{iit} > 1$: Haiti (2000, 2005), Japan (all 5 available periods), Kyrgyzstan (1995). This is because of Syncretism in Haiti (Vodou combining Yoruba and Catholicism) and the fact that many Japanese practice Shinto and Buddhism simultaneously. Hence, these observations are genuine. Yet, for Kyrgyzstan 1995, the data records a population share of 0.7229 for all three: islmsun, islmothr and total Islam, suggesting that islmothr is a faulty entry. Consequently, I set islmothr in Kyrgyzstan 1995 to zero. Moreover, in Cuba, the religious population shares also add up to >1. As for Haiti this is due to Syncretism. However, it does not lead to $relsim > 1$.

also Belot and Ederveen 2012). However, it is very likely that population shares of religious adherents change over time. For example, general tendencies such as secularization impact the population shares of religious vs nonreligious groups. Not only does this decrease the number of people who identify as believers, but it also affects religions at different rates in different regions of the world. For example, Oceania and Europe seem to secularize much quicker compared to other regions (Maoz and Henderson 2013). Using the Maoz and Henderson accounts for these developments. The data is available in 5-year intervals until 2010. In my regressions, I use values at the beginning of each 5-year period to represent religious similarity, e.g., I match religious similarity in 2010 with migration flows between 2010 and 2014.

The top left panel in Figure 3 shows a positive relationship between religious similarity and migration flows ($\rho = 0.288$), suggesting that migration flows are higher between countries with a larger overlap in religious beliefs.

5.2.3 Linguistic similarity

My second proxy of cultural similarity is linguistic similarity. Past studies in the migration literature rely on proxies based on spoken or official languages (e.g., Mayda 2010; Belot and Ederveen 2012; Adserà and Pytliková 2015). However, such proxies strongly reflect other language-related determinants of migration but not necessarily cultural determinants. Shared official languages may reflect past colonial relationships or institutional similarities, while common spoken languages may indicate that individuals from two populations speak languages such as English, French, or Spanish. For instance, the country pairs Guinea-France and Guinea-Senegal are certainly similar in official and spoken languages (all speak French) but very different in terms of similarity of cultural beliefs. A Guinean who has never been to Senegal or France can presumably adjust more quickly to the prevailing values and norms of the former. Therefore, while common official and spoken languages can facilitate communication and lower interaction costs between populations, I argue that they are not necessarily good proxies for cultural similarity regarding the “unwritten rules of the game”.

To accommodate these concerns, I use *native* language similarity as a proxy for cultural similarity. Native languages are the first language people learn and speak in their homes. They indicate cultural similarities and differences between populations more than spoken and official languages. The argument for this is analogous to dual inheritance theory mentioned earlier. Similar to genes, native (or maternal) languages are transmitted from parents to children. As language communities split and develop their own languages, they also develop their own cultural beliefs, which are also intergenerationally transmitted from parents to children (e.g., Bisin and Verdier 2001; Dohmen et al. 2012). Hence, (dis-)similarity of native languages indicates the degree of separation between two groups and their cultural beliefs. This makes native languages a good proxy for cultural similarity.

Thus, for linguistic similarity, I use the linguistic proximity index (LP_{ij}) provided by Gurevich et al. (2021), which builds on the common native language measure by Toubal and Melitz (2014).[36] This index is a weighted "Herfindahl-style" measure calculated using two components: the populations of speakers of native languages[37] and a measure of linguistic proximity between languages. More specifically, it calculates, for each country pair, the sum of the products of the population shares ($s_i^m$and s_j^n) of speakers of native languages (m,$n \in K$), weighted by the proximity of the respective native languages (p_{mn}):

$$lingsim_{ij} = \sum_m \sum_n \left(s_i^m * s_j^n\right) * p_{mn} \,, \tag{5}$$

Where p_{mn} represents the linguistic proximity between languages m and n based on their place within language trees; more specifically, how many branches of the tree m and n have in common. $p_{mn} = 1$ represents identical languages sharing the maximum amount of branches. $p_{mn} = 0$ represents languages that do not share any branch of the tree, i.e., that do not generate from the same proto-language. Consequently, $lingsim_{ij}$ represents the degree of linguistic similarity between two randomly drawn individuals from each country and captures the overlap between the language communities in the two countries. Therefore, the higher $lingsim_{ij}$, the more similar the two populations are with respect to native languages and their inherited cultural norms.

The Gurevich et al. (2021) data also includes observations of linguistic similarity within countries, where $lingsim_{ii} = \sum_m \sum_n (s_i^m * s_i^n) * p_{mn}$ represents the degree of linguistic similarity between two randomly drawn individuals from one country. A drawback of the data is that it does not vary over time. This means that the measure is constant for all of my six time periods and, therefore, absorbed by corridor fixed effects necessitating alternative econometric techniques.

The top right panel in Figure 3 suggests a positive relationship between linguistic similarity and migration flows. As expected, migration flows are positively correlated with higher linguistic similarity ($\rho = 0.413$).

[36] This means that I cannot exploit variation in native language communities over time, which could be important because one might argue that native language communities change a lot as result of globalization and internationalization. For example, indigenous languages, which are intricately linked to the cultures and customs of populations, appear to vanish and be replaced by larger languages. Children who do no longer learn the native languages of their parents but rather English, Mandarin, French, and/or Spanish may also inherit more assimilated values and norms. This convergence of languages and norms is not represented in the common native language measure.

[37] Native languages are the first language people learn and speak in their homes. The authors use information from representative population surveys, such as the Eurobarometer, and other sources, such as *Wikipedia* and *Ethnologue,* to obtain the relevant population shares.

5.2.4 Genetic similarity

To obtain my measure of genetic similarity, I use the so-called FST distance between populations, as reported by Spolaore and Waczairg (2018). More specifically, I compute

$$gensim_{ij} = 1 - d_{ij}^{\text{FST}}\,, \tag{6}$$

where d_{ij}^{FST} represents the genetic FST distance between the ethnic plurality groups in countries i and j, where $d_{ij}^{FST} = 1$ indicates maximum diversity between two populations and $d_{ij}^{FST} = 0$ maximum identity.[38] Therefore, higher values of $gensim_{ij}$ indicate higher degrees of long-term relatedness between the ethnic plurality groups in two countries. For the domestic component of genetic similarity I assume $d_{ii}^{FST} = 0$ and therefore $gensim_{ii} = 1$. This indicates complete similarity of two individuals from the same ethnic plurality group in a country. The genetic measure is time-invariant as the underlying population share measures used by Spolaore and Wacziarg do not vary over time (population shares from Alesina et al. 2003). Hence, like with the language-based proxy, I will have to use alternative techniques to the corridor-fixed effects to estimate the effect of genetic similarity.

The bottom left panel in Figure 3 suggests a positive relationship between genetic similarity and migration ($\rho = 0.430$). This means that migration flows are larger between countries with a closer ancestral relationship (among their plurality groups) and smaller with a farther relationship.

5.3 Control Variables

I use *migrant networks* to control for the influence that existing origin-country communities exert on prospective migrants by facilitating knowledge exchange about the new destination or aiding integration into the new environment. This control also holds changes in the composition of the destination population constant due to recent migrations, e.g., religious population shares or shares of native language speakers. For international corridors, migrant networks are measured by bilateral, international migrant stocks, i.e., the number of i-born individuals residing in destination j at the beginning of each 5-year interval. For domestic corridors, migrant – better: stayer – networks are measured by the number of i-born individuals (i.e., natives) in country i. To obtain this, I subtract the sum of all bilateral, international stocks in a country from its total population. Data for migrant stocks and total

[38] Ideally, I would use a population-weighted measure resembling the linguistic and religious measures. Yet, while Spolaore and Waczairg (2018) compute such a weighted measure, observations for domestic migration corridors are not readily available. Thus, I rely on their plurality measure, assigning value 1 to domestic observations. The correlation between Spolaore and Wacziarg's plurality measure and the population weighted measure is 0.92. This suggests that the plurality measure captures a considerable part of the variation in genetic similarity/distance between populations of countries.

population comes from the UN's Trends in International Migrant Stocks (UNDESA 2020) and is available for all periods. The correlations table shows that migrant networks are highly correlated with migrant flows ($\rho = .98$). The correlation with the cultural similarity variables ranges from $\rho = .48$ (linguistic similarity) to $\rho = -.29$ (coll./ind. similarity).

I also control for changing economic conditions that influence the attractiveness of migration corridors by including a measure of *income distance* in my specification.[39] I compute this as the absolute (or Euclidean) distance between destination and origin GDP per capita:

$$incdist_{ijt} = \sqrt{\left(gdp_{it} - gdp_{jt}\right)^2} = \left|gdp_{it} - gdp_{jt}\right|, \tag{7}$$

For domestic corridors, the variable takes a value of 0, indicating perfect similarity. GDP per capita (in constant 2010 $) is taken from the World Development Indicators (World Bank) and measured in the first year of each of the 5-year periods. Note that because income distance is symmetric, it does not capture directional differences, i.e., that migrants choose destinations with higher average incomes. The combination of origin-time and destination-time fixed effects already accounts for these differences. So, income distance accounts for changing economic differences between countries on top of income differentials. My variable $incdist_{ijt}$ shows a low correlation with migrant flows ($\rho = .01$) and a moderately high correlation with coll./ind. similarity ($\rho = -.48$) and distrust/trust similarity ($\rho = -.33$). It is less strongly correlated with the other cultural variables.

In addition, I include two dummy variables that indicate for each corridor whether the two countries are members of the European Union/European Free Trade Association (EU/EFTA)[40] and whether the two countries have entered a free trade agreement (FTA). EU/EFTA membership comes with the freedom of movement of workers, meaning nationals of any

[39] Note the distinction between *distances* and *differences*. Differences are symmetric and can be estimated in the 3WFE approach when they are time-varying. Differences (e.g., $gdp_j - gdp_i$) are asymmetric and are absorbed by the combination of the directional, time-varying fixed effects (origin-year and destination-year).

[40] EU/EFTA membership comes with the freedom of movement of workers, meaning that nationals of any member state can take up employment and residence in another member state on the same conditions as nationals of the latter. At the beginning of my panel in 1990, there are 12 member states: the founding countries Belgium, France, Germany, Italy, Luxembourg, and the Netherlands as well as Denmark, Ireland, the United Kingdom, Greece, Spain, and Portugal, who joined later. With the fourth enlargement in 1995, Austria, Finland, and Sweden also join the EU, making it 15 member states. This number remains stable until 2004, when the Czech Republic, Estonia, Cyprus, Latvia, Lithuania, Hungary, Malta, Poland, Slovakia, and Slovenia join. So, in my panel, the EU has 25 members from 2005-09. In the period 2010, Romania and Bulgaria are added after their EU entry in 2007. Croatia is added from 2015 onward after entering the EU in 2013. This yields a total of 28 EU countries with varying EU status over the 1990 to 2020 period. I also include the EFTA member states in this subset of countries as similar free movement conditions apply with respect to migration amongst them and the EU. Iceland, Liechtenstein, Norway, and Switzerland are EFTA members in all periods. Austria, Finland and Sweden are EFTA members in the 1990-94 period before joining the EU in the 1995-2000 period of my panel.

member state can take up employment and residence in another member state on the same conditions as nationals of the latter. FTAs represent openness between two countries. The data on EU/EFTA membership is constructed from publicly available information; the data on FTAs is from the Dynamic Gravity Dataset (Gurevich and Herman 2018)[41].

While the above controls were all time-varying, I will also use time-invariant controls for the TSFE and matching approaches. The first is an international corridor dummy with value 1 if $i \neq j$ and value 0 if $i = j$. In analogy to the trade literature, where this dummy variable captures trade integration effects (Bergstrand, Larch, and Yotov 2015), this dummy captures the ease of access of origin i's migrants to international labor markets.

Further, I use *common official language* and *common spoken language* to control for language-related confounders due to translation and language proficiency. Common official language is a dummy variable from Gurevich et al. (2021) with value 1 if two countries have at least one official language in common; common spoken language is a dummy variable from Gurevich and Herman's (2018) Dynamic Gravity Dataset indicating that residents of two countries speak at least one common language as listed by the CIA World Factbook. Both variables are time-invariant and take value 1 in domestic corridors. The correlation of common official language and common spoken language with my linguistic similarity measure is 0.324 and 0.237, respectively (see Appendix 8.3).

Finally, I also use typical gravity control variables from the Dynamic Gravity Dataset (Gurevich and Herman 2018). *Distance*, indicates the population-weighted distance between the largest cities of a country pair (in km); *common border* is a dummy variable with 1 if countries share a common border; and *ever colony of* is a dummy variable with value 1 if, for a country pair, the origin ever was a colony of the destination. For an overview of all variables, see Table 5.2.

[41] I use Version 2.1 (2021), available from https://www.usitc.gov/data/ gravity/dgd.htm.

Table 5.2 Variables, Definitions and Sources

Variable	Definition	Source
Migration flows	Estimates of bilateral migration flow between country pairs over 5-year periods from 1990 to 2020. Total of 6 periods.	*"Bilateral international migration flow estimates for 200 countries",* Abel and Cohen (2019)
Linguistic similarity	Time-invariant "Herfindahl index"-style measure of common native languages as the product of the population shares of native speakers weighted by the linguistic proximity of languages in a language tree. See equation (5). Because variable is time-invariant, it is available for all 6 periods.	*"One Nation, One Language?",* Gurevich et al. (2021)
Religious similarity	Time-varying "Herfindahl index"-style measure of religious proximity at the beginning of each period calculated by the product of the population shares of adherents of religious families. See equation (4). Available from 1990 to 2010, for a total of 5 periods in my sample.	*World Religion Dataset,* Maoz and Henderson (2013)
Genetic similarity	Time-invariant measure of genetic similarity based on FST distance between ethnic plurality groups in two countries. See equation (6). Because variable is time-invariant, it is available for all 6 periods.	*"Ancestry and development: New evidence",* Spolaore and Wacziarg (2018)
Cult. Attitudes similarity	Time-varying "Herfindahl index"-style similarity based on three cultural dimensions by Beugelsdijk and Welzel (2018) matching WVS waves to 5-year periods. Available for 6 periods between 1990 and 2020. See equation (3).	*World Values Survey (WVS),* Inglehart et al. (2020)
Migrant networks	Migrant stocks at the beginning of each 5-year period. Available for 6 periods between 1990 and 2020.	*International Migrant Stock 2020*, UNDESA (2020)
Income distance	Time-varying measure of distance of average incomes in countries based on absolute (Euclidean) distance of GDP per capita (PPP, at 2011 international $). See Equation **Error! Reference source not found.**). Available for 6 periods between 1990 and 2020.	*World Development Indi-cators (WDI)*, World Bank
EU/EFTA	Time-varying dummy variable indicating that the two countries of a corridor are members in the EU or EFTA. Available for all 6 periods. Domestic corridors take value EU/EFTA=1.	publicly available information
FTA	Time-varying dummy variable indicating that both countries of a corridor have entered a free trade agreement. Available for all 6 periods. Domestic corridors take value FTA=0.	*Dynamic Gravity Dataset,* Gurevich and Herman (2018)
Intl. corridor	Time-invariant dummy variable indicating whether a corridor is international or domestic. Takes value 1 if i≠j and value 0 if i=j.	*constructed*
Comm. official language	Time invariant dummy variable indicating that two countries have at least one official language in common. Domestic corridors take value COL==1.	*"One Nation, One Language?",* Gurevich et al. (2021)
Comm. spoken language	Time invariant dummy variable indicating that there is at least one common spoken language between the two countries. Domestic corridors take value CSL==1.	*Dynamic Gravity Dataset,* Gurevich and Herman (2018)
Distance (km)	Time-invariant, population-weighted distance between largest cities of the two countries in kilometers. For domestic corridors the distance indicates internal distances between large cities. If there is only one city, the distance is set to 1, such that log(1)=0.	*Dynamic Gravity Dataset,* Gurevich and Herman (2018)
Common border	Dummy variable indicating that two countries share a border. Domestic corridors take value =0	*Dynamic Gravity Dataset,* Gurevich and Herman (2018)
Ever Colony	Dummy variable indicating whether the origin country was ever a colony of the destination. Domestic corridors take value =0.	*Dynamic Gravity Dataset,* Gurevich and Herman (2018)

6 Analysis

6.1 Three-way fixed effects approach

This section presents the partial migration gravity effects of cultural similarity from the three-way fixed effects model (3WFE) using time-varying measures of cultural similarity. [42] The inclusion of corridor fixed effects produces estimates of *within*-corridor effects of cultural similarity. Based on theory and previous findings I expect to find a positive estimate on the parameter β on $\ln cultsim_{ijt}$.

The results in Table 6.1, which shows preferred specifications, including the full set of control variables, do not confirm this hypothesis. First, regarding the attitudinal measures, while collectivism/individualism similarity and duty/joy similarity have the expected positive sign, distrust/trust has a negative sign. However, none of these effects are statistically significant – neither when included separately (columns 1-3), nor when included simultaneously (4). Second, I find a negative and statistically significant effect of religious similarity on migration. The estimate of 0.41 indicates that a 10% increase of the overlap of religious beliefs between the countries in a corridor leads, on average, to a 4.1% decrease ($p < 0.05$) of migration.[43]

Different sets of control variables – (i) no controls, (ii) plus migrant networks, (ii) plus migrant networks and income distance – do not have a qualitative influence on the main results using the full sample (Appendix 8.5). The model with religious similarity (column 5) captures 90% of all international migration in the available periods (5 periods between 1990-2014)[44]. In contrast, the samples of the models using the attitudinal measures are smaller capturing only around 30% of all international migration due to the limited availability of the WVS data. To investigate the effects of this difference in sample size, Table 8.7 shows results that force the estimation sample to contain only observations that are non-missing across all models. While all cultural similarity variables have a negative sign, they are all imprecisely estimated and statistically insignificant.

[42] The effects are partial effects – borrowing terminology from the trade literature – as opposed to general equilibrium effects because cultural similarity and other migration determinants are not modelled as endogenous determinants.

[43] Parameter estimates from PPML on the log of continuous variables represent elasticities of migration flows with respect to the continuous variable (e.g., Yotov et al. 2016, 28). Estimates on dummy variables can be calculated in percentage terms by $\left[\exp(\hat{\beta}) - 1\right] \times 100$.

[44] $intl.mig.captured = \frac{(total \text{ in sample } intl.migration_t)}{(total \text{ overall } intl.migration_t)}$, excludes domestic migration corridors.

Table 6.1 Effects of cultural similarity on migration using three-way fixed effects approach

Dependent variable: migration flows					
	(1)	(2)	(3)	(4)	(5)
coll./ind. similarity	0.33			0.22	
	(1.77)			(3.17)	
duty/joy similarity		2.07		1.95	
		(3.18)		(4.13)	
distrust/trust similarity			-0.14	-1.23	
			(2.20)	(2.82)	
religious similarity					-0.41**
					(0.19)
migrant networks	0.13***	0.17***	0.19***	0.15**	0.22***
	(0.05)	(0.05)	(0.04)	(0.06)	(0.02)
income distance	0.04	0.06*	0.04	0.05	-0.01
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
EU/EFTA	0.55***	0.42***	0.53***	0.63***	0.39***
	(0.13)	(0.13)	(0.11)	(0.14)	(0.09)
FTA	0.12**	0.11**	0.13***	0.15***	0.07*
	(0.05)	(0.05)	(0.04)	(0.05)	(0.04)
Intercept	17.10***	15.73***	15.91***	16.53***	15.01***
	(1.31)	(1.36)	(0.93)	(1.50)	(0.42)
N	7358	8992	8203	5916	45722
intl. mig. captured	0.30	0.32	0.29	0.25	0.90
corridors	2573	3155	2743	2256	9955
origins	71	79	78	69	183
destinations	69	77	76	67	183
periods	5	5	6	5	5
pseudo R-sq.	0.9999	0.9999	0.9999	0.9999	0.9999

Notes: Table shows parameter estimates obtained from Poisson-Pseudo-Maximum-Likelihood (PPML) estimation using the ppmlhdfe routine, which automatically drops singletons and observations that are separated by a fixed effect (Correia et al. 2020,2021). Standard errors in parentheses are clustered at migration corridor level. *** p<.01, ** p<.05, * p<.1. All models include origin-year, destination-year and migration corridor fixed effects. All continuous variables enter the estimation in logs, $\ln(x+1)$ if the variable takes values smaller than 1. Parameter estimates on log of continuous variables represent elasticities of migration flows with respect to the continuous variable. Estimates on dummy variables can be calculated in percentage terms by $[\exp(\hat{\beta}) - 1] \times 100$.

The above results suggest that the effect of cultural similarity on migration is more nuanced than previously thought. They do not provide evidence for the hypothesis that cultural similarity between countries fosters migration between them. The following section further elaborates on these results and explores whether they are brought about by cultural selection into migration.

6.2 Cultural selection and sorting of migrants

Of course, the insignificant findings in the previous section regarding attitudinal measures of cultural similarity may be explained by the fact that culture simply does not have an effect on migration. However, both negative and insignificant results are also consistent with cultural selection and sorting of migrants. Cultural selection means that individuals' cultural values are a determinant of whether they become international migrants; cultural sorting means that their subsequent choice of destination is determined by their cultural values, too. As discussed earlier, if selection is powerful enough, it may depress or even reverse the average effect of cultural similarity on migration. This section explores whether this could be the case with the earlier results from the 3WFE approach.

While an aggregate-level study such as this is not suited to study the individual-level phenomenon of cultural selection directly, it is possible to derive aggregate-level hypotheses that follow from selection. To do so, consider collectivism-individualism at the origin. Previous literature suggests that migrants are more individualistic than their compatriots who do not migrate internationally (e.g., Knudsen 2022; see also "voluntary settlement hypothesis" Kitayama et al. 2006). This is intuitive as individualists incur lower costs from leaving existing social structures than their collectivistic counterparts, who derive utility from being part of a group and the support system provided by the group. Thus, migrants are likely to be selected on their cultural beliefs regarding individualism.

Furthermore, if these individualistic migrants seek destinations that are similar to their personal cultural beliefs, then, independent of whether the origin country is, on average, collectivistic or individualistic, they migrate to destinations that are more individualistic where they can expect to find like-minded people who also value individual achievement and initiative. This means that those emigrating from a collectivistic country will seek destinations that are culturally dissimilar to their home, i.e., individualistic. In contrast, those emigrating from an individualistic country will seek destinations that are culturally similar, i.e., individualistic. This yields the "selection" hypothesis that the effect of collectivism/individualism similarity is negative on migrations from collectivistic origins, while it is positive on migrations from individualistic origins.[45]

Such a "selection" hypothesis can also be derived regarding religious similarity. To do so, distinguish between countries where populations are more tolerant towards minority religions and countries in which they are *in*tolerant.[46] Religious selection into migration will be stronger in the latter than in the former, affecting the religious composition of the migration outflows: adherents of minority religions are, *ceteris paribus*, more likely to migrate from intolerant than tolerant origins. Moreover, turning to cultural sorting, those who decide to emigrate from less tolerant origins will seek out destinations that are religiously dissimilar to their origin in order to avoid a general public that is likely to be hostile against them. This means that cultural similarity will have a negative effect on migration from intolerant places and a positive effect on migration from tolerant places because, in religiously tolerant societies, the composition of outmigration flows is more representative of average cultural beliefs, and people seek religiously similar places.

[45] The other attitudinal dimensions, duty/joy and distrust/trust, do not lend themselves to derive similar hypotheses because the cultural traits of these dimensions do not yield clear predictions about the migration behavior of individuals who possess them. For instance, it is not obvious whether more distrusting or more trusting people would select into migration.

[46] For a study on the effect of more tolerant origin culture on the integration of immigrants into destination countries, see Berggren, Ljunge, and Nilsson (2023).

Table 6.2 Selection on individualism and religion

Dependent variable: migration flows		
	(1)	(2)
coll/ind sim. x LOW individualism	-0.37	
	(2.08)	
coll/ind sim. x HIGH individualism	3.58**	
	(1.69)	
religious sim. x LOW religious tolerance		-0.61***
		(0.22)
religious sim. x HIGH religious tolerance		0.92***
		(0.34)
migrant networks	0.13***	0.22***
	(0.04)	(0.02)
income distance	0.04	-0.01
	(0.03)	(0.03)
EU/EFTA	0.55***	0.38***
	(0.13)	(0.09)
FTA	0.12**	0.07*
	(0.05)	(0.04)
Intercept	17.05***	15.10***
	(1.19)	(0.43)
N	7358	45168
intl. mig. captured	0.30	0.89
corridors	2573	9838
origins	71	180
destinations	69	183
periods	5	5
pseudo R-sq.	0.9999	0.9999

Notes: Table shows parameter estimates obtained from Poisson-Pseudo-Maximum-Likelihood (PPML) estimation using the ppmlhdfe routine, which automatically drops singletons and observations that are separated by a fixed effect (Correia et al. 2020,2021) All models include origin-year, destination-year and migration corridor fixed effects. All continuous variables enter the estimation in logs, $\ln(x+1)$ if the variable takes values smaller than 1.). Standard errors in parentheses are clustered at migration corridor level. *** p<.01, ** p<.05, * p<.1.

To test these two “selection” hypotheses, I interact my measure of collectivism/individualism similarity with a dummy variable indicating high vs. low values of individualism at origin and interact my measure of religious similarity with a dummy variable indicating high vs. low levels of religious tolerance at origin.[47] Table 6.2 shows the results from estimating the 3WFE model including the respective interaction terms. Column (1) supports the “selection” hypothesis regarding selection on individualism: next to a large, statistically significant, and positive effect of cultural similarity on migrants from individualistic countries (coll/ind sim. x HIGH individualism), the negative sign on the interaction coll/ind sim. x LOW individualism is consistent with powerful selection of individualistically-minded people into migration who seek countries dissimilar to their home country.

[47] Levels of individualism are computed as averages over all respondents and waves of the five WVS items associated with collectivism/individualism (see Appendix 8.4) in 107 countries. Levels of religious tolerance are computed as averages of the Social Hostilities Index (SHI) of 198 countries by the Pew Research Center (https://www.pewresearch.org/religion/dataset/global-restrictions-on-religion-2007-2016/). The SHI is a measure of the extent to which “individuals and social groups infringe on religious beliefs and practices“ (Codebook for Pew Research Center’s global Restrictions on Religion Data).

My findings regarding religious similarity (column 2) are even more suggestive: I find a statistically significant negative effect on migrations from countries with LOW religious tolerance and a statistically significant, positive effect on migrations from countries with HIGH religious tolerance. These results are consistent with the selection of religious minorities who live in religiously intolerant societies into migration and their subsequent sorting into destinations based on religious similarity.

The results in this section suggest strong effects of cultural selection and subsequent cultural sorting of migrants on migration flows. The presence of the origin-year, destination-year, and corridor fixed effects also exclude that the effects reflect income differences or levels of religious tolerance and individualism at the destination. Thus, I maintain that the results are consistent with the view that migrants choose destinations that are similar with respect to their *individual* cultural beliefs. On the aggregate, the effects of cultural similarity between the LOW vs HIGH groups may cancel out or, if selection is strong enough, lead to a reversal of the otherwise positive effect. This may explain the recalcitrant evidence in the previous section. After all, cultural similarity may well have its theoretical positive effect on migration but it may be masked on the aggregate by other significant factors, such as cultural selection, that determine the migration process.

6.3 Alternative gravity estimations

Up to this point, the corridor fixed effects in the three-way fixed effects approach (3WFE) prevented estimating coefficients of linguistic and genetic similarity measures because they are time-invariant. This section is dedicated to methods that *can* recover the effects of time-invariant cultural similarity measures while addressing theoretical and empirical challenges of gravity estimations. In addition, implementing these alternative methods will yield further insights into the relationship between cultural similarity and migration, which cannot be drawn from the 3WFE results.

6.3.1 Two-step fixed effects

The first method I employ to estimate the effects of time-invariant cultural similarity variables is a two-step fixed effects approach (TSFE) following Honoré and Kesina (2017), with trade applications in Egger and Nigai (2015), Spornberger (2022) and Frensch, Fidrmuc, and Rindler (2023). To my knowledge, I am the first to use this approach in the field of migration.

The basic idea behind TSFE is to isolate the corridor-specific migration (cost) factors from inward and outward migration frictions and then to parameterize the former using the time-invariant variables of interest. The first step is a so-called constrained ANOVA decomposition of observed migration flows into origin-time, destination-time, and corridor fixed effects. One then recovers fitted values of the corridor fixed effects from the first stage, which represent corridor-specific migration (cost) factors. In the second step, these recovered estimates of bilateral migration costs are then regressed on the time-invariant cultural similarity variables to obtain coefficients.

Following the respective literature (Spornberger 2022), the first-stage decomposition is given by

$$m_{ijt} = exp\left[\delta_{it} + \delta_{jt} + \delta_{ij} + \boldsymbol{X_{ijt-1}}\gamma\right] * \eta_{ijt}. \tag{8}$$

This equation is similar to the 3WFE gravity equation presented earlier, with the important difference that it does not yet estimate a coefficient for cultural similarity. As before, δ_{it} are origin-time fixed effects, δ_{jt} are destination time fixed effects and δ_{ij} are time-invariant (directional) corridor fixed effects.[48] The vector $\boldsymbol{X_{ijt-1}}$ contains lags of observed time-varying covariates such as networks, income distance, and bilateral agreements.

From the first-stage decomposition, I recover fitted values of the corridor fixed effects, $\widehat{\delta_{ij}}$, which represent unbiased, bilateral corridor-specific migration costs for international and domestic corridors. As explained earlier, these corridor fixed effects capture relative migration costs and differences in cost levels across corridors. In the second stage, I use the exponentiated recovered fixed effects as the dependent variable to parameterize their observable components. The second-stage estimation equation is

$$\exp\widehat{\delta_{ij}} = exp\left[\delta_i + \delta_j + \beta_1 \ln cultsim_{ij} + \boldsymbol{X}_{ij}\gamma\right] * \eta_{ij}\ , \tag{9}$$

where $cultsim_{ij}$ represents time-invariant (or slow-moving) cultural similarity between i and j. In the second stage, $\boldsymbol{X_{ij}}$ is a vector of time-invariant covariates such common language, distance, common border, and past colonial relationships. It also contains an international corridor dummy, $intl_{ij}$, with value 1 if $i \neq j$ and value 0 if $i = j$. The second stage also includes country fixed effects, δ_i and δ_j (e.g., Spornberger 2022). I follow the literature and estimate both stages with the Poisson pseudo maximum likelihood estimator (e.g., Spornberger 2022; for PPML see Santos Silva and Tenreyro 2006).

[48] Ideally, the first stage decomposition would yield corridor fixed effects estimates for each time period. This would allow the second-stage dependent variable, bilateral migration costs, to vary not only between corridors but also over time. However, such a fully saturated decomposition is not possible because there are insufficient degrees of freedom to decompose my N^2*T panel of N countries and T=6 periods into 2*N*T directional FE (origin-year + destination-year) and 6*N^2 corridor-specific FE. To address this, Frensch and Rindler (2023) suggest letting the corridor FE vary at a lower frequency. In an annual panel they let their pair FE vary over a three-year period. However, because I only have six time periods in my panel, I do not attempt this more fine-grained decomposition. Hence, my first-stage decomposition is unsaturated.

Table 6.3 Effects of cultural similarity on migration using two-step fixed effects approach

Dependent variable:									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
coll./ind. similarity		0.05			0.11				
		(0.06)			(0.08)				
duty/joy similarity			0.81***		0.76***				
			(0.24)		(0.27)				
dis-/trust similarity				0.27***	0.23**				
				(0.10)	(0.10)				
religious similarity						0.46***			0.30***
						(0.15)			(0.12)
linguistic similarity							1.71***		1.32***
							(0.28)		(0.34)
genetic similarity								18.26***	13.83***
								(5.34)	(4.74)
intl. corridor		-5.13***	-5.26***	-5.34***	-5.11***	-5.51***	-5.19***	-5.43***	-5.09***
		(0.16)	(0.17)	(0.17)	(0.17)	(0.20)	(0.21)	(0.23)	(0.26)
comm. official lang.		0.52***	0.49***	0.52***	0.49***	0.77***	0.67***	0.80***	0.60***
		(0.10)	(0.10)	(0.10)	(0.10)	(0.13)	(0.11)	(0.14)	(0.12)
comm. spoken lang.		0.75***	0.73***	0.77***	0.75***	0.83***	0.78***	0.85***	0.78***
		(0.10)	(0.10)	(0.09)	(0.10)	(0.09)	(0.09)	(0.10)	(0.10)
distance (km)		-0.53***	-0.50***	-0.40***	-0.55***	-0.20***	-0.17***	-0.18**	-0.20***
		(0.08)	(0.09)	(0.08)	(0.08)	(0.06)	(0.04)	(0.07)	(0.07)
common border		0.28**	0.39***	0.39***	0.30**	0.75***	0.62***	0.64***	0.49**
		(0.13)	(0.13)	(0.13)	(0.13)	(0.19)	(0.18)	(0.23)	(0.22)
colony ever		1.00***	1.10***	0.98***	0.97***	0.66***	0.86***	0.71***	0.77***
		(0.16)	(0.15)	(0.16)	(0.15)	(0.15)	(0.16)	(0.16)	(0.16)
migrant networks	0.21***								
	(0.02)								
income distance	-0.04								
	(0.03)								
EU/EFTA	0.54***								
	(0.10)								
FTA	0.02								
	(0.04)								
Intercept	15.03***	1.80***	1.40***	1.07**	1.54***	0.11	-0.61**	-12.57***	-9.90***
	(0.40)	(0.40)	(0.46)	(0.42)	(0.40)	(0.37)	(0.28)	(3.81)	(3.33)
N	60130	9296	10735	10147	7887	46483	60130	51765	42169
intl. mig. captured	0.92	0.34	0.35	0.33	0.29	0.90	0.92	0.82	0.82
corridors	11026	4501	4885	4678	4222	10234	11026	9255	9179
origins	202	98	102	103	97	185	202	202	184
destinations	202	97	102	103	96	185	202	202	184
periods	6	5	5	6	5	5	6	6	5
pseudo R-sq.	0.9999	0.7592	0.7598	0.7589	0.7595	0.7892	0.7903	0.8007	0.7962

Notes: Column (1) first-stage decomposition of migration flows with corridor FE. Columns (2)-(9) gravity regressions of time-invariant and slowly moving determinants on exponentiated recovered corridor fixed effects from (1). Table shows parameter estimates obtained from Poisson-Pseudo-Maximum-Likelihood (PPML) estimation using the ppmlhdfe routine, which automatically drops singletons and observations that are separated by a fixed effect (Correia et al. 2020,2021). All continuous variables enter the estimation in logs, $\ln(x+1)$ if the variable takes values smaller than 1. Standard errors in parentheses are clustered at migration corridor level. *** p<.01, ** p<.05, * p<.1.

Table 6.3 presents the results of the two-step fixed effects (TSFE) procedure. Column (1) shows the results of the first stage ANOVA decomposition of 5-year migration flows into origin-year, destination-year, and corridor fixed effects, including time-varying bilateral control variables. The first-stage results suggest a very high fit between the three-way fixed effects model and the migration flows, which indicates that bias due to unaccounted migration determinants is negligible.

Columns (2)-(9) show results of the second stage regressions, which parameterize corridor-specific determinants of migration, using the recovered fitted-values of the corridor fixed effects from the first stage as dependent variable and using cultural similarity measures as well as time-invariant control variables as independent variables. I consecutively introduce the similarity measures, which now include linguistic and genetic similarity.

In contrast to the 3WFE results, the TSFE results suggest a positive relation between cultural similarity and migration. All measures, except coll./ind. similarity, have positive and statistically significant parameter coefficients. Let me turn to the time-invariant measures first, as they were the reason for implementing TSFE initially. The estimate of 1.71 on linguistic similarity suggests that a 10% increase in native language similarity between countries is associated with a 17.1% increase in migration between them. As this result was obtained while controlling for common official and spoken languages, linguistic similarity is associated with higher migration beyond improved communication between individuals and country officials, respectively.

Genetic similarity, the other time-invariant measure, yields very large, positive, and significant estimates. The estimate of 18.26 suggests that, on average, observing corridors that are 10% more similar with respect to genetic ancestry is associated with almost a tripling of migration. When interpreting these strong effects, one should view them in light of the unique distribution of genetic similarity, which varies only between 0.89 and 1. For instance, while the genetic similarity between the Netherlands and Belgium is measured at 1, the lowest similarity between the Netherlands and other countries is slightly above 0.93 (Solomon Islands, Paraguay). So, a 10% change in genetic similarity between populations spans roughly the entire range of genetic similarity values and, therefore, the full range of possible destinations. Although this puts the effect size into perspective, the results suggest that genetic similarity is strongly associated with migration decisions.

Turning to the time-varying measures, I find that in contrast to the 3WFE, the effect of religious similarity obtained with TSFE is positive rather than negative. Moreover, duty/joy and distrust/trust similarity are positive and significant, while collectivism/individualism similarity is positive but not statistically significant. These results strongly suggest that there is a difference between the effects of cultural similarity obtained from different methods (see the discussion in Section 7).

The TSFE results in Table 6.3 are robust to including year fixed effects in the second stage to account for common time trends (results not reported). They are also robust to estimating the second stage on a sample of 3,260 non-missing observations for all similarity measures (Table 8.8, Appendix 8.6). The only qualitative difference is that, in this smaller sample, coll/ind similarity also has a statistically significant, positive effect.

Table 8.9 (Appendix 8.5) shows the influence of different (sets of) control variables on the TSFE main results. The traditional gravity variables, distance, common border and colony ever, and the intl. corridor dummy have the strongest impact on model fit and the size of the cultural similarity estimates. The common spoken and common official language controls have a smaller impact. The importance of the intl. corridor dummy should be noted when estimating migration gravity with domestic corridors. First, the strong and highly significant negative effect suggests that, in line with common sense, people generally tend to stay at home and that international migration is not a prominent option in their choice set. After all, international migration is an expensive investment in one's human capital. Second, the inclusion of the intl. corridor dummy "corrects" the sign of the common border variable from negative to positive.

6.3.2 The role of historical cultural similarity

As discussed, one of the advantages of the 3WFE approach is that the corridor fixed effects account for unobserved migration costs between countries. The TSFE approach presents a way to re-parameterize these unobserved migration costs by estimating the effects of, for example, time-invariant explanatory variables on migration. However, this re-parameterization may not account for all unobserved migration costs. Therefore, if some unexplained costs correlate with the cultural variables of interest, their estimates may be subject to omitted variable bias.

This section addresses one possible source of omitted variable bias in the TSFE estimates of cultural similarity: cultural persistence and the historical effects of cultural similarity. The TSFE estimates from the previous section may be decomposed into historical and contemporaneous effects of cultural similarity. Historical effects are brought about by the levels of cultural similarity between countries in the time before the period under investigation (before 1990); contemporaneous effects are brought about by similarity levels during the period under investigation (1990-2019). To disentangle these effects, I implement additional specifications of the TSFE approach proposed by Frensch, Fidrmuc, and Rindler (2023). Note: the 3WFE approach produces contemporaneous effects as the corridor fixed effects absorb historical levels of cultural similarity.

To disentangle long-term, historical effects from contemporaneous effects, I follow Frensch, Fidrmuc and Rindler (2023) and use the recovered corridor fixed effects from early periods in the panel as a regressor in the second stage of later periods.[49] More specifically I conduct the first stage decomposition on two separate subsamples denoted by T1 (decade 1: 1990-1999)

[49] See discussion of a pre-sample mean estimator in Blundell, Griffith, and Windmeijer (2002). The typical way of addressing persistence is to implement dynamic gravity approaches by including the lagged dependent variable among the regressors and controlling for Nickel bias using appropriate difference or system GMM estimators. Yet, this approach may not sufficiently mitigate the bias in my "small T, large N" panel (Roodman 2009). Additionally, the longer-term effects of culture may not be appropriately captured by including lagged migration flows (cf. Frensch, Fidrmuc, and Rindler 2023). Besides, my "migrant network" control variable (in levels, bilateral migrant stocks at $t-1$) already accounts for the more recent migration history between countries.

and T2 (decades 2+3: 2000-2019). This yields two sets of estimates of bilateral, corridor-specific migration determinants, $\exp(\hat{\delta}_{ij}^{T1})$ and $\exp(\hat{\delta}_{ij}^{T2})$. I then perform the second stage parametrization by regressing $\exp(\hat{\delta}_{ij}^{T2})$ on my (time-invariant) explanatory variables including the log of the estimated T1 determinants, i.e., $\hat{\delta}_{ij}^{T1}$, as an additional regressor. Thus, the second stage equation becomes:

$$\exp \hat{\delta}_{ij}^{T2} = exp\left[\delta_i + \delta_j + \hat{\delta}_{ij}^{T1} + \beta_1 \ln cultsim_{ij} + \boldsymbol{X}_{ij}\gamma\right] * \eta_{ij} , \quad (10)$$

The idea behind this is that the estimated migration determinants $\hat{\delta}_{ij}^{T1}$ contain the entire history of migration determinants until 1999, including bilateral cultural similarity between countries. They therefore account for the predetermined component of the regressors in the second step estimation. This allows us to disentangle historical from contemporaneous effects: any impact of the remaining explanatory variables, including $cultsim_{ij}$, represent contemporaneous effects; the difference between the estimates in the previous subsection and the new estimates represent historical effects.

The results in Table 6.4 show that the T1 fixed effects, $\hat{\delta}_{ij}^{T1}$, obtained from the first-stage decomposition in column (1), have a positive and significant effect on migration across all second-stage specifications (columns 3-10). They represent the effect of historical drivers of migration (see discussion in Section 7). The estimates on all other variables denote their contemporaneous impact. For better comparison, I report TSFE results without inclusion of $\hat{\delta}_{ij}^{T1}$ in the period 2000-2019 in Table 8.10 (Appendix 8.6). There, both collectivism/individualism and duty/joy similarity have positive and significant effects on migration. However, after including the control for historical drivers, the contemporaneous effects on migration are statistically insignificant. This suggests that the initial TSFE results were largely driven by historical levels of coll/ind and duty/joy similarity. Distrust/trust similarity, keeps its positive and significant effect, although it is smaller in size. Also the estimates on the proxy measures of cultural similarity are smaller in size compared to before and the results in the Appendix.

Table 6.4 Contemporaneous effects of cultural similarity on migration using two-step fixed effects approach

Dependent variable:										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
coll./ind. similarity			0.02			-0.01				
			(0.02)			(0.03)				
duty/joy similarity				0.04		-0.12				
				(0.05)		(0.08)				
dis-/trust similarity					0.11***	0.13***				
					(0.04)	(0.05)				
religious similarity							0.15*			0.09
							(0.09)			(0.10)
linguistic similarity								0.39**		0.34*
								(0.17)		(0.20)
genetic similarity									4.01	2.59
									(3.09)	(3.37)
$\widehat{\delta_{ij}^{T1}}$, historical drivers			0.65***	0.65***	0.65***	0.65***	0.72***	0.71***	0.70***	0.69***
			(0.03)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)
intl. corridor			-1.24***	-1.37***	-1.36***	-1.33***	-0.74***	-0.74***	-0.77***	-0.73***
			(0.23)	(0.21)	(0.22)	(0.25)	(0.24)	(0.24)	(0.24)	(0.23)
comm. official lang.			0.29**	0.33**	0.32**	0.31*	0.34***	0.30***	0.24***	0.19**
			(0.14)	(0.14)	(0.15)	(0.17)	(0.08)	(0.08)	(0.08)	(0.08)
comm. spoken lang.			0.18	0.12	0.16	0.20	0.08	0.09	0.13	0.14
			(0.13)	(0.12)	(0.13)	(0.15)	(0.09)	(0.09)	(0.09)	(0.09)
distance (km)			-0.09**	-0.05	-0.06	-0.07*	-0.07**	-0.07**	-0.10***	-0.10***
			(0.04)	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)
common border			-0.00	0.06	0.07	0.04	-0.01	-0.01	-0.13	-0.15
			(0.11)	(0.11)	(0.11)	(0.12)	(0.13)	(0.13)	(0.13)	(0.13)
colony ever			-0.11	-0.05	-0.14	-0.09	-0.17	-0.10	-0.10	-0.09
			(0.11)	(0.12)	(0.12)	(0.12)	(0.11)	(0.11)	(0.12)	(0.12)
migrant networks	-0.09*	0.02								
	(0.05)	(0.02)								
income distance	-0.03	-0.04								
	(0.05)	(0.04)								
EU/EFTA		0.38***								
		(0.13)								
FTA	0.28***	-0.04								
	(0.05)	(0.07)								
Intercept	20.65***	18.66***	0.10	-0.05	-0.10	-0.02	0.25*	0.14	-2.35	-1.43
	(0.89)	(0.40)	(0.21)	(0.21)	(0.23)	(0.24)	(0.13)	(0.14)	(2.13)	(2.30)
N	13288	42872	4373	5336	4559	3634	19152	26472	23136	17331
intl. mig. captured	0.75	0.95	0.28	0.30	0.27	0.24	0.77	0.76	0.69	0.71
corridors	6644	10994	2558	2832	2630	2354	6384	6618	5784	5777
origins	154	202	72	75	75	71	147	154	154	147
destinations	154	202	71	75	75	70	147	154	154	147
periods	2	4	4	4	4	4	3	4	4	3
pseudo R-sq.	0.9999	0.9999	0.7698	0.7745	0.7660	0.7634	0.8514	0.8534	0.8563	0.8537

Notes: Column (1) first-stage decomposition of migration flows with corridor FE in period T1 (1990-1999). Column (2) first-stage decomposition of migration flows with corridor FE in period T2 (2000-2019). Columns (3)-(10) gravity regressions of time-invariant determinants on exponentiated recovered corridor fixed effects from (2) using log of T1 fixed effects, $\widehat{\delta_{ij}^{T1}}$, as additional regressor. Table shows parameter estimates obtained from Poisson-Pseudo-Maximum-Likelihood (PPML) estimation using the ppmlhdfe routine, which automatically drops singletons and observations that are separated by a fixed effect (Correia et al. 2020,2021). All continuous variables enter the estimation in logs, $\ln(x+1)$ if the variable takes values smaller than 1. Standard errors in parentheses are clustered at country pairs. *** p<.01, ** p<.05, * p<.1.

6.3.3 *The effects of cultural similarity using matching econometrics*

In this section, I use matching econometrics as an alternative to the 3WFE and TSFE approaches to estimating the effects of bilateral (cost) factors on international migration flows. Following a methodology developed for trade flows by Baier and Bergstrand (2009; see also Kohl and Trojanowska 2015), the idea is to match similar corridors, except with respect to culture, and then compare the effect that cultural similarity has on migration in these corridors. The methodological details are described in Appendix 8.7.

Table 6.5 shows the results from nearest neighbor matching (NNM) with $k = 3$ neighbors and replacement in each period separately. The coefficients represent the average treatment effect (ATE) on migration of being culturally similar vs culturally dissimilar with respect to the similarity measure in the column headers. They were obtained after matching on a set of familiar control variables[50] as well as on recovered origin-year and destination-year fixed effects from a first-stage ANOVA decomposition. I use the recovered fixed effects to account for multilateral resistance (origin-year FE) and unobserved heterogeneity at destination (destination-year) in the matching process. This is a departure from the original "gravity" matching approach by Baier and Bergstrand (2009), who used a Taylor-series decomposition to approximate inward and outward multilateral resistance terms.

The results are as follows. First, the estimates on similarity regarding coll/ind, duty/joy, and distrust/trust attitudes reflect the mixed evidence on the similarity hypothesis reported for the 3WFE approach. Only a few estimates in the first three columns are statistically significant; if they are, they have negative signs. Second, the effects regarding religious similarity are not in line with the 3WFE results. Here, I find positive and significant effects throughout, although the effect becomes weaker over time. Finally, linguistic and genetic similarity show positive and significant estimates in almost all periods.

[50] The covariates are migrant networks, income distance, geodesic distance (in km), common language, common border, colony ever, dummy variables for bilateral agreements, and the international corridor dummy

Table 6.5 Cultural effects using nearest neighbor matching

Dependent variable: Migration flows, ln(x+1)						
	coll/ind	duty/joy	dis-/trust	linguistic	religious	genetic
1990			-0.10	-0.03	0.10**	0.10***
			(0.15)	(0.08)	(0.04)	(0.04)
			N=465	N=7,217	N=6,737	N=6,295
			n=218	n=4,764	n=4,217	n=3,573
1995	0.06	0.05	0.06	0.04	0.10***	-0.01
	(0.10)	(0.10)	(0.08)	(0.06)	(0.03)	(0.08)
	N=1,403	N=1,417	N=1,479	N=9,992	N=9,520	N=8,773
	n=649	n=745	n=757	n=6,817	n=5,815	n=5,540
2000	-0.30***	-0.05	-0.05	0.17***	0.10***	0.09**
	(0.09)	(0.09)	(0.07)	(0.06)	(0.03)	(0.04)
	N=1,443	N=1,397	N=1,555	N=10,389	N=9,939	N=9,129
	n=561	n=778	n=860	n=7,068	n=5,818	n=5,837
2005	-0.10	-0.14***	0.10	0.08	0.06*	0.03
	(0.12)	(0.06)	(0.07)	(0.06)	(0.03)	(0.04)
	N=1,677	N=2,605	N=1,470	N=10,633	N=10,103	N=9,200
	n=604	n=1,467	n=679	n=7,265	n=5,889	n=5,902
2010	-0.28***	-0.03	-0.04	0.13***	0.06*	0.07*
	(0.05)	(0.05)	(0.04)	(0.04)	(0.03)	(0.04)
	N=3,408	N=3,687	N=3,509	N=11,064	N=10,245	N=9,281
	n=1,345	n=2,140	n=1,703	n=7,613	n=5,995	n=5,974
2015	-0.28***	0.06	0.03	0.16***		0.13***
	(0.09)	(0.15)	(0.07)	(0.05)		(0.04)
	N=1,368	N=1,631	N=1,672	N=11,440		N=9,371
	n=541	n=921	n=837	n=7,864		n=6,061

Notes: Non-parametric estimates of the ATE of $cultsim = 1$ vs $cultsim = 0$ using $k = 3$ nearest-neighbor matching in the periods indicated in the left-most column. Treatment variable created by splitting the sample at the median of each similarity measure. SEs in parenthesis are adjusted for continuous variables bias. N denotes the matched sample; n indicates the number of treated corridors with $cultsim = 1$. The dependent variable is in logs (ln x+1). Likewise, all continuous matching variables enter in logs. *** p<.01, ** p<.05, * p<.1

Similar to the TSFE estimates, the estimates from matching may be affected by historical drivers. They may include effects of historical, pre-1990 similarity and contemporaneous similarity in the 1990-2019 period. To disentangle these effects, I follow a similar strategy as in the TSFE approach earlier and add the T1 fixed effects, $\hat{\delta}_{ij}^{T1}$, which I retrieved earlier for the TSFE approach, to the list of matching covariates and re-estimate the ATEs for the remaining periods from 2000-2019.

The results in Table 6.6 show that controlling for historical drivers, as captured by $\hat{\delta}_{ij}^{T1}$, does not affect the estimates of attitudinal similarity greatly but attenuates many of the estimates of the proxies of cultural similarity. In particular, it renders estimates of religious similarity, which showed positive and significant estimates earlier, statistically insignificant. This implies, for instance, that linguistic similarity has a more significant contemporaneous impact on migration than religious and genetic similarity. Collectivism/individualism similarity continues to have a negative and significant impact after controlling for longer-term cultural

Table 6.6 Contemporaneous effects using nearest neighbor matching

Dependent variable: Migration flows, ln(x+1)						
	coll/ind	duty/joy	dis-/trust	linguistic	religious	genetic
2000	-0.27***	-0.04	-0.01	0.17***	0.04	0.16***
	(0.10)	(0.09)	(0.07)	(0.05)	(0.03)	(0.04)
	N=1,354	N=1,303	N=1,452	N=9,291	N=8,899	N=8,192
	n=516	n=724	n=785	n=6,493	n=5,293	n=5,428
2005	-0.02	-0.12**	0.02	0.02	0.03	-0.02
	(0.11)	(0.06)	(0.07)	(0.05)	(0.03)	(0.04)
	N=1,585	N=2,489	N=1,434	N=9,500	N=9,032	N=8,257
	n=555	n=1,405	n=652	n=6,669	n=5,343	n=5,489
2010	-0.27***	-0.02	-0.09**	0.13***	0.03	-0.03
	(0.06)	(0.05)	(0.04)	(0.05)	(0.04)	(0.04)
	N=3,049	N=3,309	N=3,144	N=9,601	N=9,032	N=8,257
	n=1,132	n=1,918	n=1,466	n=6,754	n=5,387	n=5,489
2015	-0.26***	0.09	0.08	0.13***		0.08*
	(0.10)	(0.15)	(0.07)	(0.05)		(0.04)
	N=1,271	N=1,522	N=1,560	N=9,780		N=8,331
	n=494	n=871	n=788	n=6,880		n=5,561

Notes: Non-parametric estimates of the ATE of $cultsim = 1$ vs $cultsim = 0$ using $k = 3$ nearest-neighbor matching in the periods indicated in the left-most column. Treatment variable created by splitting the sample at the median of each similarity measure. SEs in parenthesis are adjusted for continuous variables bias. N denotes the matched sample; n indicates the number of treated corridors with $cultsim = 1$. The dependent variable is in logs (ln x+1). Likewise, all continuous matching variables enter in logs. *** p<.01, ** p<.05, * p<.1

history. Overall, the matching estimates support the view that the effect of cultural similarity on migration is not always positive.

7 Discussion and Conclusion

This study presents a state-of-the-art assessment of whether cultural similarity increases international migration using various empirical methods to estimate a theory-consistent gravity model of migration and recent data on a variety of cultural similarity measures. Earlier studies in the field were not equipped to estimate gravity models that allow for causal interpretation of their estimates. To improve upon these findings, I use three different approaches – 3WFE, TSFE, and a matching estimator – to estimate theory-consistent gravity models of international migration, with domestic flows, while minimizing omitted variables bias and simultaneity bias. The results I find differ quite drastically between methods. In what follows, I elaborate on several potential explanations for these difference placing my findings in the broader context.

Let me begin with addressing the apparently contradicting results between the 3WFE and TSFE approaches – in particular the TSFE approach without control for historical drivers. To interpret the results correctly, one needs to consider that the 3WFE estimates can be interpreted causally, while the second-stage TSFE estimates merely report correlations. By using corridor (asymmetric pair) fixed effects, the 3WFE approach allows for an estimation of temporal panel effects showing how, *within* corridors, migration flows adjust on average to changes in cultural similarity between origin and destination countries over time. In contrast, the second stage of the TSFE approach pools the observations across all periods in the sample. This means that the resulting estimates do not just capture variation *within* corridors, but also variation *between* corridors. The difference between the TSFE and 3WFE effects is then analogous to the difference between, say, pooled OLS and linear fixed effects panel regressions. Thus, the second-stage TSFE results are not suited for a causal interpretation of the effect of cultural similarity on migration.

Based on the 3WFE results I conclude that, on average, cultural similarity between countries does not increase migration. This is in contrast to evidence reported by earlier studies in the field, which suggested a positive effect (e.g., Belot and Ederveen 2012; Adserà and Pytliková 2015; Wang, De Graaff, and Nijkamp 2016). Thus, initially, the country-level evidence I provide in this study is not in line with the theoretical hypothesis that cultural similarity increases migration. Yet, as my analysis has also shown, these non-positive average effects could be brought about by cultural selection and sorting of migrants. In fact, the results of estimating the effects of cultural similarity at different levels of individualism and religious tolerance in the origin country are consistent with the idea that migrants are attracted by destinations that are culturally similar to their *personal* cultural beliefs, but not by destinations that are similar to the *average* cultural beliefs in their home country. Therefore, cultural similarity increases migration, but the effect is not straightforward and can, on the aggregate, be masked by (cultural) selection and sorting. This conclusion is broadly in line with previous literature documenting (cultural) selection and sorting as powerful determinants of migration processes (e.g., Docquier, Tansel, and Turati 2020; Grogger and Hanson 2011; Belot and Hatton 2012).

What, then, can be learned from the TSFE results? First, as established earlier, the results of the parameterization in the second stage capture a large share of *between* corridor variation in cultural similarity and migration. While between variation may not be relevant in other applications, it has relevance in the case of migration because migrants' location decisions do not just depend on temporal comparisons within corridors, but largely also on comparisons across corridors. In other words, they are not independent of alternative destinations. Thus, the TSFE results suggest that there is a positive *cross-corridor* element in the effect of almost all measures of cultural similarity on migration – except for individualism, which did not yield significant coefficients (Table 6.3).

Second, we can learn about the role of historical drivers in the relation between cultural similarity and migration. Findings showed that controlling for historical drivers attenuates the strong positive correlations of the initial TSFE results, and more moderate, contemporaneous correlations remain (Table 6.4). For an explanation of what may drive these results consider that the recovered corridor fixed effects from the 1990-1999 period, which I used as controls of historical drivers in the second period of my panel (2000-2019), capture the entire migration history of each corridor before 1999. This history includes all bilateral drivers and costs of migration that have accumulated between countries over time and that are not captured yet by the first-stage covariates, including the directional time-varying fixed effects. Candidates are, for instance, long-standing labor agreements between countries but also persistent levels of cultural similarity between countries which may go back even deeper in time. I will briefly focus on the latter.

Cultural similarity is persistent, because culture itself is persistent. Culture, the customary beliefs about which values and norms are socially desirable, is transmitted "fairly unchanged from generation to generation" (Guiso, Sapienza, and Zingales 2006, 23). There are numerous examples of present-day differences in economic outcomes that have roots deeply entrenched in a society's cultural history.[51] Although on the one hand exogenous institutional shocks can have long-term impacts on economic outcomes through individuals' cultural beliefs[52], and on the other hand cultural change over time is a worldwide phenomenon (e.g. Inglehart 1997), unique cultural differences between countries persist, suggesting the presence of remote historical drivers (e.g. Beugelsdijk and Welzel 2018).

Coming back to this study, it is plausible to think that, in historic times, when labor markets were globally less connected, travel distances were shorter, and global welfare levels much lower, migration was predominantly achievable for the affluent (cf., The Mobility Transition Hypothesis; Zelinsky 1971). Moreover, the affluent who migrated likely belonged to the majority group in their home country, seeking destinations that are culturally similar to their home and contributing to positive average effects of cultural similarity. These historical preferences could be handed down from one migrant generation to the next and, thus, be still reflected in contemporaneous ideas about what constitutes a "good" destination country. This

[51] For example, present-day economic performance in several African countries is related to mistrust that can be traced back to the transatlantic and Indian Ocean slave trades (Nunn 2008; Nunn and Wantchekon 2011). Other examples include that traditional agricultural practices have shaped historical gender roles and so influence the evolution of present-day gender norms (Alesina, Giuliano, and Nunn 2013); or that cultural differences of the origin countries of their parents can determine second-generation migrants' economic decisions (e.g., Alesina and Giuliano 2010; 2011; Fernández and Fogli 2006; Kleinhempel, Klasing, and Beugelsdijk 2022). For a review, see e.g., Spolaore and Wacziarg (2013).

[52] For instance, the slave-trades investigated by Nunn (2008) and Nunn and Wantchekon (2011) are seen as exogenously inflicted, institutional shocks that deeply affected societies for centuries. Findings along similar lines are reported by Guiso, Sapienza and Zingales (2016), who show how the institutional environment of self-governance of Italian free-city states has generated long term persistence in their economic development through civic attitudes until today. Moreover, change in cultural beliefs and attitudes might occur relatively quickly, as for example Alesina and Fuchs-Schündeln (2007) show in their investigation of on the redistribution preferences of people in East-Germany after reunification. In their paper they suggest that preferences of West and East Germans disappear within 20 to 40 years.

is how historically persistent cultural similarity could drive the strong and positive correlations in the initial TSFE procedure without the control of historical drivers – even in a modern panel as mine.

Since then, however, migration became increasingly available also for members of minority groups through rising welfare levels, longer travel distances, and globalization of labor markets. As a consequence, the effects of cultural selection and sorting may have become much stronger, resulting in the attenuated TSFE correlations when accounting for historical drivers. This line of argument is consistent with the contemporaneous effects estimated with the 3WFE model. All the measures of cultural similarity in my study allow for this interpretation of their TSFE results, except individualism-collectivism similarity which did not show significant correlations initially. However, given evidence for selection on individualism during mass emigration from Scandinavia to North America in the 19th century (Knudsen 2022), it could be that selection and sorting on this particular cultural trait has a longer historical past than selection and sorting on other traits. Such arguments are, of course, highly speculative and so I leave a thorough investigation of globalization effects on the influence of cultural similarity on migration for future research. Nonetheless, the above discussion shows that, even though the TSFE results do not have a causal interpretation, they should not be rashly dismissed.

Finally let me turn to the results from the matching estimator. The idea behind matching is to compare corridors that are similar to each other except with respect to their level of cultural similarity. To achieve estimation of theory-consistent gravity effects, I account for multilateral resistance by including recovered directional fixed effects (origin-year, destination-year) among the matching covariates. This improves on the original proposal to include Taylor approximations of multilateral resistance terms (Baier and Bergstrand 2009) and ensures that the matched corridors are "gravity-similar" by accounting for the influence exerted by the attractiveness of other destinations. As the corridors are similar except for their level of cultural similarity, the results can be interpreted causally. The difference to the 3WFE estimates is that the matching estimates do not explain temporal changes *within* corridors but rather variation *between* corridors. So, instead of capturing cultural similarity effects *over time*, the matching estimator captures similarity effects *across corridors at distinct points in time*. Additionally, matching also allows to recover estimates for the time-invariant cultural similarity variables, linguistic and genetic similarity.

The results (after controlling for historical drivers) show that individualism-collectivism, duty-joy, distrust-trust, and religious similarity do not increase migration flows across similar corridors on average. In fact, country-level similarity regarding individualism-collectivism even decreases migration between countries. Although not tested in the case of the matching estimator, I speculate that these non-positive average effects are subject to cultural selection and sorting of migrants too, just as the earlier results. In contrast, the two time-invariant measures, linguistic and genetic similarity, lend more support to the hypothesis of a positive effect of cultural similarity, even on the average. In particular, linguistic similarity, which captures how similar countries are with respect to their native language communities, has a strong positive effect on migration.

Overall, the matching results confirm the earlier observation that the relationship between cultural similarity and migration is more nuanced than previously thought: average effects do not unambiguously support the theoretical prediction that cultural similarity increases migration. The matching estimates also show that the sign of average effects depends on the specific aspect of culture that is considered (e.g., religion and cultural attitudes vs. native languages).

One of the main insights of this study is that cultural selection and sorting of migrants may, to a large extent, affect the sign and size of average effects. However, as my study only speculates about how personal cultural beliefs (values) interact with average cultural beliefs in the migration choice, I believe that there is room for future research on cultural selection and sorting. Such work could be empirical, exploiting data that follows migrants over time and borders (e.g., TRANSMIT at HU Berlin). On the other hand, I also see room for theoretical work. This could be in line with typical selection approaches (Borjas 1987; Grogger and Hanson 2011), but could also include different forms of utility maximization and reasons for migration than currently presented in the standard model of migration. For example, it has been argued that people actively seek out ties to similar others because these ties give meaning to their lives, by creating and reinforcing social identities (e.g., Stets et al. 2021). If cultural beliefs make up at least part of people's social identities, migrants may choose their eventual destination among other potential destinations based on how well the cultural beliefs at this particular destination correspond to the beliefs associated with their social identity. So, in addition to being a human capital investment where culturally similar destinations are instrumental to higher returns to human capital, migration could also be seen as an investment in social identity, where cultural similarity is an end in itself.[53]

[53] There is a large literature in behavioral economics according to which economic behaviors, to which I also count migration decisions, can be interpreted as social identity investments (e.g., Akerlof and Kranton 2010; Bénabou and Tirole 2011; Bursztyn, Fujiwara, and Pallais 2017).

References

Abadie, Alberto, and Guido W. Imbens. 2006. "Large Sample Properties of Matching Estimators for Average Treatment Effects." *Econometrica* 74 (1): 235–67. https://doi.org/10.1111/j.1468-0262.2006.00655.x.

———. 2011. "Bias-Corrected Matching Estimators for Average Treatment Effects." *Journal of Business & Economic Statistics* 29 (1): 1–11.

Abel, Guy J. 2019. "Bilateral International Migration Flow Estimates for 200 Countries (Data)." figshare. https://doi.org/10.6084/m9.figshare.c.4470464.

Abel, Guy J., and Joel E. Cohen. 2019. "Bilateral International Migration Flow Estimates for 200 Countries." *Scientific Data* 6 (82): 1–13. https://doi.org/10.1038/s41597-019-0089-3.

Adserà, Alícia, and Mariola Pytliková. 2015. "The Role of Language in Shaping International Migration." *Economic Journal* 125 (586): F49–81. https://doi.org/10.1111/ECOJ.12231.

Agnosteva, Delina E., James E. Anderson, and Yoto V. Yotov. 2019. "Intra-National Trade Costs: Assaying Regional Frictions." *European Economic Review* 112: 32–50. https://doi.org/10.1016/j.euroecorev.2018.11.008.

Akerlof, George A., and Rachel E. Kranton. 2010. *Identity Economics - How Our Identities Shape Our Work, Wages and Well-Being*. Princeton: Princeton University Press. https://doi.org/10.1515/9781400834181.

Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat, and Romain Wacziarg. 2003. "Fractionalization." *Journal of Economic Growth* 8 (2): 155–94.

Alesina, Alberto, and Nicola Fuchs-Schündeln. 2007. "Good-Bye Lenin (or Not?): The Effect of Communism on People's Preferences." *American Economic Review* 97 (4): 1507–28. https://doi.org/10.1257/aer.97.4.1507.

Alesina, Alberto, and Paola Giuliano. 2010. "The Power of the Family." *Journal of Economic Growth* 15: 93–125. https://doi.org/10.1007/s10887-010-9052-z.

———. 2011. "Family Ties and Political Participation." *Journal of the European Economic Association* 9 (5): 817–39. https://doi.org/10.1111/j.1542-4774.2011.01024.x.

———. 2015. "Culture and Institutions." *Journal of Economic Literature* 53 (4): 898–944. https://doi.org/10.1257/jel.53.4.898.

Alesina, Alberto, Paola Giuliano, and Nathan Nunn. 2013. "On the Origins of Gender Roles: Women and the Plough." *The Quarterly Journal of Economics* 128 (2): 469–530. https://doi.org/10.1093/qje/qjt005.

Anderson, James E. 2011. "The Gravity Model." *Annual Review of Economics* 3 (1): 133–60. https://doi.org/10.1146/annurev-economics-111809-125114.

Anderson, James E., and Eric van Wincoop. 2003. "Gravity with Gravitas: A Solution to the Border Puzzle." *The American Economic Review* 93 (1): 170–92.

Azose, Jonathan J., and Adrian E. Raftery. 2019. "Estimation of Emigration, Return Migration, and Transit Migration between All Pairs of Countries." *Proceedings of the National Academy of Sciences of the United States of America* 116 (1): 116–22. https://doi.org/10.1073/pnas.1722334116.

Baier, Scott L., and Jeffrey H. Bergstrand. 2009. "Estimating the Effects of Free Trade Agreements on International Trade Flows Using Matching Econometrics." *Journal of International Economics* 77 (1): 63–76. https://doi.org/10.1016/j.jinteco.2008.09.006.

Bakker, Dik, André Müller, Viveka Velupillai, Søren Wichmann, Cecil H. Brown, Pamela Brown, Dmitry Egorov, Robert Mailhammer, Anthony Grant, and Eric W. Holman. 2009. “Adding Typology to Lexicostatistics: A Combined Approach to Language Classification.” *Linguistic Typology* 13 (1): 169–81. https://doi.org/10.1515/LITY.2009.009.

Baldwin, Richard, and Daria Taglioni. 2006. “Gravity for Dummies and Dummies for Gravity Equations.” *NBER Working Paper Series* No. 12516. https://doi.org/10.3386/w12516.

Beine, Michel, Simone Bertoli, and Jesús Fernández-Huertas Moraga. 2016. “A Practitioners’ Guide to Gravity Models of International Migration.” *The World Economy* 39 (4): 496–512. https://doi.org/10.1111/twec.12265.

Beine, Michel, Frédéric Docquier, and Çağlar Özden. 2011. “Diasporas.” *Journal of Development Economics* 95 (1): 30–41. https://doi.org/10.1016/j.jdeveco.2009.11.004.

Beine, Michel, and Christopher Parsons. 2015. “Climatic Factors as Determinants of International Migration.” *Scandinavian Journal of Economics* 117 (2): 723–67. https://doi.org/10.1111/sjoe.12098.

Belot, Michèle, and Sjef Ederveen. 2012. “Cultural Barriers in Migration between OECD Countries.” *Journal of Population Economics* 25 (3): 1077–1105. https://doi.org/10.1007/s00148-011-0356-x.

Belot, Michèle, and Timothy J. Hatton. 2012. “Immigrant Selection in the OECD.” *Scandinavian Journal of Economics* 114 (4): 1105–28. https://doi.org/10.1111/j.1467-9442.2012.01721.x.

Bénabou, Roland, and Jean Tirole. 2011. “Identity, Morals, and Taboos: Beliefs as Assets.” *The Quarterly Journal of Economics* 126 (2): 805–55. https://doi.org/10.1093/qje/qjr002.

Berggren, Niclas, Martin Ljunge, and Therese Nilsson. 2023. “Immigrants from More Tolerant Cultures Integrate Deeper into Destination Countries.” *Journal of Comparative Economics* 51 (4): 1095–1108. https://doi.org/10.1016/j.jce.2023.06.005.

Bergstrand, Jeffrey H., Mario Larch, and Yoto V. Yotov. 2015. “Economic Integration Agreements, Border Effects, and Distance Elasticities in the Gravity Equation.” *European Economic Review* 78: 307–27. https://doi.org/10.1016/j.euroecorev.2015.06.003.

Bertoli, Simone, and Jesús Fernández-Huertas Moraga. 2013. “Multilateral Resistance to Migration.” *Journal of Development Economics* 102: 79–100. https://doi.org/10.1016/j.jdeveco.2012.12.001.

———. 2015. “The Size of the Cliff at the Border.” *Regional Science and Urban Economics* 51: 1–6. https://doi.org/10.1016/j.regsciurbeco.2014.12.002.

Beugelsdijk, Sjoerd, and Chris Welzel. 2018. “Dimensions and Dynamics of National Culture: Synthesizing Hofstede With Inglehart.” *Journal of Cross-Cultural Psychology* 49 (10): 1469–1505. https://doi.org/10.1177/0022022118798505.

Beverelli, Cosimo. 2022. “Pull Factors for Migration: The Impact of Migrant Integration Policies.” *Economics & Politics* 34 (1): 171–91. https://doi.org/10.1111/ecpo.12191.

Bisin, Alberto, and Thierry Verdier. 2001. “The Economics of Cultural Transmission and the Dynamics of Preferences.” *Journal of Economic Theory* 97 (2): 298–319. https://doi.org/10.1006/jeth.2000.2678.

Blundell, Richard, Rachel Griffith, and Frank Windmeijer. 2002. "Individual Effects and Dynamics in Count Data Models." *Journal of Econometrics* 108 (1): 113–31. https://doi.org/10.1016/S0304-4076(01)00108-7.

Bodvarsson, Örn B., Nicole B. Simpson, and Chad Sparber. 2015. "Migration Theory." In *Handbook of the Economics of International Migration*, edited by Barry R. Chiswick and Paul W. Miller, 3–51. Amsterdam: North Holland.

Bodvarsson, Örn B., and Hendrik Van den Berg. 2013. "The Determinants of International Migration: Theory." In *The Economics of Immigration: Theory and Policy*, 27–57. New York: Springer. https://doi.org/10.1007/978-1-4614-2116-0.

Borjas, George J. 1987. "Self-Selection and the Earnings of Immigrants." *The American Economic Review* 77 (4): 531–53.

Bove, Vincenzo, and Gunes Gokmen. 2020. "Cultural Distance and Income Divergence over Time." *Economics Letters* 194: 1–3. https://doi.org/10.1016/j.econlet.2020.109348.

Boyd, Robert, and Peter J. Richerson. 1988. *Culture and the Evolutionary Process*. Chicago: University of Chicago Press.

Bredtmann, Julia, Klaus Nowotny, and Sebastian Otten. 2020. "Linguistic Distance, Networks and Migrants' Regional Location Choice." *Labour Economics* 65: 101863. https://doi.org/10.1016/j.labeco.2020.101863.

Bursztyn, Leonardo, Thomas Fujiwara, and Amanda Pallais. 2017. "'Acting Wife': Marriage Market Incentives and Labor Market Investments." *American Economic Review* 107 (11): 3288–3319. https://doi.org/10.1257/aer.20170029.

Caragliu, Andrea, Chiara Del Bo, Henri L.F. de Groot, and Gert Jan M. Linders. 2013. "Cultural Determinants of Migration." *Annals of Regional Science* 51 (1): 7–32. https://doi.org/10.1007/s00168-012-0531-3.

Collier, Paul, and Anke Hoeffler. 2018. "Migration, Diasporas and Culture: An Empirical Investigation." *Kyklos* 71 (1): 86–109. https://doi.org/10.1111/kykl.12163.

Correia, Sergio, Paulo Guimarães, and Thomas Zylkin. 2021. "Verifying the Existence of Maximum Likelihood Estimates for Generalized Linear Models." arXiv. http://arxiv.org/abs/1903.01633.

Correia, Sergio, Paulo Guimarães, and Tom Zylkin. 2020. "Fast Poisson Estimation with High-Dimensional Fixed Effects." *The Stata Journal: Promoting Communications on Statistics and Stata* 20 (1): 95–115. https://doi.org/10.1177/1536867X20909691.

Czaika, Mathias, and Hein de Haas. 2017. "The Effect of Visas on Migration Processes." *International Migration Review* 51 (4): 893–926. https://doi.org/10.1111/imre.12261.

Czaika, Mathias, and Christopher R. Parsons. 2017. "The Gravity of High-Skilled Migration Policies." *Demography* 54 (2): 603–30. https://doi.org/10.1007/s13524-017-0559-1.

Dao, Thu Hien, Frédéric Docquier, Chris Parsons, and Giovanni Peri. 2018. "Migration and Development: Dissecting the Anatomy of the Mobility Transition." *Journal of Development Economics* 132: 88–101. https://doi.org/10.1016/j.jdeveco.2017.12.003.

Disdier, Anne Célia, Silvio H.T. Tai, Lionel Fontagné, and Thierry Mayer. 2010. "Bilateral Trade of Cultural Goods." *Review of World Economics* 145 (4): 575–95. https://doi.org/10.1007/s10290-009-0030-5.

Docquier, Frédéric, Giovanni Peri, and Ilse Ruyssen. 2014. "The Cross-Country Determinants of Potential and Actual Migration." *International Migration Review* 48 (s1): S37–99. https://doi.org/10.1111/imre.12137.

Docquier, Frédéric, Aysit Tansel, and Riccardo Turati. 2020. “Do Emigrants Self-Select Along Cultural Traits? Evidence from the MENA Countries.” *International Migration Review* 54 (2): 388–422. https://doi.org/10.1177/0197918319849011.

Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde. 2012. “The Intergenerational Transmission of Risk and Trust Attitudes.” *The Review of Economic Studies* 79 (2): 645–77. https://doi.org/10.1093/restud/rdr027.

Dyen, Isidore, Joseph B. Kruskal, and Paul Black. 1992. “An Indoeuropean Classification: A Lexicostatistical Experiment.” *Transactions of the American Philosophical Society* 82 (5): iii–132. https://doi.org/10.2307/1006517.

Egger, Peter H., and Andrea Lassmann. 2015. “The Causal Impact of Common Native Language on International Trade: Evidence from a Spatial Regression Discontinuity Design.” *The Economic Journal* 125 (584): 699–745. https://doi.org/10.1111/ecoj.12253.

Egger, Peter H., and Sergey Nigai. 2015. “Structural Gravity with Dummies Only: Constrained ANOVA-Type Estimation of Gravity Models.” *Journal of International Economics* 97 (1): 86–99. https://doi.org/10.1016/j.jinteco.2015.05.004.

Fearon, James D. 2003. “Ethnic and Cultural Diversity by Country.” *Journal of Economic Growth* 8 (2): 195–222.

Felbermayr, Gabriel J., and Farid Toubal. 2010. “Cultural Proximity and Trade.” *European Economic Review* 54 (2): 279–93. https://doi.org/10.1016/j.euroecorev.2009.06.009.

Fernández, Raquel, and Alessandra Fogli. 2006. “Fertility: The Role of Culture and Family Experience.” *Journal of the European Economic Association* 4 (2/3): 552–61.

Frensch, Richard, Jarko Fidrmuc, and Michael Rindler. 2023. “Topography, Borders, and Trade across Europe.” *Journal of Comparative Economics* 51 (3): 816–32. https://doi.org/10.1016/j.jce.2023.02.002.

Frensch, Richard, and Michael Rindler. 2023. “Time-Varying Trade Cost (Terms) and the Distance Puzzle.” *IOS Working Paper Series* No. 399: 1–31.

Gorodnichenko, Yuriy, and Gerard Roland. 2017. “Culture, Institutions, and the Wealth of Nations.” *Review of Economics and Statistics* 99 (3): 402–16. https://doi.org/10.1162/REST.

Grogger, Jeffrey, and Gordon H. Hanson. 2011. “Income Maximization and the Selection and Sorting of International Migrants.” *Journal of Development Economics* 95: 42–57. https://doi.org/10.1016/j.jdeveco.2010.06.003.

Grohmann, Tobias, and Hanna Fromell. 2023. “Measures of International Migration Flows: Which One Should I Use?” University of Groningen - unpublished.

Gröschl, Jasmin. 2012. “Climate Change and the Relocation of Population.” In *Beiträge Zur Jahrestagung Des Vereins Für Socialpolitik 2012: Neue Wege Und Herausforderungen Für Den Arbeitsmarkt Des 21. Jahrhunderts - Session: Migration II, No. D03-V1*. ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz- Informationszentrum Wirtschaft.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2006. “Does Culture Affect Economic Outcomes?” *Journal of Economic Perspectives* 20 (2): 23–48. https://doi.org/10.1257/jep.20.2.23.

———. 2016. “Long-Term Persistence.” *Journal of the European Economic Association* 14 (6): 1401–36. https://doi.org/10.1111/jeea.12177.

Gurevich, Tamara, and Peter Herman. 2018. “The Dynamic Gravity Dataset: 1948–2016.” *U.S. International Trade Commission Economics Working Paper Series* No. 2018–02–A.

Gurevich, Tamara, Peter Herman, Farid Toubal, and Yoto V. Yotov. 2021. “One Nation, One Language? Domestic Language Diversity, Trade and Welfare.” SSRN Scholarly Paper. Rochester, NY. https://doi.org/10.2139/ssrn.3774667.

Henrich, Joseph, and Richard McElreath. 2003. “The Evolution of Cultural Evolution.” *Evolutionary Anthropology: Issues, News, and Reviews* 12 (3): 123–35. https://doi.org/10.1002/evan.10110.

Hofstede, Geert. 2001. *Culture’s Consequences: Comparing Values, Behaviors, Institutions, and Organizations Across Nations*. Thousand Oaks: SAGE Publications.

Honoré, Bo E., and Michaela Kesina. 2017. “Estimation of Some Nonlinear Panel Data Models With Both Time-Varying and Time-Invariant Explanatory Variables.” *Journal of Business & Economic Statistics* 35 (4): 543–58. https://doi.org/10.1080/07350015.2015.1123635.

Inglehart, Ronald. 1997. *Modernization and Postmodernization: Cultural, Economic and Political Change in 43 Societies*. Princeton: Princeton University Press.

Kitayama, Shinobu, Keiko Ishii, Toshie Imada, Kosuke Takemura, and Jenny Ramaswamy. 2006. “Voluntary Settlement and the Spirit of Independence: Evidence from Japan’s ‘Northern Frontier.’” *Journal of Personality and Social Psychology* 91 (3): 369–84. https://doi.org/10.1037/0022-3514.91.3.369.

Kleinhempel, Johannes, Mariko J. Klasing, and Sjoerd Beugelsdijk. 2022. “Cultural Roots of Entrepreneurship: Evidence from Second-Generation Immigrants.” *Organization Science* 34 (5): 1800–1819. https://doi.org/10.1287/orsc.2022.1645.

Knudsen, Anne Sofie Beck. 2022. “Those Who Stayed: Selection and Cultural Change in the Age of Mass Migration.” *Working Paper*, August. https://annesofiebeckknudsen.com/wp-content/uploads/2022/08/thosewhostayed.pdf.

Kohl, Tristan, and Sofia Trojanowska. 2015. “Heterogeneous Trade Agreements, WTO Membership and International Trade: An Analysis Using Matching Econometrics.” *Applied Economics* 47 (33): 3499–3509. https://doi.org/10.1080/00036846.2015.1016211.

Krieger, Tim, Laura Renner, and Jens Ruhose. 2018. “Long-Term Relatedness between Countries and International Migrant Selection.” *Journal of International Economics* 113: 35–54. https://doi.org/10.1016/j.jinteco.2018.04.002.

Lanati, Mauro, and Alessandra Venturini. 2021. “Cultural Change and the Migration Choice.” *Review of World Economics* 157 (4): 799–852. https://doi.org/10.1007/s10290-021-00418-1.

Larch, Mario, Joschka Wanner, Yoto V. Yotov, and Thomas Zylkin. 2018. “Currency Unions and Trade: A PPML Re-Assessment with High-Dimensional Fixed Effects.” *Oxford Bulletin of Economics and Statistics* 81 (3): 487–510. https://doi.org/10.1111/obes.12283.

Larch, Mario, and Yoto V. Yotov. 2023. “Estimating the Effects of Trade Agreements: Lessons from 60 Years of Methods and Data.” *CESifo Working Paper Series* No. 10267. https://doi.org/10.2139/ssrn.4359371.

Lawrence, Barbara S., and Neha Parikh Shah. 2020. “Homophily: Measures and Meaning.” *Academy of Management Annals* 14 (2): 513–97. https://doi.org/10.5465/annals.2018.0147.

Lazarsfeld, Paul F., and Robert K. Merton. 1954. “Friendship as a Social Process: A Substantive and Methodological Analysis.” In *Freedom and Control in Modern Society*, edited by Morroe Berger, Theodore Abel, and Charles H. Page, 18–66. Toronto: D. van Nostrand Company, Inc.

Lewis, M. Paul. 2009. *Ethnologue: Languages of the World*. Dallas, TX: SIL International. https://unesdoc.unesco.org/ark:/48223/pf0000130250.

Lucas, Robert E. B. 2015. “African Migration.” In *Handbook of the Economics of International Migration*, edited by Barry R. Chiswick and Paul W. Miller, 1445–1596. Amsterdam: North Holland. https://doi.org/10.1016/B978-0-444-53768-3.00026-6.

Maoz, Zeev, and Errol A. Henderson. 2013. “The World Religion Dataset, 1945-2010: Logic, Estimates, and Trends.” *International Interactions* 39 (3): 265–91. https://doi.org/10.1080/03050629.2013.782306.

Mayda, Anna Maria. 2010. “International Migration: A Panel Data Analysis of the Determinants of Bilateral Flows.” *Journal of Population Economics* 23 (4): 1249–74. https://doi.org/10.1007/sOOl.

Mayer, Thierry, and Soledad Zignago. 2011. “The GeoDist Database on Bilateral Geographical Information.” *CEPII Working Paper Series* No. 2011-25. https://doi.org/10.2139/ssrn.1994531.

McFadden, Daniel L. 1984. “Econometric Analysis of Qualitative Response Models.” In *Handbook of Econometrics*, edited by Z. Griliches and M.D. Intriligator, 1396–1457. Amsterdam: North Holland.

McPherson, Miller, Lynn Smith-Lovin, and James M Cook. 2001. “Birds of a Feather: Homophily in Social Networks.” *Annual Review of Sociology* 27 (415–44).

Melitz, Jacques, and Farid Toubal. 2014. “Native Language, Spoken Language, Translation and Trade.” *Journal of International Economics* 93 (2): 351–63. https://doi.org/10.1016/j.jinteco.2014.04.004.

Nunn, Nathan. 2008. “The Long-Term Effects of Africa’s Slave Trades.” *The Quarterly Journal of Economics* 123 (1): 139–76. https://doi.org/10.1162/qjec.2008.123.1.139.

Nunn, Nathan, and Leonard Wantchekon. 2011. “The Slave Trade and the Origins of Mistrust in Africa.” *American Economic Review* 101 (7): 3221–52. https://doi.org/10.1257/aer.101.7.3221.

Ortega, Francesc, and Giovanni Peri. 2013. “The Effect of Income and Immigration Policies on International Migration.” *Migration Studies* 1 (1): 47–74. https://doi.org/10.1093/migration/mns004.

Pedersen, Peder J., Mariola Pytlikova, and Nina Smith. 2008. “Selection and Network Effects-Migration Flows into OECD Countries 1990-2000.” *European Economic Review* 52 (7): 1160–86. https://doi.org/10.1016/j.euroecorev.2007.12.002.

Roland, Gerard. 2015. “Economics and Culture.” In *Emerging Trends in the Social and Behavioral Sciences*, edited by Robert Scott and Stephen Kosslyn, 1–18. Hoboken: John Wiley & Sons. https://doi.org/10.21098/bemp.v22i1.1035.

Roodman, David. 2009. “A Note on the Theme of Too Many Instruments.” *Oxford Bulletin of Economics and Statistics* 71 (1): 135–58.

Roy, Andrew Donald. 1951. “Some Thoughts on the Distribution of Earnings.” *Oxford Economic Papers* 3 (2): 135–46.

Ruyssen, Ilse, and Glenn Rayp. 2014. "Determinants of Intraregional Migration in Sub-Saharan Africa 1980-2000." *Journal of Development Studies* 50 (3): 426–43. https://doi.org/10.1080/00220388.2013.866218.

Ruyssen, Ilse, and Sara Salomone. 2018. "Female Migration: A Way out of Discrimination?" *Journal of Development Economics* 130: 224–41. https://doi.org/10.1016/j.jdeveco.2017.10.010.

Santos Silva, J. M. C., and Silvana Tenreyro. 2006. "The Log of Gravity." *The Review of Economics and Statistics* 88 (4): 641–58.

Sjaastad, Larry A. 1962. "The Costs and Returns of Human Migration." *Journal of Political Economy* 70 (5): 80–93.

Spolaore, Enrico, and Romain Wacziarg. 2009. "The Diffusion of Development." *The Quarterly Journal of Economics* 124 (2): 469–529. https://doi.org/10.1162/qjec.2009.124.2.469.

———. 2013. "How Deep Are the Roots of Economic Development?" *Journal of Economic Literature* 51 (2): 325–69. https://doi.org/10.1257/jel.

———. 2018. "Ancestry and Development: New Evidence." *Journal of Applied Econometrics* 33 (5): 748–62. https://doi.org/10.1002/jae.2633.

Spornberger, Julia. 2022. "EU Integration and Structural Gravity: A Comprehensive Quantification of the Border Effect on Trade." *Review of International Economics* 30 (4): 915–38. https://doi.org/10.1111/roie.12589.

Stets, Jan E., John Aldecoa, Quinn Bloom, and Joel Winegar. 2021. "Using Identity Theory to Understand Homophily in Groups." In *Identities in Action: Developments in Identity Theory*, edited by Philip S. Brenner, Jan E. Stets, and Richard T. Serpe, 285–302. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-76966-6_14.

UNDESA. 2020. *International Migrant Stock 2020*. United Nations, Department of Economic and Social Affairs, Population Division. https://www.un.org/development/desa/pd/content/international-migrant-stock.

Wang, Zhiling, Thomas De Graaff, and Peter Nijkamp. 2016. "Cultural Diversity and Cultural Distance as Choice Determinants of Migration Destination." *Spatial Economic Analysis* 11 (2): 176–200. https://doi.org/10.1080/17421772.2016.1102956.

White, Roger, and David Buehler. 2018. "A Closer Look at the Determinants of International Migration: Decomposing Cultural Distance." *Applied Economics* 50 (33): 3575–95. https://doi.org/10.1080/00036846.2018.1430337.

Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data*. Cambridge: MIT Press.

Yotov, Yoto V. 2012. "A Simple Solution to the Distance Puzzle in International Trade." *Economics Letters* 117 (3): 794–98. https://doi.org/10.1016/j.econlet.2012.08.032.

———. 2022a. "Gravity at Sixty: The Workhorse Model of Trade." *CESifo Working Paper Series* No. 9584. https://www.cesifo.org/en/publications/2022/working-paper/gravity-sixty-workhorse-model-trade.

———. 2022b. "On the Role of Domestic Trade Flows for Estimating the Gravity Model of Trade." *Contemporary Economic Policy* 40 (3): 526–40. https://doi.org/10.1111/coep.12567.

Yotov, Yoto V, Roberta Piermartini, José-Antonio Monteiro, and Mario Larch. 2016. *An Advanced Guide to Trade Policy Analysis: The Structural Gravity Model*. WTO iLibrary. http://vi.unctad.org/tpa.

Zelinsky, Wilbur. 1971. “The Hypothesis of the Mobility Transition.” *Geographical Review* 61 (2): 219–49.

8 Appendix

8.1 Previous literature on cultural similarity and migration

Table 8.1 Overview of previous studies in migration literature

Publication	Cultural similarity/difference	Effect on migration	Dependent variable	Empirical model/estimation
Belot and Ederveen (2012)	Linguistic proximity	Positive	Annual bilateral migration flows; 1990-2003; among 22 OECD countries	Gravity model; Negative binomial; destination FE; account for multilat. resistance: no
	Religious proximity	Positive		
	Attitudes-based similarity using Hofstede dimensions, Kogut-Singh index	Positive		
Adserà and Pytliková (2015)	*linguistic proximity* between official and spoken languages using a variety of proximity indices	Positive	Annual bilateral migration flows; 1980-2010, based on inflows to 30 OECD countries	Gravity model; linear regression and Poisson; destination FE, origin FE and year FE; account for multilat. resistance: no
	genetic distance (as control variable)	Negative		
Bredtmann, Nowotny and Otten (2020)	*Linguistic distance* using ASJP data	Negative; effect becomes smaller the larger size of networks	Share of origin specific foreign-born population living in EU-14 regions, 2007,	Gravity model; Random parameters (mixed) logit and Poisson with origin and destination FE; account for multilat. resistance: yes (RPL model)
	Genetic distance (as control variable) using Pemberton et al. data	Negative		
Collier and Hoefler (2018)	*Genetic distance* using Cavalli -Sforza et al data	Insignificant, but positive interaction with migrant networks	Bilateral stock differences over 10-year intervals, 1960-2000;	Gravity model; Linear regression; origin FE, destination FE and year FE account for multilat. resistance: yes (relative income term)
	Linguistic distance (as control variable) based on relatedness	Negative, but insignificant interaction with migrant networks		
Krieger, Renner and Ruhose (2018)	*Genetic distance* using Cavalli -Sforza et al data	Non-linear effect on migrant skill mix	Migrant skill mix using, education specific migrant stocks for 15 destinations and 85 origins, 2000	Linear model; OLS and Instrumental variables
Caragliu, Del Bo, de Groot and Linders (2013)	*Attitudes-based distance* using generalized trust, post-materialist values and traditional vs secular-rational values from WVS	Trust distance: negative Post-materialist distance: positive Trad. Vs secular distance: negative	Annual bilateral migration flows from EUROSTAT 2002-2007 and OECD 1998-2007.	Gravity model; Linear regression and Poisson (upon request); year FE account for multilat. resistance: no
Wang, de Graaff and Nijkamp (2016)	*Attitudes-based distance* obtained from performing PCA on ESS items	Negative	Attractiveness of European regions for EU and non-EU migrants, ESS 2008 and 2010, (no direct measure of migrant flows)	Equilibrium sorting model; two stage procedure
White and Buehler (2018)	*Attitudes-based distance* using Hofstede dimensions, Inglehart dimensions and GLOBE	Negative	Annual migration flows; 1983-2013, 36 destinations and 102 origins	Gravity model, negative binomial; origin FE, destination FE and year FE account for multilat. resistance: no
Lanati and Venturini (2021)	*Cultural trade* using BACI dataset of CEPII	Positive	Bilateral migrant flows on 30 OECD destinations and 185 origins, 2004-2013; similar to Adserà and Pytliková (2015)	Gravity model; Linear regression; IV; Poisson with origin-year, destination-year and corridor FE account for multilat. resistance: yes (origin -year FE)

Notes: List may not be complete; includes studies known to the author at the moment of writing. Studies are included that specifically investigate either linguistic, religious, genetic, and attitudes-based or trade based cultural similarity or difference. Studies are excluded when merely control for common official language etc.

8.2 RUM foundations of the gravity model of migration

The gravity model of migration is a model of location choice derived from a c The framework is very popular and has been applied to a variety of questions. For example, Grogger and Hanson (2011) investigate the selection into international migration and the sorting of migrants across destinations based on income maximization; Ortega and Peri (2013) investigate how income shapes migration under different immigration policy schemes; Gröschl (2012) and Beine and Parsons (2015) investigate how climatic factors affect migrants' location decisions; Beine, Docquier and Özden (2011) investigate how diasporas, i.e., the presence of migrant networks at destination, shape the size and human-capital structure of migrant flows; Czaika and Parsons (2017) investigate how immigration policies aimed at attracting and selectiong high skilled migrants affect where people move; and Lanati and Venturini (2021) investigate the impact of culutral change on migration.

All of these have in common that either the migrants' decision to move or their decision on where to settle is based on maximizing utility. Potential migrants have the choice between their home country i and all other countries (including their home country) as a potential destination j. Migration decsions are then modelled as evaluating, for each ij-pair, whether moving from i to any j maximizes utility and which of these would yield the highest return to human capital. This evaluation is based on the expected benefits V and costs C associated with moving. Some people decide to stay home, which means that the costs moving away outweigh the potential benefits that they could achieve elsewhere. Others decide to migrate and they choose the place which yields, on balance, the highest benefits with the lowest possible moving costs.

Following the literature, an individual's utility of moving from i to j is modelled by

$$u_{ijt} = V_{jt}(\cdot) - C_{ijt}(\cdot) + \varepsilon_{ijt}\,, \tag{11}$$

where $V_j(\cdot) - C_{ij}(\cdot)$ represents the deterministic component of migrant utility and ε_{ij} represent its stochastic component. More specifically, V_{jt} represents all j-specific expected benefits of living in j. These determine the *attractiveness* of destination and depend on the instantaneous wage, w_j, that an individual is expected to earn in j, as well as on other amenities, A_j, that destination j has to offer. The latter comprise, for example, employment conditions, the political system, institutional quality or ecological factors. Hence, V_j can be expressed as

$$V_{jt} = f\left(w_{jt}, A_{jt}\right).$$

$C_{ijt}(\cdot)$ in equation (11) represents the expected monetary and non-monetary costs associated with moving from i to j. Migration costs are the frictions that apply to migration flows and determine the *accessibility* of destination j from origin i. They are a function of a variety of factors, which can be bilateral or unilateral and they can vary over time or be entirely time invariant. Bilateral time-varying factors, c_{ijt}, express changing relationships between origin and destination countries. These could either increase or decrease the costs of migration. For instance, migrant networks reduce the costs of migration by providing information to prospective migrants and new arrivals, while visa requirements increase costs of migration by regulating entry. Bilateral and time-invariant costs, c_{ij}, represent unchanging relationships between origins and destinations. For example, while geo-physical distance between two countries is likely to increase travel costs, the same official language between two countries will reduce costs by facilitating communication at the new destination. Unilateral, time-varying origin-country frictions to migration, c_{it}, are, for example, credit constraints that migrants experience at origin, while changing attitudes towards immigration in general would represent time-varying costs at destination, c_{jt}. The geographical location of origin and destination countries would represent time-invariant cost factors c_i and c_j, respectively. Finally, c_t would represent common trends over time that apply to all migrants, e.g., changes of the oil price that makes travelling more expensive. So, C_{ijt} can be expressed as

$$C_{ijt}(\cdot) = f\left(c_{ijt}, c_{ij}, c_{it}, c_{jt}, c_i, c_j, c_t\right).$$

So far, individuals at origin i decide whether to move to destination j or 'move' to destination i, their origin country. With migration costs $C_{iit} = 0$ (staying at home) equation (11) reduces to $u_{ii} = V_i + \varepsilon_{ii}$. This utility can be easily compared to u_{ij}; if $u_{ii} > u_{ij}$ then the individual stays at home and if $u_{ii} < u_{ij}$ the individual emigrates to destination j. However, this decision model does not account for alternative migration destinations K.

Assuming that the stochastic term, ε_{ij}, from equation (11) follows an i.i.d. extreme-value distribution, one can follow the literature (e.g. Beine, Bertoli, and Fernández-Huertas Moraga 2016) and apply the McFadden (1984) results pertaining to discrete choice models. Then the probability that an individual migrates from i to j in the presence of all alternatives K can be expressed as

$$pr\left[u_{ijt} = \max_k u_{ikt}\right] = \frac{\exp\left[\mathrm{V}_{jt} - C_{ijt}\right]}{\sum_k \exp[\mathrm{V}_{kt} - C_{ikt}]}. \tag{12}$$

Multiplying this probability with the population at origin p_i we obtain the expected number of migrants (the expected migrant flow) from i to j

$$E\left[m_{ijt}\right] = \frac{\exp\left[\mathrm{V}_{jt} - C_{ijt}\right]}{\sum_k \exp[\mathrm{V}_{kt} - C_{ikt}]} * p_{it}. \tag{13}$$

Substituting $y_j = \exp[V_j]$, $\phi_{ij} = \exp[-C_{ij}]$ and $\Omega_i = \sum_k y_k \phi_{ik} = \sum_k \exp[\mathrm{V}_k - C_{ik}]$ yields the gravity model of migration as used in the literature (Beine, Bertoli, and Fernández-Huertas Moraga 2016):

$$E[m_{ijt}] = \frac{\phi_{ijt} y_{jt}}{\Omega_{\mathrm{it}}} p_{it}. \tag{14}$$

Equation (14) models bilateral migration flows as origin i's ability to send migrants at a given time, p_{it}, multiplied by the attractiveness, y_{jt}, and accessibility, ϕ_{ijt}, of destination j relative to the attractiveness and accessibility of all other options, Ω_{it}. Therefore, Ω_{it} represents the influence that alternative destinations have on bilateral migration flows. As the set of alternative destinations and their differential attractiveness depends on the origin county and its characteristics, Ω varies only over origin i characteristics. Following the trade literature, this influence of alternative destinations has been termed multilateral resistance to migration (MRM) (Bertoli and Fernández-Huertas Moraga 2013). Just like the multilateral resistance terms in the trade gravity model, Ω_{it} is a theoretical construct and not observed in practice. In order to estimate the above model in way that is consistent with the theoretical structure imposed in equation (14), one needs to properly control for Ω_{it}.

For clarity of exposition, I find it useful to re-write equation (14) in its exponential form. I also substitute again for y_{jt}, ϕ_{ijt} and Ω_{it}, and I add parameters as well as a well-behaved error term with $E[\eta_{ijt}] = 1$. This yields

$$m_{ijt} = \exp[\alpha_1 V_{jt} - \alpha_2 C_{ijt} + \alpha_3 \ln p_{it} - \alpha_4 MRM_{it}] * \eta_{ijt}\ , \tag{15}$$

with $MRM_{it} = \sum_k (V_{kt} - C_{ikt})$. This represents a generic and theory-consistent econometric version of the gravity model of migration, which I adjust to my specific purposes by adding the respective variables.

8.3 Spearman's rank correlations

Spearman correlations

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1. migration flows	-																	
2. coll./ind. similarity	-0.294	-																
3. duty/joy similarity	0.271	-0.314	-															
4. distrust/trust similarity	0.043	0.280	0.015	-														
5. religious similarity	0.288	-0.067	0.264	0.275	-													
6. linguistic similarity	0.413	-0.229	0.199	0.251	0.362	-												
7. genetic similarity	0.430	-0.300	0.287	0.081	0.254	0.577	-											
8. migrant networks	0.979	-0.280	0.282	0.048	0.291	0.418	0.433	-										
9. income distance	0.110	-0.638	0.079	-0.368	-0.121	-0.013	0.048	0.093	-									
10. distance (km)	-0.347	0.106	-0.162	-0.128	-0.257	-0.386	-0.533	-0.342	0.085	-								
11. EU/EFTA	0.297	-0.360	0.364	-0.000	0.193	0.252	0.389	0.295	0.038	-0.455	-							
12. BLA	0.287	-0.161	0.147	0.022	0.126	0.145	0.165	0.292	0.023	-0.136	0.196	-						
13. FTA	0.280	-0.331	0.277	0.010	0.194	0.276	0.354	0.274	0.122	-0.513	0.640	0.179	-					
14. intl. corridor	-0.270	-0.090	-0.158	-0.142	-0.206	-0.254	-0.222	-0.270	0.266	0.249	-0.069	0.042	0.101	-				
15. comm. official lang.	0.323	0.069	0.137	0.155	0.283	0.311	0.197	0.331	-0.143	-0.350	0.072	0.099	0.080	-0.328	-			
16. comm. spoken lang.	0.230	0.117	0.195	0.055	0.214	0.177	0.098	0.244	-0.084	-0.077	-0.031	0.065	-0.024	-0.258	0.462	-		
17. comm. border	0.267	0.037	0.060	0.095	0.174	0.197	0.198	0.279	-0.129	-0.291	0.111	0.205	0.160	0.036	0.324	0.249	-	
18. colony ever	0.158	0.005	0.033	0.035	0.087	0.101	0.064	0.160	-0.017	-0.026	-0.022	0.060	0.003	0.019	0.162	0.141	0.185	-

8.4 Attitudinal measures: Details

Table 8.2 WVS items for Beugelsdijk and Welzel (2018) dimensions

Dimension	WVS Item	Label	Response options
Collectivism/ Individualism	D054	One of main goals in life has been to make my parents proud	4 point scale (agree strongly, agree, disagree, disagree strongly)
	E036	Private vs state ownership of business	10 point scale (1- private ownership should be increased, …, 10 government ownership should be increased)
	F118	Justifiable: Homosexuality	10 point scale (1 never justifiable, … , always justifiable)
	F120	Justifiable: Abortion	10 point scale (1 never justifiable, … , always justifiable)
	C002	Jobs scarce: Employers should give priority to (nation) people than immigrants	3 options (1 agree, 2 disagree, 3 neither)
Duty/Joy	A038	Important child qualities: thrift saving money and things	2 options (0 not mentioned, 1 important)
	A003	Important in life: Leisure time	4 point scale (1 very important, 2 rather important, 3 not very important, 4 not at all important)
	A008	Feeling of happiness	4 point scale (1 very happy, 2 quite happy, 3 not very happy, 4 not at all happy)
	A173	How much freedom of choice and control	10 point scale (1 none at all, …, 10 a great deal)
	Y002	Post-Materialist index 4-item	3 point scale (1 materialist, 2 mixed, 3 postmaterialist)
Distrust/ Trust	A165	Most people can be trusted	2 options (1 most people can be trusted, 2 Can't be to careful)
	E069_12	Confidence: The Political Parties	4 point scale (1 a great deal, 2 quite a lot, 3 not very much, 4 none at all)
	E069_17	Confidence: Justice System/Courts	4 point scale (1 a great deal, 2 quite a lot, 3 not very much, 4 none at all)

Table 8.3 Match between 5-year intervals and WVS waves

5-year interval	WVS wave	EVS wave
	1 (1981-1984)	1 (1981-1984)
1990-1994	2 (1989-1993)	2 (1990-1993)
1995-1999	3 (1994-1998)	-
2000-2004	4 (1999-2004)	3 (1999-2001)
2005-2009	5 (2005-2009)	4 (2008-2009)
2010-2014	6 (2010-2016)	-
2015-2020	7 (2017-2020)	5 (2017-2020)

8.5 3WFE: additional results

Table 8.4 3WFE model – (i) no controls

Dependent variable: migration flows					
	(1)	(2)	(3)	(4)	(5)
coll./ind. similarity	0.17			-0.52	
	(1.80)			(3.08)	
duty/joy similarity		1.92		2.13	
		(3.05)		(4.05)	
distrust/trust similarity			-0.59	-1.63	
			(1.88)	(2.56)	
religious similarity					-0.37*
					(0.20)
	19.73***	19.08***	19.78***	19.83***	19.16***
Intercept	(0.66)	(1.01)	(0.69)	(1.20)	(0.08)
N	7358	8992	8203	5916	45722
intl. mig. captured	0.30	0.32	0.29	0.25	0.90
corridors	2573	3155	2743	2256	9955
origins	71	79	78	69	183
destinations	69	77	76	67	183
periods	5	5	6	5	5
pseudo R-sq.	1.00	1.00	1.00	1.00	1.00

Table 8.5 3WFE model– (ii) plus migrant networks

Dependent variable: migration flows					
	(1)	(2)	(3)	(4)	(5)
coll./ind. similarity	-4.37***			-6.42***	
	(1.54)			(1.83)	
duty/joy similarity		1.31		4.38*	
		(1.64)		(2.29)	
distrust/trust similarity			-1.73	-0.52	
			(1.63)	(1.38)	
religious similarity					0.07
					(0.19)
migrant networks	0.30***	0.29***	0.30***	0.27***	0.23***
	(0.04)	(0.03)	(0.03)	(0.04)	(0.02)
Intercept	15.35***	13.35***	14.33***	15.25***	14.54***
	(1.00)	(0.81)	(0.81)	(1.32)	(0.39)
N	11068	12862	11604	9611	49928
intl. mig. captured	0.36	0.38	0.37	0.33	0.93
corridors	3446	3939	3649	3246	10513
origins	81	88	88	81	189
destinations	78	85	85	78	189
periods	6	6	6	6	5
pseudo R-sq.	1.00	1.00	1.00	1.00	1.00

Table 8.6 3WFE model - (iii) plus income distance

Dependent variable: migration flows					
	(1)	(2)	(3)	(4)	(5)
coll./ind. similarity	-4.33***			-6.59***	
	(1.56)			(1.86)	
duty/joy similarity		1.61		4.69*	
		(1.71)		(2.40)	
distrust/trust similarity			-1.50	-0.24	
			(1.68)	(1.42)	
religious similarity					-0.31*
					(0.18)
migrant networks	0.30***	0.29***	0.30***	0.27***	0.23***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.02)
income distance	0.07**	0.08**	0.07**	0.08**	-0.01
	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)
Intercept	15.30***	13.29***	14.26***	15.15***	14.79***
	(1.03)	(0.85)	(0.84)	(1.36)	(0.40)
N	10404	12231	10990	9003	44026
intl. mig. captured	0.36	0.38	0.36	0.33	0.85
corridors	3319	3853	3568	3126	9642
origins	79	86	86	79	178
destinations	76	83	83	76	178
periods	6	6	6	6	5
pseudo R-sq.	1.00	1.00	1.00	1.00	1.00

Table 8.7 3WFE same sample

	Dependent variable: migration flows				
	(1)	(2)	(3)	(4)	(5)
coll./ind. similarity	-3.48			-3.37	
	(3.34)			(3.59)	
duty/joy similarity		-2.41		-1.78	
		(3.63)		(3.90)	
distrust/trust similarity			-0.14	0.04	
			(2.46)	(2.52)	
religious similarity					-1.49
					(1.28)
migrant networks	0.25***	0.24***	0.25***	0.24***	0.27***
	(0.06)	(0.06)	(0.06)	(0.06)	(0.05)
income distance	0.02	0.03	0.03	0.02	0.03
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
EU/EFTA	0.40***	0.44***	0.43***	0.41***	0.45***
	(0.14)	(0.14)	(0.15)	(0.14)	(0.15)
FTA	0.12**	0.12**	0.12**	0.12**	0.12***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Intercept	16.13***	15.74***	14.91***	16.73***	15.04***
	(2.13)	(1.39)	(1.19)	(1.52)	(1.15)
N	4612	4612	4612	4612	4612
intl. mig. captured	0.25	0.25	0.25	0.25	0.25
corridors	1886	1886	1886	1886	1886
origins	58	58	58	58	58
destinations	57	57	57	57	57
periods	4	4	4	4	4
pseudo R-sq.	1.00	1.00	1.00	1.00	1.00

*** p<.01, ** p<.05, * p<.1

8.6 TSFE: additional results

Table 8.8 TSFE - same sample

Dependent variable:									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
coll./ind. similarity		0.14*			0.17*				
		(0.08)			(0.09)				
duty/joy similarity			0.72**		0.77***				
			(0.29)		(0.30)				
dis-/trust similarity				0.24**	0.22**				
				(0.10)	(0.10)				
religious similarity						0.45***			0.31**
						(0.14)			(0.14)
linguistic similarity							1.73***		1.56***
							(0.24)		(0.24)
genetic similarity								31.76***	24.62***
								(6.08)	(4.72)
intl. corridor		-4.89***	-4.89***	-4.89***	-4.89***	-4.85***	-4.40***	-4.80***	-4.32***
		(0.16)	(0.16)	(0.16)	(0.16)	(0.15)	(0.16)	(0.16)	(0.16)
distance (km)		0.45***	0.44***	0.44***	0.44***	0.41***	0.37***	0.48***	0.38***
		(0.11)	(0.11)	(0.11)	(0.11)	(0.11)	(0.10)	(0.11)	(0.10)
comm. official lang.		0.62***	0.62***	0.63***	0.62***	0.61***	0.53***	0.63***	0.54***
		(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)
comm. spoken lang.		-0.71***	-0.71***	-0.71***	-0.71***	-0.70***	-0.69***	-0.66***	-0.66***
		(0.08)	(0.07)	(0.08)	(0.07)	(0.07)	(0.07)	(0.08)	(0.07)
common border		0.12	0.12	0.12	0.12	0.12	0.01	0.08	-0.01
		(0.13)	(0.13)	(0.13)	(0.13)	(0.12)	(0.13)	(0.13)	(0.13)
colony ever		1.02***	1.03***	1.02***	1.02***	1.02***	0.87***	1.01***	0.84***
		(0.18)	(0.18)	(0.18)	(0.18)	(0.18)	(0.18)	(0.18)	(0.18)
migrant networks	0.21***								
	(0.02)								
income distance	-0.04								
	(0.03)								
EU/EFTA	0.54***								
	(0.10)								
FTA	0.02								
	(0.04)								
Intercept	15.03***	2.94***	2.74***	2.89***	2.57***	2.78***	2.07***	-19.31***	-15.24***
	(0.40)	(0.46)	(0.45)	(0.46)	(0.46)	(0.45)	(0.44)	(4.30)	(3.38)
N	60130	5799	5799	5799	5799	5799	5799	5799	5799
intl. mig. captured	0.92	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26
corridors	11026	3260	3260	3260	3260	3260	3260	3260	3260
origins	202	81	81	81	81	81	81	81	81
destinations	202	80	80	80	80	80	80	80	80
periods	6	4	4	4	4	4	4	4	4
pseudo R-sq.	0.9999	0.7259	0.7259	0.7259	0.7259	0.7259	0.7262	0.7260	0.7263

Notes: Column (1) first-stage decomposition of migration flows with corridor FE by decade. Columns (2)-(9) gravity regressions of time-invariant and slowly moving determinants on exponentiated recovered corridor fixed effects from (1). Table shows parameter estimates obtained from Poisson-Pseudo-Maximum-Likelihood (PPML) estimation using the ppmlhdfe routine, which automatically drops singletons and observations that are separated by a fixed effect (Correia et al. 2020,2021). All continuous variables enter the estimation in logs, $\ln(x+1)$ if the variable takes values smaller than 1. Standard errors in parentheses are clustered at migration corridor level. *** p<.01, ** p<.05, * p<.1.

Table 8.9 TSFE - sensitivity to control variables

Panel A: Attitudes

Dependent variable:						
	(1)	(2)	(3)	(4)	(5)	(6)
coll./ind. similarity	31.72***	1.41***	19.56***	0.63***	0.93***	0.28***
	(4.32)	(0.32)	(3.05)	(0.19)	(0.22)	(0.10)
duty/joy similarity	103.78***	4.64***	64.87***	2.40***	2.89***	1.26***
	(11.87)	(1.07)	(8.95)	(0.63)	(0.69)	(0.35)
dis-/trust similarity	30.60***	1.63***	19.95***	1.31***	0.76***	0.33***
	(4.86)	(0.41)	(3.76)	(0.30)	(0.24)	(0.12)
intl. corridor					-6.82***	-5.67***
					(0.06)	(0.18)
comm. official lang.			2.54***	0.84***		
			(0.14)	(0.13)		
comm. spoken lang.			2.35***	1.51***		
			(0.17)	(0.13)		
distance (km)		-3.38***		-2.66***		-0.64***
		(0.06)		(0.06)		(0.08)
common border		-2.30***		-2.59***		0.70***
		(0.13)		(0.12)		(0.15)
colony ever		1.10***		0.81***		1.33***
		(0.25)		(0.22)		(0.15)
Intercept	-58.83***	13.87***	-41.73***	9.26***	-1.02***	2.97***
	(4.15)	(0.69)	(3.28)	(0.58)	(0.33)	(0.42)
	7887	7887	7887	7887	7887	7887
N	0.29	0.29	0.29	0.29	0.29	0.29
intl. mig. captured	4222	4222	4222	4222	4222	4222
corridors	97	97	97	97	97	97
origins	96	96	96	96	96	96
destinations	5	5	5	5	5	5
periods	0.5053	0.7452	0.5923	0.7495	0.7549	0.7587
pseudo R-sq.	31.72***	1.41***	19.56***	0.63***	0.93***	0.28***

Notes: First-stage decomposition omitted (the same as in main results). Columns (1)-(6) gravity regressions of time-invariant and slowly moving determinants on exponentiated recovered corridor fixed effects from first stage decomposition. Table shows parameter estimates obtained from Poisson-Pseudo-Maximum-Likelihood (PPML) estimation using the ppmlhdfe routine, which automatically drops singletons and observations that are separated by a fixed effect (Correia et al. 2020,2021). Standard errors in parentheses are clustered at migration corridor level. *** p<.01, ** p<.05, * p<.1..

Panel B: Proxies

Dependent variable:						
	(1)	(2)	(3)	(4)	(5)	(6)
linguistic similarity	5.79***	0.96***	5.66***	0.84***	0.56***	0.42***
	(1.34)	(0.26)	(1.33)	(0.23)	(0.17)	(0.14)
religious similarity	16.06***	5.77***	11.90***	4.17***	3.33***	2.52***
	(1.05)	(0.41)	(1.00)	(0.33)	(0.35)	(0.44)
genetic similarity	286.38***	29.75***	293.57***	33.57***	16.34***	9.02*
	(28.03)	(7.06)	(28.19)	(6.68)	(4.79)	(4.61)
intl. corridor					-5.47***	-5.27***
					(0.20)	(0.27)
comm. official lang.			1.20***	0.55***		
			(0.20)	(0.13)		
comm. spoken lang.			1.20***	1.14***		
			(0.19)	(0.11)		
distance (km)		-1.73***		-1.61***		-0.24***
		(0.11)		(0.10)		(0.08)
common border		-2.22***		-2.22***		0.66***
		(0.15)		(0.15)		(0.24)
colony ever		0.78***		0.22		1.24***
		(0.17)		(0.18)		(0.16)
	-208.20***	-16.00***	-213.44***	-19.99***	-12.14***	-5.65*
Intercept	(19.28)	(5.16)	(19.38)	(4.90)	(3.24)	(3.24)
	42169	42169	42169	42169	42169	42169
N	0.82	0.82	0.82	0.82	0.82	0.82
intl. mig. captured	9179	9179	9179	9179	9179	9179
corridors	184	184	184	184	184	184
origins	184	184	184	184	184	184
destinations	5	5	5	5	5	5
periods	0.6838	0.7826	0.6919	0.7851	0.7931	0.7949
pseudo R-sq.	5.79***	0.96***	5.66***	0.84***	0.56***	0.42***

Notes: First-stage decomposition omitted (the same as in main results). Columns (1)-(6) gravity regressions of time-invariant and slowly moving determinants on exponentiated recovered corridor fixed effects from first stage decomposition. Table shows parameter estimates obtained from Poisson-Pseudo-Maximum-Likelihood (PPML) estimation using the ppmlhdfe routine, which automatically drops singletons and observations that are separated by a fixed effect (Correia et al. 2020,2021). Standard errors in parentheses are clustered at migration corridor level. *** p<.01, ** p<.05, * p<.1.

Table 8.10 TSFE in T2 sample (2000-2019)

Dependent variable:									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
L5lncollind_herf		0.12**			0.14**				
		(0.05)			(0.07)				
L5lndutyjoy_herf			0.67***		0.62***				
			(0.21)		(0.22)				
L5lndistrust_herf				0.31***	0.24***				
				(0.10)	(0.09)				
religious similarity						0.55**			0.38**
						(0.27)			(0.19)
linguistic similarity							2.45***		2.20***
							(0.42)		(0.51)
genetic similarity								7.08	-0.73
								(9.54)	(8.21)
intl. corridor		-5.76***	-5.98***	-6.00***	-5.74***	-6.18***	-5.65***	-6.11***	-5.58***
		(0.21)	(0.22)	(0.22)	(0.20)	(0.34)	(0.34)	(0.39)	(0.43)
comm. official lang.		0.61***	0.57***	0.57***	0.56***	1.01***	0.83***	1.03***	0.68***
		(0.13)	(0.14)	(0.14)	(0.13)	(0.19)	(0.15)	(0.22)	(0.18)
comm. spoken lang.		1.05***	1.01***	1.11***	1.08***	1.21***	1.12***	1.22***	1.14***
		(0.16)	(0.14)	(0.14)	(0.15)	(0.13)	(0.13)	(0.14)	(0.14)
distance (km)		-0.77***	-0.70***	-0.61***	-0.80***	-0.22***	-0.22***	-0.23**	-0.26***
		(0.11)	(0.12)	(0.12)	(0.11)	(0.08)	(0.05)	(0.09)	(0.09)
common border		0.31*	0.48***	0.46***	0.34**	0.86***	0.63**	0.73*	0.50
		(0.16)	(0.15)	(0.16)	(0.15)	(0.33)	(0.30)	(0.40)	(0.39)
colony ever		1.18***	1.41***	1.21***	1.21***	0.48**	0.91***	0.61***	0.66***
		(0.20)	(0.19)	(0.21)	(0.20)	(0.20)	(0.24)	(0.23)	(0.22)
migrant networks	0.02								
	(0.02)								
income distance	-0.04								
	(0.04)								
EU/EFTA	0.38***								
	(0.13)								
FTA	-0.04								
	(0.07)								
Intercept	18.66***	2.52***	2.08***	1.76***	2.39***	0.22	-0.61	-4.61	0.23
	(0.40)	(0.51)	(0.58)	(0.58)	(0.51)	(0.48)	(0.39)	(6.76)	(5.80)
N	42872	7891	9315	8202	6614	30167	42872	36684	27238
intl. mig. captured	0.95	0.35	0.36	0.34	0.30	0.94	0.95	0.85	0.86
corridors	10994	4490	4881	4646	4192	10205	10994	9232	9156
origins	202	98	102	103	97	185	202	202	184
destinations	202	97	102	103	96	185	202	202	184
periods	4	4	4	4	4	3	4	4	3
pseudo R-sq.	0.9999	0.8266	0.8254	0.8284	0.8263	0.8739	0.8711	0.8760	0.8771

Notes: Column (1) first-stage decomposition of migration flows with corridor FE by decade. Columns (2)-(9) gravity regressions of time-invariant and slowly moving determinants on exponentiated recovered corridor fixed effects from (1). Table shows parameter estimates obtained from Poisson-Pseudo-Maximum-Likelihood (PPML) estimation using the ppmlhdfe routine, which automatically drops singletons and observations that are separated by a fixed effect (Correia et al. 2020,2021). All continuous variables enter the estimation in logs, $\ln(x+1)$ if the variable takes values smaller than 1. Standard errors in parentheses are clustered at migration corridor level. *** p<.01, ** p<.05, * p<.1.

8.7 Nearest neighbor matching

Matching estimators (e.g., Abadie and Imbens 2006) recover the effect of a treatment from observational data by matching similar observations across the two treatment levels. Here, the unit of analysis are migration corridors and the "treatment" conditions are (i) corridors between culturally similar countries and (ii) corridors between dissimilar countries. I create the two conditions by splitting the full sample of corridors at the median of each of my cultural measures. The dummy variable $cultsim_{ijt} = 1$ ($cultsim_{ij} = 0$) then indicates corridors with culturally similar (dissimilar) origin and destination countries in period t.

The effect of this binary "treatment" is estimated by taking the difference of the observed outcome, i.e., migration flows, of corridors in one group and a counterfactual outcome of these corridors, which is imputed from the observed outcomes of matched corridors in the other group. The matching of corridors is based on observed covariates such that the two groups are similar to each other except in the treatment variable. The details of the matching procedure are explained later.

Following Wooldridge (2010), for each ij-corridor, matching estimators impute values for the counterfactuals m_{ij}^1 and m_{ij}^0 but use the observed values whenever possible. Hence, $\widehat{m}_{ij}^1$ and $\widehat{m}_{ij}^0$ denote the imputed values with $\widehat{m}_{ij}^1 = m_{ij}$ when a corridor between similar countries ($cultsim_{ij} = 1$) is observed and $\widehat{m}_{ij}^0 = m_{ij}$ when a corridor between dissimilar countries ($cultsim_{ij} = 0$) is observed.

The estimator of the average treatment effect (ATE), which takes the average over all observations,[55] takes the form

$$\hat{\tau}_t^{ate} = \frac{1}{N}\sum_{ij=1}^{N}\left[\widehat{m}_{ijt}^1 - \widehat{m}_{ijt}^0\right]. \tag{16}$$

Estimating treatment effects from observational data (using matching) requires that three conventional assumptions are made (e.g. Wooldridge 2010). The first is *conditional independence*, which means that treatment and outcomes are uncorrelated conditional on observed covariates. This ensures that the treatment is random and exogenous after controlling for the covariates. In addition to familiar control variables, I will use recovered origin-year and destination-year fixed effects from an ANOVA decomposition as conditioning

[55] One could also compute the average treatment effect on the treated (ATT), which only considers corridors between culturally similar countries. The ATT estimator takes the form $\hat{\tau}_t^{att} = \frac{1}{N}\sum_{ij=1}^{N} cultsim_{ijt} * \left[m_{ijt} - \widehat{m}_{ijt}^0\right]$ Yet, for sake of brevity I focus on reporting the ATE but results on the ATT are qualitatively similar.

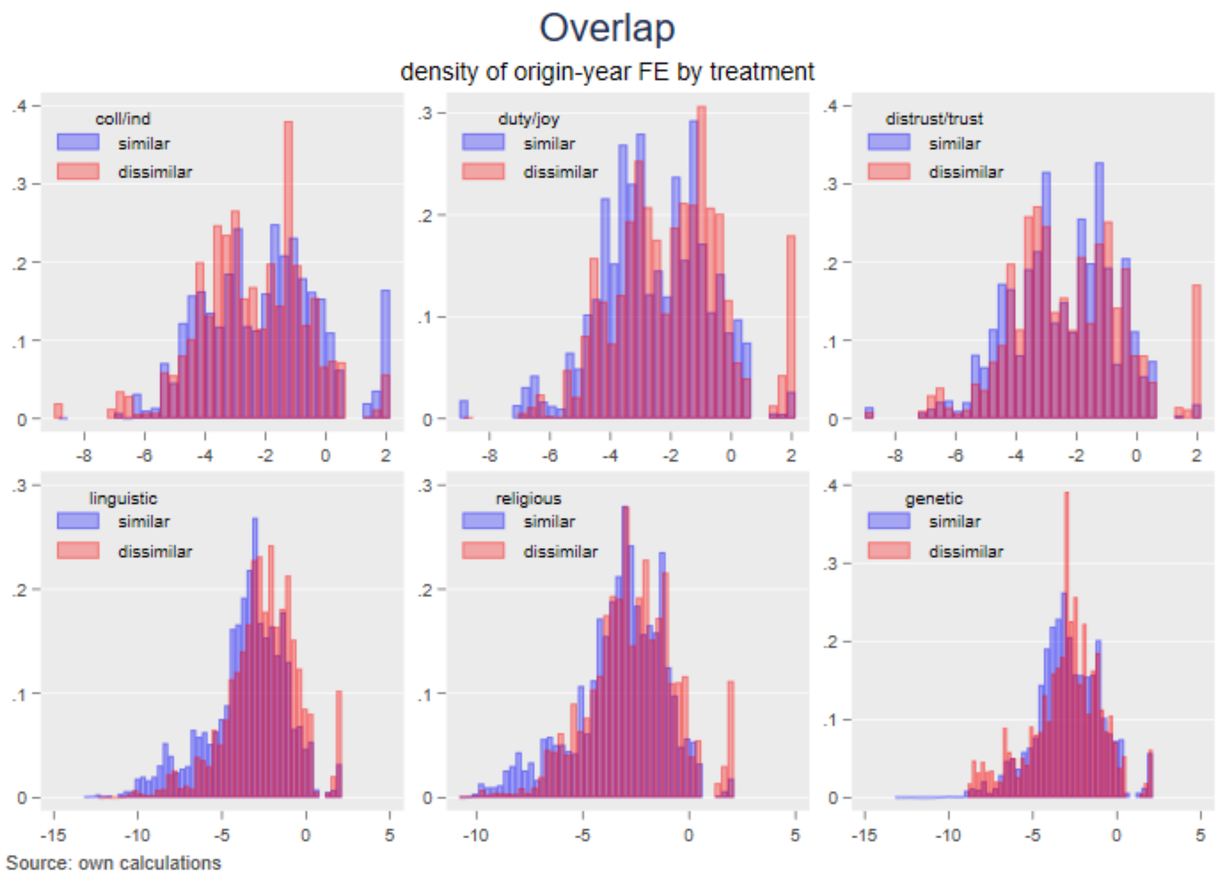

Figure 4: Overlap of treatment groups with respect to recovered origin-year FE (outward multilateral resistance)

covariates. These capture multilateral resistance to migration and thus ensure consistency with theory.

The second assumption is the overlap assumption and implies that, conditional on the covariates, there are treated and non-treated corridors for each outcome, i.e., size of migration flows. Following the reasoning in the trade literature, because the data set is very detailed and includes a large number of observations it is likely that corridors can be matched across the two treatment groups (e.g., Kohl and Trojanowska 2015).

The third assumption, the so-called stable unit treatment value assumption (SUTVA), has two components: first, the 'no multiple versions' assumption requires that the treatment is identical for each treated observation; second, the 'non-interference' assumption requires that the cultural similarity of one corridor does not influence the migration flows in other corridors – particularly those in the other group. The latter non-interference assumption is met by including the recovered origin-year and destination-year fixed effects, which account for the general equilibrium effects. However, the 'no multiple versions' assumption is unlikely to hold in the current setting because although corridors are labeled 'culturally similar' ($cultsim_{ijt} = 1$) they still differ in degree of similarity.

The first step to implement the matching estimator step is to impute the missing counterfactual outcomes m_{ijt}^1 and m_{ijt}^0 using matching corridors from the two treatment conditions. As mentioned earlier, the matching is done according to observed covariates to ensure that the treatment is exogenous. To maintain the 'non-interference' assumption, i.e., to account for multilateral resistance, Baier and Bergstrand (2009) use a first-order log-linear Taylor-series expansion of the trade gravity model to obtain a reduced-form function of linear combinations of 'exogenous' variables including terms that capture multilateral resistance (cf. Kohl and Trojanowska 2015). However I expand on their method and, instead of using constructed exogenous variables, I recover origin-year and destination-year fixed effects, $\widehat{\delta_{it}}$ and $\widehat{\delta_{jt}}$, from the following unsaturated ANOVA decomposition:

$$m_{ijt} = \ exp\left[\delta_{it} + \delta_{jt} + \delta_{ij}\right] * \eta_{ijt} \ , \tag{17}$$

I use the fitted values, $\widehat{\delta_{it}}$ and $\widehat{\delta_{jt}}$ that capture inward and outward multilateral resistance, and the set of familiar covariates[56] to match migrant corridors in the 5-year periods from 1990 to 2019. Following Baier and Bergstand (2009) and Kohl and Trojanowska (2015) I use 'nearest neighbor' matching with replacement and $k = 3$ nearest neighbors while adjusting for continuous variables bias (Abadie and Imbens 2006; 2011).

For completeness, the Baier and Bergstrand (2009) first-order log-linear Taylor-series expansion of the trade gravity model takes the form

$$\begin{aligned} \ln TF_{ijt} = \beta_0 + \beta_1 TA_{ijt} + \beta_2 SGDP_{ijt} + \beta_3 BVD_{ij} \\ + \beta_4 BVC_{ij} + \beta_5 BVL_{ij} + \epsilon_{ijt} \end{aligned} \tag{18}$$

Where TF_{ijt} are trade flows between i and j, and TA_{ijt} is a dummy variable indicating a trade agreement between i and j. Moreover,

$$SGDP_{ijt} = \ln gdp_{it} + \ln gdp_{jt}$$

$$BVD_{ij} = \ln D_{ij} - \frac{1}{N}\sum_{i=1}^{N} \ln D_{ij} - \frac{1}{N}\sum_{j=1}^{N} \ln D_{ij} + \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} \ln D_{ij}$$

$$BVC_{ij} = \ln C_{ij} - \frac{1}{N}\sum_{i=1}^{N} \ln C_{ij} - \frac{1}{N}\sum_{j=1}^{N} \ln C_{ij} + \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} \ln C_{ij}$$

[56] migrant networks, income distance, geodesic distance (in km), common language, common border, colony ever, dummy variables for bilateral agreements and the international corridor dummy

$$BVL_{ij} = \ln L_{ij} - \frac{1}{N}\sum_{i=1}^{N} \ln L - \frac{1}{N}\sum_{j=1}^{N} \ln L_{ij} + \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} \ln L_{ij}$$

where D_{ij} is the bilateral distance between countries in a corridor, C_{ij} is a dummy variable for contiguity (shared border), L_{ij} is a dummy variable for a common language and N is the number of countries. The $SGDP_{ijt}$ term is an expression of the trade potential between two countries, where the GDP of one country indicates how many goods/how much value it could export, and the BV-terms ($BVD_{ij}, BVC_{ij}, BVL_{ij}$) represent the frictions of trade *including* the influence that other exporters and importers have on the trade between i and j.[57]

[57] BV stands for *bonus vetus* ("good old") because Baier and Bergstrand initially derive their exogenous terms for "good old" OLS estimations of the gravity model.