Departament d'Economia Aplicada
Universitat Autònoma de Barcelona

WORKING PAPER 25-17

Immigration, Identity Choices, and Cultural Diversity

Yasmine Elkhateeb
Riccardo Turati
Jérôme Valette

Departament d'Economia Aplicada
Edifici B
Universitat Autònoma de Barcelona
08193 Bellaterra, Barcelona, Spain.
d.econ.aplicada@uab.cat
www.uab.cat/en/applied-economics
@ecapuab.bsky.social

UAB
Universitat Autònoma
de Barcelona

Immigration, Identity Choices, and Cultural Diversity

Yasmine Elkhateeb[1], Riccardo Turati[2], and Jérôme Valette[3]

[1] *J-PAL MENA, Faculty of Economics & Political Science, Cairo University (Egypt)*
[2] *DEA, UAB (Spain), IZA (Germany), RFBerlin (Germany)*
[3] *CEPII, IC Migrations (France), IZA (Germany), RFBerlin (Germany)*

November, 2025

Abstract

Does immigration challenge the identities, values, and cultural diversity of receiving societies? This paper addresses this question by analyzing the impact of immigration on cultural diversity in Europe between 2004 and 2018. It combines regional cultural diversity indices derived from the European Social Survey with immigration shares from the European Labor Force Survey. The results indicate that immigration increases the salience of birthplace identity along cultural lines, fostering a shift toward nativist identities among the native population. These identity shifts, in turn, trigger a process of cultural homogenization among natives. This effect is stronger in regions receiving culturally distant immigrants. It reflects a process of convergence toward the values of highly skilled liberal natives and divergence from those of low-skilled conservative immigrants.

JEL classification: F22, D03, D72, Z10.
Keywords: Immigration, Social Identity, Cultural Diversity.

This paper has benefited from valuable suggestions and comments from Simone Bertoli, Mario Carillo, Lisa Chauvet, Melchior Clerc, Anthony Edo, Flore Gubert, Anna Maria Mayda, Marion Mercier, Stelios Michalopoulos, Hillel Rapoport, Enrico Spolaore, Marco Tabellini, Farid Toubal, Romain Wacziarg, and Dean Yang. We also wish to thank seminar participants at the 13th Annual Conference on Immigration in OECD countries, at the 2nd Istanbul Applied Micro Workshop, King's College London and Özyegin University, at the 8th "Understanding Voluntary and Forced Migration" Workshop at Université de Lille, at the Paris Migration Economics Seminar at PSE, at the University of Salerno, at the 5th Workshop on the Economics and Politics of Migration in King's College London, at the DIMIG workshop in Paris, at the 14th International Conference "Economics of Global Interactions" at University of Bari Aldo Moro, the 4th CReAM/RWI Workshop on the Economics of Migration, at the International Trade Lunch Seminar, Georgetown University, at the Growth Lab Seminar, Brown University, and at the Economic Development Seminar, University of Michigan, for helpful discussions. ✉ Corresponding author: Jérôme Valette, CEPII, 20 avenue de Ségur, 75007, Paris ; email:jerome.valette@cepii.fr.

1 Introduction

Immigration is one of the main concerns of Western societies, as evidenced by the rising electoral success of populist parties with strong anti-immigration platforms (Guriev and Papaioannou, 2022). While earlier debates on immigration focused on immigrants' economic impacts, concerns now seem to have increasingly shifted toward their consequences for the receiving communities' culture and identity (Alesina and Tabellini, 2024; Gennaioli and Tabellini, 2023). In particular, the inflow of people from geographically, economically, and culturally distant countries has intensified concerns about the assimilation of diverse norms and values into host societies (Collier, 2013), raising a fundamental question: Does immigration challenge the social identities, values, and cultural diversity of receiving countries?

Recent studies, which have documented shifts along the socioeconomic class axis (Gethin et al., 2022), have explored the potential effects of immigration on polarization along this dimension (Bonomi et al., 2021). However, although immigration is a major driver of cultural change (e.g., Fernández, 2025), the question of how immigration could have shaped the distribution of norms and preferences along the cultural dimension, and its overall effect on cultural diversity remains ambiguous. Indeed, while immigrants bring distinct cultural norms and values, which increase cultural diversity in the host society (Rapoport et al., 2021; Bazzi and Fiszbein, 2025), they also induce a re-categorization of identities in the receiving population, redefining in-group and out-group boundaries (Fouka et al., 2022; Fouka and Tabellini, 2022). Ultimately, the changes in cultural values induced by the redefinition of identity boundaries may, depending on their direction and intensity, lead to cultural convergence or greater diversity in destination countries.

This paper addresses this question by relying on model of endogenous social identity (Shayo, 2009) to investigate how immigration shapes individuals' social identity choices, norms, and values. This model posits that individuals endogenously identify with social groups as a function of the groups' relative status and the perceived distance between their own characteristics and those of other groups' members. Within this framework, immigration, particularly originating from culturally distant countries (Fouka and Tabellini, 2025), acts as a cultural shock. By changing the existing balance of cultural attributes within the host society, immigration increases the salience of birthplace identity in the society.[1] It contributes to identity re-categorization, with natives prioritizing birthplace over other pre-existing social divisions, and fosters a cultural realignment of individuals within their newly adopted social group.

This paper empirically tests these predictions by focusing on European countries, examining the extent to which immigration has influenced the overall cultural diversity of European regions over the past two decades. It takes advantage of European Social Surveys (ESS) data from 2004 to 2018 to measure the regional evolution of cultural diversity over 175 European regions along several cultural dimensions.[2] Building on Desmet and Wacziarg (2021), we compute regional measures of cultural diversity defined as the likelihood that two randomly selected individuals from the entire resident population of a given NUTS-2 region hold a different variant

[1]This echoes empirical observations that birthplace is one of the strongest markers of identity and values. Obradovich et al. (2022), using data from two billion Facebook users across 225 countries, demonstrate the importance of national borders in shaping culture.

[2]We follow Alesina et al. (2017) and exploit the richness of the ESS by selecting 46 different cultural traits on religiosity, sexual morality, the role of the state, cultural capital, political engagement, trust in institutions, attitudes toward immigration, and general openness.

of a randomly selected memetic trait. This measure reflects a definition of culture, known for its complex and multifaceted nature, that echoes the seminal work of Kroeber and Kluckhohn (1952). Then, we decompose the overall diversity index into within-group and between-group components, using various identity markers to split the population. When birthplace serves as the identity marker, changes in these diversity measures are interpreted as reflecting variations in cultural diversity within the native and immigrant populations and changes in the salience of birthplace, respectively.

The empirical analysis regresses these measures of cultural diversity on the share of foreign-born over the total 2004 population and a full vector of controls at the regional level. Immigration stocks are taken from the European Labor Force Survey (EU-LFS). To ensure a causal interpretation of the estimates, the benchmark specification first includes wave and regional fixed effects, which control for common aggregated changes over time and time-invariant regional characteristics.[3] Then, to minimize concerns related to self-selection and the non-random sorting of immigrants, we rely on 2SLS estimates with a modified shift-share instrument to predict exogenous immigration stocks by origin based on the initial spatial sorting of immigrants and the growth of their diasporas at the national level over time (Card, 2001). To enhance the validity of the identifying assumption based on the exogeneity of aggregate immigration flows (Borusyak et al., 2022), the overall stock of immigrants from each origin is predicted using a zero-stage gravity equation that only includes exogenous push factors, such as conflicts and natural disasters, in migrants' origin countries. Furthermore, we provide evidence that our results are robust to a leave-one-out version of the instrument, and we do not find evidence of natives' mobility response to migration, ensuring that natives' sorting does not confound our estimates.

The empirical results are consistent with the model's predictions. First, we find that immigration is associated with a significant increase in the salience of birthplace in society, i.e, a greater predictability of responses to questions on cultural norms, attitudes, and preferences based only on an individual's country of birth. Then, our main finding shows that rising immigration leads to a significant decline in cultural diversity within the native population. Specifically, we find that a one percentage point increase in the immigrant share is associated with a 0.16 percentage point decrease in the likelihood that two randomly selected natives hold different views on a randomly chosen cultural trait.[4] Notably, natives tend to converge toward similar norms across a broad range of dimensions, and the effect is not driven by any particular set of cultural traits. Overall, this paper shows that immigration, by challenging natives' social identities, shapes the distribution of their cultural values, generating a reduction of cultural diversity among the native population. It also underscores the role of birthplace as

[3]Our empirical setting will exploit the arrival of new immigrants as a source of variation, rather than the persistent presence of immigrants. Hence, it does not allow us to assert anything about the potential temporal dynamics of immigration, despite the plausible conjecture that the effect stemming from the natives' reaction is more likely to manifest over a longer duration compared to the direct impact triggered by the arrival of new immigrants.

[4]This effect must be interpreted as a decrease in the cultural diversity of the native population since we find no evidence of increasing polarization at both sides of the distribution among natives and between different sociodemographic groups. Moreover, we do not find any effect of new inflows of immigrants on the values of foreign-born residents. Also, with no similar effects when focusing on second-generation immigrants, or with respect to other identity cleavages, we provide evidence that the effect pertains to birthplace and does not extend to parental background or other socioeconomic identities.

a key identity marker for culture.

We conduct several additional analyses to investigate the underlying mechanisms at play. First, the heterogeneity analysis shows that convergence within the native population is stronger in regions with higher concentrations of low-skilled immigrants from culturally distant countries. This pattern is consistent with these groups contributing more to increased birthplace salience, as their norms differ more sharply from those of natives, making them more socially visible and generating greater utility costs for exposed natives due to increased perceived cultural distance. Second, using individual-level data, we show that immigration is indeed associated with higher national pride and increased support for nationalist parties, providing suggestive evidence of potential identity changes among natives toward nativism. Finally, by estimating the impact of immigration on the Euclidean cultural distance between each native and several potential reference groups in the population, we find that cultural homogenization among natives reflects an overall convergence toward the cultural values of high-skilled liberal natives and a divergence from those of low-skilled immigrants, who hold relatively more conservative cultural attributes. These last results echo those of Fouka and Tabellini (2022), who shows that inflows of Mexicans in the United States contributed to a shift of white Americans towards more liberal policy stances.

This paper contributes to two main strands of literature. First, it contributes to the burgeoning literature exploring the drivers of cultural change and divides in Western societies (e.g., Desmet and Wacziarg, 2021; Bertrand and Kamenica, 2023; Fernández, 2025), with specific focus on the salience and re-categorization of social identities (e.g., Akerlof and Kranton, 2000; Shayo, 2020; Grossman and Helpman, 2021; Bonomi et al., 2021; Fouka et al., 2022; Gethin et al., 2022). Our results directly speak to the contributions of Bonomi et al. (2021), who highlight that individuals tend to identify with the most salient group in society and subsequently adopt the stereotypical views associated with this group (Abrams and Hogg, 2006; Bordalo et al., 2016). In their model, Bonomi et al. (2021) suggest that immigration, by increasing the salience of cultural divisions within society, can lead to identity switches from class-based to culture-based identification within the native population, potentially explaining shifts in norms and values such as changes in preferences for redistribution.[5] While polarization in these papers occurs along the socioeconomic class axis, US historical data suggests that identity re-categorization can also occur along other ethnic-based identity markers (Fouka et al., 2022; Fouka and Tabellini, 2022, 2025). Our paper, therefore, contributes to this literature by highlighting birthplace as a relevant identity marker for studying cultural change within host countries and by providing empirical evidence of immigration as a shock to individuals' identity choices. In addition, while most of this research focuses on the United States (e.g., Desmet and Wacziarg, 2021; Fouka et al., 2022; Bertrand and Kamenica, 2023), our analysis provides the first empirical evidence that immigration may generate cultural convergence among natives in the European context.[6]

This paper also contributes to the literature examining the cultural impact of immigration in receiving countries (see Bazzi and Fiszbein, 2025, for a recent review).[7] Prior research

[5] Gennaioli and Tabellini (2023) extend this analysis to the supply side by modeling political parties' reactions to increased salience of cultural divergence.

[6] A notable exception is Alesina et al. (2017), which reports on the evolution of culture in Europe across four waves of the European Values Survey (EVS) between 1980 and 2008, although it does not test immigration's implications for cultural evolution.

[7] At the international level, Rapoport et al. (2021) show that migration increases cultural

suggests that immigration impacts the distribution of values in host societies through a direct compositional effect, as immigrants hold distinct values and norms compared to the native population. This effect depends on the distribution of values among newly arrived immigrants relative to that of the host society (Desmet and Wacziarg, 2021) and on immigrants' initial patterns of cultural self-selection at origin (Docquier et al., 2020a; Knudsen, 2022).[8] Our results confirm that significant cultural differences between immigrants and natives increase the relevance of birthplace in predicting culture as immigration rises. In addition to this compositional effect, immigration may also affect the distribution of values at destination by inducing attitudinal changes within the native population (see Dustmann and Preston, 2007; Dustmann et al., 2018; Edo et al., 2019; Steinmayr, 2021; Alesina et al., 2022; Alesina and Tabellini, 2024; Keita et al., 2023; Schneider-Strawczynski and Valette, 2025, among others), due to transmission of values from immigrants to natives (see Fisman and Miguel, 2007; Giuliano and Tabellini, 2021; Miho et al., 2024; Bazzi et al., 2023) and re-categorization of social identity group boundaries (Fouka and Tabellini, 2025). Our paper underscores the relevance of the latter mechanism, providing new empirical evidence of immigration-induced cultural convergence among natives. Therefore, this paper contributes to the literature by highlighting the salience of birthplace identity and natives' responses as key to understanding immigration's impact on host societies' cultural diversity.

The remainder of this paper is organized as follows. Section 2 derives our main testable hypotheses from Shayo (2009)'s model. Section 3 presents the data and measures of cultural diversity, along with preliminary evidence on birthplace as a predictor of cultural diversity. Section 4 describes our empirical strategy, and Section 5 reports the main results and robustness checks. Section 6 presents heterogeneity analysis, Section 7 examines potential mechanisms, and Section 8 concludes.

2 Social Identity and Cultural Diversity: A Theoretical Framework

This paper builds on Shayo (2009, 2020)'s theoretical framework, developed to understand how individuals choose to belong to a specific social group or identity and how this choice influences their values. We present the baseline setup of the model, and we connect it with recent theoretical and measurement developments in the literature on cultural diversity. Then, we explain how this setting applies to the context of immigration, and we formalize new testable

proximity between home and host countries through the transmission of norms from diasporas. However, in contrast to our paper, their findings do not focus on the impact of immigration on within-country cultural diversity.

[8] The extent to which this compositional effect persists over time is strongly related to the rate of cultural assimilation (Algan et al., 2012; Abramitzky et al., 2014; Galli and Russo, 2019; Gonnot and lo Polito, 2021; Fouka et al., 2022; Abramitzky and Boustan, 2022; Gonnot and lo Polito, 2023) and intergenerational transmission of cultural traits (Bisin and Verdier, 2001; Desmet et al., 2017; Rapoport et al., 2021; Desmet and Wacziarg, 2021; Abramitzky et al., 2020). Although a small but significant selection has been found at origin, aspiring emigrants and actual migrants still exhibit large cultural differences from the destination country's native population (Obradovich et al., 2022). Within our setting, we find that the effect of immigration on cultural diversity is driven by new inflows of immigrants only, suggesting that this effect likely dissipates over time.

predictions regarding the impact of immigration on cultural diversity in the host society.

2.1 General setup

Consider a society of N individuals and of G given identity-groups. Each individual i has a vector of personal identities G_i, which are given, and has a set of available actions A_i. The action profile is denoted by $a = (a_i)_{i \in N}$. Among the available personal identities, each individual chooses the one that is associated with his behavior with others, defined as social identity. For simplicity, and following Shayo (2009), we assume here that each individual can have only one social identity $J \in G_i$.

Two main forces influences individuals' choices in the model: *i)* conformity, measured by the perceived distance from other group members, as individuals tend to value being in a group that shares similar attributes (Tajfel and Turner, 1986), and *ii)* the social status of the group, measured by the economic payoff of the group relative to a reference group (Shayo, 2009).

Perceived distance from the group. Each individual is characterized by a vector of attributes $q_i = (q_i^1, ...q_i^H)$ defined for each trait $h \in \{1, ..., H\}$. In our setting, traits have to be understood as cultural values or memes as defined in Desmet and Wacziarg (2021), namely individual attitudes and preferences covering a wide range of dimensions such as religiosity, cultural capital, trust in institutions, among others. Attributes are the expression, or manifestation, of each trait. For instance, if the trait is religiosity, being extremely religious or completely agnostic are distinct manifestations (attributes) of it.[9] A given group J is characterized by its group members' average attributes $q_J = E[q_i | i \in J]$. Indeed, groups are assumed to be sufficiently large so that q_J remains constant, and the average attributes of the group J are not affected by the inclusion of a new individual i who identifies with J.

The perceived distance between individual i and J's average group member is defined as the weighted Euclidean distance between each q_i and q_J such that:[10]

$$d_{iJ} = \left[\sum_{h=1}^{H} w_h (q_i^h - q_J^h)^2 \right]^{0.5} \tag{1}$$

where w_h is the attention weight that is placed to trait h (with $w_h \geq 0$ and $\sum_h^H w_h = 1$). The weights capture the salience of each trait, namely, how attention is divided between all traits. Thus, perceived distance from the representative member of group J can be influenced either by changes in individual attributes ($\partial d_{iJ} / \partial (q_i^h - q_J^h) > 0$) or by the salience of specific traits relative to others (w_h).

Status of the group. The status of a given group J (S_J) is defined by a set of exogenous factors σ_J, such as its historical prestige, but also by social comparison with other groups.

[9] Attributes are not exogenous since individuals can change them with varying degrees of ease. However, for some traits, which are not the object of our study, attributes are given, and they may include distinctive features of the immigrant population, such as skin color or accent, in contrast to attributes like language proficiency or names, which can often be adjusted more readily (Biavaschi et al., 2017). These attributes are still of major importance since, unlike more interpersonal attitudes, they are more directly observable to the broader population and therefore can contribute more strongly to the salience of immigration.

[10] As detailed by Shayo (2020), this echoes the definition of selective attention by Nosofsky (1986) as "differential weighting of the dimensions in the conceptual space".

In economic applications, social comparison can be modeled by comparing the material or economic payoffs of each group J with its reference group $R(J)$. For simplicity, it could be modeled through a linear function such that the status of the group J is given by:

$$S_J = \sigma_J + \Pi_J(a) - \Pi_{R(J)}(a) \tag{2}$$

where Π_J represents the average payoffs of J, (i.e., the average individual payoffs of group J's members $\Pi_J = E[\pi_i | i \in J]$). Equation (2) makes explicit that group status is an increasing function of the average payoff of the group ($\partial S_J / \partial \Pi_J(a) > 0$) and a decreasing function of the average payoffs of members of the reference group ($\partial S_J / \partial \Pi_{R(J)}(a) < 0$), which aligns with the seminal work by Tajfel and Turner (1986) on the in-group and out-group bias.

Individual Utility Maximization Problem. Drawing on micro and experimental evidence from the social identity literature, Shayo (2020) assumes that individuals derive utility not only from material payoffs, but also from social status and from their perceived distance to the groups with which they identify. Thus, the utility of an agent i that identifies with J is:

$$U_{iJ}(a) = \pi(a) - \beta_i d_{iJ}(a) + \gamma_i S_J(a) \tag{3}$$

Again, the negative sign in front of $\beta_i > 0$, reflects that individuals derive utility from conformism, specifically by reducing the perceived distance from members of the group they belong to. Conversely, the positive sign of the $\gamma_i > 0$ reflects the utility gain associated with the status of group J. This utility function clarifies that individuals can increase their utility through different types of actions, which are not mutually exclusive and can happen simultaneously. On the one hand, they can enhance the social status of their group S_J by taking actions that maximize the group's payoff (or by trying to reduce the payoff of the reference group) and they can adjust their attributes q_i to better conform with other group members;[11] both actions operating through a constant identity choice. On the other hand, they can maximize their utility by changing their identity and selecting a new group that offers a higher status and/or a lower perceived distance. The endogenous nature of identities can therefore be summarized under the following maximization problem for an individual i, which has to chose his social identity J and a set of actions a_i, which includes attributes q_i, such that:

$$max_{J \in G_i, a_i \in A_i} \{\pi(a) - \beta_i d_{iJ}(a) + \gamma_i S_J(a)\} \tag{4}$$

2.2 From individual choices to aggregated diversity

We connect the model defined by Shayo (2009) to the literature on cultural diversity (Desmet et al., 2017; Desmet and Wacziarg, 2021) by shifting from individual preferences to their collective expression as aggregated cultural diversity indices. Indeed, a key objective of this paper is to understand how immigration affects the cultural fragmentation of host country populations. This is fundamentally a question of heterogeneity (variance) rather than the specific direction of cultural change (mean). As a result, focusing on second-order moments seems to be more appropriate with the theoretical framework developed by Shayo (2009), as it captures the ef-

[11] Conformism arises as individuals derive a premium from coordinating on the same values as the majority and/or simply because people do not like to differ from mainstream views (Alba and Nee, 2009; Desmet and Wacziarg, 2021).

fects on the homogenization of the native group without making prior assumptions about the directional shift in cultural norms, which can be trait-specific (Desmet and Wacziarg, 2021).

The individual maximization problem defined in Equation (4) predicts, at the equilibrium, a vector of attributes q_i that maximizes each individual's utility. At the aggregated level, one can define s^{q^h} as the share of the population that holds the q manifestation of the trait h.[12] Following Desmet and Wacziarg (2021), this allows us to define an index of overall heterogeneity in the resident population (CF), which represents the average diversity of the whole population across the various cultural traits h. It is defined as:

$$CF = \frac{1}{H}\sum_{h=1}^{H} CF^h = 1 - \frac{1}{H}\sum_{h=1}^{H}\sum_{q^h=1}^{Q^h}\left(s^{q^h}\right)^2 \tag{5}$$

CF captures the likelihood that two randomly chosen individuals from the entire population hold a different attribute of a randomly chosen trait. To derive a measure of cultural diversity that does not focus only on the overall population but also highlights the relevance of social identity cleavages, such as birthplace, we decompose our index of overall heterogeneity in its within-groups and between-groups components.[13] By focusing on a subset of groups $g \in G$, we can can first compute the average within-group heterogeneity for a trait h is such that:

$$CF_g^{W_h} = \sum_g share_g CF_g^h = \sum_g share_g \left(1 - \sum_{q^h=1}^{Q^h}(s_J^{q^h})^2\right) \tag{6}$$

where $share_g$ is the share of individuals from group g in the overall population and CF_g^h the within-group g heterogeneity for trait h. It is important to note that in a setting characterized by an unbalanced distribution of the groups, as in the case of natives and immigrants, changes within the largest group (the native population here) account for the majority of the variation of the CF^{W_h} index. As for the previous case, we can average this index over all traits h to obtain the overall within-group heterogeneity:

$$CF_g^W = \frac{1}{H}\sum_{h=1}^{H} CF_g^{W_h} \tag{7}$$

Finally, the between-group component F_{ST}^g corresponds to a measure of cultural fixation (Wright's fixation index), namely the share of the total population's cultural diversity that is not due to within-group diversity:

$$F_{ST}^g = \frac{CF - CF_g^W}{CF} \tag{8}$$

It is worth noting that F_{ST}^g equals one when there is no within-group heterogeneity g, hence there is a perfect overlap between attributes and groups. Conversely, F_{ST}^g equals zero suggests that the group an individual identifies with provides no information on his/her cultural traits.

By computing Equations (6) and (8) focusing on groups defined over birthplace ($g = B \in \{N, I\}$), we can interpret F_{ST}^B as a measure of the salience of birthplace. An higher value of F_{ST}^B

[12] For each trait h, there are $q^h \in \{1^h, .., Q^h\}$ attributes or manifestations. The number of different attributes are trait-specific.

[13] For instance, focusing on birthplace as identity marker, would then describe the population over two groups: N (natives) and I (immigrants).

indicates that birthplace is more predictive of individuals cultural stances, thereby making it a more salient social identity. Conversely, a lower value of F_{ST}^B suggests that cultural diversity is largely explained by variation within groups, and that birthplace provides little information about individual norms. In that case the salience of birthplace is low.

2.3 How does immigration affect individuals' social identity choice and cultural diversity?

Within the aforementioned framework, we hypothesize a society composed of four groups, with $G = \{R, P, N, I\}$. Groups R and P are categorized based on their income attribute, distinguishing between the rich and the working class (poor), while N and I differentiate individuals based on their birthplace attribute, dividing natives and immigrants.[14] Assume that, initially, the society contains a marginal share of foreign-born individuals such that *i)* the salience attached to immigration-related traits is very low, and *ii)* natives share similar immigration-related characteristics such that the distance between each native in those attributes and the average native is close to zero.

Some individuals may have an immigrant background as second-, third-, or later-generation descendants. Still, due to cultural assimilation and their long-standing presence in the population, we assume that they are considered part of the native population, such that only newly arrived foreign-born individuals can increase the salience of immigration.[15] In this society, given the low number of foreign-born, birthplace is not relevant to identity traits, and thus individuals sort themselves based on other characteristics, such as income, for instance, ultimately identifying as either part of the rich or the working class.

Howeover, with a recurrent inflow of foreign-born individuals, immigration begins to increase. Given that immigrants have distinct attributes from the initial native population, this raises the salience of birthplace. An individual i exposed to this immigration inflow ceases to identify with her initial group J and begins to identify as a native (group N) as long as:

$$U_{iJ} < U_{iN} \Leftrightarrow S_N - S_J > (d_{iN} - d_{iJ})\frac{\beta_i}{\gamma_i} \tag{9}$$

This shift can occur through either of the two mechanisms previously explained. The inflow of foreign-born may hold immigration-specific attributes (such as ethnicity, accent, skin color, or distinct cultural traits), which create divisions within the group J. This provides us with a first testable hypothesis such that:

H1: *A rise in immigration increases the salience of birthplace (increase F_{ST}^B), as foreign-born individuals introduce immigration-specific attributes into the host society.*

As immigration increases, it amplifies the perceived distance d_{iJ} between a given native i

[14] It is important to note that the focus on income as the initial partition of the population aligns with Shayo's original model, which emphasizes redistribution. However, alternative partitions could be considered, such as religiosity, urbanicity, gender, and others. Additionally, for simplicity, we just focus on the distinction between natives and foreign-borns, without increasing the complexity of the model accounting for country-of-origin specific identities.

[15] This echoes recent work by Fouka et al. (2022), which shows that, in certain contexts, the arrival of new minority groups may increase the likelihood that existing minorities are perceived as part of the majority community.

and the average member of his group J, generating a utility cost for the latter. On top of that, immigration may also increase the salience of specific immigration-related traits (Tabellini, 2020; Fouka et al., 2022). By increasing the attention weights w_h on these traits, the relative importance of other traits, which may have been crucial for the native i to initially identify with J, is reduced. If the perceived distance d_{iJ} from the initial group becomes too large, the individual may switch to another social identity, starting to identify herself as native N. Through the described mechanisms, immigration contributes to a rise in the national sentiment within the native population: by getting a higher relative benefit, individuals are more likely to choose a birthplace-related social identity, rather than an income-related one. As a result, they sort into a group N which they perceive as more homogeneous than their initial group J, and they plausibly change some of their cultural attributes to minimize the distance from the average member of this new group. The force of conformism leads to convergence in the attributes of the native population, eroding the distinctiveness of initial groups and drawing individuals toward a common identity N. This provides us with a second testable hypothesis such that:

H2: *In response to the perceived challenge of cultural diversity, natives increasingly align their norms and values with those of the broader native-identified population, inducing a process of cultural convergence (CF^N decreases, and ultimately CF, which is largely driven by changes in the native population).*

In addition, immigration may also induce a re-categorization of the initial population by affecting the social status of the original groups through compositional effects or labor market competition, for instance. The literature on the labor market effect of immigration indeed suggests that immigration can negatively impact the wages of natives who are the closest substitutes for immigrant labor (Borjas, 2003; Card, 1990), while positively affecting the wages of natives whose skills complement those of immigrants (Ottaviano and Peri, 2006). For instance, an inflow of low-skilled immigrants, by exerting downward pressure on the wages of the working class, may negatively affect the average economic payoff of this group (Π_P) while positively impacting the average payoff of the rich group (Π_R), increasing the relative social-status distance of the income groups ($S_R - S_P$). Such a rising difference in the social groups negatively affects the utility of agents belonging to P, making the identification with this group less attractive.

Shifts in identities within the native population also likely depends on both the characteristics of immigrants and natives, as well as the interactions between the two groups. We expect, therefore, stronger shifts among natives who are directly exposed to immigrant inflows, as birthplace should become particularly salient for them. For instance, given that a large share of immigrants in Europe are low-skilled (Dorn and Zweimüller, 2021) and that they likely present larger cultural differences with the native population of receiving countries, identity shifts should be more pronounced among low-skilled natives. Second, immigrants who are culturally distant are likely to introduce more distinctive attributes into the population, which should make them more visible to the native population and produce larger identity shifts within the latter. This provides us with a third testable hypothesis such that:

H3: *Identity shifts and cultural convergence among natives are stronger with inflows of culturally distant immigrants, and among natives with similar socioeconomic characteristics to immigrants.*

3 Data

This section outlines our two main sources of data: (i) the *European Social Survey* (ESS) in Section 3.1 to compute various indices of cultural diversity; and (ii) the *European Labor Force Survey* (EU-LFS) in Section 3.2 to measure immigration stocks at the regional level.[16] Section 3.3 reports descriptive statistics on the final sample of analysis as well as preliminary evidence on the relationship between immigration and overall cultural diversity.

3.1 Cultural Diversity Data

The European Social Survey (ESS) is a multi-country individual-level survey conducted every two years since 2002 to track the distribution and evolution of values and attitudes across European countries. In each country-wave, the ESS selects a representative sample of approximately 1,500 individuals who are surveyed at home by trained interviewers.[17] The survey collects a rich set of personal and household socioeconomic characteristics, such as education, age, birthplace, and parents' background as well as several answers on cultural values.

The ESS encompasses data from 39 European countries, although not all countries participated in every wave of the survey. To ensure an adequate dataset for panel analysis, we exclude countries surveyed in fewer than five waves (Albania, Croatia, Kosovo, Latvia, Luxembourg, Montenegro, North Macedonia, Romania, Serbia, and Turkey) and countries not belonging to the European Union (Iceland, Israel, Norway, the Russian Federation, and Ukraine). Our final sample includes 23 European countries, which are composed of 175 NUTS-2 regions.

Regarding the selection of cultural traits for constructing the measure of regional cultural diversity presented in Equation (5), we adopt a comprehensive approach commonly used in the literature. This approach entails analyzing a broad array of traits and memes that can be linked to culture, as evidenced in previous studies (Alesina et al., 2017; Rapoport et al., 2021; Desmet and Wacziarg, 2021; Jaschke et al., 2022). Two criteria lead our selection of variables. First, whether possible, these traits should overlap the ones used in other studies (e.g., Alesina et al. 2017). Second, the traits should be part of the core module of the ESS, hence asked in every wave. Based on these criteria, we carefully select 46 cultural traits, detailed in Table A-2 in the Appendix.

[16] The analysis at the regional level relies on the data matched from the European Social Survey (ESS) and the European Labor Force Survey (EU-LFS). To properly match the data and to be able to track consistent regions over time, we made some methodological choices. These choices mainly relate to the small number of observations of specific regions and the way regional units are defined in the different datasets. These choices are reported in Appendix A. Specifically, we primarily use the NUTS 2 level for most countries. However, due to data limitations, the analysis is conducted at the NUTS 0 level for Cyprus, Estonia, Lithuania, and the Netherlands, and the NUTS 1 level for Austria, Germany, and the United Kingdom. The list of regions and countries is reported in Table A-1 in the Appendix.

[17] It is worth noting that the ESS first introduced self-completion surveys in Round 10 (2020–2022), mainly in response to the COVID-19 pandemic, when some countries were unable to conduct face-to-face interviews. In that self-completion surveys some questions have not been elicited, such us OP1 to OP14, SM2 and SM3 of Table A-2. Although face-to-face interviewing remained the dominant mode in Round 11 (2023–2024), several countries started to implement parallel self-completion interviews. Because these methodological changes may affect comparability across waves, and lack of some cultural variables in ESS10, we adopt a cautious approach and exclude these rounds from our benchmark analysis.

From Individual observations to Aggregate measures. Following the measurement framework presented in section 2.2, we first compute for each region-year the overall measure of cultural diversity ($CF_{r,t}$), using individual weights provided by the European Social Surveys to make our measures representative. We then decompose the overall measure by its within- and between-group components relying on different identity marker to split the population. Our main identity marker is birthplace (native vs. foreign-born), which will allow us to explore its salience after the exposure to immigration (F_{ST}^{B}) and analyze the response within the native and immigrant populations (CF_{B}^{W}). Regarding its within- and between-group dimensions, the correlation between the within component and the overall heterogeneity in our sample is close to one (0.938) as depicted in Figure D-3 in the Appendix, confirming a strong potential co-movement between the overall and within-origin components. This is not surprising, observing that the within-component is a weighted average of cultural diversity computed among either natives or immigrants, and the weights associated with natives always largely dominate those for immigrants in all regions. The correlation between overall cultural diversity and the between-group dimension stands at -0.343. Nonetheless, we also decompose the overall diversity relying on alternative identity markers highlighted as relevan from the literature (Desmet and Wacziarg, 2021), such as education (low- vs. high-skilled), gender, urbanicity (urban vs. rural) and belonging to a religious confession.

3.2 Immigration Data and Regional Characteristics

This study leverages immigration inflows to assess variations in the salience of birthplace. While increases in the salience of birthplace can also be influenced by factors such as political discourse (Card et al., 2022; Bhatiya, 2024) or media coverage (Keita et al., 2023; Schneider-Strawczynski and Valette, 2025), immigration inflows offer the advantage of greater comparability across regions and over time.[18]

We retrieve information on the size and composition of the immigrant population at the regional level from the *European Labor Force Survey* (EU-LFS), which collects information on a representative sample of the population above 15 years old. From 2004 on, it provides information on respondents' birthplace over fourteen broad regions.[19] We aggregate this information at the region-by-year level to obtain the stock of foreign-born ($k_{r,t}$), and we then decompose it by migrant population characteristics, such as education (tertiary and not tertiary educated), and length of stay in the host country (less than 5 years, between 5 and 10 years, more than 10 years). We then define the share of foreign-born over the total 2004 population as follows:

$$m_{r,t} = \frac{k_{r,t}}{Pop_{r,2004}} \tag{10}$$

where $Pop_{r,2004}$ is the 2004 total population of the region r. By keeping the population

[18] Indeed, if increased media attention on immigration generates similar predictions within the aforementioned model (even at constant immigration stocks), media data are generally less accessible, less comparable across regions and time, and are subject to additional endogenous biases due to their selective reporting on immigration news.

[19] The fourteen birthplace regions are: EU15 country, another EU country included with the 2004 expansion, another EU country included with the 2007/2013 expansion, EFTA, Other Europe, North Africa, Other Africa, Near and Middle East, East Asia, South and South East Asia, North America, Central America and Caribbean, South America, and Australia and Oceania.

in the denominator fixed at its 2004 value, we avoid empirical results from being affected by potentially endogenous native population growth over the period (Moriconi et al., 2022; Orefice and Peri, 2024). By replacing the total stock of migrants with their decomposed counterparts by education and length of stay, we provide the share of migrants of different types over the 2004 population. Again, all our measures are computed taking into account individual weights provided by the EU-LFS.

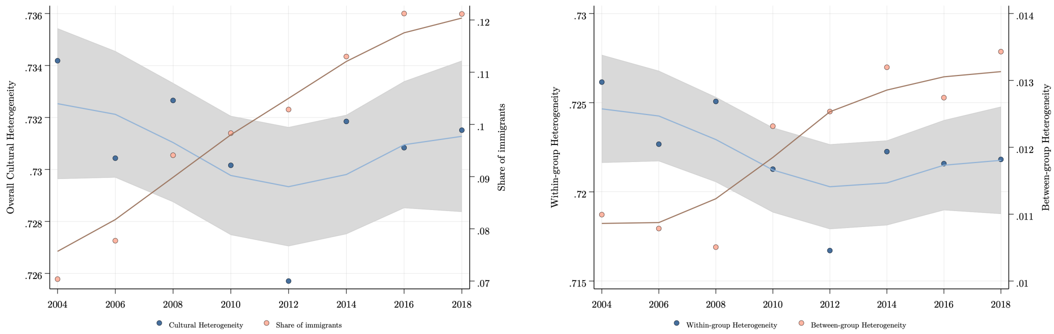
It is important to note that our main analysis focuses exclusively on first-generation immigrants, excluding second-generation individuals (native-born with at least one parent born abroad) from the main sample. Throughout this paper, we conduct several robustness checks in this regard. Notably, we show that our results remain unaffected by the reintegration of second-generation individuals into our sample and that the effects we uncover operate only through inflows of first-generation migrants.

3.3 Sample of Analysis and descriptive evidence

Sample of analysis. Combining cultural, immigration data, and other regional relevant characteristics, we end up with a sample of 1,235 regional-year observations, which corresponds to an unbalanced panel of 175 distinct regions from 23 countries across 8 waves (even years between 2004 and 2018). Summary statistics associated with this sample are reported in Table A-3 in the Appendix. The average degree of overall cultural diversity is around 0.731. Dissecting this measure into its two primary components - within-group and between-group heterogeneity - confirms that a sizeable share of cultural diversity is attributed to the within-group component (0.722), consistently with findings from previous studies on US data (Desmet and Wacziarg, 2021). The average share of migrants is around 10% of the total population and is mainly driven by low-skilled immigrants (7%) or immigrants coming from outside the European borders (6.1%). The average geographical distribution of our sample is presented in Figure A-1 in the Appendix A.4. It shows that regions with high cultural diversity often overlap with areas of high immigrant concentrations.

Relevance of birthplace as a cultural cleavage. We quantify the relevance of birthplace, alongside traditional identity markers, exploring the incremental explanatory power (R^2) of each trait across 46 cultural traits. The findings reported in Appendix B.1 report that birthplace is as influential as gender or marital status, for instance, and has gained importance over time, with its explanatory power tripling over our period of analysis. Additionally, Appendix B.2 takes advantage of the individual-level dimension available in the European Social Survey to show that, on average, immigrants indeed differ significantly from natives across almost all cultural traits. In detail, immigrants, on average, introduce significantly more conservative values to their destination country on sexual morality and religiosity: immigrants tend to be more religious, hold more conservative views on gay rights, and are more inclined to believe that traditions and customs must be followed, for instance. They are also less likely to be politically engaged at the destination. On the other hand, immigrants tend to lean more toward left-wing political views compared to the native population, and they report a higher level of trust and more positive attitudes toward immigrants. This aligns with previous theoretical predictions that foreign-born individuals exhibit distinctive cultural features, which can raise the salience of birthplace and induce shifts in the identities, attributes, and cultural diversity within the native population.

Figure 1: Cultural diversity and immigration - Evolution over time.



(a) Overall CF and Immigration (b) Between and Within Components

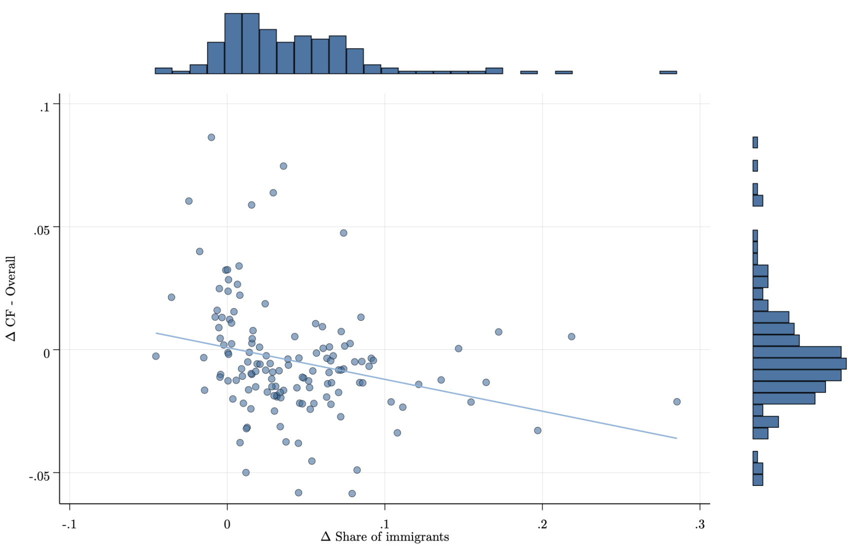
Notes: Figure (a) plots the average cultural diversity index and share of immigrants at the NUTS 2 level. Figure (b) plots the average between and within components of cultural diversity. Source: Authors' calculations on ESS and EU-LFS data (2004-2018).

Preliminary Evidence. Descriptive statistics on the evolution of cultural diversity and immigration are reported in Figure 1, which illustrates the average trends of our primary variable of interest. Similar to Desmet and Wacziarg (2021) in the US, Figure 1(a) shows that overall cultural diversity in Europe exhibited a U-shaped pattern. There was a mild decline in the early part of the period, which could be partially attributed to the high degree of economic insecurity following the financial crisis. This decline was followed by a period of positive growth, ultimately bringing the overall degree of cultural diversity back closer to its initial level.[20] The share of immigrants evolved with a positive and stable trend, moving from around 7% to 12% of the 2004 population over the 2004-2018 period. Concerning the decomposition of the two dimensions of cultural diversity, Figure 1(b) shows that the within-group components experienced a similar trend to the overall cultural diversity, while the average cultural fixation component, which is a proxy of the salience of birthplace as identity marker, experienced a positive trend starting from 2008 onward.

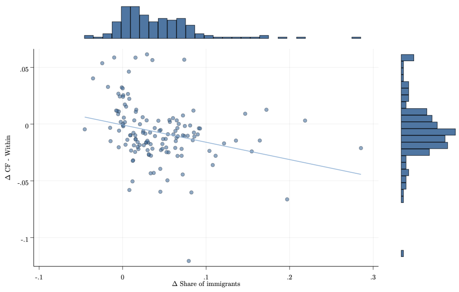
To get even closer to our empirical analysis, we also report in Figure 2 the partial correlation between average long-term variations in our measures of cultural diversity and the share of immigration between 2004 and 2018. Figure 2(a) suggests that regions that experienced larger inflows of immigrants have also experienced a larger decrease in their overall cultural diversity. At the same time, the decomposition of the within- and between-group components reveals opposite relationships. On the one hand, Figure 2(b) shows that the variation of the within-group component is negatively related to the variation in the share of immigrants. On the other hand, Figure 2(c) depicts a slightly positive correlation between immigration and the

[20]Stewart et al. (2020) find an association between economic insecurity and polarization, driven by risk aversion regarding interactions with out-groups.Additionally, several other studies have demonstrated links between economic insecurity in the aftermath of the 2008-2009 financial crisis, the decline in political trust (Wroe, 2016; Algan et al., 2017; Foster and Frieden, 2017; Tormos, 2019), and the rise in demand for populism and far-right voting (Funke et al., 2016; Ausserladscheider, 2019; Guiso et al., 2020; Guriev and Papaioannou, 2022; Ivanov, 2023); all these potentially leading to reduced cultural diversity by fostering a more homogeneous identity.

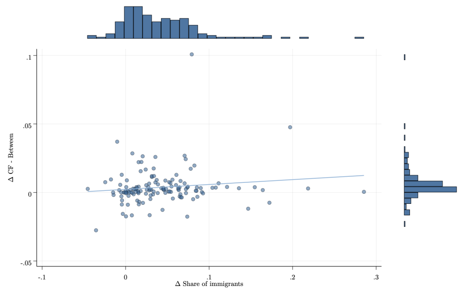
Figure 2: Cultural diversity and immigration - Long-term variations.



(a) Overall



(b) Within-group



(c) Between-group

Notes: These scatterplots illustrate the relationship between the change in cultural diversity (Overall, Within- and Between-group), as defined in Eq. (5), and the change in the Share of Immigrants, as defined in Eq. (10), between 2004 and 2018 for different regions. Histograms display the distribution of the change in the two variables.
Source: Authors' elaboration on ESS and EU-LFS data (2004-2018).

extent to which birthplace is a good predictor of cultural diversity. Plausibly contaminated by the non-random allocation of immigrants across regions, this first set of observations calls for a more formal empirical analysis accounting for both unobserved factors and the non-random distribution of immigrants across European regions.

4 Empirical Strategy

This section outlines the empirical strategy of the paper, with the objective of assessing the impact of immigration on the evolution of the cultural diversity of recipient countries. We begin by introducing the benchmark specification in Section 4.1. Then, Section 4.2 discusses the potential threats to identification and describes our identification strategy and identifying assumptions, which are based on a Shift-Share (Bartik) instrument.

4.1 Benchmark Specification

The benchmark specification features $I_{r,t} \in \{CF_{r,t},\ CF^W_{r,t},\ F_{ST_{r,t}}\}$ an index of of cultural diversity (as described in Section 2.2) in the region r at time t as a dependent variable and $m_{r,t}$ as the share of foreign-born over the total 2004 population as the main variable of interest, such that:

$$I_{r,t} = \alpha + \beta_1 m_{r,t} + \boldsymbol{\beta' X_{r,t}} + \gamma_t + \gamma_r + \varepsilon_{r,t} \tag{11}$$

where $\boldsymbol{X_{r,t}}$ is a parsimonious vector of time-varying controls at the regional level including the log of population density, the log of GDP per capita, the share of high-skilled in the population, and the unemployment rate.[21] The parameters γ_t and γ_r stand for year and regional fixed effects, respectively, which control for common aggregated change over time as well as time-invariant regional characteristics.[22] Standard errors are clustered at the regional level since regions are our treated units (Abadie et al., 2023).

The coefficient β_1 is our main coefficient of interest. It measures the marginal impact of immigration on a given cultural index $I_{r,t}$ of the receiving country. When using $F_{ST_{r,t}}$ as the dependent variable, we expect $\widehat{\beta_1}$ to be positive in line with our previous hypothesis (H1). This happens if immigration introduces attributes that increase the salience of birthplace and challenge the identity of natives who are directly exposed to immigrant inflows. When using $CF_{r,t}$ and $CF^W_{r,t}$ as dependent variables, we test the hypothesis (H2), and the predictions of the model in Section 2.3, that the arrival of foreign-born individuals decreases cultural diversity among the native population. In that case, we expect $\widehat{\beta_1}$ to be negative.

4.2 Identification Strategy

Estimating Equation (11) with OLS provides a first insight into the partial correlation between immigration and cultural diversity. Still, immigrants' location choice is not random; therefore, this specification may suffer from endogeneity bias, and the estimated coefficients cannot be interpreted in causal terms under two conditions: i) time-varying specific regional shocks drive the correlation between immigration and cultural diversity, or ii) immigrants select their locations of residence based on the prevalent cultural diversity. Specifically, if immigrants choose to live in regions with higher levels of multiculturalism rather than randomly, it could create a spurious positive correlation between immigration and cultural diversity.[23] These endogeneity threats, which can be recasted in the form of omitted variable bias and reverse causality,

[21]Controls are taken from the harmonized *Eurostat* data, which provides time-varying regional characteristics over time.

[22]It is worth noting that this benchmark equation does not include country-year fixed effects. Indeed, four countries lack sub-regional data, and given that immigration dynamics over time are largely shared across regions within each country, the variability of the instrument would be limited under this configuration. However, we report in Section 5 that our main conclusions remain unchanged when time-fixed effects are interacted with broader groups of countries based on geography, the 2004 EU enlargement, or differences in welfare systems.

[23]Immigrants may be a self-selected sample of the origin population (Docquier et al., 2020a). This aspect can be an issue as long as this cultural self-selection drives their destination choice.

are rather common in empirical studies on immigration, particularly those investigating labor market effects (see Edo (2019) for a review of this literature).

To tackle this issue, we adopt an instrumental variable (IV) strategy, relying on a shift-share approach (Card, 2001). Such an approach has been widely used in the migration literature (e.g., Ottaviano and Peri, 2006; Docquier et al., 2020b; Derenoncourt, 2022) and it builds on the well-documented empirical observation that contemporaneous inflows of migrants from a given origin allocate across different destinations based on the historical geographical distribution of migrants from the same origin. Thus, using information on the initial breakdown by immigrants' origins across regions, one can predict exogenous stocks of immigrants by applying the same allocation scheme to subsequent aggregated inflows. Such an approach then provides a source of variation of immigration that is only driven by the historical distribution of immigrants and by the total inflows by origin, and not by other factors that may drive immigrants' destination selection, such as the region-specific changes in cultural diversity or other unobserved factors. Under the assumption that the predicted immigration flows are orthogonal to omitted characteristics that are correlated with changes in cultural diversity after 2004, the newly generated and as good as random allocation of immigrants can allow for a causal interpretation of our estimates.

Recent developments of the shift-share literature point out that the validity of the instrument relies either on this exogeneity of the initial distribution of immigrants by origin (Goldsmith-Pinkham et al., 2020) or on the exogeneity of the aggregate shocks (Borusyak et al., 2022). Given our empirical setting, our approach better matches the identifying assumption of the exogeneity of the aggregate shocks by origin. To put it differently, we assume the variation of the aggregate inflows of immigrants by origin to be exogenous to the variation of regional-specific cultural diversity.[24] If true, the shift-share approach provides a source of exogenous variation of the immigrant population with respect to cultural diversity. To assuage the concerns about the validity of this identifying assumption, we first present the standard shift-share approach and then propose a modified shift-share approach with predicted aggregate flows through exclusively origin region-specific shocks. By purging out destination-specific pull factors, such an approach is more likely to satisfy our main identifying assumption.[25]

Standard shift-share based instrument. We define $Sk_{o,r,2004}$ as the initial presence of foreign-born from origin o in the hosting region r in 2004 as the share of the total immigrants from the same origin country as follows:

$$Sk_{o,r,2004} = \frac{k_{o,r,2004}}{\sum_r k_{o,r,2004}} \tag{12}$$

where $k_{o,r,2004}$ is the stock of foreign-born from origin o living in region r in 2004. Our initial year is 2004 since it is the first year in which the EU-LFS provides the fourteen disaggregated birthplace regions. Then, we compute $Tk_{o,t}$ the total stock of foreign-born for each origin o

[24]For instance, the inflows of immigrants from North Africa in our whole sample of European regions should be orthogonal to the changes in cultural diversity in the Brussels-Capital (B10) region.

[25]It is worth noting that our main conclusions remain unchanged when using a simpler version of the shift-share instrument without using immigration stocks estimated from the zero-stage bilateral migration gravity equation.

and year t such as:

$$Tk_{o,t} = \sum_r k_{o,r,t} \tag{13}$$

This allows us to construct a predicted stock of foreign-born from origin o in the region r at year t based on their initial distribution in 2004 as the interaction between $Tk_{o,t}$ and $Sk_{o,r,2004}$, such as:

$$\tilde{k}_{o,r,t} = Sk_{o,r,2004} \times Tk_{o,t} \tag{14}$$

The aggregate time-variant stocks by origin are then distributed across the regions of our sample based on the 2004 distribution. Finally, we compute the region r and year t predicted migration share ($\tilde{m}_{r,t}$) by simply taking the sum of all $\tilde{k}_{o,r,t}$ predicted stocks across origin, as follows:

$$\tilde{m}_{r,t} = \frac{\sum_o \tilde{k}_{o,r,t}}{Pop_{r,2004}}. \tag{15}$$

Modified shift-share based instrument. To enhance the validity of our identifying assumption, which relies on the exogeneity of the aggregate shocks (Borusyak et al., 2022), we modify our shift-share approach by replacing $Tk_{o,t}$ with its predicted version obtained from a zero-stage bilateral migration gravity equation that includes *exclusively* push factors such as conflicts and natural disasters in migrants' origin countries as explanatory variables. This novel approach of combining gravity models with shift-share instruments has gained traction in recent migration literature (see Ortega and Peri, 2014; Docquier et al., 2020b; Orefice et al., 2025, among others). Our gravity equation looks as follows:

$$k_{o,d,t} = \alpha_1 \ln(Deaths_{o,t}) + \alpha_2 Disaster_{o,t} + \beta_t \ln(distance_{o,d}) + \theta_{d,t} + \theta_{o,d} + \varepsilon_{o,d,t} \tag{16}$$

where $k_{o,d,t}$ is the bilateral stock of immigrants from origin country o to destination d at year t sourced from the United Nations (UN, 2020).[26] The $\ln(Deaths_{o,t})$ and $Disaster_{o,t}$ correspond to origin-specific and time-varying push factors and stand for the logarithm of the cumulative five-year count of total fatalities due to armed conflicts and the cumulative five-year count of natural disasters, respectively.[27] The parameter $\beta_t \ln(distance_{o,d})$ is the time-varying effect of distance on immigration following, which captures the differential impact of changes in technology over time across pairs of countries (Feyrer, 2019; Docquier et al., 2020b). Finally, $\theta_{d,t}$ and $\theta_{o,d}$ are destination-year and origin-destination fixed effects, respectively. While destination-year fixed effects are not used to obtain the predicted exogenous stocks, their inclusion in the gravity model allows us to enhance the precision of our estimates.[28] Equation (16) is estimated using a Poisson Pseudo Maximum Likelihood estimator (PPML), which performs well under

[26] To enhance the precision of our estimates, we maintain the complete 214×214 matrix of origin-destination pairs. The gravity model encompasses exclusively 5-year data spanning from 1990 to 2020. For the years in between, immigrant stocks are interpolated before the aggregation of the projected immigration figures.

[27] We sourced data from the UCDP/PRIO Armed Conflict Dataset for deaths (Gleditsch et al., 2002; Davies et al., 2022) and from the Emergency Events Database (EM-DAT) for natural disasters (EMDAT, 2022). Natural disasters include biological (epidemic), climatological (drought, wildfire), geophysical (mass movement, earthquake, volcanic activity), meteorological (storm, fog, extreme temperature), and hydrological events (flood, landslide).

[28] Results for the gravity model are reported in Table C-4 in the Appendix. As predicted by the theory, an increase in the number of natural disasters or fatalities due to conflicts increases international migration, while the influence of distance is found to decrease over time.

various heteroskedasticity patterns, rounding errors for the dependent variables, and a large number of zeroes (Silva and Tenreyro, 2006, 2010). Standard errors are clustered at the pair level. The total predicted stock of foreign-born for each origin o and year t is such as:

$$\begin{aligned}\widehat{Tk_{o,t}} &= \sum_d \widehat{k_{o,d,t}} \\ &= \sum_d e^{\widehat{\alpha}_1 \ln(Deaths_{o,t}) + \widehat{\alpha}_2 Disaster_{o,t} + \widehat{\beta}_t \ln(distance_{o,d}) + \widehat{\theta}_{o,d}}\end{aligned} \tag{17}$$

Relying predominantly on origin-specific time-varying shocks and purging out the variation generated by destination-specific pull factors, the predicted stocks by origin computed in Equation (17) are more likely to satisfy our identifying assumption, which assumes that the variation of the aggregate stocks should be exogenous with respect to changes of the outcome variable and unobserved factors at regional level (Borusyak et al., 2022). We aggregate them in the fourteen broad origins available in EU-LFS and use them to calculate the predicted stocks (Eq. 14), necessary to compute the modified predicted migration share ($\widehat{m}_{r,t}^{Mod}$).

To gain a deeper insight into the underlying variability that our modified shift-share approach leverages, Figure C-2 in the Appendix illustrates the variation in predicted stocks when aggregated into the fourteen broad origins available in EU-LFS. Not surprisingly, within-EU immigrants account for the biggest part of the total stock of migrants in our setting. However, upon further examination of group-specific variation, we observe that the predicted growth among different origin groups is rather similar, suggesting that our predicted variation is not driven by any specific origin group. This evidence allows us to interpret our results as being driven by the variation in the immigrant population as a whole, rather than by the variation of specific origin groups.

Finally, Appendix C provides a series of empirical checks suggested by the literature to support the validity of our identifying assumption. First, we show that our results hold with a leave-one-out version of our shift-share instrument, hence removing from the total stocks those related to each region-year observation, minimizing the presence of any correlation driven by the construction of the instrument (Autor and Duggan, 2003). More precisely, the leave-one-out estimator excludes own-destination i predicted stock of foreign-born when calculating the total predicted stock of foreign-born for each origin-year across all destinations. This allows us to enhance the exogeneity of the instrument by eliminating any remaining mechanical relationships when computing the total predicted stocks for each origin-year observation. We also provide evidence that our main conclusions remain unchanged with the more conventional approach of using the stocks of foreign-born as obtained from the EU-LFS, instead of the predicted stocks derived from the gravity model. Second, to mitigate concerns related to pre-trends, we provide evidence of no correlation between the growth of predicted total stocks and previous region-specific characteristics such as GDP per capita, population density, unemployment rate, and the share of the tertiary educated population (Moriconi et al., 2022). Third, we show that the precision of our estimates is not driven by a similar initial distribution of origin groups across regions, which could potentially bias the estimated error terms by inducing spatial correlation of shocks across regions (Adão et al., 2019). Although not essential for our identifying assumption, we additionally provide a series of tests to alleviate potential concerns associated with the time closeness between the initial distribution of our historical shares and our period of analysis. Specifically, we show that our results hold by excluding sequentially from our sample those years close to the initial share, hence increasing the gap between our initial distribution by origin and the period of analysis. Finally, the short-term variation exploited in our empirical setting is less prone to the concern of conflating short- and long-term effects of immigration on our results, as suggested by Jaeger et al. (2018).

5 Main Results

This section reports the results associated with our first two main testable hypotheses. Section 5.1 focuses on whether a rise in immigration is associated with an increase in the salience of birthplace and national identity in host societies, measured with F_{ST}. Section 5.2 provides evidence that immigration is associated with cultural convergence within the overall and native population, measured with CF and CF^W, respectively.

5.1 Immigration and the Salience of Birthplace

This section tests our first theoretical prediction (H1), that a rise in immigration increases the salience of birthplace, as foreign-born individuals introduce immigration-specific attributes into the host society. Put differently, we test whether immigration increases the relevance of birthplace as an identity marker (F_{ST}), therefore as a relevant predictor for cultural diversity.

Table 1: Immigration and the Salience of Birthplace (F_{ST}^B)
OLS and 2SLS Results.

	(1) OLS	(2) OLS	(3) 2SLS	(4) 2SLS
m_{rt}	0.032* (0.019)	0.040* (0.022)	0.027* (0.016)	0.038** (0.019)
Regional controls	No	Yes	No	Yes
Observations	1,235	1,235	1,235	1,235
Mean Cultural Index	0.012	0.012	0.012	0.012
Mean Immig. Share	0.101	0.101	0.101	0.101
First-stage			1.465	1.292
KP F-Test			82.958	48.568

Notes: *** p <0.01, ** p <0.05, * p < 0.1. Standard errors in parentheses are clustered at the regional level. The dependent variable is F_{ST}, the measure of cultural diversity in the region r at time t. The independent variable m_{rt} is the share of foreign-born in the total 2004 population. Regional controls include the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled individuals in the total population. All estimates include regional and year fixed effects.

Sources: Authors' elaboration on ESS and EU-LFS data (2004-2018).

Table 1 reports the results of our benchmark specification (Eq. 11) with the fixation index (F_{ST}) as the dependent variable, using birthplace as the identity cleavage. Cols. (1) and (2) report OLS estimates of the relationship between immigrant share and the between-component of cultural diversity (Eq. 8), controlling for regional and year fixed effects, without and with controls respectively. Cols. (3) and (4) replicate these results using 2SLS. This first set of results supports our hypothesis that a rise in the share of first-generation immigrants is significantly

Table 2: Immigration and the Salience of Identity Groups (F_{ST}^{g})
2SLS Results.

Identity Marker	(1) Birthplace	(2) Education	(3) Gender	(4) Urbanicity	(5) Religion
m_{rt}	0.038**	-0.029	-0.036**	-0.064**	-0.021
	(0.019)	(0.019)	(0.016)	(0.030)	(0.018)
Observations	1,235	1,235	1,235	1,222	1,235
Mean Cultural Index	0.012	0.016	0.014	0.012	0.023
Mean Immig. Share	0.101	0.101	0.101	0.101	0.101
First-stage	1.292	1.292	1.292	1.304	1.292
KP F-Test	48.568	48.568	48.568	49.078	48.568

Notes: *** p <0.01, ** p <0.05, * p < 0.1. Standard errors in parentheses are clustered at the regional level. The dependent variable is F_{ST}, the measure of between-group cultural diversity in the region r at time t. It is computed for immigrants and natives in Col. (1), high-skilled and low-skilled in Col. (2), male and female in Col. (3), urban and rural in Col. (4), and religious and non-religious individuals in Col. (5). The independent variable m_{rt} is the share of foreign-born in the total 2004 population. We control for the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled individuals in the total population. All estimates include regional and year fixed effects.

Sources: Authors' elaboration on ESS and EU-LFS data (2004-2018).

associated with an increase in the salience of birthplace within the society, as $\widehat{\beta_1}$ is positive and significant at conventional levels. This implies that increasing immigrants' inflows are associated with greater predictability of responses to questions on cultural norms, attitudes, and preferences based only on a respondent's country of birth. As far as the magnitude is concerned, a 10 percentage-points increase in the share of immigrants (one standard deviation) is associated with a 0.004 percentage-point increase in F_{ST} (30% of its standard deviation).

We replicate this analysis by computing alternative fixation indices based on other identity cleavages, such as education (college vs. non-college graduates), gender (male vs. female), urbanicity (urban vs. rural), and religion (religious vs. non-religious). Table 2 confirms that a rise in immigration is associated only with an increase in salience of birthplace, since we observe a decline in the predictive power of other relevant identity cleavages. These 2SLS estimates not only support the theoretical predictions of the model proposed in Section 2, but also reveal that immigration, while fostering the relevance of birthplace as an identity marker, hampered the predictive power of alternative socioeconomic cleavages.

Overall, the results presented in this section provide empirical support for the first hypothesis (H1) that immigration contributes to a rise in salience of birthplace in the society. Immigration increases the predictability of responses to questions on cultural norms, attitudes, and preferences based only on an individual's country of birth.

5.2 Immigration and Cultural Diversity

This section tests our second and main hypothesis (H2), which conjectures that a rise in immigration is associated with a convergence in cultural norms within the native population following a rise in the salience of immigration.

Table 3 presents the results of our benchmark specification (Eq. 11). Col. (1) reports the OLS estimate of the relationship between the share of first-generation immigrants and overall cultural diversity, excluding control variables, but accounting for time-varying common shocks

Table 3: Immigration and Cultural diversity (CF) - OLS and 2SLS Results.

	(1) OLS	(2) OLS	(3) 2SLS	(4) 2SLS
m_{rt}	-0.097***	-0.107***	-0.140***	-0.165***
	(0.029)	(0.029)	(0.046)	(0.061)
Regional controls	No	Yes	No	Yes
Observations	1,235	1,235	1,235	1,235
Mean Cultural Index	0.731	0.731	0.731	0.731
Mean Immig. Share	0.101	0.101	0.101	0.101
First-stage			1.465	1.292
KP F-Test			82.958	48.568

Notes: *** p <0.01, ** p <0.05, * p < 0.1. Standard errors in parentheses are clustered at the regional level. The dependent variable is CF_{rt}, the measure of cultural diversity in the region r at time t. The independent variable m_{rt} is the share of foreign-born in the total 2004 population. Regional controls include the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled individuals in the total population. All estimates include regional and year fixed effects.

Sources: Authors' elaboration on ESS and EU-LFS data (2004-2018).

and time-invariant unobserved regional heterogeneity with year and region fixed effects. This analysis reveals that an increase in the share of immigrants is associated with a significant decrease in overall cultural diversity, which is robust to the inclusion of regional controls (Col. 2). Col. (3) and (4) provide consistent results after employing 2SLS estimation, addressing potential endogeneity concerns through the use of our modified shift-share instrument described in Section 4.2. Furthermore, compared to OLS estimates, the coefficient with 2SLS becomes bigger in magnitude, aligning with the hypothesis that OLS estimates may be upward biased if immigrants choose their residency based on the region-specific high degree of multiculturalism. In terms of magnitude, the coefficient of the benchmark specification in Col. (4) reports that a 10 percentage-points increase in the share of immigrants (one standard deviation) corresponds to a 1.65 percentage-point decrease in overall cultural diversity.[29]

To gain a comprehensive understanding of the driving forces behind the negative association between immigration and overall cultural diversity in our benchmark specification, we decompose our index of heterogeneity in Table 4 into its within and between components, following the approach of Desmet and Wacziarg (2021). Col. (1) reports the benchmark result of Table 3, while the between-group (F_{ST}) estimated in the previous section and the within-group components are reported in Col. (2) and (3), respectively.

As already highlighted in Section 5.1, the positive impact of immigration on the between-group component (F_{ST}) indicates that immigrants directly contribute to cultural diversity by introducing new norms and values to their host country, thereby enhancing the salience of birthplace in the receiving society. However, this positive effect on the overall cultural diversity is largely offset by the negative association between immigration and the within-group

[29]We replicated this analysis using political rather than cultural diversity. Based on survey questions about voting in the most recent elections, we construct an index of political diversity, which we use as an alternative dependent variable in Equation 11. In contrast to cultural diversity, we find that immigration does not affect political diversity. Results are available upon request.

component of the cultural diversity index (CF^W), as reflected by the negative and statistically significant coefficient in Col. (3). Taken together, these results reveal that immigration: (i) makes birthplace a more salient identity trait, and (ii) there is a reduction of cultural diversity within the groups of natives and immigrants.

Table 4: Immigration and Cultural diversity
Decomposition - 2SLS Results.

	(1)	(2)	(3)	(4)	(5)
	Overall	Between	Within	Within Natives	Within Immig.
	CF	F^B_{ST}	CF^W_B	CF^N	CF^I
m_{rt}	-0.165***	0.038**	-0.189***	-0.206***	0.136
	(0.061)	(0.019)	(0.068)	(0.067)	(0.391)
Observations	1,235	1,235	1,235	1,235	1,111
Mean Cultural Index	0.731	0.731	0.731	0.731	0.732
Mean Immig. Share	0.101	0.101	0.101	0.101	0.110
First-stage	1.292	1.292	1.292	1.292	1.265
KP F-Test	48.568	48.568	48.568	48.568	36.954

Notes: *** p <0.01, ** p <0.05, * p < 0.1. Standard errors in parentheses are clustered at the regional level. The dependent variable in Col. (1) is CF_{rt}, the measure of cultural diversity in the region r at time t. The dependent variable in Col. (2) is F_{ST}, the measure of between-group cultural diversity in the region r at time t. The dependent variable in Col. (3) is CF^W_{rt}, the measure of within-group cultural diversity in the region r at time t. The dependent variables in Col. (4) and (5) are the measure of cultural diversity in the region r at time t for native and immigrants, respectively. The independent variable m_{rt} is the share of foreign-born in the total 2004 population. We control for the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the total population. All estimates include regional and year fixed effects.

Sources: Authors' elaboration on ESS and EU-LFS data (2004-2018).

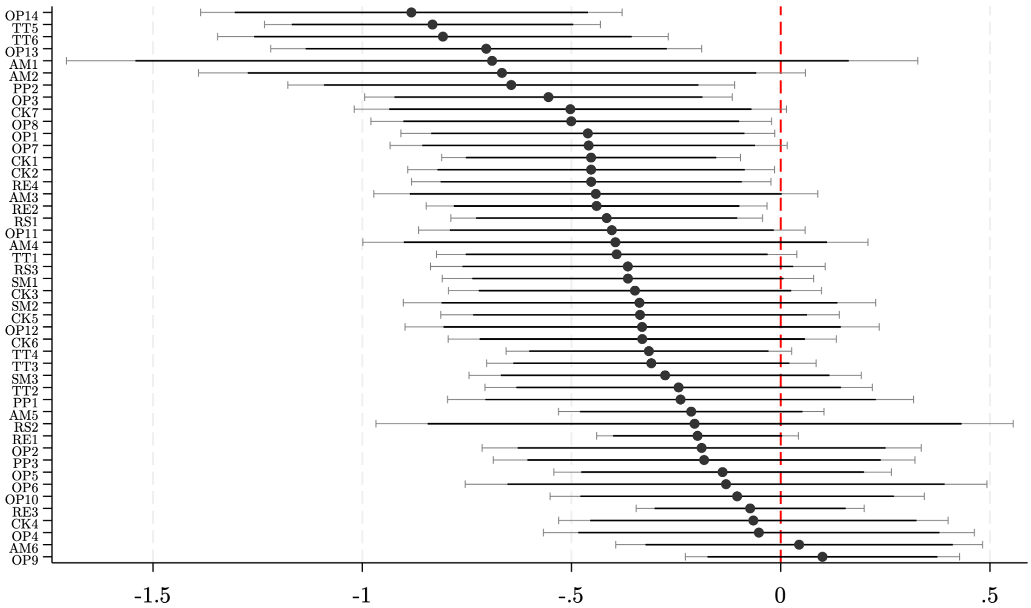
Which group is driving such a reduction in cultural diversity? Given the unbalanced distribution of natives and immigrants in the resident population, changes within the native population account for the majority of the variation of the CF^W index, thus it is not surprising to see that this negative effect remains when isolating the effect of an increased immigrant share on cultural diversity within the native population only in Col. (4) but not when replicating this analysis with immigrants only in Col. (5).[30] Overall, these findings align closely with the predictions of Shayo (2009): immigration increases the salience of one's birthplace within society, thereby increasing the pressure on natives to conform, as they perceive benefits in aligning with the predominant values of a group whose salience has been amplified by the arrival of immigrants.

One could argue that the reduction in cultural diversity within the native population could be interpreted as either a more concentrated distribution of values (i.e., unimodal distribution) or instead a polarization of values (i.e., bimodal distribution). To disentangle between these two interpretations of the estimated effect, Appendix E.2 first shows the absence of an effect of immigration on a cultural polarization index (Montalvo and Reynal-Querol, 2005) computed on the native population only. Moreover, we test whether immigration increases the salience of other identity traits among natives (such as education, gender, domicile, and religion).

[30] Although not statistically significant, the positive effect observed for immigrants could reflect the arrival of new cohorts whose cultural values differ from those of older diasporas, which may have already begun a process of cultural assimilation toward the native population.

Our results show that immigration does not foster the relevance of socioeconomic identities among natives, while there is a general decrease of cultural diversity among the various groups. Overall, these results suggest that the decrease in cultural diversity within the native population has to be interpreted as a more concentrated rather than a polarized distribution of values.

Figure 3: Convergence Within Natives by Cultural Traits (Within Component)



Notes: Standard errors are clustered at the regional level. Over the x-axis are reported the standardized estimated coefficients. We include 95% and 90% confidence intervals around the estimated coefficients. The dependent variable is CF^W_{rt} the measure of within-group cultural diversity in the region r at time t, computed for each cultural trait separately. The list of cultural trait is available at Table A-2 in the Appendix. The independent variable m_{rt} is the share of foreign-born in the total 2004 population. We control for the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the total population at the regional-level. All estimates include regional and year fixed effects.
Sources: Authors' elaboration on ESS and EU-LFS data (2004-2018).

Finally, we acknowledge that the presented results may be influenced by the multidimensional nature of our cultural diversity index. While following the existing work of the literature, such index may hide heterogeneous and even contrasting patterns across cultural traits. We explore this issue by estimating our benchmark regression using as dependent variable indexes of within-group cultural diversity computed on each trait in isolation. The 46 estimated coefficients, one per each cultural trait, are presented in Figure 3. To facilitate the interpretation of the differences in magnitudes between these estimates, the figure reports standardized beta coefficients for both the dependent and independent variables. Although this approach introduces a loss of precision due to the sequential focus on each trait, Figure 3 conveys the key message that natives tend to converge towards the same norm across a large set of dimensions, suggesting that a particular set of traits does not drive the effect.[31] This aligns with our theoretical predictions that natives derive utility from becoming closer to the average attributes

[31]When focusing on each trait separately, our estimates lose statistical power and significance. Nevertheless, the estimated coefficient is negative for almost all traits (44 out of 46) and statistically significant at conventional levels for half of the latter.

of the individuals in their group with which they socially identify.

Robustness checks

We perform additional robustness checks on our benchmark specification presented in Col. (4) of Table 4, detailed below, with all tables and figures reported in Appendix D.

Grouped fixed effects. Table D-1 reports the main results, providing a different structure of year fixed effects interacted with aggregated regions to capture time-varying regional shocks. Given that our sample includes four countries lacking sub-regional data, namely Cyprus, Estonia, Lithuania (due to the size of the countries), and the Netherlands (due to data availability), we first report in Col. (2) that our results are not affected by their removal. Then, Col. (3) reports that the inclusion of country-year fixed effects removes the statistical significance of our coefficient of interest at conventional levels. This could be attributed to the limited variability in our data and the instrument, as cultural and immigration dynamics over time are largely shared across regions within each country. The incorporation of country-year fixed effects indeed results in an additional reduction of 8 and 7 percentage points in the standard deviation of immigrant shares and overall cultural diversity, respectively, and the F-stat of the first-stage equation in the IV is divided by 2. However, reassuringly, the significance of our main effect is restored when time-fixed effects are interacted with broader groups of countries based on geographical regions (Col. 4, including Eastern Europe, Central Europe, Southern Europe, Western Europe, and Northern Europe), the 2004 EU enlargement (Col. 5, including EU15, NMS10, NMS3, and EFTA), or welfare systems (Col. 6, including Nordic, Continental, Anglo-Saxon, Mediterranean, and Eastern Europe).

Excluding regions with no immigrants. Table D-2 provides the main results after excluding from the sample regions with no recorded immigrants in our data sources. Our main conclusions remain unchanged when excluding such regions, suggesting that our results are not driven by some specific outliers with no immigrants recorded in the EU-LFS, ESS, or both.

2nd generations immigrants. Table D-3 in the Appendix shows that the inclusion of second-generation immigrants in the construction of our outcome variables does not affect our main results. Interestingly, Table D-4 shows that the effect on the between-group component is not statistically significant once we exclude first-generation from the analysis. This result suggests that inflows of immigrants foster birthplace as an identity marker, but not immigrant parental background.

Number of observations. Appendix D.4 tackles concerns about the potentially small number of observations within each region-year cell used to compute our dependent variables. First, Table D-5 presents results excluding regions with a small number of observations. Second, Table D-6 reports additional findings interacting our main effect with either a dummy variable for regions with fewer than 50 or 100 respondents. Such a test aims to capture potential heterogeneous effects driven by the number of observations in each region-year cell without modifying the sample of analysis. In both cases, we find that our results remain unaffected once accounting for regions with a limited number of observations.

Alternative cultural indices. Appendix D.5 discusses extensively the properties of alternative indices of diversity in the literature. We provide evidence that our results are robust to alternative definitions of cultural diversity including: i) an augmented cultural diversity index that assigns higher weights to answers that deviate further from the region-year average answer for a given cultural trait, ii) a discretized version of all cultural trait variables to compute our diversity indices, iii) two alternative diversity measures such as the Rosenbluth index (Hall and Tideman, 1967) and the Entropy index (Shannon, 1948), and iv) polarization index (Montalvo and Reynal-Querol, 2005).[32] All robustness checks consistently indicate a negative effect of immigration on all cultural diversity.

Natives' mobility. We investigate whether our main effect could be attributed to selective mobility among natives in response to immigration. Drawing from methodologies outlined in Edo et al. (2019) and Moriconi et al. (2022), Appendix D.6 explores the impact of lagged regional migration on changes in the native population or on changes in the inflow of natives to the region. We find no effect, even after exploring potential heterogeneous effects based on native education levels or examining the contemporaneous reactions of natives to immigration. These results suggest that the mechanism of selective mobility among natives is unlikely to drive our main conclusions.

6 Heterogeneity analysis

This section further explore our main results by conducting heterogeneity analysis based on immigrants' characteristics. Our theoretical framework predicts that more salient immigrant groups should have a stronger influence on reshaping native identity and increasing homogeneity among natives (H3). Thus, section 6.1 investigates the influence of individual characteristics such as education and duration of stay, while section 6.2 investigates the impact of cultural and economic distances between immigrants and natives, which is measured at the country pair level.

6.1 Immigrants' individual characteristics

To understand which immigrant groups increase the salience of birthplace, we compute between-group cultural diversity indexes (F_{ST}) that sequentially isolate natives and specific immigrant sub-groups based on different characteristics such as education, and years of residency, thus modifying the left-hand side of Equation (11).[33] Therefore, our analysis explores the effect of an overall immigration shock on the relevance of birthplace in explaining cultural diversity between natives and a specific group of migrants.

Table 5 reports the results of these new estimates. In Col. (2) and (3), we regress the overall percentage of foreign-born individuals within the total 2004 population against the between-group cultural diversity of two distinct hypothetical populations: one composed exclusively of natives and high-skilled immigrants, and the other composed of natives and low-skilled immigrants, respectively. Table 5 reveals a strong effect of immigration on the cultural fixation of the population with low-skilled immigrants, while no significant effect is observed for college

[32] As Figure D-3 in the Appendix shows, the polarization index is negatively correlated with our measure of cultural diversity.

[33] This prevents us from altering immigration flows in our variable of interest, which would reduce comparability across specifications that would report different first-stage equation.

Table 5: Immigrants' characteristics and the Salience of Birthplace (F^B_{ST})
Natives and Immigrants Sub-groups - 2SLS Results.

	All	Education		Duration of stay		
	(1)	(2)	(3)	(4)	(5)	(6)
Immig. charac.:	All	HS	LS	ST	MT	LT
m_{rt}	0.038**	0.004	0.050**	0.066**	0.042*	0.014
	(0.019)	(0.016)	(0.019)	(0.027)	(0.023)	(0.023)
Observations	1,235	1,235	1,235	1,235	1,235	1,235
Mean Cultural Index	0.012	0.007	0.012	0.006	0.006	0.010
Mean Immig. Share	0.101	0.101	0.101	0.101	0.101	0.101
First-stage	1.292	1.292	1.292	1.292	1.292	1.292
KP F-Test	48.568	48.568	48.568	48.568	48.568	48.568

Notes: *** p <0.01, ** p <0.05, * p < 0.1. Standard errors in parentheses are clustered at the regional level. The dependent variable is F_{ST}, the measure of between-group cultural diversity in the region r at time t for natives and first-generation immigrants' group reported in each column. The independent variable m_{rt} is the share of foreign-born in the total 2004 population. We control for the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled individuals in the total population. All estimates include regional and year fixed effects.

Sources: Authors' elaboration on ESS and EU-LFS data (2004-2018).

graduates.[34] This aligns with the observation that low-skilled immigrants represent the majority of immigration inflows in Europe but also that they are more likely to introduce novel attributes within host societies, contributing to an increase in the salience of birthplace and a rise in nationalism among low-skilled natives.

Then, Col. (4) to (6) focus on immigrants' years of residency. We find that an inflow of immigrants is associated with an increase in the relevance of birthplace as an identity marker for migrants with less than 5 years of residency, but this effect strongly decreases for immigrants between 6 and 10 years of residency and ultimately disappears entirely after 10 years of residency. This may reflect the higher salience of newly arrived immigrants, cultural assimilation - if immigrants adopt the values of the native population over time - or selective return migration, if less assimilated immigrants are more likely to return. However, we cannot rule out concerns that categorizing immigrants by years of residency may also introduce cohort effects, as different immigrant cohorts may have varying characteristics (Borjas, 1985).

We now focus on the impact of immigrants' characteristics on cultural convergence within the native population in Table 6 by focusing sequentially on specific immigrant inflows as independent variables.[35] In line with previous results, Col. (2) and (3) show that the response

[34] Additional results, available upon request, suggest that the results among low-skilled immigrants are stronger when focusing on non-European immigrants rather than European ones. These results suggest that skill-specific results depend on immigrants' origin, as we test in the next Section.

[35] There is no alternative method available here, that would allow us to keep a consistent first-stage across estimates, as immigrants have a marginal impact on the variability in within-group heterogeneity, by construction. Still, our instrument can be replicated for any migrants' characteristics, as long as we restrict $Tk_{o,t}$ to a given characteristic. It implies that the variability of such an instrument between high- and low-skilled groups, for instance, would differ mainly due to variations in skill-specific inflows from various origin countries. This difference

Table 6: Immigrants' characteristics and Natives' Response
Outcome: CF_B^W - 2SLS Results.

	All	Education		Duration of stay		
	(1)	(2)	(3)	(4)	(5)	(6)
Immig. charac.:	All	HS	LS	ST	MT	LT
m_{rt}	-0.206***	-0.269***	-0.409	-0.290***	-0.281*	-0.100
	(0.067)	(0.100)	(0.250)	(0.087)	(0.162)	(0.065)
Observations	1,235	1,235	1,235	1,235	1,235	1,235
Mean Cultural Index	0.728	0.728	0.728	0.728	0.728	0.728
Mean Immig. Share	0.101	0.027	0.074	0.020	0.018	0.061
First-stage	1.292	1.228	0.925	0.971	0.946	1.040
KP F-Test	48.568	84.600	5.935	17.036	24.544	45.407

Notes: *** p <0.01, ** p <0.05, * p < 0.1. Standard errors in parentheses are clustered at the regional level. The dependent variable is CF_{rt}^W, the measure of within-group cultural diversity in the region r at time t. The independent variable m_{rt}^g is the share of foreign-born, belonging to group g reported in each column, in the total 2004 population. We control for the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the total population at the regional-level. All estimates include regional and year fixed effects.

Sources: Authors' elaboration on ESS and EU-LFS data (2004-2018).

of natives is stronger to low-skilled immigrants than to college-graduated immigrants, although the coefficient for low-skilled immigrants is imprecisely estimated due to a weaker first stage. Moreover, the results in Col. (4) to (6) confirm that these effects are enhanced by newly arrived immigrants.

Overall, the results of this section align with theoretical predictions (H3) that immigrants' characteristics matter to explain the influence of immigration on the salience of birthplace. Particularly, low-educated immigrants and newly-arrived immigrants, who are more likely to be perceived as distant from the native population, magnify these effects.

6.2 Immigrants' Cultural Distances

This section provides additional insight into the relevance of cultural distances between immigrants and natives on the estimated relationship between immigration and natives' cultural diversity (H3).

We first compute a weighted average index of the distance between immigrants and natives as reported in Eq. (18). Weights are the share that each origin represents in the overall immigration stock of the NUTS-2 region r in year t, while $D_{r(d),o}$ represents the distance between the origin country o and the region r within the destination country d.

$$\overline{Dist}_{r,t} = \sum_o \frac{k_{r,o,t}}{\sum_o k_{r,o,t}} \times D_{r(d),o} \qquad (18)$$

may be quite low if the initial distribution of both groups across regions in 2004 was fairly similar. Additionally, the instrument's strength may vary across groups, making comparisons harder. This is the case for low-skilled immigrants in our analysis, which warrants a cautious interpretation of the 2SLS coefficients associated with this specific group.

Table 7: Cultural distance between immigrants and natives
Outcome: CF^W_B - 2SLS Results.

	(1) Economic	(2) Genetic	(3) Linguistic	(4) Religious	(5) Cultural	(6) Geodesic
m_{rt}	-0.091	-0.079	-0.124	-0.130	-0.044	-0.070
	(0.110)	(0.090)	(0.106)	(0.086)	(0.097)	(0.105)
$\overline{Dist}_{rt}$	-0.003	-0.001	0.003	0.005	-0.002	-0.002
	(0.007)	(0.008)	(0.006)	(0.007)	(0.007)	(0.007)
$m_{rt} \times \overline{Dist}_{rt}$	-0.211*	-0.439**	-0.135	-0.194	-0.282***	-0.369**
	(0.125)	(0.182)	(0.122)	(0.157)	(0.108)	(0.184)
Observations	1,235	1,235	1,235	1,235	1,235	1,235
Mean Cultural Index	0.731	0.731	0.731	0.731	0.731	0.731
Mean Immig. Share	0.101	0.101	0.101	0.101	0.101	0.101
First-stage	1.222	1.211	1.187	1.264	1.214	1.172
KP F-Test	21.006	15.229	16.959	18.033	20.518	16.512

Notes: *** p <0.01, ** p <0.05, * p < 0.1. Standard errors in parentheses are clustered at the regional level. The dependent variable is CF^W_{rt}, the measure of within-group cultural diversity in the region r at time t. The independent variable m_{rt} is the share of foreign-born in the total 2004 population. $\overline{Dist}_{r,t}$ is a weighted average index of the distance between immigrants and natives as reported in Eq. 18. All estimates include a vector of control with the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the total population at the regional level, and regional and year fixed effects.
Sources: Authors' elaboration on ESS and EU-LFS data (2004-2018).

Then we interact this variable with our main variable of interest as follows:

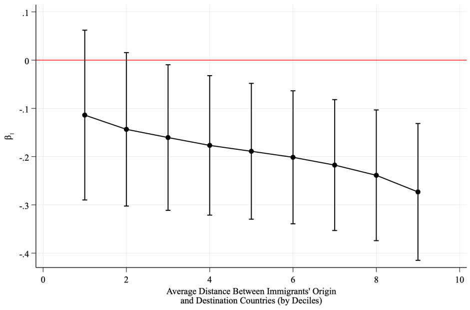
$$CF_{r,t} = \alpha + \beta_1 m_{r,t} + \beta_2 \overline{Dist}_{r,t} + \beta_3 m_{r,t} \times \overline{Dist}_{r,t} + \boldsymbol{\beta}' \mathbf{X_{r,t}} + \gamma_t + \gamma_r + \varepsilon_{r,t} \quad (19)$$

While $m_{r,t}$ controls for the size effect of immigration, $\overline{Dist}_{r,t}$ proxies for its composition in terms of attributes. The parameter β_3 represents the extent to which the marginal impact of immigration on the cultural diversity of the native population depends on the composition of the immigrant group, and more specifically, on the distances between immigrants and natives. Our theoretical framework predicts that greater distances should foster the convergence of norms within the native population, as more salient or culturally distant diasporas should disproportionately increase the salience of immigration at the destination and challenge, more importantly, the homogeneity of the native groups exposed to it ($\beta_3 < 0$).

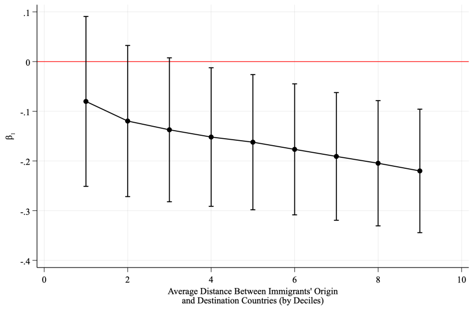
We sequentially use several measures to proxy distance between immigrants' origin country and their destination, which include geographic, cultural, genetic, linguistic, religious, and economic distances (Conte et al., 2022; Spolaore and Wacziarg, 2016; Pellegrino et al., 2025). Each distance definition is reported in Appendix A.5 and normalized for each destination country such that the minimum distance equals zero and the maximum equals one. Table 7 reports the estimates of Eq. (19), while Figure 4 reports, for each distance, the marginal impact of immigration on cultural diversity of the native population for different deciles of $\overline{Dist}_{r,t}$. Overall, we find that the greater the cultural distance, the larger the negative impact of immigration on natives' cultural diversity, as the coefficient of the interaction term is always negative and also statistically significant, specifically for economic, genetic, cultural, and geographic distance.

These results suggest that immigration affects the cultural homogenization of natives primarily through culturally distant immigrants whose attributes are highly visible, making im-

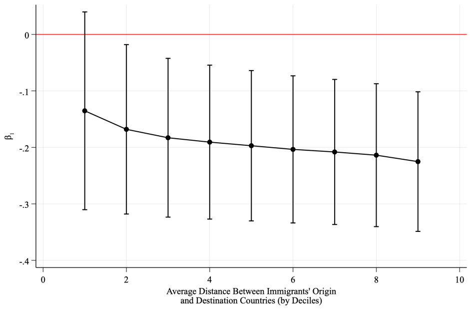
Figure 4: Cultural distance between immigrants and natives
Outcome: CF_B^W - 2SLS Results - Marginal Effects.



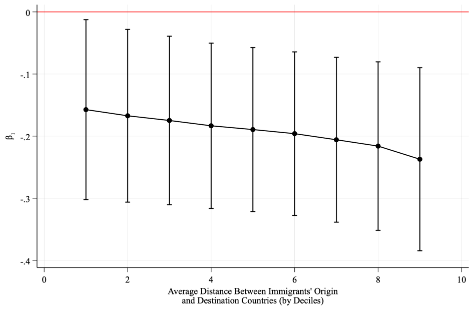
(a) Geodesic



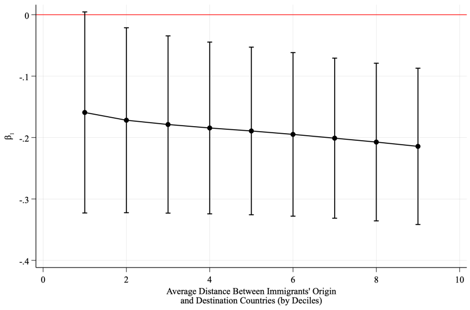
(b) Cultural



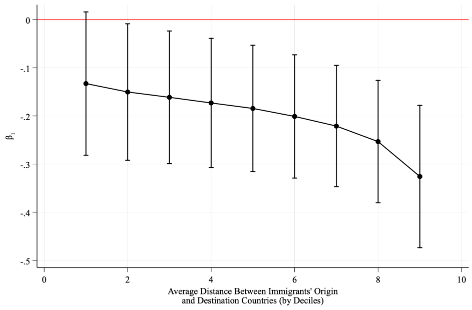
(c) Economic



(d) Religious



(e) Linguistic



(f) Genetic

Notes: These figures plot the marginal impact of immigration on natives' cultural diversity conditional on cultural distance between immigrants and natives. Coefficients are based on the estimated reported in Table 7. The dependent variable is CF_{rt}, the measure of cultural diversity in the region r at time t. All estimates include a vector of control with the log of population density, the log of GDP per capita, the unemployment rate, the share of high-skilled in the total population, and regional and year fixed effects. Standard errors are clustered at the regional level. We include 95% confidence intervals around the estimated coefficients. Source: Authors' elaboration on ESS and EU-LFS data (2004-2018).

migration salient enough to trigger shifts in native identities. Notably, the interaction term is not statistically significant only for linguistic and religious distances, which plausibly capture dimensions less directly linked to an immigrant background and visible from the natives' perspective. Figure 4 shows that across various deciles of language and religious distance, the marginal effect remains constant and barely significant at standard levels. By contrast, cultural and geographic distance may correlate with visible or easily identifiable traits, such as skin color, dress, or other physical features, that are immediately observable and thus more directly linked to the salience of birthplace. To reinforce this interpretation, Figure 4 shows a rise in the magnitude of the estimated marginal effect of the interaction term from the first to the last decile. Overall, these results support the intuition that distances matter, particularly when they are visible.

To confirm the influence of immigrants' cultural distance, we use an alternative strategy by computing a Greenberg index (Greenberg, 1956) applied to immigration (Alesina et al., 2016; Docquier et al., 2020b), which enables us to assign greater weight to immigrants from specific origin countries in our benchmark specification. Specifically, the Greenberg index allows us to weigh our measure of the share of immigrants with the aforementioned proxies of cultural distances between origin and destination countries, such as:

$$m^G_{r(d),t} = \sum_o \frac{k_{r,o,t}}{Pop_{r,2004}} \times D^{\theta}_{r(d),o} \quad (20)$$

where θ is a factor ranging from 1 to infinity. By sequentially increasing θ, we create a set of Greenberg indices that we use in a horse race within our baseline model. Given that our distances are normalized between zero and one, an increase in θ means an overweighting of immigrants from more distant countries compared to relatively closer ones. This alternative strategy has the advantage of maintaining the same first stage as the benchmark specification and avoiding the need to instrument two endogenous variables, as in the previous estimates. Figure E-2 in the Appendix confirms that placing greater weight on culturally distant groups magnifies the convergence of norms among natives, as evidenced by the increasingly negative coefficient as distance increases.

Overall, these findings corroborate our third theoretical hypothesis, indicating that immigration has a stronger impact when immigrants originate from countries with more culturally distant backgrounds.

7 Mechanisms

The empirical results thus far show that immigration increases both birthplace salience and cultural convergence among natives, effects that are amplified when immigrants are culturally distant, have lower education levels, and have recently arrived. This section investigates the mechanisms through which a shift in the salience of birthplace translates into cultural homogenization among the native population. To do so, we shift the focus of the analysis from aggregate measures of diversity to individual-level data on cultural values and test two implications of the model presented in Section 2: first, that immigration is plausibly associated with identity changes toward nativism, and second, that in response to this new identity, natives adjust their cultural values toward those of the a representative native of this expanding native-identified population.

7.1 Immigration, identity changes and nativism

To explore the implications in terms of identity choices of an increase in the salience of birthplace for the natives, who are directly exposed to immigration, we test the hypothesis that immigration is associated with more national pride and/or greater support for nationalist parties. This echoes previous results in the literature on the political economy of immigration, including on electoral outcomes (Moriconi et al., 2022; Alesina and Tabellini, 2024)

To do so, we first rely on two additional questions from the ESS: one from the 2014 wave, which asks respondents whether they feel close to their country of residence ("*How close do you feel to [country]?*"), and another from the 2016 and 2018 waves, which measures the extent of their emotional attachment to their country ("*How emotionally attached are you to [country]?*"). To simplify the interpretation of the coefficients, we standardize both variables to have a mean of 0 and a standard deviation of 1. This analysis mimics Equation (11), using observations at the individual level for the dependent variable, but only relying on cross-sectional variations due to the limited availability of these questions over time. Consequently, caution must be exercised when interpreting these estimates, as they cannot be conclusively regarded as causal. Still, we add to the vector of regional control defined Equation (11) a second vector of individual control with age, age squared, education, gender, children, urbanicity, marital, and employment status, and country fixed effects.

Table 8 presents the 2SLS estimates on the relationship between immigration and national pride among natives. Col. (1) and (2) show that an increase in the share of immigrants is associated with higher nationalist sentiment, as evidenced by natives reporting feeling closer to and more emotionally attached to their country. In terms of magnitude, a one standard deviation increase in the share of immigration (0.10) is associated with a 0.06 standard deviation increase in the likelihood of feeling closer to their country and a 0.10 standard deviation increase in emotional attachment to their country.

In Col. (3) and (4), we test the extent to which this relationship is skill-specific by interacting the share of immigrants with a dummy variable distinguishing tertiary-educated natives from others. As indicated by the negative and statistically significant coefficient for the interaction term, we find that the impact of immigration on nationalism is much stronger for low-skilled natives. Specifically, the effect is insignificant for high-skilled natives in Col. (3), while in Col. (4), it is 30% lower than the effect observed for low-skilled natives. This aligns with the third prediction (H3) of the theoretical framework in Section 2.3, which suggests that a rise in the salience of birthplace should be higher for low-skilled natives, since low-skilled immigrants make up a disproportionately higher share of immigration flows (70% of the overall immigration stock in our sample). This may also reflect the higher likelihood of interaction among low-skilled natives (compared to high-skilled natives) due to their greater proximity to low-skilled immigrants in sectoral employment or residential location (Dustmann et al., 2018).

Finally, to confirm the association between immigration and a rise in nationalism, and given the limited data on national pride, we also take advantage of two further questions that record whether respondents voted in the last election and, if so, for which party. This also allows us to examine whether immigration influences electoral turnout and shifts votes toward nationalist parties.

To quantify nationalism, we extract party positions from political manifestos sourced from the Manifesto Project Database (Klingemann, 2006). The MPD provides quantitative measures

Table 8: Immigration and the Rise of Nationalism
National pride - 2SLS Results

	(1) Feel Close	(2) Emotionally Attached	(3) Feel Close	(4) Emotionally Attached
m_{rt}	0.602*	1.156***	0.925**	1.346***
	(0.337)	(0.424)	(0.401)	(0.412)
High-skilled	-0.020	0.056***	0.069*	0.113***
	(0.025)	(0.014)	(0.042)	(0.027)
m_{rt} × High-Skilled			-0.789**	-0.444**
			(0.311)	(0.202)
Year FE	No	Yes	No	Yes
Observations	30,168	65,686	30,168	65,686
Mean Cultural Index	0.038	0.029	0.038	0.029
Mean Immig. Share	0.112	0.120	0.112	0.120
First-stage	0.600	0.623	0.598	0.623
KP F-Test	634.588	779.726	407.846	419.663
Total effect HS			0.136	0.902
P-value			0.638	0.047

Notes: *** p <0.01, ** p <0.05, * p < 0.1. Standard errors in parentheses are clustered at the regional level. The dependent variable In Col. (1) and (3) is a dummy variable equal to 1 if the respondent feels close to [Country] and 0 otherwise. The dependent variable In Col. (2) and (4) is a scale ranging from 0 to 10, where 0 indicates no emotional attachment to [Country] and 10 represents a firm emotional attachment. The independent variable m_{rt} is the share of foreign-born in the total 2004 population. We control for the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the total population. All estimates include a vector of individual controls with age, age squared, education, gender, children, urbanicity, marital, and employment status. All estimates include country fixed effects.

Sources: Authors' elaboration on ESS and EU-LFS data (2004-2018).

of parties' political stances over 1,093 parties across 715 parliamentary elections, covering all countries and years in our benchmark sample. These measures are the results of a content analysis and a precise counting of the share of quasi-sentences that are associated with a specific political issue.

Following Moriconi et al. (2022), we construct a nationalism index through a Principal Component Analysis (PCA), which reflects party positive stances on the national way of life and negative stances towards the European Union and its expansion. Using this measure, we construct a region-election-year panel by imputing individual votes to the corresponding year's election and measuring the stock of immigrants and other regional controls for that year. Thus, this additional analysis uses variations over time between elections within a region.

Table 9 provides two main evidence in line with the predictions of our theoretical framework and the results of the literature. First, by estimating the effect on the likelihood to vote, Col. (1) to (3) show that immigration has consequences on electoral outcomes (Alesina and Tabellini, 2024). Col. (1) shows that individuals are more likely to vote in regions highly exposed to immigration, and this effect is mainly driven by low-skilled voters (Col. 2). In addition, Col. (4) to (6) confirm the direction of such skill-specific shift in voting preferences, which aligns both with our theoretical model and the literature (Edo et al., 2019; Moriconi

Table 9: Immigration and the Rise of Nationalism
Voting Outcomes - 2SLS Results

	(1) Voted All	(2) Voted Low-skilled	(3) Voted High-skilled	(4) Nationalism All	(5) Nationalism Low-skilled	(6) Nationalism High-skilled
m_{rt}	0.756**	1.221***	0.141	0.163	2.525*	-2.205
	(0.295)	(0.423)	(0.344)	(0.885)	(1.328)	(1.413)
Observations	238,481	176,567	61,914	145,334	100,031	45,303
Mean Dep. Var.	0.736	0.693	0.858	1.000	1.000	1.000
Mean Immig. Share	0.097	0.093	0.109	0.102	0.098	0.112
First-stage	0.962	0.960	0.939	0.973	0.982	0.940
KP F-Test	85.876	71.832	110.392	90.105	76.857	103.766

Notes: *** p <0.01, ** p <0.05, * p < 0.1. Standard errors in parentheses are clustered at the regional level. The dependent variable In Col. (1) to (3) is a dummy variable equal to 1 if the respondent voted in the last elections and 0 otherwise. The dependent variable in Col. (4) to (6) is the synthetic measure of nationalism (decreasing with the party's support for the European Union and increasing with emphasis on patriotism and pride in citizenship). The independent variable m_{rt} is the share of foreign-born in the total 2004 population. We control for the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the total population. All estimates include regional and year fixed effects. All estimates include a vector of individual controls with age, age squared, gender, children, urbanicity, marital, and employment status.

Sources: Authors' elaboration on ESS and EU-LFS data (2004-2018).

et al., 2022). Following a rise in immigration, low-skilled natives, who are more exposed to immigrants with similar socioeconomic backgrounds, become more likely than highly educated natives to support nationalist parties.

Overall, these results show that immigration is associated with stronger nationalism and native pride, particularly among low-skilled natives. providing suggestive evidence of identity change within this group. As described in the model in Section 2, this identity change is the main force driving adjustments in cultural attributes and the incentives to converge toward the representative individual of the native-identified population. Although we cannot directly test for identity changes, this provides suggestive evidence that immigration may lead more natives to identify with this particular dimension of their identity.

7.2 To Whom and Toward Which Cultural Values Are Natives Converging?

This section investigates the reference groups and values toward which natives are converging, and those from which they are diverging.

To answer this question, we compute for each native i in region r the average standardized cultural Euclidean distance with respect to the cultural norms of a given reference group J within region r.[36] This measure serves two purposes. First, it directly relates to the perceived distance within social identity groups (Shayo, 2009), as previously described in Eq. (1), assuming that each trait holds equal attention weight. Second, it allows us to directly test whether natives are converging toward or diverging from a reference group J when exposed to

[36]We first normalize each cultural trait h between 0 (minimum value) and 1 (maximum value) to enable comparison across traits.

immigration. We compute the average standardized Euclidean distance as follows:

$$d_{iJr} = \left[\frac{1}{H} \sum_{h=1}^{H} \frac{(q_{ir}^h - q_{Jr}^h)^2}{\sigma_{Jr}^h} \right]^{0.5} \tag{21}$$

with q_{Jr}^h the average cultural norms of the reference group J within region r and σ_{Jr}^h its standard deviation. We compute d_{iJr} for several reference groups and use the resulting distance measures as the dependent variable. Since the regressions in this setting rely on individual level observations, we include in the main analysis the same vector of individual socioeconomic controls as described in Section 7.1. Additionally, we include the same set of region and year fixed effects as Eq. (11), therefore exploiting within-region variations.

The results are reported in Table 10. Panel A reports the results using natives as the reference group, whereas Panel B uses immigrants. Then, each column reports the results for a specific native or immigrant reference group defined by specific socioeconomic characteristics such as education, gender, place of residence, and religiosity. Reference groups are first defined using observations from the baseline year 2004 only, to avoid capturing subsequent changes in cultural values induced by immigration.

Several new findings emerge from these estimates. First, Panel A shows that immigration in a European region is associated with a reduction in cultural distance between a randomly selected native and the average native in the same region, consistent with the results from the previous sections. However, this process of cultural convergence is much stronger toward high-skilled natives (col. 2) than toward the other reference groups tested, particularly low-skilled natives. This suggests that college-educated natives may serve as new cultural reference points, or role-model, for natives exposed to immigration. Similarly, Panel B shows that immigration induces an increased cultural distance between the average native and immigrants with specific characteristics, mainly low-skilled, male, and urban residents. These results, suggest therefore that rising immigration not only fosters cultural convergence among natives but also generates cultural divergence from immigrants.

Table F-1 in the Appendix shows that these results are robust to excluding the 2004 wave from the sample, which could otherwise drive the results since the reference group is directly included in the analysis, or to computing the reference group over the entire period of analysis rather than using only 2004. In addition, Table F-2 in the Appendix shows that immigrant inflows are more likely to increase cultural divergence between natives and immigrants from specific origin groups, particularly those from African countries.

Overall, these results indicate that immigration fosters cultural convergence toward high-skilled natives while increasing cultural distance with low-skilled immigrants. These findings suggest that the education gap between immigrants and natives plays a central role in shaping the observed dynamics of convergence and divergence.

A legitimate question is what cultural stances characterize college-educated natives and low-skilled immigrants, and toward which values natives are converging or diverging. Figure 5 reports the average cultural norms of natives and immigrants by education level, as broadly presented in Appendix B.2. We first recode all cultural trait variables such that higher values always correspond to more liberal views. For instance, higher values on religiosity indicate less attachment to religion, higher values or sexual morality more tolerance towards homosexuality and nontraditional sexual behaviors, or higher level of trust towards institutions. We then

Table 10: Cultural distance between natives and a given reference group

	Panel A: Reference Group = Average Native (in 2004) Sample: All Years								
	(1) Overall	(2) High Skill	(3) Low Skill	(4) Female	(5) Male	(6) Urban	(7) Rural	(8) Religious	(9) Not religious
m_{rt}	–0.114**	–0.197***	–0.091*	–0.115*	–0.125**	–0.093*	–0.099	–0.085**	–0.051
	(0.057)	(0.064)	(0.053)	(0.061)	(0.057)	(0.053)	(0.071)	(0.042)	(0.046)
Observations	257,166	257,166	257,166	257,166	257,166	257,166	257,166	257,166	257,166
Regions	175	175	175	175	175	175	175	175	175
KP F-Test	26.711	26.711	26.711	26.711	26.711	26.711	26.711	26.711	26.711
	Panel B: Reference Group = Average Migrant (in 2004) Sample: All Years								
	(1) Overall	(2) High Skill	(3) Low Skill	(4) Female	(5) Male	(6) Urban	(7) Rural	(8) Religious	(9) Not religious
m_{rt}	0.043	–0.012	0.092***	0.055	0.085**	0.079**	0.042	0.069*	0.062*
	(0.034)	(0.047)	(0.035)	(0.043)	(0.036)	(0.037)	(0.039)	(0.037)	(0.032)
Observations	257,166	257,166	257,166	257,166	257,166	257,166	257,166	257,166	257,166
Regions	175	175	175	175	175	175	175	175	175
KP F-Test	26.711	26.711	26.711	26.711	26.711	26.711	26.711	26.711	26.711

Notes: *** p <0.01, ** p <0.05, * p < 0.1. The dependent variable is the cultural distance between each native and a reference group of natives (Panel A) or immigrants (Panel B) computed in 2004, with certain characteristics. The independent variable is the share of foreign-born in the total 2004 population. All estimates include a vector of regional controls, including the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the total population. All estimates include a vector of individual control with age, age squared, education, gender, children, urbanicity, marital, and employment status. All estimates include regional and year fixed effects. Estimates are weighted using individual weights. Standard errors in parentheses are clustered at the regional level.

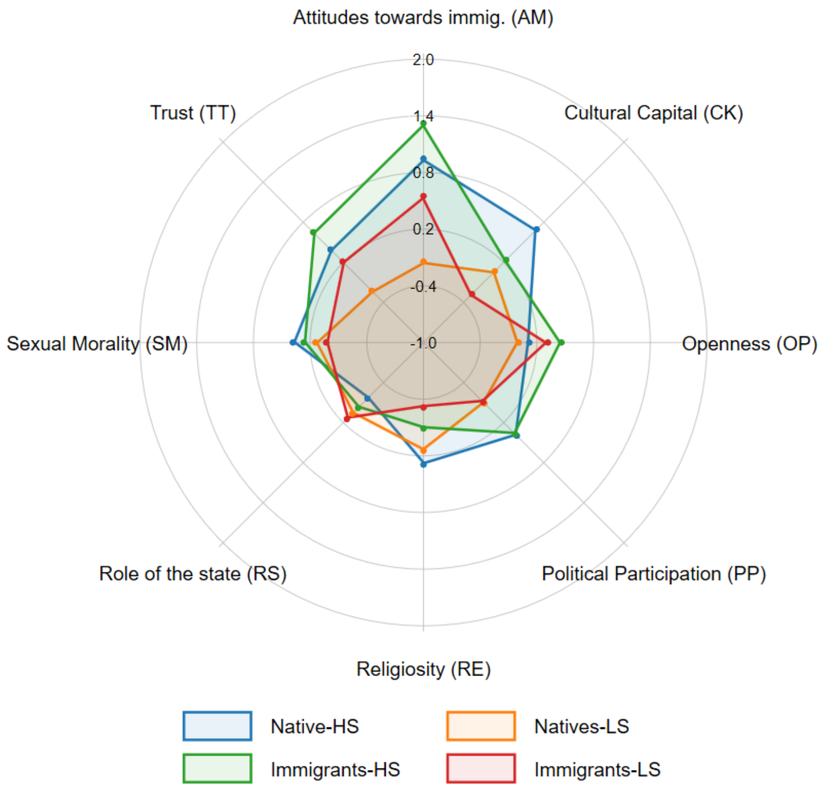
Sources: Authors' elaboration on ESS and EU-LFS data (2004-2018).

extract the first component from a principal component analysis conducted on each broad set of cultural traits an we plot itts average on Figure 5.

Focusing on the reference groups of interest, highly educated natives and low-skilled immigrants, Figure 5 provides a clear depiction of their average stances and highlights the predominance of education over birthplace. Highly educated natives, on average, hold more liberal values (e.g., are less religious, more accepting of nontraditional sexual behaviors, and more trusting of institutions), whereas low-skilled immigrants tend to hold more conservative views on these dimensions for instance.

In response to immigration, natives more exposed to immigration therefore adopts and converges toward the more liberal cultural orientations of the highs-killed natives. These results echoes those of Fouka and Tabellini (2022), who shows that inflows of Mexicans in the United States also contributed to a shift of white Americans towards more liberal values. This set of evidence therefore suggest that the cultural implications of immigration are fundamentally shaped by educational stratification, with natives gravitating toward the liberal values of their highly educated peers while distancing themselves from the conservative orientations prevalent among low-skilled immigrants.

Figure 5: Natives and Immigrants Cultural Stances by Education



Notes: This graph displays the average value of the first principal component score of each cultural block, disaggregated by education. The principal component scores are obtained from a principal component analysis (PCA) of all cultural traits within each block. All traits were recoded so that lower values consistently represent more conservative attitudes and higher values more liberal ones. Averages are computed using individual weights.

Source: Authors' elaboration on ESS data (2004-2018).

8 Conclusions

This paper investigates the impact of immigration on the cultural diversity of host populations in European regions. Building on the theoretical framework of endogenous social identity introduced by Shayo (2009), we provide empirical evidence that immigration, particularly low-skilled and from culturally distant countries, affects the distribution of cultural norms and values in host societies. To do so, we combine regional data from the European Social Survey, which tracks the evolution of cultural diversity across multiple dimensions, with immigration data from the European Labor Force Survey between 2004 and 2018.

Our findings show that immigration challenges the social identities and values of the host population. An increase in the share of immigrants enhances birthplace-driven identity as a relevant predictor of cultural preferences, fostering a rising national sentiment among natives, who feel more attached to the nation. Simultaneously, natives increasingly align their norms and values with those of the broader native-identified population, reinforcing a process of cultural convergence along national lines, reducing cultural diversity at the regional level.

Our results highlight that social identities and natives' responses to immigration are key to understanding the cultural dynamics of societies facing rising immigration flows (Bazzi and Fiszbein, 2025; Fernández, 2025; Fouka and Tabellini, 2025). This paper shows that immigration reshapes social identities beyond economic or labor market channels, emphasizing the cultural mechanisms driving social change. It also underscores birthplace as a crucial cleavage in explaining cultural divides in modern European societies. Finally, our findings contribute to debates on nationalism and identity-driven preferences, suggesting that immigration-induced identity realignment may have lasting effects on cultural diversity among natives.

Future research should dig deeper in the cultural traits' specific direction in which natives shift due to immigration, a central question beyond the scope of this paper. Indeed, each cultural trait may be affected differently by immigration, depending on the initial distribution of cultural norms in the host population and the preexisting partition of the population across salient identities (Desmet and Wacziarg, 2021).

These results call for a broader perspective on immigration, moving beyond an exclusive focus on economic costs and benefits, and also highlighting natives' response to immigration as driving factors of societal changes. Therefore, policymakers should consider the role of immigration in shaping social identity dynamics, specifically among natives, which are not necessarily static, and its implications for social cohesion and electoral outcomes in increasingly diverse societies.

References

Abadie, A., Athey, S., Imbens, G. W., and Wooldridge, J. M. (2023). When should you adjust standard errors for clustering? *The Quarterly Journal of Economics*, 138(1):1–35.

Abramitzky, R. and Boustan, L. (2022). *Streets of gold: America's untold story of immigrant success*. PublicAffairs.

Abramitzky, R., Boustan, L., and Eriksson, K. (2020). Do Immigrants Assimilate More Slowly Today Than in the Past? *American Economic Review: Insights*, 2(1):125–41.

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

Abrams, D. and Hogg, M. A. (2006). *Social identifications: A social psychology of intergroup relations and group processes*. Routledge.

Adão, R., Kolesár, M., and Morales, E. (2019). Shift-Share Designs: Theory and Inference. *The Quarterly Journal of Economics*, 134(4):1949–2010.

Akerlof, G. A. and Kranton, R. E. (2000). Economics and identity. *The Quarterly Journal of Economics*, 115(3):715–753.

Alba, R. D. and Nee, V. (2009). *Remaking the American mainstream: Assimilation and contemporary immigration*. Harvard University Press.

Alesina, A., Harnoss, J., and Rapoport, H. (2016). Birthplace diversity and economic prosperity. *Journal of Economic Growth*, 21(2):101–138.

Alesina, A., Miano, A., and Stantcheva, S. (2022). Immigration and Redistribution. *The Review of Economic Studies*, 90(1):1–39.

Alesina, A., Tabellini, G., and Trebbi, F. (2017). Is Europe an Optimal Political Area? *Brookings Papers on Economic Activity*, pages 169–213.

Alesina, A. and Tabellini, M. (2024). The political effects of immigration: Culture or economics? *Journal of Economic Literature*, 62(1):5–46.

Algan, Y., Bisin, A., Manning, A., and Verdier, T. (2012). *Cultural integration of immigrants in Europe*. Oxford University Press.

Algan, Y., Guriev, S., Papaioannou, E., and Passari, E. (2017). The European trust crisis and the rise of populism. *Brookings Papers on Economic Activity*, 2017(2):309–400.

Ausserladscheider, V. (2019). Beyond economic insecurity and cultural backlash: Economic nationalism and the rise of the far right. *Sociology Compass*, 13(4):e12670.

Autor, D. H. and Duggan, M. G. (2003). The rise in the disability rolls and the decline in unemployment. *The Quarterly Journal of Economics*, 118(1):157–206.

Bazzi, S., Ferrara, A., Fiszbein, M., Pearson, T. P., and Testa, P. A. (2023). The confederate diaspora. *NBER Working Paper*, (No. 31331).

Bazzi, S. and Fiszbein, M. (2025). When do migrants shape culture? *NBER Working Paper*, (No. 34001).

Bertrand, M. and Kamenica, E. (2023). Coming Apart? Cultural Distances in the United States over Time. *American Economic Journal: Applied Economics*.

Bhatiya, A. (2024). Do enfranchised immigrants affect politicians' behavior? *The Center for Growth and Opportunity*.

Biavaschi, C., Giulietti, C., and Siddique, Z. (2017). The economic payoff of name Americanization. *Journal of Labor Economics*, 35(4):1089–1116.

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

Bonomi, G., Gennaioli, N., and Tabellini, G. (2021). Identity, beliefs, and political conflict. *The Quarterly Journal of Economics*, 136(4):2371–2411.

Bordalo, P., Coffman, K., Gennaioli, N., and Shleifer, A. (2016). Stereotypes. *The Quarterly Journal of Economics*, 131(4):1753–1794.

Borjas, G. J. (1985). Assimilation, changes in cohort quality, and the earnings of immigrants. *Journal of Labor Economics*, 3(4):463–489.

Borjas, G. J. (2003). The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *The Quarterly Journal of Economics*, 118(4):1335–1374.

Borusyak, K., Hull, P., and Jaravel, X. (2022). Quasi-experimental shift-share research designs. *The Review of Economic Studies*, 89(1):181–213.

Card, D. (1990). The impact of the Mariel boatlift on the Miami labor market. *ILR Review*, 43(2):245–257.

Card, D. (2001). Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration. *Journal of Labor Economics*, 19(1):22–64.

Card, D., Chang, S., Becker, C., Mendelsohn, J., Voigt, R., Boustan, L., Abramitzky, R., and Jurafsky, D. (2022). Computational analysis of 140 years of us political speeches reveals more positive but increasingly polarized framing of immigration. *Proceedings of the National Academy of Sciences*, 119(31):e2120510119.

Collier, P. (2013). *Exodus: Immigration and multiculturalism in the 21st century*. Penguin UK.

Conte, M., Cotterlaz, P., Mayer, T., et al. (2022). The CEPII gravity database.

Davies, S., Pettersson, T., and Öberg, M. (2022). Organized violence 1989–2021 and drone warfare. *Journal of Peace Research*, 59(4):593–610.

Derenoncourt, E. (2022). Can you move to opportunity? Evidence from the Great Migration. *American Economic Review*, 112(2):369–408.

Desmet, K., Ortuño-Ortín, I., and Wacziarg, R. (2017). Culture, Ethnicity, and Diversity. *American Economic Review*, 107(9):2479–2513.

Desmet, K. and Wacziarg, R. (2021). The cultural divide. *The Economic Journal*, 131(637):2058–2088.

Docquier, F., Tansel, A., and Turati, R. (2020a). Do emigrants self-select along cultural traits? Evidence from the MENA countries. *International Migration Review*, 54(2):388–422.

Docquier, F., Turati, R., Valette, J., and Vasilakis, C. (2020b). Birthplace diversity and economic growth: evidence from the US states in the Post-World War II period. *Journal of Economic Geography*, 20(2):321–354.

Dorn, D. and Zweimüller, J. (2021). Migration and labor market integration in Europe. *Journal of Economic Perspectives*, 35(2):49–76.

Dustmann, C. and Preston, I. P. (2007). Racial and economic factors in attitudes to immigration. *The BE Journal of Economic Analysis & Policy*, 7(1):1–41.

Dustmann, C., Vasiljeva, K., and Piil Damm, A. (2018). Refugee Migration and Electoral Outcomes. *The Review of Economic Studies*, 86(5):2035–2091.

Edo, A. (2019). The impact of immigration on the labor market. *Journal of Economic Surveys*, 33(3):922–948.

Edo, A., Giesing, Y., Öztunc, J., and Poutvaara, P. (2019). Immigration and electoral support for the far-left and the far-right. *European Economic Review*, 115:99 – 143.

Edo, A. and Özgüzel, C. (2023). The impact of immigration on the employment dynamics of European regions. *Labour Economics*, 85:102433.

EMDAT (2022). The Emergency Events Database. Technical report, Université catholique de Louvain (UCL), Centre for Research on the Epidemiology of Disasters (CRED).

Fernández, R. (2025). Understanding cultural change. *NBER Working Paper*, (No. 34077).

Feyrer, J. (2019). Trade and income—exploiting time series in geography. *American Economic Journal: Applied Economics*, 11(4):1–35.

Fisman, R. and Miguel, E. (2007). Corruption, norms, and legal enforcement: Evidence from diplomatic parking tickets. *Journal of Political Economy*, 115(6):1020–1048.

Foster, C. and Frieden, J. (2017). Crisis of trust: Socio-economic determinants of Europeans' confidence in government. *European Union Politics*, 18(4):511–535.

Fouka, V., Mazumder, S., and Tabellini, M. (2022). From Immigrants to Americans: Race and assimilation during the Great Migration. *The Review of Economic Studies*, 89(2):811–842.

Fouka, V. and Tabellini, M. (2022). Changing in-group boundaries: The effect of immigration on race relations in the United States. *American Political Science Review*, 116(3):968–984.

Fouka, V. and Tabellini, M. (2025). Migration, Identity and Political Economy in a Globalizing World. In *Handbook of Political Economy. Edited by James Robinson, Daron Acemoglu (Forthcoming)*. Elsevier.

Funke, M., Schularick, M., and Trebesch, C. (2016). Going to extremes: Politics after financial crises, 1870–2014. *European Economic Review*, 88:227–260.

Galli, F. and Russo, G. (2019). Immigration restrictions and second-generation cultural assimilation: Theory and quasi-experimental evidence. *Journal of Population Economics*, 32:23–51.

Gennaioli, N. and Tabellini, G. (2023). Identity Politics. *CEPR Discussion Paper*, (DP18055):1–86.

Gethin, A., Martínez-Toledano, C., and Piketty, T. (2022). Brahmin left versus merchant right: Changing political cleavages in 21 western democracies, 1948–2020. *The Quarterly Journal of Economics*, 137(1):1–48.

Giuliano, P. and Tabellini, M. (2021). The Seeds of Ideology: Historical Immigration and Political Preferences in the United States. *Harvard Business School BGIE Unit Working Paper*, (20-118).

Gleditsch, N. P., Wallensteen, P., Eriksson, M., Sollenberg, M., and Strand, H. (2002). Armed conflict 1946-2001: A new dataset. *Journal of Peace Research*, 39(5):615–637.

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8):2586–2624.

Gonnot, J. and lo Polito, F. (2021). The Political Assimilation of Immigrants: Migrant-to-Native Differences in Western Europe. *Robert Schuman Centre for Advanced Studies Migration Policy Centre No. 2021-87.*

Gonnot, J. and lo Polito, F. (2023). Cultural Transmission and Political Attitudes: Explaining Differences between Natives and Immigrants in Western Europe. Technical report, CEPII, Documents de Travail.

Greenberg, J. H. (1956). The measurement of linguistic diversity. *Language*, 32(1):109–115.

Grossman, G. M. and Helpman, E. (2021). Identity politics and trade policy. *The Review of Economic Studies*, 88(3):1101–1126.

Guiso, L., Herrera, H., Morelli, M., Sonno, T., et al. (2020). Economic insecurity and the demand of populism in Europe. *Einaudi Institute for Economics and Finance*, 3.

Guriev, S. and Papaioannou, E. (2022). The political economy of populism. *Journal of Economic Literature*, 60(3):753–832.

Hall, M. and Tideman, N. (1967). Measures of concentration. *Journal of the American Statistical association*, 62(317):162–168.

Ivanov, D. (2023). Economic insecurity, Institutional trust and populist voting across Europe. *Comparative Economic Studies*, pages 1–22.

Jaeger, D. A., Ruist, J., and Stuhler, J. (2018). Shift-share instruments and the impact of immigration. *NBER Working Paper*, (No. 24285).

Jaschke, P., Sardoschau, S., and Tabellini, M. (2022). Scared Straight? Threat and Assimilation of Refugees in Germany. *NBER Working Paper*, (No. 30381).

Keita, S., Renault, T., and Valette, J. (2023). The Usual Suspects: Offender Origin, Media Reporting and Natives' Attitudes Towards Immigration. *The Economic Journal*, 134(657):322–362.

Klingemann, H.-D. (2006). *Mapping policy preferences II: estimates for parties, electors, and governments in Eastern Europe, European Union, and OECD 1990-2003*, volume 2. Oxford University Press, USA.

Knudsen, A. S. B. (2022). Those who stayed: Selection and cultural change in the age of mass migration.

Kroeber, A. L. and Kluckhohn, C. (1952). Culture: A critical review of concepts and definitions. *Papers. Peabody Museum of Archaeology & Ethnology, Harvard University.*

Miho, A., Jarotschkin, A., and Zhuravskaya, E. (2024). Diffusion of gender norms: Evidence from Stalin's ethnic deportations. *Journal of the European Economic Association*, 22(2):475–527.

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

Moriconi, S., Peri, G., and Turati, R. (2022). Skill of the immigrants and vote of the natives: Immigration and nationalism in European elections 2007–2016. *European Economic Review*, 141:103986.

Moriconi, S., Peri, G., and Turati, R. (2025). Analyzing political preferences of second-generation immigrants across the rural–urban divide. *Journal of Urban Economics*, 146:103740.

Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General*, 115(1):39.

Obradovich, N., Özak, Ö., Martín, I., Ortuño-Ortín, I., Awad, E., Cebrián, M., Cuevas, R., Desmet, K., Rahwan, I., and Cuevas, Á. (2022). Expanding the measurement of culture with a sample of two billion humans. *Journal of the Royal Society Interface*, 19(190):20220085.

Orefice, G. and Peri, G. (2024). Immigration and worker-firm matching. *The Review of Economics and Statistics*, pages 1–45.

Orefice, G., Rapoport, H., and Santoni, G. (2025). How do immigrants promote exports? *Journal of Development Economics*, 174:103443.

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

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

Pellegrino, B., Spolaore, E., and Wacziarg, R. (2025). Barriers to global capital allocation. *The Quarterly Journal of Economics*, page qjaf031.

Rapoport, H., Sardoschau, S., and Silve, A. (2021). Migration and cultural change. *IZA Working Paper*, 14772:1–64.

Schneider-Strawczynski, S. and Valette, J. (2025). Media coverage of immigration and the polarization of attitudes. *American Economic Journal: Applied Economics*, 17(1):337–368.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423.

Shayo, M. (2009). A model of social identity with an application to political economy: Nation, class, and redistribution. *American Political Science Review*, 103(2):147–174.

Shayo, M. (2020). Social identity and economic policy. *Annual Review of Economics*, 12(1):355–389.

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

Silva, J. S. and Tenreyro, S. (2010). On the existence of the maximum likelihood estimates in Poisson regression. *Economics Letters*, 107(2):310–312.

Spolaore, E. and Wacziarg, R. (2016). Ancestry, language and culture. In *The Palgrave Handbook of Economics and Language*, pages 174–211. Springer.

Steinmayr, A. (2021). Contact versus Exposure: Refugee Presence and Voting for the Far Right. *The Review of Economics and Statistics*, 103(2):310–327.

Stewart, A. J., McCarty, N., and Bryson, J. J. (2020). Polarization under rising inequality and economic decline. *Science Advances*, 6(50):eabd4201.

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

Tajfel, H. and Turner, J. (1986). The social identity theory of intergroup behavior (vol. 2). *Chicago: Nelson-Hall*.

Tormos, R. (2019). Measuring personal economic hardship and its impact on political trust during the great recession. *Social Indicators Research*, 144:1209–1232.

UN (2020). Population Division, International Migrant Stock 2020. Technical Report POP/DB/MIG/Stock/Rev.2020, United Nations Department of Economic and Social Affairs.

Wroe, A. (2016). Economic insecurity and political trust in the United States. *American Politics Research*, 44(1):131–163.

Immigration, Identity Choices, and Cultural Diversity

Yasmine Elkhateeb, Riccardo Turati, and Jérôme Valette

Appendix – For Online Publication

A Data construction

A.1 Regional Harmonization

Our regional-level analysis hinges on the integration of data sourced from both the European Social Survey (ESS) and the European Labor Force Survey (EU-LFS). We made a series of methodological decisions to ensure the full comparability of the regions between surveys and across time. These choices were mainly prompted by the relatively limited number of observations associated with particular regions, as well as the distinct manners in which regional entities are defined across the various datasets.

Austria - EU-LFS provides information only at NUTS 1 level, hence we aggregate the observations available in ESS to match the same NUTS 1 administrative units.

Finland - The NUTS 2 Åland region (FI20) appears in the ESS data only on four waves, given the small size of the region.

France - We exclude from our analysis the *territoire d'outre-mer*. Moreover, ESS does not provide enough observations to have a representative sample of the region FR83 (Corsica).

Germany - EU-LFS provides information only at NUTS 1 level, hence we aggregate the observations available in ESS to match the same NUTS 1 administrative units.

Ireland - We follow EU-LFS NUTS 2 classification, which splits Ireland into two regions: the Border, Midland and Western (IE01) and the Southern and Eastern (IE02).

Italy - We merge together the observations belonging to the region of Trento (ITH1) and Sud-Tirol (ITH2). These two areas are part of the same region, named Trentino Alto-Adige, which appears in our dataset only in four waves, compared to the rest of the Italian regions, in which we have over five different waves. Moreover, we merge the region Molise (ITF2) with Abruzzo (ITF1) and the region Valle D'Aosta (ITC2) with Pidemont (ITC1), given the small number of observations associated to these regions ITF2 and ITC2, characterized by a reduced population.

Spain - We merge in one unique region the information associated with the two autonomous cities Ceuta (ES63) and Melilla (ES64), which appear in only eight waves of the ESS, compared to the rest of the regions that are defined over the whole ESS dataset. Moreover, information on La Rioja (ES23) are available only from 2004, hence we merge the few observations associated with this region with the ones from the Aragon region (ES24).

United Kingdom - EU-LFS provides information only at NUTS 1 level, hence we aggregate the observations available in ESS to match the same NUTS 1 administrative units.

Table A-1: List or regions and countries.

Region	Country	Nb. waves	Region	Country	Nb. waves	Region	Country	Nb. waves
AT1	Austria	7	FR82	France	8	PL61	Poland	8
AT2	Austria	7	DE1	Germany	8	PL62	Poland	8
AT3	Austria	7	DE2	Germany	8	PL63	Poland	8
BE10	Belgium	8	DE3	Germany	8	PT11	Portugal	8
BE21	Belgium	8	DE4	Germany	8	PT15	Portugal	8
BE22	Belgium	8	DE5	Germany	8	PT16	Portugal	8
BE23	Belgium	8	DE6	Germany	8	PT17	Portugal	8
BE24	Belgium	8	DE7	Germany	8	PT18	Portugal	8
BE25	Belgium	8	DE8	Germany	8	SK01	Slovak Republic	6
BE31	Belgium	8	DE9	Germany	8	SK02	Slovak Republic	6
BE32	Belgium	8	DEA	Germany	8	SK03	Slovak Republic	6
BE33	Belgium	8	DEB	Germany	8	SK04	Slovak Republic	6
BE34	Belgium	8	DEC	Germany	8	SI03	Slovenia	8
BE35	Belgium	8	DED	Germany	8	SI04	Slovenia	8
BG31	Bulgaria	5	DEE	Germany	8	ES11	Spain	8
BG32	Bulgaria	5	DEF	Germany	8	ES12	Spain	8
BG33	Bulgaria	5	DEG	Germany	8	ES13	Spain	8
BG34	Bulgaria	5	HU10	Hungary	8	ES21	Spain	8
BG41	Bulgaria	5	HU21	Hungary	8	ES22	Spain	8
BG42	Bulgaria	5	HU22	Hungary	8	ES24	Spain	8
CY00	Cyprus	5	HU23	Hungary	8	ES30	Spain	8
CZ01	Czech Republic	7	HU31	Hungary	8	ES41	Spain	8
CZ02	Czech Republic	7	HU32	Hungary	8	ES42	Spain	8
CZ03	Czech Republic	7	HU33	Hungary	8	ES43	Spain	8
CZ04	Czech Republic	7	IE01	Ireland	8	ES51	Spain	8
CZ05	Czech Republic	7	IE02	Ireland	8	ES52	Spain	8
CZ06	Czech Republic	7	ITC1	Italy	3	ES53	Spain	8
CZ07	Czech Republic	7	ITC3	Italy	3	ES61	Spain	8
CZ08	Czech Republic	7	ITC4	Italy	3	ES62	Spain	8
DK01	Denmark	5	ITF1	Italy	3	ES70	Spain	8
DK02	Denmark	5	ITF3	Italy	3	SE11	Sweden	8
DK03	Denmark	5	ITF4	Italy	3	SE12	Sweden	8
DK04	Denmark	5	ITF5	Italy	3	SE21	Sweden	8
DK05	Denmark	5	ITF6	Italy	3	SE22	Sweden	8
EE00	Estonia	8	ITG1	Italy	3	SE23	Sweden	8
FI18	Finland	8	ITG2	Italy	3	SE31	Sweden	8
FI19	Finland	8	ITH2	Italy	3	SE32	Sweden	8
FI1D	Finland	8	ITH3	Italy	3	SE33	Sweden	8
FI20	Finland	4	ITH4	Italy	3	CH01	Switzerland	6
FR10	France	8	ITH5	Italy	3	CH02	Switzerland	6
FR21	France	8	ITI1	Italy	3	CH03	Switzerland	6
FR22	France	8	ITI2	Italy	3	CH04	Switzerland	6
FR23	France	8	ITI3	Italy	3	CH05	Switzerland	6
FR24	France	8	ITI4	Italy	3	CH06	Switzerland	6
FR25	France	8	LT00	Lithuania	6	CH07	Switzerland	6
FR26	France	8	NL00	Netherlands	8	UKC	United Kingdom	8
FR30	France	8	PL11	Poland	8	UKD	United Kingdom	8
FR41	France	8	PL12	Poland	8	UKE	United Kingdom	8
FR42	France	8	PL21	Poland	8	UKF	United Kingdom	8
FR43	France	8	PL22	Poland	8	UKG	United Kingdom	8
FR51	France	8	PL31	Poland	8	UKH	United Kingdom	8
FR52	France	8	PL32	Poland	8	UKI	United Kingdom	8
FR53	France	8	PL33	Poland	8	UKJ	United Kingdom	8
FR61	France	8	PL34	Poland	8	UKK	United Kingdom	8
FR62	France	8	PL41	Poland	8	UKL	United Kingdom	8
FR63	France	8	PL42	Poland	8	UKM	United Kingdom	8
FR71	France	8	PL43	Poland	8	UKN	United Kingdom	8
FR72	France	8	PL51	Poland	8			
FR81	France	8	PL52	Poland	8			

Source: Authors' elaboration on ESS and EU-LFS data (2004-2018).

A.2 Cultural questions - Definitions and Proxies

Table A-2: Variables and Descriptions.

Questions	Scale
RE1 - Do you belong to a religious group?	0-1
RE2 - How religious are you?	0-10
RE3 - How often do you attend religious services?	0-6
RE4 - How often do you pray?	0-6
SM1 - Gays and lesbians free to live the life they wish	0-4
SM2 - Important to seek fun and things that give pleasure	0-5
SM3 - Important to follow tradition and customs	0-5
RS1 - Government should reduce income differences	0-4
RS2 - Self-positioning left-right scale	0-10
RS3 - Government should be strong and should ensure safety	0-5
CK1 - Most people can be trusted	0-10
CK2 - Most people try to be fair	0-10
CK3 - Most of the time, people try to be helpful	0-10
CK4 - Important to make own decision and be free	0-5
CK5 - Important to be successful and people recognize you	0-5
CK6 - Important to do what is told and follow the rules	0-5
CK7 - Important to help people and care for others	0-5
PP1 - How interested would you say you are in politics?	0-3
PP2 - Did you vote in the last national election?	0-1
PP3 - Is there a particular political party you feel closer to than all the other parties?	0-1
TT1 - Do you trust the United Nations?	0-10
TT2 - Do you trust the European Parliament?	0-10
TT3 - Do you trust politicians?	0-10
TT4 - Do you trust the police?	0-10
TT5 - Do you trust the legal system?	0-10
TT6 - Do you trust the country's parliament?	0-10
AM1 - Do you think [country] should allow people of the same race or ethnic group as most [country] people to come and live here?	0-3
AM2 - How about people of a different race or ethnic group from most [country] people?	0-3
AM3 - How about people from the poorer countries outside Europe?	0-3
AM4 - Would you say it is generally bad or good for [country]'s economy that people come to live here from other countries?	0-10
AM5 - [country]'s cultural life is generally undermined or enriched by people coming to live here from other countries?	0-10
AM6 - Is [country] made a worse or a better place to live by people coming to live here from other countries?	0-10
OP1 - Important to think new ideas and being creative	0-5
OP2 - Important to be rich, have money and expensive things	0-5
OP3 - Important that people are treated equally and have equal opportunities	0-5
OP4 - Important to show abilities and be admired	0-5
OP5 - Important to live in secure and safe surroundings	0-5
OP6 - Important to try new and different things in life	0-5
OP7 - Important to understand different people	0-5
OP8 - Important to be humble, modest and not draw attention	0-5
OP9 - Important to have a good time	0-5
OP10 - Important to seek adventures and have an exciting life	0-5
OP11 - Important to behave properly	0-5
OP12 - Important to get respect from others	0-5
OP13 - Important to be loyal to friends and devote to people close	0-5
OP14 - Important to care for nature and environment	0-5

Note: All the questions selected from ESS are available in all the ESS waves.
Source: Authors' elaboration on ESS data (2004-2018).

A.3 Summary Statistics

Table A-3: Summary statistics

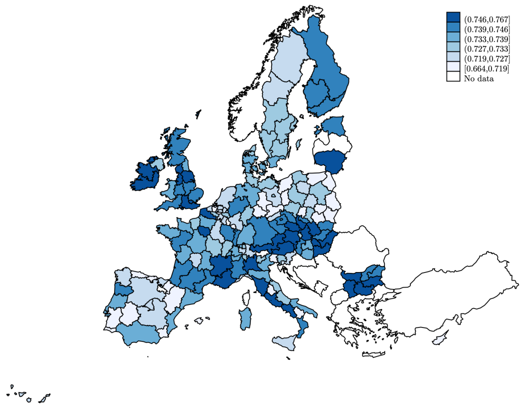
	Mean	Std. Dev.	Min.	Max.
Cultural Diversity:				
Overall (CF)	0.731	0.021	0.574	0.784
Within (CF^W)	0.722	0.025	0.547	0.780
Betwenn (F_{ST})	0.012	0.013	0.000	0.158
Share of immigrants:				
All (m_{rt})	0.101	0.094	0.000	0.654
High-skilled	0.027	0.034	0.000	0.286
Low skilled	0.074	0.064	0.000	0.407
Outside Europe	0.061	0.058	0.000	0.390
Within Europe	0.040	0.046	0.000	0.322
Less than 6 years	0.020	0.024	0.000	0.193
From 6 to 10 years	0.018	0.021	0.000	0.142
More than 10 years	0.061	0.059	0.000	0.363
Controls:				
ln(Density)	5.620	1.134	1.909	9.612
ln(GDP per capita)	10.688	0.643	8.516	12.114
Unemployment rate	8.955	5.007	1.193	34.800
Share of High-skilled	27.753	8.999	6.800	58.400

Source: Authors' elaboration on ESS, EU-LFS, and Eurostat data (2004-2018).

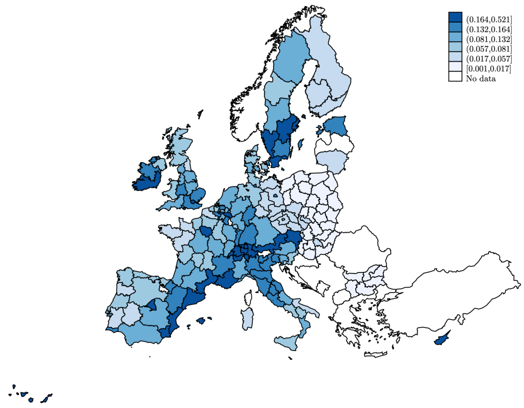
A.4 Geographical distribution

The geographical distribution of our sample is presented in Figure A-1, showing the average values of cultural diversity and immigrant populations across European regions from 2004 to 2018. In Figure A-1(a), we observe the dispersion of average cultural diversity. Regions in Central-Eastern Europe, such as Eastern Austria (AT1), Central Slovakia (SK03), Ireland, and Île-de-France (FR10), exhibit a high degree of cultural diversity. Conversely, Polish regions and central Spain display the lowest levels of heterogeneity. Figure A-1(b) provides a descriptive representation of immigrant distribution across regions. Predictably, coastal areas in France, Spain, and Italy exhibit the largest concentration of immigrants, as do regions housing major metropolitan areas like London and Brussels. The overlap between these two distributions is visualized in Figure A-1(c), where regions characterized by both a high immigrant population and significant cultural diversity are shaded in dark colors. Once more, coastal regions demonstrate distinctive patterns in both cultural diversity and immigrant populations, as do regions hosting capital cities. Notably, the Iberian Peninsula is primarily characterized by a high immigrant population and low cultural diversity, while Eastern European countries tend to display the opposite trend.

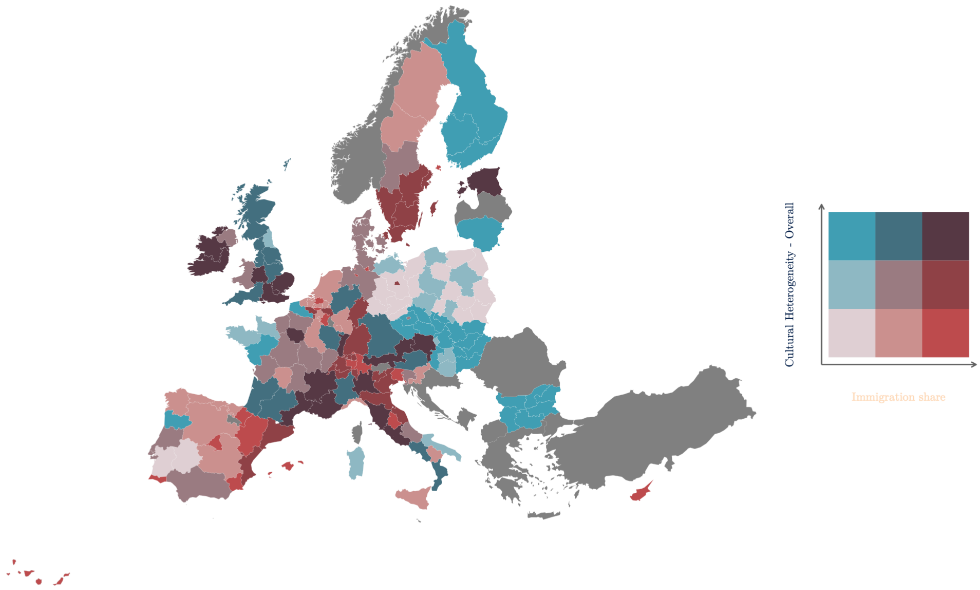
Figure A-1: Overall cultural diversity and the share of immigrants - distribution.



(a) Cultural Diversity



(b) Share of Immigrants



(c) Combined Distributions

Note: This figure depicts the average overall cultural diversity across European regions as defined in Eq. (5) between 2004 and 2018 and its associated overall share of immigrants as defined in Eq. (10).
Source: Authors' elaboration on ESS and EU-LFS data (2004-2018).

A.5 Cultural distances: Definitions

We first construct a bilateral dataset of 199 origin countries and 23 destination countries. We follow a simple approach to impute missing distance values between country pairs. First, we replace missing distances with the average distance between the destination country d and other countries within the same region of origin o, where regions are defined according to the detailed UN classification (22 regions in total). If all distances between d are missing for a region, we substitute the missing value with the global average distance between d and all origin countries. Finally, to align with the classification used in the European Social Survey (ESS), we aggregate distances for origin countries within each ESS-defined area by taking the average distance of all origin countries in the respective area. The list of distances that we use is reported below:

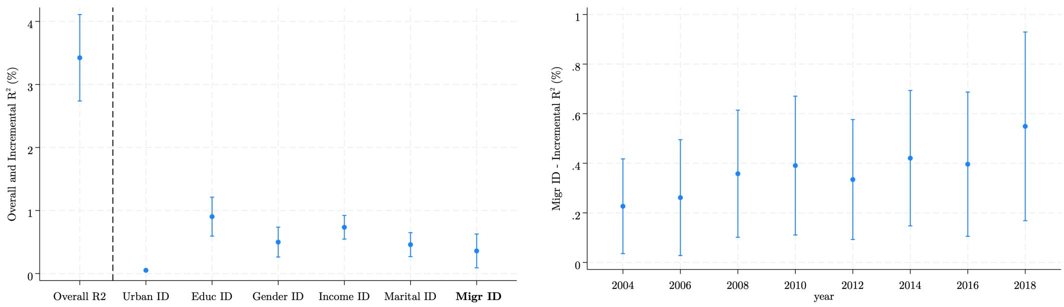
- Log of geodesic distance: Log of the population-weighted average distance between any pairs of cities from two countries. (Pellegrino et al., 2025).
- Cultural distance: Overall cultural distance capturing the average expected disagreement on a question of the World Values Survey by two individuals randomly drawn from those two countries (disagreement = 1 for different answers, 0 for identical answers). (Pellegrino et al., 2025).
- Economic distance: Differences in GDP per capita PPP (current thousands international dollars) (Conte et al., 2022).
- Religious distance: Expected normalized tree distance between the religions of two individuals randomly drawn from the populations of two countries. (Pellegrino et al., 2025).
- Linguistic distance: Expected normalized tree distance between the languages spoken by two individuals randomly drawn from the population of two countries. (Pellegrino et al., 2025).
- Genetic distance: Weighted F_{ST} genetic distance (Spolaore and Wacziarg, 2016).

B Birthplace and Traits

B.1 On the Relevance of Birthplace as Cleavage

One of our main contributions to the literature lies in introducing birthplace as a new cleavage to study the evolution of the cultural divide in modern societies. To further motivate our analysis and to grasp the role played by migration as an identity marker or cleavage, we first estimate a simple linear regression model over the 46 cultural traits, which include a series of dummy regressors highlighted by Desmet and Wacziarg (2021) as potentially relevant identity markers and our birthplace identity marker.[37] We then re-estimate the models, excluding one identity marker after the other, and record the different R^2 of the estimated models. Finally, we compute the incremental R^2 contribution of each identity marker for each cultural trait i by taking the difference between the estimated overall R^2 once we include all the identity markers and the conditional(s) R^2 once we exclude identity markers one by one. We average these results over the different traits, and report them in Figure B-1(a).

Figure B-1: Migration status as identity marker - Overall and incremental R^2.



(a) Overall R^2 and Incremental by ID (b) Incremental R^2 of Migr ID

Notes: Figure (a) plots the average overall R^2 of linear regressions over the 46 cultural traits and also including all the identity cleavages together, and the incremental R^2 due to the inclusion of one identity cleavage at a time. Figure (b) plots the average incremental R^2 of linear regressions over the 46 cultural traits of the migration identity cleavage over time. The figures report the average value and the 95% CI over the 46 cultural traits.
Source: Authors' calculations on ESS data (2004-2018).

The average overall R^2 is small, around 3.5%, in line with previous empirical analysis on the impact of identity markers on individual preferences (Desmet and Wacziarg, 2021; Moriconi et al., 2025). Focusing on the relevance of the different identity markers, income, and tertiary education appear to be the most relevant in explaining individual cultural traits. Migration status fares relatively well among the set of identity markers, being as relevant as marital status or gender, and reporting higher explanatory power than living in an urban area (vs. rural area). Figure B-1(b) plots the evolution of the incremental R^2 of migration status as an identity marker. Over time, migration status as an identity marker increases its explanatory power almost threefold. This suggestive evidence confirms the importance of focusing on the

[37] These identity markers are (i) gender, (ii) college education, (iii) living in an urban area, (iv) belonging to the top two quantiles of the income distribution, (v) marital status. We add immigration status as an additional identity marker.

role of immigration as a potential contributor to the evolution of the diffusion of cultural values.

B.2 Individual Analysis

This appendix takes advantage of the individual-level dimension available in the European Social Survey to examine the extent to which first-generation immigrants indeed exhibit differences in their cultural traits compared to the native population (with second-generation immigrants excluded from this analysis).

For the sake of interpretability, all cultural traits are first recoded so that lower values consistently represent greater conservatism and higher values greater liberalism. We also group cultural traits by blocks related to broader topics such as religiosity (RE), sexual morality (SM), role of the state (RS), cultural capital (CK), political participation (PP), trust toward the institution (TT), attitudes towards immigrants (AM) and openness (OP). This allows us to summarize the information contained in each block by extracting the first principal component score from a principal component analysis (PCA) of all cultural traits within each block. The interpretation of each trait and first principal component scores is provided in Table B-1.

Table B-1: Interpretation of traits.

Cultural Block	Higher Value	Lower value
RE	Less religiosity	More religiosity
SM	Liberalism	Conservatism
RS	Liberalism	Conservatism
CK	Liberalism	Conservatism
PP	Higher engagement	Lower engagement
TT	More trust	Less Trust
AM	Positive attitudes	Negative attitudes
OP	More Open	Less Open

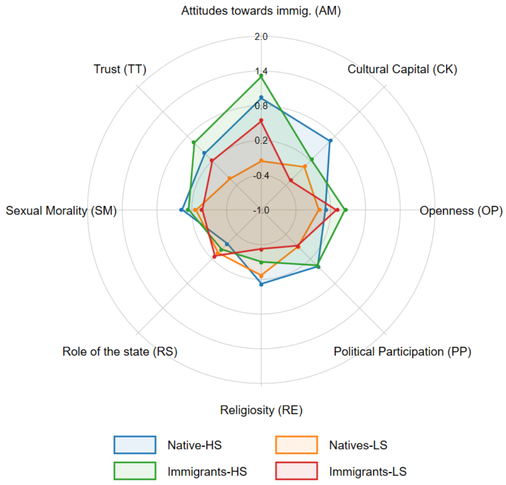
Source: Authors' elaboration on ESS data (2004-2018).

Figure B-2 depicts the average value of the first principal component score for each cultural block, for natives and immigrants, disaggregated by cleavages. Again, lower values indicate greater conservatism and higher values indicate greater liberalism, in line with Table B-1.

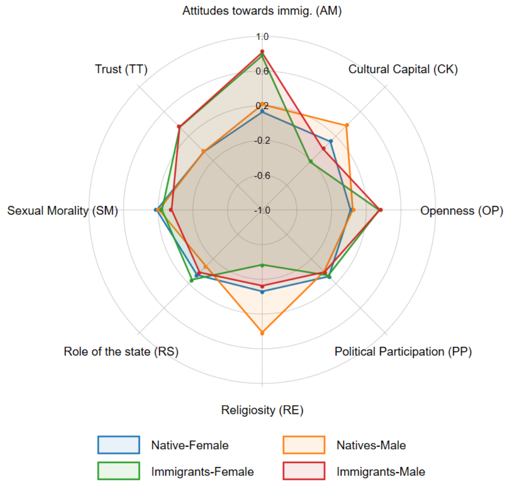
Figure B-2(a) reports that low-skilled immigrants and natives exhibit lower levels of liberalism (i.e., greater conservatism) on almost every trait except on the role of the state dimension compared to high-skilled individuals. Figures B-2(b) and (c) depict some differences between men and women, as well as between urban and rural residents, but the latter appear as relatively modest compared to those observed by educational attainment. Figure B-2(d) shows that more religious individuals display, mechanically, greater conservatism in religiosity, but also in sexual morality, with again less striking differences between immigrants and natives. Finally, we replicate the analysis for immigrants from different origin countries in Figure B-3, which again displays relatively modest differences between origins. Still, one can observe that immigrants from African origins are more religious, more conservative on sexual morality, and less likely to engage in politics than other immigrants.

It is important to acknowledge that raw differences between immigrants and natives can be confounded by other individual characteristics. To account for this, we conduct an additional individual-level analysis that allows us to isolate the effect of immigration status from other

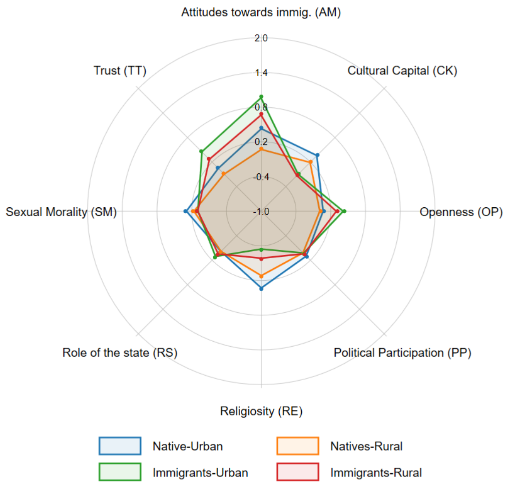
Figure B-2: Immigrants and Natives' average cultural values by cleavages



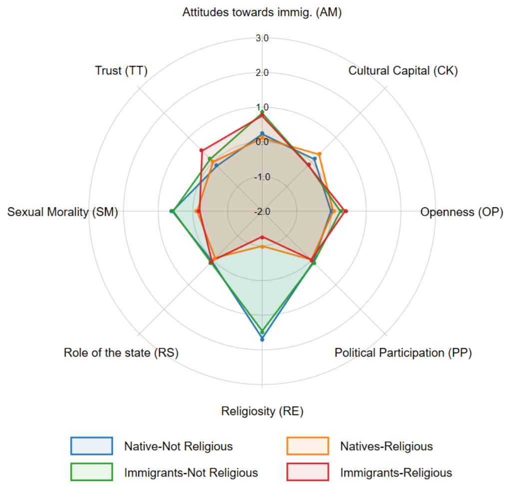
(a) Education



(b) Gender



(c) Urbanicity



(d) Religiosity

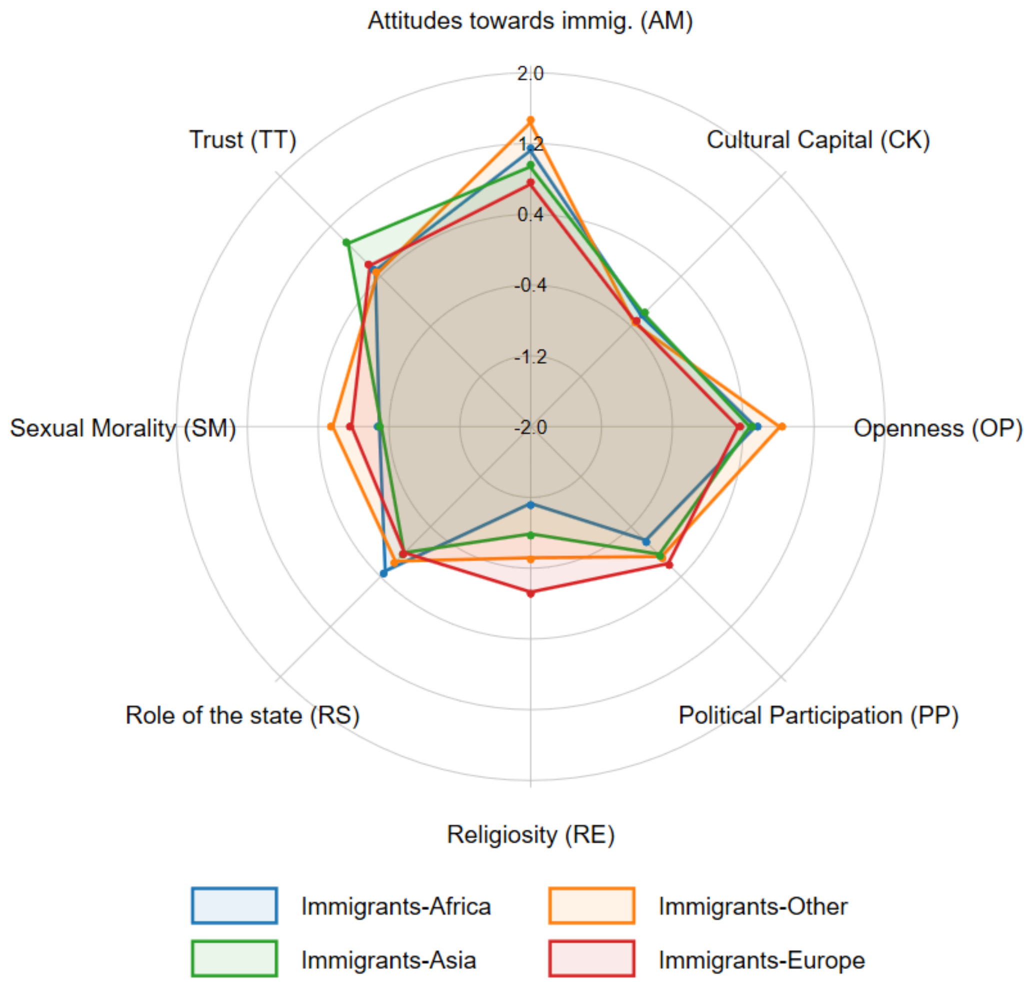
Notes: Each graph displays, for a given cleavage, the average value of the first principal component score of each cultural block, disaggregated by birthplace. Principal component scores are obtained from a principal component analysis (PCA) of all cultural traits within each block. All traits were first recoded so that lower values consistently represent more conservative attitudes and higher values more liberal ones. Averages are computed using individual weights. Source: Authors' elaboration on ESS data (2004-2018).

factors. Estimates at the individual level also have the advantage of mitigating the impact of limited observations in the European Social Survey (ESS) when constructing cultural diversity indices at the region-year level. We estimate the following specification for each memetic trait:

$$Y_{i,r,t} = \alpha + \delta_1 Firstgen^g_{i,r,t} + \boldsymbol{\sigma' Z_{i,r,t}} + \boldsymbol{\phi' X_{r,t}} + \gamma_t + \gamma_r + \varepsilon_{i,r,t} \tag{B-1}$$

where $Y_{i,r,t}$ is the standardized (mean zero and standard deviation of one) individual cultural traits of individual i in region r at year t. $Firstgen^g_{i,r,t}$ is a dummy variable equal to 1 if the

Figure B-3: Immigrants' average cultural traits by origin

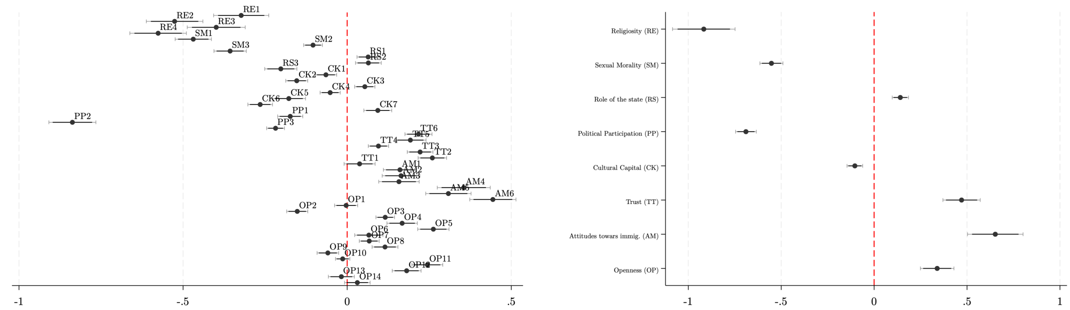


Notes: This graph displays the average value of the first principal component score of each cultural block, disaggregated by origin. The principal component scores are obtained from a principal component analysis (PCA) of all cultural traits within each block. All traits were recoded so that lower values consistently represent more conservative attitudes and higher values more liberal ones. Averages are computed using individual weights.
Source: Authors' elaboration on ESS data (2004-2018).

individual is a first-generation immigrant of a subgroup g and 0 if the individual is a native. Second-generations remain excluded from the analysis here. $\boldsymbol{Z_{i,r,t}}$ is a vector of individual-level controls including age, age-squared, gender, employment, educational attainment, marital status, presence of children, and urbanization. $\boldsymbol{X_{r,t}}$ is the same vector of regional-level controls, including the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the population. γ_t and γ_r stand for year and regional fixed effects, respectively. Standard errors are clustered at the regional level. Estimates are weighted using individual weights. It is worth noting that these OLS estimates are by no means causal but should be viewed only as suggestive additional evidence supporting the notion of evolving cultural traits associated with specific immigrant characteristics in destination countries.

Figure B-4(a) depicts that, on average, immigrants differ significantly from natives across almost all cultural traits. In detail, immigrants, on average, introduce significantly more conservative values to their destination country on sexual morality and religiosity: immigrants tend to be more religious, hold more conservative views on gay rights, and are more inclined to believe that traditions and customs must be followed, for instance. As expected, they are also less likely to be politically engaged at the destination. On the other hand, immigrants tend to lean more toward left-wing political views compared to the native population, and they report a higher level of trust towards the institutions and more positive attitudes toward immigrants. Replicating this analysis on the first principal component score of each cultural block leads to the same conclusions as reported in Figure B-4(b).

Figure B-4: Immigrants' compositional effect - individual analysis.



(a) Cultural traits (b) Cultural blocks

Notes: Each coefficient on Figure (a) represents the estimate obtained from a separate regression of first-generation immigrant dummies on each memetic trait. Each coefficient on Figure (b) represents the estimate obtained from a separate regression of first-generation immigrant dummies on the first component of each cultural block. All estimates include a vector of individual control with age, age squared, gender, children, urbanicity, marital, and employment status. All estimates include a vector of regional controls, including the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the total population. All estimates include regional and year fixed effects. Estimates are weighted using individual weights. Standard errors are clustered at the regional level.
Source: Authors' elaboration on ESS data (2004-2018).

C Validity of the Instrument

This appendix addresses several potential concerns associated with the use of a shift-share IV strategy. First, we assess the robustness of our findings by employing two alternative approaches for computing the "shift" component in our shift-share instrument. We then verify the presence or absence of any correlation between pre-existing regional characteristics and the variability in our instrument. We examine the potential bias in our standard errors arising from the correlation in the error term across regions with similar origin-specific shares, as pointed out by Adão et al. (2019). Finally, we provide a robustness analysis excluding years close to the initial distribution of immigrants by origin.

Standard & Leave-one-out shift-share. The shift-share instrument described in Section 4.2 consists in combining the initial distribution of foreign-born for each origin-region pair in 2004 ($SK_{o,r,2004}$) with the predicted total stock of foreign-born for each origin-year ($\widehat{Tk_{o,t}}$). Our identifying assumption hinges on the exogeneity of these predicted immigrants' stocks (Borusyak et al., 2022). We test the robustness of our benchmark result presented in Table 3 with two alternative methods for computing the total immigrants' stocks. First, we consider the more conventional approach, using the actual total stock of foreign-born from origin o in year t in our overall sample of 23 EU countries, as obtained from the EU-LFS. Second, we implement a leave-one-out version of our primary shift-share instrument, originally proposed by Autor and Duggan (2003). The leave-one-out estimator excludes own-destination i predicted stock of foreign-born when calculating the total predicted stock of foreign-born for each origin o and year t across all destinations d. Hence, we can rewrite Equation (17) as follows:

$$\widehat{Tk_{i,o,t}} = \sum_{d-i} \widehat{k_{o,d,t}} \tag{C-2}$$

The rationale behind using this leave-one-out version of the shift-share is to enhance the exogeneity of our instrument by eliminating any remaining mechanical relationships when computing the total predicted stocks for each origin-year observation. The results are presented in Table C-1 and remain robust to using these two alternative versions of the instrument.

Pre-trend analysis. We check that the variation in the predicted immigrant stock is not associated with pre-existing regional trends, which could be correlated with cultural diversity. To test that, we estimate the correlation between the growth of several regional indicators over the three years leading up to our initial sample year and the growth of the regional predicted stock of foreign-born over the subsequent three years. In other words, we regress the growth of our shift-share instrument over the period 2004-2007 on the 2000-2003 growth of GDP per capita, population density, unemployment rate, and the share of the tertiary educated population, while controlling for country fixed effects. Results are reported in Table C-2, with Col. (1) displaying the results for all immigrants, Col. (2) for high-skilled immigrants, Col. (3) for low-skilled immigrants, Col. (4) for immigrants from EU28 countries, and Col. (5) for immigrants from Non-EU28 countries. Our findings indicate that, overall, there is no significant correlation between the pre-2004 trend growth in regional indicators and the variation in predicted immigrant stocks as measured by our instrument. One exception relates to population density's correlation with overall immigration growth, although it is only statistically significant at the 10% level.

Inference à la Adão et al. (2019). Another concern raised by Adão et al. (2019) in the

Table C-1: 2SLS estimates using alternative shift-share instruments.

	(1) Standard	(2) Standard	(3) Leave-one-out	(4) Leave-one-out
m_{rt}	-0.196***	-0.249***	-0.180***	-0.219***
	(0.058)	(0.083)	(0.048)	(0.065)
Regional FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Regional controls	No	Yes	No	Yes
Observations	1,235	1,235	1,235	1,235
Mean Cultural Index	0.728	0.728	0.728	0.728
Mean Immig. Share	0.101	0.101	0.101	0.101
First-stage	0.739	0.633	1.754	1.551
KP F-Test	77.882	39.922	130.665	63.022

Notes: *** p <0.01, ** p <0.05, * p < 0.1. Standard errors in parentheses are clustered at the regional level. The dependent variable is CF_{rt}, the measure of cultural diversity in the region r at time t. The independent variable m_{rt} is the share of foreign-born in the total 2004 population. Regional controls include the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the resident population.

Source: Authors' elaboration on ESS and EU-LFS data (2004-2018).

shift-share setting is the potential spatial correlation of shocks across regions with similar shares. This spatial correlation, if present, could lead to a downward bias in standard errors due to estimation noise in the error terms. To address this concern, we follow the approach proposed by Adão et al. (2019) and compute standard errors that account for the correlation in the error terms between regions with a similar initial distribution of immigrants. Table C-3 reports the standard errors, p-values, and confidence intervals obtained using robust standard errors, clustered standard errors (as in our benchmark specification), and the inference procedure described in Adão et al. (2019) (referred to as AKM). Reassuringly, the precision of our estimates remains unaffected when employing any of the aforementioned inference methods.

Historical shares. Although not crucial for our identifying assumption, which is based on the exogeneity of the shocks by origin (Borusyak et al., 2022), one might raise concerns regarding the proximity of the shares we use to construct our instrument to the initial year of estimation in our sample. Therefore, as a robustness check, we adopt a reverse approach by maintaining shares defined in 2004 but sequentially excluding each year in our sample from 2004 onwards. This introduces a time gap between the year our shares are defined in and the initial year of our estimation sample. Figure C-1 demonstrates that until the exclusion of the 2004-2012 period, when our sample size starts becoming very small, our main conclusions remain unaffected by the exclusion of the preceding years.[38]

[38] An alternative approach would be to use initial shares obtained before 2004. For instance, Edo and Özgüzel (2023) digitized census data from the early 1990s and merged it with EU-LFS. Although relevant, this data would cover only 13 countries from our sample, hence generating concerns due to the restriction of our sample to 23 countries.

Table C-2: Pre-trend analysis.

	Instrument growth				
	(1) All	(2) HS	(3) LS	(4) EU28	(5) NEU28
GDP per capita growth	-0.004	0.048	-0.021	-0.001	0.006
	(0.008)	(0.050)	(0.017)	(0.004)	(0.006)
P-value	0.621	0.364	0.222	0.722	0.431
t-stat	-0.506	0.956	-1.209	-0.379	0.888
Population density growth	0.058*	-0.062	0.088	0.020	0.037
	(0.031)	(0.137)	(0.067)	(0.013)	(0.034)
P-value	0.243	0.806	0.501	0.349	0.604
t-stat	1.858	-0.450	1.324	1.461	1.112
Unemployment rate growth	0.002	-0.004	0.004	0.000	-0.000
	(0.002)	(0.003)	(0.003)	(0.000)	(0.002)
P-value	0.354	0.146	0.206	0.448	0.812
t-stat	0.993	-1.452	1.439	0.834	-0.225
Tertiary education growth	-0.006	0.005	-0.009	0.001	-0.003
	(0.003)	(0.011)	(0.007)	(0.001)	(0.003)
P-value	0.029	0.862	0.206	0.674	0.249
t-stat	-1.763	0.429	-1.286	0.423	-1.117
Country FE	✓	✓	✓	✓	✓

Notes: *** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at the country level. This table shows the coefficients of regressing the regional predicted migration growth (shift-share instrument) over the period 2004 and 2007 on the growth rate of regional economic indicators between 2000 and 2003. We report the p-value and t-stat of the wild cluster bootstrap (999 replications) with Webb weights test.

Source: Authors' elaboration on ESS, EU-LFS, and Eurostat data (2000-2007).

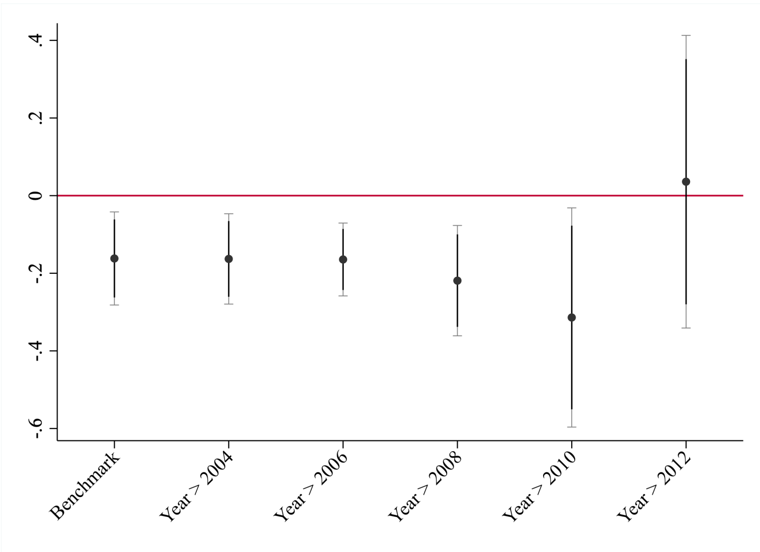
Table C-3: Adão et al. (2019) inference procedure.

	Coefficient	Std. error	P-value	Confidence Interval
Second-stage				
Robust	-0.206	0.048	0.0000	[-0.300,-0.112]
Cluster	-0.206	0.067	0.0024	[-0.338,-0.074]
AKM	-0.203	0.041	0.0000	[-0.282,-0.123]
First-stage				
Robust	1.292	0.122	0.0000	[1.053,1.531]
Cluster	1.292	0.185	0.0000	[0.928,1.656]
AKM	1.299	0.130	0.0000	[1.044,1.554]

Notes: This table reports the first and second stages benchmark coefficients, standard errors, p-values, and confidence intervals using various inference methods. Robust refers to robust standard errors. Cluster refers to clustered standard errors at the regional level. AKM refers to the inference procedure described in Adão et al. (2019). The dependent variable is CF_{rt} the measure of cultural diversity in the region r at time t. The independent variable is the share of foreign-born in the total 2004 population. We control for the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the resident population. All estimates include regional and year fixed effects.

Source: Authors' elaboration on ESS and EU-LFS data (2004-2018).

Figure C-1: 2SLS estimates dropping years sequentially.



Notes: This figure depicts the coefficients obtained from estimating Equation (11) dropping years sequentially until 2012. The dependent variable is CF_{rt}, the measure of cultural diversity in the region r at time t. The independent variable is the share of foreign-born in the total 2004 population. We include 95% and 90% confidence intervals around the estimated coefficients. We control for the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the total population. All estimates include regional and year fixed effects. Standard errors are clustered at the regional level.
Source: Authors' elaboration on ESS and EU-LFS data (2004-2018).

Table C-4: Zero-stage bilateral migration gravity model estimates.

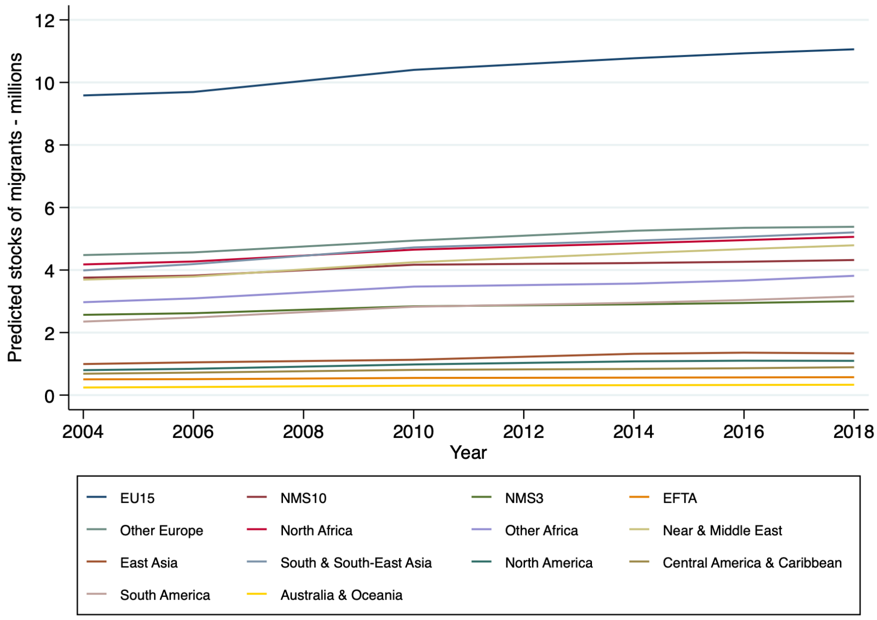
	(1) Stock of immigrants
ln(Deaths)	0.026***
	(0.006)
Disasters	0.004***
	(0.001)
ln(Distance) × 1995	-0.085***
	(0.031)
ln(Distance) × 2000	-0.113***
	(0.036)
ln(Distance) × 2005	-0.059**
	(0.025)
ln(Distance) × 2010	-0.030
	(0.020)
ln(Distance) × 2015	-0.018**
	(0.009)
Year FE	Yes
Destination × Year FE	Yes
Origin × Destination FE	Yes
Observations	78,561
R-squared	0.987

Notes: *** p <0.01, ** p <0.05, * p < 0.1. Standard errors in parentheses are clustered at the country-pair level. The dependent variable is the stocks of immigrants from origin o in destination d at year t.

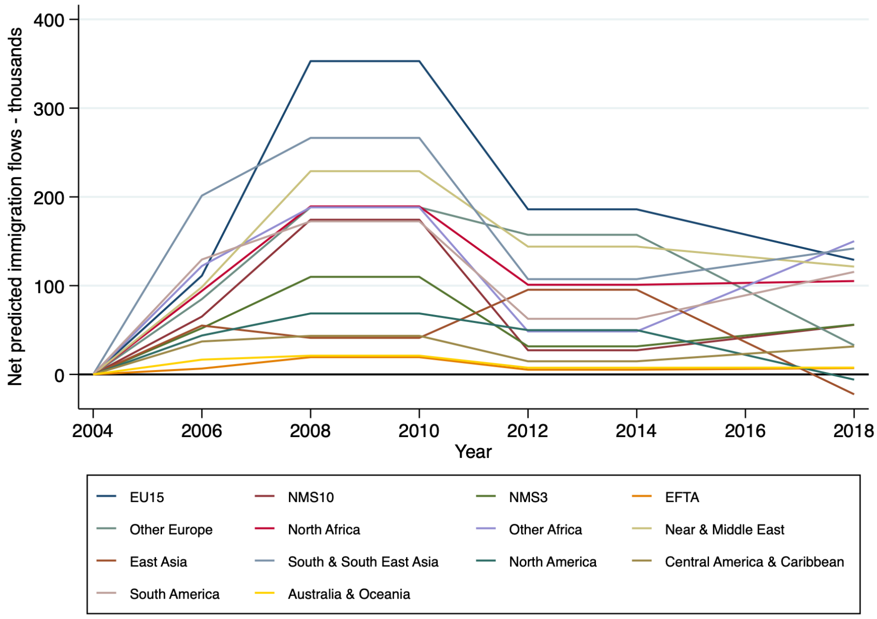
Source: Authors' elaboration on data from the United Nations, the UCDP/PRIO Armed Conflict Dataset, and the Emergency Events Database (EM-DAT) (1990-2020).

Figure C-2: The variation of the predicted stocks of migrants by origin-group.

(a) Predicted immigration stocks



(b) Net predicted immigration flows



Notes: These figures depict the variation of the predicted stocks of migrants for each origin group. The predicted immigration stocks are displayed in Figure (a), while the predicted net immigration flows with two years lagged are displayed in Figure (b).
Source: Authors' elaboration on the United Nations Population Division data (2004-2018).

D Robustness Checks

This appendix conducts a series of robustness checks and tests to check whether our benchmark results are sensitive to various methodological choices and data.

D.1 Group-Year fixed effects

Table D-1: 2SLS estimates using alternative FE

	(1) Benchmark	(2) Dropping NUTS0	(3) Country-Year FE	(4) Geography-Year FE	(5) Enlargement-Year FE	(6) Welfare-Year FE
m_{rt}	-0.206***	-0.197***	-0.055	-0.181***	-0.200**	-0.186***
	(0.067)	(0.063)	(0.070)	(0.066)	(0.096)	(0.060)
Observations	1,235	1,208	1,208	1,235	1,235	1,234
Mean Cultural Index	0.728	0.728	0.728	0.728	0.728	0.728
Mean Immig. Share	0.101	0.100	0.100	0.101	0.101	0.101
First-stage	1.292	1.378	1.147	1.235	1.118	1.264
KP F-Test	48.568	60.236	24.157	40.047	22.922	41.840

Notes: *** p <0.01, ** p <0.05, * p < 0.1. Standard errors in parentheses are clustered at the regional level. The dependent variable is CF_{rt}, the measure of cultural diversity in the region r at time t. The independent variable m_{rt} is the share of foreign-born in the total 2004 population. Regional controls include the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the resident population. We add regional and year fixed effects.
Source: Authors' elaboration on data from the ESS and the EU-LFS (2004 to 2018).

Sources: Authors' elaboration on ESS and EU-LFS data (2004-2018).

D.2 Robustness to regions with no immigrants

Table D-2 also reports that our main conclusions remain unchanged when excluding regions with no immigrants in the EU-LFS, ESS, or both.

Table D-2: 2SLS estimates excluding regions with no migrants.

		Excluding no mig.		
	(1) Benchmark	(2) EULFS	(3) ESS	(4) EULFS & ESS
m_{rt}	-0.206***	-0.206***	-0.168**	-0.168**
	(0.067)	(0.067)	(0.067)	(0.067)
Observations	1,235	1,234	1,111	1,111
Mean Cultural Index	0.728	0.728	0.729	0.729
Mean Immig. Share	0.101	0.101	0.110	0.110
First-stage	1.292	1.292	1.265	1.265
KP F-Test	48.568	48.552	36.954	36.954

Notes: *** p <0.01, ** p <0.05, * p < 0.1. Standard errors in parentheses are clustered at the regional level. The dependent variable is CF_{rt}, the measure of cultural diversity in the region r at time t. The independent variable m_{rt} is the share of foreign-born in the total 2004 population. Regional controls include the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the total population. We add regional and year fixed effects.

Source: Authors' elaboration on ESS and EU-LFS data (2004-2018).

D.3 Including second-generations immigrants

Table D-3: 2SLS estimates including second-generation immigrants in the immigration group.

	(1) Overall	(2) Between	(3) Within	(4) Within Natives	(5) Within Immig.
m_{rt}	-0.173***	0.029*	-0.190***	-0.206***	-0.372
	(0.060)	(0.015)	(0.065)	(0.067)	(0.362)
Observations	1,235	1,235	1,235	1,235	1,196
Mean Cultural Index	0.731	0.731	0.731	0.731	0.731
Mean Immig. Share	0.101	0.101	0.101	0.101	0.103
First-stage	1.292	1.292	1.292	1.292	1.283
KP F-Test	48.568	48.568	48.568	48.568	46.678

Notes: *** p <0.01, ** p <0.05, * p < 0.1. Standard errors in parentheses are clustered at the regional level. The dependent variable in Col. (1) is CF_{rt}, the measure of cultural diversity in the region r at time t. The dependent variable in Col. (2) is F_{ST}, the measure of between-group cultural diversity in the region r at time t. The dependent variable in Col. (3) is CF^W_{rt} the measure of within-group cultural diversity in the region r at time t. The dependent variable in Col. (4) and (5) are CF^W_{rt}, the measure of within-group cultural diversity in the region r at time t for native and immigrants, respectively. The independent variable m_{rt} is the share of foreign-born in the total 2004 population. We control for the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the total population. All estimates include regional and year fixed effects.

Source: Authors' elaboration on ESS and EU-LFS data (2004-2018).

Table D-4: 2SLS estimates excluding first-generation immigrants.

	(1) Overall	(2) Between	(3) Within	(4) Within Natives	(5) Within Immig.
m_{rt}	-0.198***	0.019	-0.210***	-0.206***	-0.411
	(0.066)	(0.023)	(0.064)	(0.067)	(0.358)
Observations	1,235	1,235	1,235	1,235	1,089
Mean Cultural Index	0.731	0.731	0.731	0.731	0.733
Mean Immig. Share	0.101	0.101	0.101	0.101	0.103
First-stage	1.292	1.292	1.292	1.292	1.290
KP F-Test	48.568	48.568	48.568	48.568	45.244

Notes: *** p <0.01, ** p <0.05, * p < 0.1. Standard errors in parentheses are clustered at the regional level. The dependent variable in Col. (1) is CF_{rt}, the measure of cultural diversity in the region r at time t. The dependent variable in Col. (2) is F_{ST}, the measure of between-group cultural diversity in the region r at time t. The dependent variable in Col. (3) is CF^W_{rt} the measure of within-group cultural diversity in the region r at time t. The dependent variable in Col. (4) and (5) are CF^W_{rt}, the measure of within-group cultural diversity in the region r at time t for native and immigrants, respectively. The independent variable m_{rt} is the share of foreign-born in the total 2004 population. We control for the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the total population. All estimates include regional and year fixed effects.

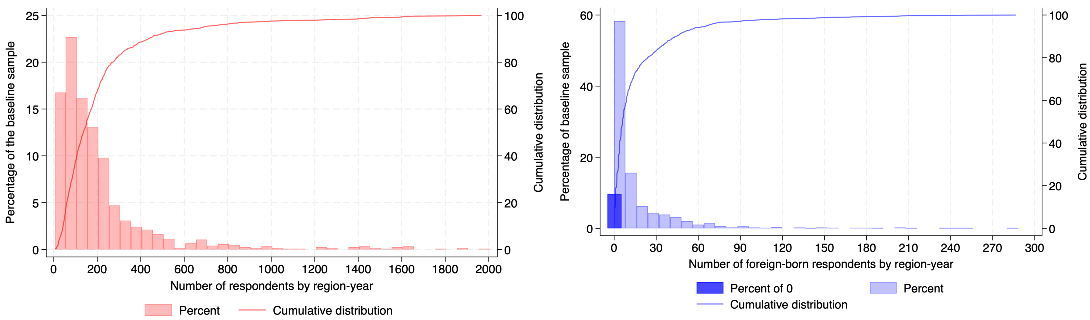
Sources: Authors' elaboration on ESS and EU-LFS data (2004-2018).

D.4 Number of observations.

Figure D-1 illustrates the distribution of the average number of observations at the region-year level in the ESS, depicted in Figure (a) for the native population and Figure (b) specifically for the foreign-born. It is important to highlight that 10% of region-year observations report zero foreign-born individuals according to the European Social Survey (ESS), and the overall distribution skews significantly to the right.[39] These observations raise concerns about the potentially small number of observations within each region-year cell used to compute our dependent variables.[40] We first check in Table D-5 whether our results are not overly sensitive to sequentially excluding region-year where cultural indices are based on fewer than 50 and 100 respondents. While our effect remains robust to the exclusion of regions with less than 50 respondents, we notice a significant decrease in magnitude and precision for the 100-respondents threshold. Still, it is plausible that this coefficient drop reflects a significant sample size change rather than the effect of regions with few observations. Hence, we conduct additional checks in Table D-6, where we report additional findings interacting our main effect with either a dummy variable for regions with fewer than 50 or 100 respondents. Such a test aims to capture potential heterogeneous effects driven by the number of observations in each region-year cell without modifying the sample of analysis. In both cases, we find that our results remain unaffected by regions with a limited number of observations.[41]

Figure D-1: Distribution of the number of observations - ESS.

(a) Native (b) Foreign-born



Notes: These figures depict the distribution of the number of observations in the European Social Survey (ESS) at the region-year level for the native population in Figure (a) and the foreign-born population in Figure (b).
Source: Authors' elaboration on ESS data (2004-2018).

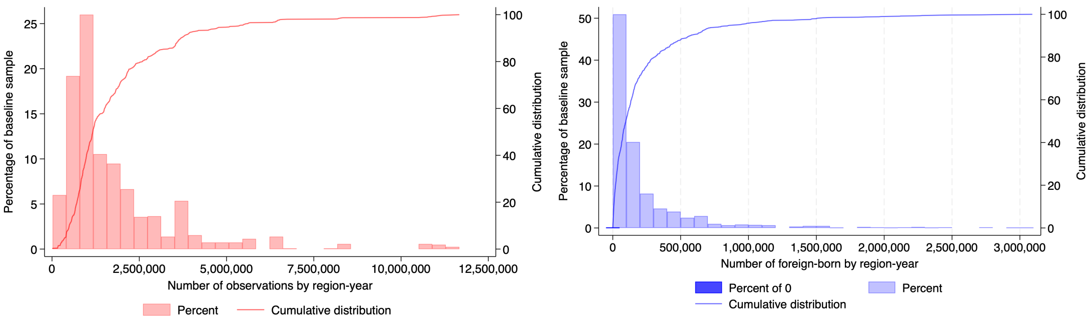
[39]Similar patterns are observed in Figure D-2 for EU-LFS data, while the number of absolute zeros is substantially lower.

[40]Due to the varying number of respondents to the different cultural traits questions in the ESS, we compute the number of observations at the region-year level using two approaches. First, we compute the number of observations for each cultural trait question in each region-year. Then, we consider both the maximum and the average number of observations across all cultural traits.

[41]We obtain similar results when we consider continuous or categorical measures of observation count by region-year. Our main conclusions also remain robust to weighted estimates using regions' population size, despite a slight decrease in the precision of the estimates.

Figure D-2: Distribution of the number of observations - EU-LFS.

(a) Native (b) Foreign-born



Notes: These figures depict the distribution of the number of observations in the European Labor Force Survey (EU-LFS) at the region-year level for the native population in Figure (a) and the foreign-born population in Figure (b).
Source: Authors' elaboration on EU-LFS data (2004-2018).

Table D-5: 2SLS estimates excluding regions with small number of observations

	Benchmark	Excluding maximum non-missing		Excluding average non-missing	
	(1)	(2)	(3)	(4)	(5)
	All obs.	Obs.<50	Obs.<100	Obs.<50	Obs.<100
m_{rt}	-0.206***	-0.178***	-0.129***	-0.156***	-0.119**
	(0.067)	(0.061)	(0.049)	(0.057)	(0.048)
Observations	1,235	1,073	771	1,055	757
Mean Cultural Index	0.728	0.732	0.735	0.732	0.735
Mean Immig. Share	0.101	0.098	0.095	0.098	0.096
First-stage	1.292	1.266	1.239	1.266	1.233
KP F-Test	48.568	55.531	49.477	55.070	48.726

Notes: *** p <0.01, ** p <0.05, * p < 0.1. Standard errors in parentheses are clustered at the regional level. The dependent variable is CF_{rt}, the measure of cultural diversity in the region r at time t. The independent variable m_{rt} is the share of foreign-born in the total 2004 population. Regional controls include the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the resident population. We add regional and year fixed effects.
Source: Authors' elaboration on ESS and EU-LFS data (2004-2018).

Table D-6: 2SLS estimates - Interaction with small regions dummies.

	Benchmark	Maximum non-missing		Average non-missing	
	(1)	(2)	(3)	(4)	(5)
m_{rt}	-0.206***	-0.203***	-0.232***	-0.208***	-0.231***
	(0.067)	(0.066)	(0.071)	(0.067)	(0.073)
$<$ 50 obs.		-0.014**		-0.015***	
		(0.006)		(0.004)	
m_{rt} $\times$ $<$ 50 obs.		0.013		0.029	
		(0.039)		(0.029)	
$<$ 100 obs.			-0.015***		-0.013**
			(0.005)		(0.005)
m_{rt} $\times$ $<$ 100 obs.			0.051*		0.042
			(0.027)		(0.028)
Observations	1,235	1,235	1,235	1,235	1,235
Mean Cultural Index	0.728	0.728	0.728	0.728	0.728
Mean Immig. Share	0.101	0.101	0.101	0.101	0.101
First-stage (m_{rt})	1.292	1.266	1.271	1.259	1.272
First-stage ($m_{rt} \times obs.$)		0.897	0.698	0.902	0.696
KP F-Test	48.568	28.102	24.554	28.173	24.724

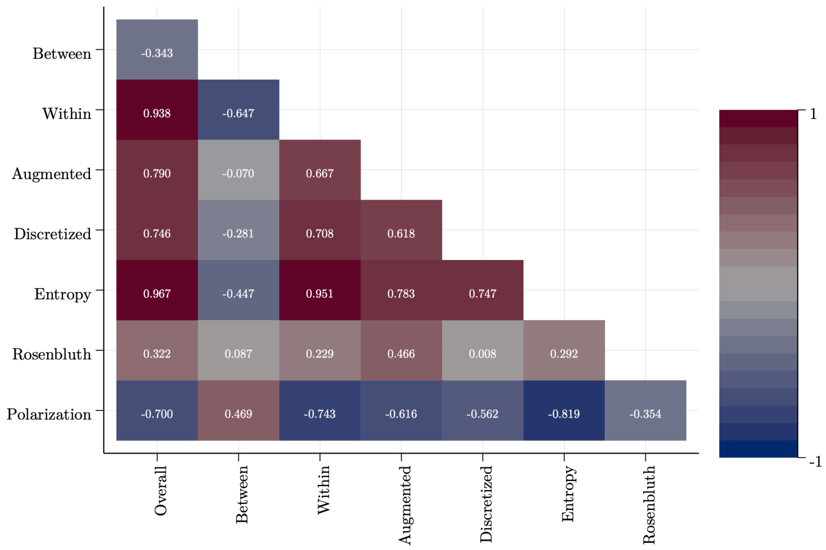
Notes: *** p <0.01, ** p <0.05, * p < 0.1. Standard errors in parentheses are clustered at the regional level. The dependent variable is CF_{rt}, the measure of cultural diversity in the region r at time t. The independent variable m_{rt} is the share of foreign-born in the total 2004 population. Regional controls include the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the resident population. We add regional and year fixed effects.

Source: Authors' elaboration on ESS and EU-LFS data (2004-2018).

D.5 Alternative indices

In our benchmark specification, we measure regional cultural diversity with a cultural diversity index. This index is a widely accepted and reliable measure with desirable measurement properties as outlined by Hall and Tideman (1967). Also, it has the advantage that it can be broken down into within and between components once an identity cleavage is identified (Desmet and Wacziarg, 2021). Furthermore, its extensive use in the literature makes it easy to understand and compare with alternative studies (Desmet et al., 2017). Still, its variation can be influenced by the number of questions and available answers for each question, and it does not consider that the contribution of each answer to the overall cultural diversity might differ based on prevailing norms. Therefore, in this appendix, we explore the construction and properties of alternative cultural diversity indices, which we subsequently use as alternative dependent variables in Tables D-7 and E-1. Figure D-3 displays the correlations between these different indices for reference.

Figure D-3: Cross-correlations across alternative indices.



Source: Authors' elaboration on ESS data (2004-2018).

Augmented Cultural Diversity - As previously noticed, our benchmark index of cultural diversity assumes that each answer to each question provides the same degree of contribution to the overall extent of cultural diversity. However, this is not necessarily the case. One way to account for this is to weigh the contribution of each answer based on the distance from the prevailing norm in a given region and year. Hence, following Greenberg (1956) for each trait $q = 1, ..., Q$ in region r at year t we compute the augmented cultural diversity index as follows:

$$CFA^q_{r,t} = \sum_{i_q=1}^{I_q} s^{i_q}_{r,t}(1 - s^{i_q}_{r,t})d^{i_q}_{r,t} \tag{D-3}$$

Compared to the benchmark definition, the augmented version weighs each answer $i_q = 1, ..., I_q$ of trait q by the relative distance from the prevailing norm in the region ($d^{i_q}_{r,t}$), which is defined as follows:

$$d^{i_q}_{r,t} = \frac{|\overline{i_{q,r,t}} - i_{q,r,t}|}{Max(i_q) - Min(i_q)} \tag{D-4}$$

$\overline{i_{q,r,t}}$ is the average norm computed for each region r and year t. By construction, the measure of distance from the prevailing norm ($d^{i_q}_{r,t}$) spans from 0 to 1, and higher values imply a further distance from the prevailing norm. By computing the augmented cultural diversity index across all the traits Q and averaging them out, we get the overall augmented cultural diversity index.

Discretized Cultural Diversity - The span of available answers for each cultural trait is quite heterogeneous, from traits that allow only two answers (e.g., RE1) to traits that allow eleven answers (e.g, TT1). To assess whether the variation in regional cultural diversity primarily arises from the construction of these traits rather than genuine shifts in respondents' views and values, we reduce the dimensionality of the set of available answers for traits with more than two options using the following criteria.[42] First, for traits that offer four answers ("Strongly Agree", "Agree", "Disagree", "Strongly Disagree"), we discretize them by combining responses into two categories: those who "Agree" and those who "Disagree".[43] Second, for traits providing answers on the frequency of certain activities, we discretize them with a dummy equal to one if they do it once or more per month, and zero otherwise.[44] Third, for traits that provide answers on a scale from 0 to 10, we reduce their dimensionality by categorizing responses into three groups: those answering from 0 to 3, those from 4 to 6, and those from 7 to 10.[45] Fourth, for traits asking whether something is important/not important for the respondent, we discretize them with a dummy equal to one if it is important/very important for the respondent, and zero otherwise.[46] Finally, for those traits providing four answers, we discretize them by combining the answers in just two blocks.[47] By computing our overall measure of cultural diversity using these discretized answers, we are less susceptible to capturing variability driven only by the measurement framework underlying each trait.

Rosenbluth Index of cultural diversity - As noted by Hall and Tideman (1967), in the cultural diversity index each answer within each trait is weighted by the share of the population holding that specific answer, implying that the relative share of respondents is more important than the absolute number of available answers in determining the degree of cultural diversity. Nonetheless, the number of available answers is indeed a relevant aspect to take into account. We do partially account for this issue by discretizing the set of available answers for each trait in the Discretized Cultural Diversity index. An alternative way to deal with this issue is to rely on the so-called Rosenbluth Index or Hall and Tideman index (Hall and Tideman, 1967). For each cultural trait $q = 1, ..., Q$, we construct the Rosenbluth Index as follows:

$$RB^q = \frac{1}{\left(2\sum_{i_q=1}^{I} r^{i_q} s^{i_q}\right) - 1} \tag{D-5}$$

[42]It is worth noting that this issue is strongly mitigated by the use of panel data, which compares variations across waves, and thus should be less affected by the definition of the variables.

[43]Traits: SM1 and RS1.

[44]Traits: RE3 and RE4.

[45]Traits: RE2, RS2, CK1, CK2, CK3, TT1, TT2, TT3, TT4, TT5, TT6, AM4, AM5 and AM6.

[46]Traits: SM2, SM3 and all the OP labeled traits.

[47]Traits: PP1, AM1, AM2, AM3.

The Rosenbluth index accounts for the rank of each answer (r^{i_q}) from the least used ($r^{i_q} = 1$) to the one that is mostly diffused in our setting. It is important to notice that the ranking is generated from the least to the most diffused answer. We then construct the average overall Rosenbluth index by averaging out the trait-specific Rosenbluth Indices.

Entropy Index of cultural diversity - An alternative measure of cultural diversity can be derived from the Entropy Index proposed by Shannon (1948). Such a measure aims to capture the degree of chaos of a specific system: the higher the value, the higher the uncertainty or the complexity of the system. Translating this type of measurement in our setting implies that higher values are associated with more cultural diversity. We then compute the average overall degree of cultural entropy (CE) index across the various cultural traits $q = 1, ..., Q$ as follows:

$$CE = \frac{1}{Q}\sum_{q=1}^{Q} CE^q = \frac{1}{Q}\sum_{q=1}^{Q}\Big(-\sum_{i_q=1}^{I} s^{i_q} ln(s^{i_q})\Big) \tag{D-6}$$

As Figure D-3 shows, our measure of cultural entropy is positive and highly correlated with our measure of cultural diversity.

Cultural Polarization - The measure of cultural diversity captures the overall degree of cultural diversity within a region. Another relevant index that can be computed is the overall degree of cultural polarization within a region. By relying on Montalvo and Reynal-Querol (2005), we construct a cultural polarization index for each cultural trait $q = 1, ..., Q$ which captures the closeness to a bimodal distribution of the trait q in each region r at year t. The Polarization Index is computed as follows:

$$PL^q = 1 - \sum_{i_q=1}^{I}\Big(\frac{0.5 - s^{i_q}}{0.5}\Big)^2 s^{i_q} \tag{D-7}$$

By averaging out the cultural polarization indices across our 46 variables, we then get an average overall measure of cultural polarization at the regional level. This index is negatively correlated with the overall index of diversity.

Robustness to alternative indices. Table D-7 challenges the robustness of our results to alternative definitions of the dependent variable, including those derived from the indices detailed in this appendix. We first adopt a more stringent selection of cultural traits for constructing the cultural diversity measure, retaining only those employed by Alesina et al. (2017). Then, we address the possibility that the contribution to overall cultural diversity may vary based on whether an individual's response is close to the prevailing norm in the region. To account for this, we follow Greenberg (1956) and construct an augmented cultural diversity index that assigns higher weights to answers that deviate further from the region-year average answer for a given cultural trait. Another concern involves the varying number of possible answers to each cultural trait question, which could impact the overall heterogeneity measure. While this concern is mitigated by our panel data structure, which explores within-region variation, we also recompute the overall cultural diversity measure using a discretized version of all cultural trait variables. Furthermore, we consider two alternative diversity measures: the Rosenbluth index (Hall and Tideman, 1967) and the Entropy index (Shannon, 1948), and a polarization index (Montalvo and Reynal-Querol, 2005). The results provided consistently indicate a negative effect of immigration on all cultural diversity indices.

Table D-7: 2SLS estimates using alternative indices.

	(1) Benchmark	(2) Augmented	(3) Discretized	(4) Entropy	(5) Rosenbluth	(6) Polarization
m_{rt}	-0.165***	-0.074***	-0.222***	-0.444***	-0.014	0.083**
	(0.061)	(0.027)	(0.072)	(0.160)	(0.009)	(0.038)
Observations	1,235	1,235	1,235	1,235	1,235	1,235
Mean Cultural Index	0.731	0.161	0.456	1.545	0.136	0.670
Mean Immig. Share	0.101	0.101	0.101	0.101	0.101	0.101
First-stage	1.292	1.292	1.292	1.292	1.292	1.292
KP F-Test	48.568	48.568	48.568	48.568	48.568	48.568

Notes: *** p <0.01, ** p <0.05, * p < 0.1. Standard errors in parentheses are clustered at the regional level. The dependent variables are the overall cultural diversity in (1), the augmented cultural diversity in (2), the discretized cultural diversity in (3), the Entropy diversity index in (4), the Rosenbluth diversity index in (5), and the polarization index in (6). The independent variable m_{rt} is the share of foreign-born in the total 2004 population. Regional controls include the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the resident population. We add regional and year fixed effects.

Source: Authors' elaboration on data from the ESS and the EU-LFS (2004 to 2018).

D.6 Natives' mobility response to Migration.

Our benchmark results show that immigration contributes to a stronger homogenization of values across European regions; an effect that is driven by the natives' response to immigration as shown in Table 4. To explore whether this effect can be driven by a selection mechanism, where natives with distinct values and preferences move out from the region after the arrival of immigrants, we test whether immigration contributed to an internal net-migration of natives (Edo et al., 2019). In line with Moriconi et al. (2022), we estimate the following equation:

$$\Delta Natives_{r,t} = \alpha + \beta_2 m_{r,t-1} + \boldsymbol{\beta' X_{r,t}} + \gamma_t + \gamma_r + \varepsilon_{r,t}, \quad \text{(D-8)}$$

where $\Delta Natives_{r,t}$ is the standardized variation of native population between year t and $t-1$ defined either as the share of the total population ($Share$) or as the inflow of new native residents to the region ($Inflow$).[48] The variable of interest is $m_{r,t-1}$, the lagged share of migrants in region r. We estimate equation (D-8) with 2SLS, relying on the same instrument used for the benchmark regression. Table D-8 reports no effect of immigration in year $t-1$ on the variation of natives (either in share or inflow) between year t and $t-1$. Additionally, we find no statistically significant effect when estimating the impact of immigration in year t or $t-2$, nor when exploring potential heterogeneous responses of natives by education level. Overall, these findings support that the potential mechanism driven by natives' selection is unlikely in our context.

[48]EU-LFS provides information on the respondent's region of residence in the previous year. Hence, we calculate the inflow of new native residents by identifying those who were living in a different region compared to their current region of residence.

Table D-8: Natives' Mobility Response to Immigration.

	Migration $(t-2)$		Migration $(t-1)$		Migration(t)	
	(1)	(2)	(3)	(4)	(5)	(6)
	$Share^n_t$-$Share^n_{t-1}$	$Inflow^n_t$-$Inflow^n_{t-1}$	$Share^n_t$-$Share^n_{t-1}$	$Inflow^n_t$-$Inflow^n_{t-1}$	$Share^n_t$-$Share^n_{t-1}$	$Inflow^n_t$-$Inflow^n_{t-1}$
Panel A - All Natives						
$m_{r,t-2}$	2.773	4.830				
	(2.026)	(3.272)				
$m_{r,t-1}$			–0.945	5.059		
			(2.396)	(4.354)		
$m_{r,t}$					–0.499	2.034
					(2.372)	(2.555)
Region FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
Observations	954	855	1095	954	1212	1040
K-Paap F-stat	58.17	22.60	68.34	37.47	47.21	43.39
Panel B - LS Natives						
$m_{r,t-2}$	2.404	–0.125				
	(2.443)	(2.870)				
$m_{r,t-1}$			–0.581	1.279		
			(3.008)	(3.510)		
$m_{r,t}$					0.180	1.236
					(2.715)	(2.602)
Region FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
Observations	954	855	1095	954	1212	1040
K-Paap F-stat	58.17	22.60	68.34	37.47	47.21	43.39
Panel C - HS Natives						
$m_{r,t-2}$	1.773	9.304				
	(2.701)	(7.211)				
$m_{r,t-1}$			–1.121	7.923		
			(2.598)	(6.152)		
$m_{r,t}$					–1.646	2.252
					(1.985)	(2.589)
Region FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
Observations	954	855	1095	954	1212	1040
K-Paap F-stat	58.17	22.60	68.34	37.47	47.21	43.39

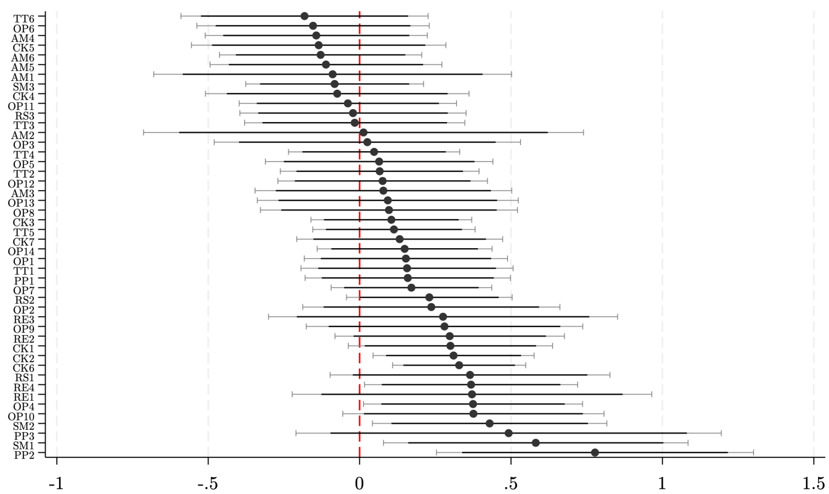
Notes: *** p <0.01, ** p <0.05, * p < 0.1. Standard errors in parentheses are clustered at the regional level. The dependent variable is the standardized change in the share of native over the total population or the standardized change in the share of new native residents between time t and *t-1*. The independent variable is $m_{r,t-2}$, $m_{r,t-1}$ or m_{rt}, which is the four-year lagged, two-year lagged share or the share of foreign-born in the total 2004 population. The results are presented for the overall native population, and splitting between college-educated and low-educated natives. Regional controls include the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the total population.

Source: Authors' elaboration on EU-LFS data (2004-2018).

E Additional Results

E.1 Salience of birthplace by cultural trait

Figure E-1: Salience of birthplace by Cultural Trait (Between component)



Notes: Standard errors are clustered at the regional level. We include 95% and 90% confidence intervals around the estimated coefficients. The dependent variable is F_{ST} the measure of between-group cultural diversity in the region r at time t. The independent variable m_{rt} is the share of foreign-born in the total 2004 population. We control for the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the total population at the regional-level. All estimates include regional and year fixed effects. Sources: Authors' elaboration on ESS and EU-LFS data (2004-2018).

E.2 Concentration vs. Polarization of Values Among Natives

The reduction in cultural diversity within the native population could be interpreted as either a more concentrated distribution of values (i.e., unimodal distribution) or, instead, a polarization of values (i.e., bimodal distribution). To disentangle between these two interpretations of the estimated effect, we push our analysis further by examining the impact of immigration on a polarization index computed on the native population only. Table E-1 shows that immigration does not contribute to a rise of a bimodal distribution of values among natives. Moreover, we explore the effect of immigration on within-group (CF^W) and between-group (F_{ST}) measures computed for the native population across various characteristics, including education, urbanicity, gender, and religiosity. Indeed, suppose natives react differently to the arrival of immigrants, inducing a polarization of their attitudes, then heterogeneity should be rooted in their sociodemographic characteristics, associated with various reactions to immigrants in the literature. Results reported in Col. (2) to (5) of Table E-1 indicate a decrease in within-native group heterogeneity, while Col. (6) to (9) reveal no effect on between-native groups heterogeneity. Therefore, these results support neither the interpretation of the effect of immigration aligned with a rise of a bimodal and polarized distribution of values among natives, nor the rise of salience of certain sociodemographic identity traits (Gennaioli and Tabellini, 2023). Thus, the decrease in cultural diversity within the native population triggered by immigration has to be interpreted as a more concentrated rather than a polarized distribution of values.

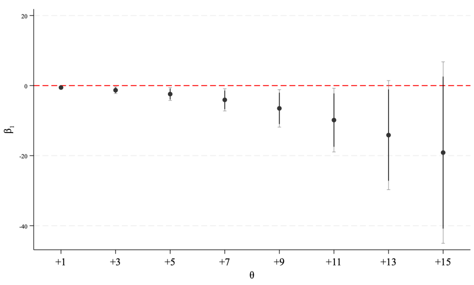
Table E-1: Polarization among Natives.

		Within-native heterogeneity				Between-native heterogeneity			
	(1) Polarization	(2) Education	(3) Urbanicity	(4) Gender	(5) Religion	(6) Education	(7) Urbanicity	(8) Gender	(9) Religion
m_{rt}	0.071	-0.214***	-0.198***	-0.210***	-0.215***	0.017	-0.009	0.012	0.024
	(0.051)	(0.067)	(0.060)	(0.076)	(0.074)	(0.021)	(0.027)	(0.024)	(0.022)
Observations	1,235	1,235	1,235	1,235	1,235	1,235	1,235	1,235	1,235
Mean Cultural Index	0.671	0.716	0.719	0.717	0.711	0.017	0.013	0.015	0.024
Mean Immig. Share	0.101	0.101	0.101	0.101	0.101	0.101	0.101	0.101	0.101
First-stage	1.292	1.292	1.292	1.292	1.292	1.292	1.292	1.292	1.292
KP F-Test	48.568	48.568	48.568	48.568	48.568	48.568	48.568	48.568	48.568

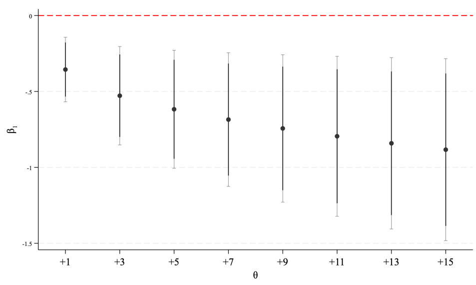
Notes: *** p <0.01, ** p <0.05, * p < 0.1. Standard errors in parentheses are clustered at the regional level. The dependent variables are the polarization index computed among natives in Col. (1), the within-native groups cultural diversity in Col. (2) to (4), and the between-native groups cultural diversity in Col. (5) to (7). The independent variable m_{rt} is the share of foreign-born in the total 2004 population. We control for the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the resident population. All estimates include regional and year fixed effects.
Source: Authors' elaboration on ESS and EU-LFS data (2004-2018).

E.3 Greenberg Index Applied to Immigration

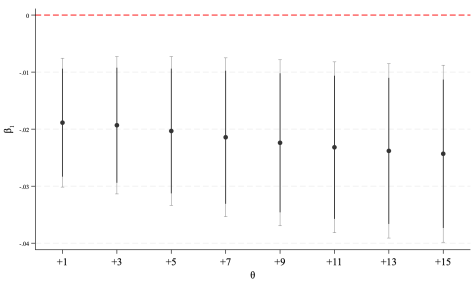
Figure E-2: Greenberg index Applied to Immigration



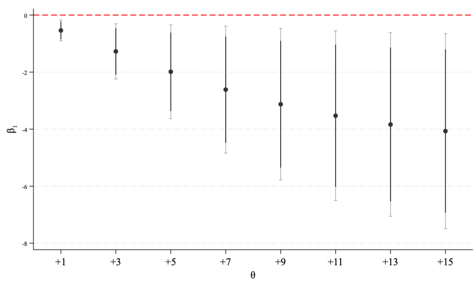
(a) Geodesic



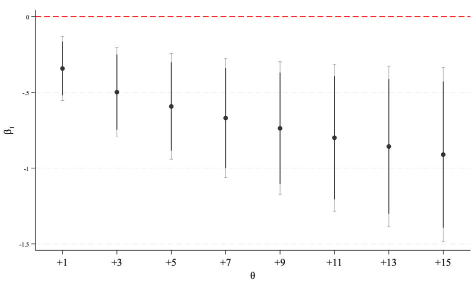
(b) Cultural



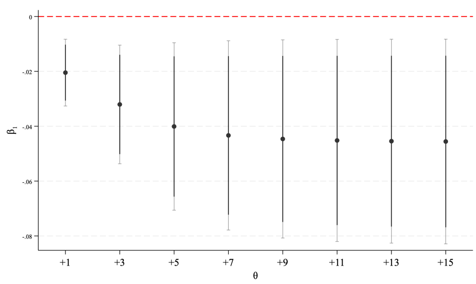
(c) Economic



(d) Religious



(e) Linguistic



(f) Genetic

Notes: These figures plot the marginal impact of a Greenberg index applied to immigration on natives' cultural diversity as defined in Equation (20). The dependent variable is CF_{rt} the measure of cultural diversity in the region r at time t. All estimates include a vector of controls, with the log of population density, the log of GDP per capita, the unemployment rate, the share of high-skilled in the total population, and regional and year fixed effects. Standard errors are clustered at the regional level. We include 95% confidence intervals around the estimated coefficients.

Source: Authors' elaboration on ESS and EU-LFS data (2004-2018).

F Convergence: To Whom and Toward Which Cultural Values?

Table F-1: Cultural distance between natives and a given reference group - Robustness Checks

	(1) Overall	(2) High Skill	(3) Low Skill	(4) Female	(5) Male	(6) Urban	(7) Rural	(8) Religious	(9) Not religious
Panel A: Reference Group = Average Native (in 2004); Sample: No 2004									
m_{rt}	–0.115**	–0.193***	–0.097**	–0.118**	–0.112**	–0.103**	–0.104	–0.098**	–0.052
	(0.050)	(0.059)	(0.049)	(0.056)	(0.046)	(0.046)	(0.067)	(0.044)	(0.043)
Observations	235,738	235,738	235,738	235,738	235,738	235,738	235,738	235,738	235,738
Regions	175	175	175	175	175	175	175	175	175
KP F-Test	40.077	40.077	40.077	40.077	40.077	40.077	40.077	40.077	40.077
Panel B: Reference Group = Average Migrant (in 2004); Sample: No 2004									
m_{rt}	0.046	–0.008	0.083**	0.057	0.079**	0.085**	0.033	0.066	0.051
	(0.043)	(0.060)	(0.039)	(0.055)	(0.037)	(0.042)	(0.052)	(0.045)	(0.031)
Observations	235,738	235,738	235,738	235,738	235,738	235,738	235,738	235,738	235,738
Regions	175	175	175	175	175	175	175	175	175
KP F-Test	40.077	40.077	40.077	40.077	40.077	40.077	40.077	40.077	40.077
Panel C: Reference Group = Average Native; Sample: All Years									
m_{rt}	–0.107**	–0.193***	–0.066	–0.115**	–0.102**	–0.118**	–0.081	–0.080**	–0.071*
	(0.045)	(0.063)	(0.042)	(0.050)	(0.041)	(0.048)	(0.052)	(0.041)	(0.037)
Observations	265,278	265,278	265,278	265,278	265,278	265,278	265,278	265,278	265,278
Regions	175	175	175	175	175	175	175	175	175
KP F-Test	26.659	26.659	26.659	26.659	26.659	26.659	26.659	26.659	26.659
Panel D: Reference Group = Average Migrant; Sample: All Years									
m_{rt}	–0.009	–0.047	0.018	–0.023	–0.004	0.011	–0.008	0.005	0.023
	(0.045)	(0.059)	(0.045)	(0.052)	(0.042)	(0.039)	(0.052)	(0.046)	(0.034)
Observations	265,278	265,278	265,278	265,278	265,278	265,278	265,278	265,278	265,278
Regions	175	175	175	175	175	175	175	175	175
KP F-Test	26.659	26.659	26.659	26.659	26.659	26.659	26.659	26.659	26.659

Notes: *** p <0.01, ** p <0.05, * p < 0.1. The dependent variable is the cultural distance between each native and a reference group of natives (Panels A and C) or immigrants (Panels B and D) computed either as the average immigrant with certain characteristics in 2004 (Panels A and B) or over the entire period of analysis (Panels C and D). Panels A and B exclude the 2004 wave from the analysis. The independent variable is the share of foreign-born in the total 2004 population. All estimates include a vector of regional controls, including the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the total population. All estimates include a vector of individual control with age, age squared, education, gender, children, urbanicity, marital, and employment status. All estimates include regional and year fixed effects. Estimates are weighted using individual weights. Standard errors in parentheses are clustered at the regional level.

Sources: Authors' elaboration on ESS and EU-LFS data (2004-2018).

Table F-2: Cultural distance between natives and a given reference group of immigrants

	Panel A: Reference Group = Average Immigrant (in 2004) Sample: All Years			
	(1) Africa	(2) Asia	(3) Europe	(4) Other
m_{rt}	0.111**	0.045*	0.046	0.030
	(0.056)	(0.025)	(0.039)	(0.039)
Observations	265,278	265,278	265,278	265,278
Regions	175	175	175	175
KP F-Test	26.659	26.659	26.659	26.659
	Panel B: Reference Group = Average Immigrant Sample: All Years			
	(1) Africa	(2) Asia	(3) Europe	(4) Other
m_{rt}	0.055	0.047	–0.006	–0.066
	(0.062)	(0.040)	(0.050)	(0.052)
Observations	265,278	265,278	265,278	265,278
Regions	175	175	175	175
KP F-Test	26.659	26.659	26.659	26.659

Notes: *** p <0.01, ** p <0.05, * p < 0.1. The dependent variable is the cultural distance between each native and a reference group of immigrants computed either as the average immigrant with a certain origin in 2004 (Panel A) or over the entire period of analysis (Panel B). The independent variable is the share of foreign-born in the total 2004 population. All estimates include a vector of regional controls, including the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the total population. All estimates include a vector of individual control with age, age squared, education, gender, children, urbanicity, marital, and employment status. All estimates include regional and year fixed effects. Estimates are weighted using individual weights. Standard errors in parentheses are clustered at the regional level.

Sources: Authors' elaboration on ESS and EU-LFS data (2004-2018).