Journal of Comparative Economics xxx (xxxx) xxx

Contents lists available at ScienceDirect

Journal of Comparative Economics

journal homepage: www.elsevier.com/locate/jce

Societal diversity, group identities and their implications for tax morale☆

Fabian ten Kate [a], Mariko J. Klasing [b], Petros Milionis [b,*]

[a] International Business School, Hanze University of Applied Sciences, Zernikeplein 7, 9747 AS Groningen, The Netherlands
[b] Faculty of Economics & Business, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands

ARTICLE INFO

JEL classification:
H26
K42
O17
R50
Z13

Keywords:
Tax evasion
Tax morale
Social cleavages
Diversity
Identity
Cultural values

ABSTRACT

We study how the tax morale of individuals is influenced by societal diversity in their place of residence. Using data from the World Value Survey, we compare the effects that diversity has on self-reported measures of tax morale at the national, sub-national and individual level. We show first that, both across countries and within countries across sub-national regions, greater diversity is associated with lower average levels of tax morale. We then document that within countries and regions tax morale is lower among individuals who are less similar to others and this effect operates more strongly in places characterized by higher levels of diversity. This pattern applies to diversity in terms of different social cleavages, including income, ethnicity, language or religion, but is particularly pronounced when it comes to diversity in terms of cultural values. This suggests that social identification is important for how people perceive their responsibility of paying taxes.

1. Introduction

Tax evasion has been the subject of an extensive literature in public economics. This literature started out by applying ideas from decision-making under uncertainty and economics of deviant behavior to the decision of how much income to report to the tax authorities (Allingham and Sandmo, 1972; Srinivasan, 1973). In these early studies on tax evasion, individual taxpayers were assumed to optimally choose what fraction of their income they should be officially reporting by balancing out their own costs and benefits. Beyond these private pecuniary considerations, however, there is increasing evidence that the decisions of individuals to evade or not to evade taxes is influenced also by non-pecuniary considerations, which are referred to collectively in the literature with the term tax morale (Feld and Frey, 2007; Luttmer and Singhal, 2014; Alm, 2019).

To understand these non-pecuniary motives for tax compliance, many authors in the literature have shifted their attention to the broader socioeconomic environment in which individual taxpayers operate. This is because taxpayers should recognize that the income taken away from them constitutes their own contribution towards the various public goods and social support programs provided by the government (Cowell and Gordon, 1988; Slemrod and Yitzhaki, 2002). Moreover, the exact way in which these tax contributions are valued should vary across individual taxpayers, depending on how much they identify and sympathize with others

☆ This paper has benefited from useful comments and suggestions by Sjoerd Beugelsdijk, Klaus Desmet, Oded Galor, Jakob de Haan, Anastasia Litina, Pierre-Guillaume Meon, the editor Timur Kuran and two anonymous reviewers. The authors would also like to thank conference participants at the 2018 ASREC Europe Conference, the 2018 CREA Workshop, the 2019 EPCS Conference, the 2020 Econometric Society World Congress and the 2021 ERSA Congress for fruitful discussions. Part of the results included in the paper were previously circulated as a working paper with the title "Diversity, Identity and Tax Morale".

* Corresponding author.
E-mail addresses: f.ten.kate@pl.hanze.nl (Fabian ten Kate), m.j.klasing@rug.nl (M.J. Klasing), p.milionis@rug.nl (P. Milionis).

https://doi.org/10.1016/j.jce.2023.04.005
Received 2 August 2021; Received in revised form 14 April 2023; Accepted 25 April 2023

outside their own household. As several recent studies have documented, individuals strongly characterized by such sentiments are more likely to pay their taxes (Christian and Alm, 2014; Dulleck et al., 2016; Dwenger et al., 2016; Hallsworth et al., 2017). Similarly, it has been shown that individuals tend to be more supportive of public goods and redistribution programs by governments if they feel that this will benefit people they identify with (Luttmer, 2001; Dahlberg et al., 2012).

These findings suggest that tax compliant behavior should be influenced by both the overall level of societal diversity and the own identity considerations of each individual taxpayer. This dual relationship is the focus of our analysis in this paper. Specifically, we study how the inclination of individuals to pay their taxes depends on the overall level of diversity as well as on how similar they are to others in society. Individuals who are more similar to others are likely to place more weight on how their tax contributions benefit broadly the rest of society rather than people in their own narrow circle. These individuals would then be more likely to pay their taxes and this should lead to higher levels of tax compliance overall in places where individuals are more similar.

At the same time, the extensive literature on social identity has demonstrated that individuals typically perceive their social environment through their own eyes. They associate more strongly with people who are like them and their considerations about others relate to their degree of social identification. Moreover, in societies that are more diverse the social identities of individuals are likely to become more salient and individuals will identify themselves more closely with their own social group, rather than with society at large. This suggests that in a more diverse social environment the group identities of individuals should exert a stronger influence on their tax morale.

In order to test these hypotheses, we use data from the World Values Survey, a commonly-used source to measure tax morale. Building on previous work in the literature, we quantify tax morale as the extent to which survey respondents find the act of cheating on taxes morally justifiable. From these responses we can measure both the tax morale of a given individual and the average level of tax morale in a given country or region. We relate these measures to a number of indicators of how similar survey respondents are in terms of various social cleavages. These include their income, their ethnicity, their main language, their religion and their cultural values. This allows us to assess whether similarity in terms of these cleavages, which should enhance social identification, tends to foster tax morale.

Our findings provide strong support for this hypothesis. Comparing first tax morale across countries and sub-national regions within the same country, we find that tax morale is weaker in places that are more diverse along each of the five aforementioned social cleavages. We then document that individual respondents who are less similar to others living in the same country or region, in terms of these cleavages, exhibit lower tax morale. Furthermore, we show that these two effects complement each other: the level of tax morale of individuals becomes more sensitive to their degree of similarity with others along each cleavage, as the overall level of societal diversity in terms of that cleavage increases.

Looking at the magnitudes of these effects, we find that they vary across cleavages and the effect is strongest for similarity in terms of cultural values. We further show that this effect of similarity in terms of cultural values also operates across individuals who share the same ethnicity, language or religion and it dominates the effects of ethnic, linguistic or religious similarity. This variation in cultural values across individuals within specific social groups can be thought as reflecting varying degrees of identification of individuals with the group. It further suggests that what matters is not necessarily whether individuals belong to the same ethnic, linguistic or religious group as others around them. It is how strongly individuals identify themselves with these groups, as reflected in their shared values.

The above results hold conditionally on a variety of individual and regional characteristics, including standard individual and regional determinants of tax morale, such as social trust, preferences for equality or the quality of governance. Moreover, the results hold conditional on various combinations of fixed effects applied at the country, region and survey-wave level to account for a wide range of potential sources of differences in tax morale. We further demonstrate that our results are not sensitive to the way in which we measure tax morale or the way in which we quantify the degree of similarity between individuals. Our results also do not hinge on the employed econometric specification and they hold even when we employ an instrumental variable strategy to account for the potential endogeneity of the employed similarity measures.

These findings are consistent with previous work in the literature that has shown that societal diversity is an important factor influencing the willingness of individuals to share resources with others. While most of the literature has studied the implications of this effect for income redistribution or public goods provision (Alesina et al., 1999; Desmet et al., 2009; Gruendler and Koellner, 2020), in this paper we look more carefully at how it influences the proclivity of individuals to pay their taxes. In this respect our analysis, as we explain in more detail in the next section, is most closely related with those of Li (2010) and Alm et al. (2016), who provide evidence that higher levels of diversity weaken tax morale and result in lower rates of tax filing respectively. Yet, a key novelty of our analysis is demonstrating that it is not just the overall level of diversity that affects individual tax morale. It is also how individuals identify themselves with others in society that matters.

In this respect, our empirical analysis complements several theoretical papers that have modeled how tax compliance can be affected by social interactions and coordination (Myles and Naylor, 1996; Fortin et al., 2007; Traxler, 2010; Litina and Palivos, 2016). These horizontal types of relationships between citizens have been shown to matter above and beyond the vertical relationships that citizens have with the state, which has been often emphasized in the literature (Frey and Torgler, 2007; Hug and Spoerri, 2011; Filippin et al., 2013; Besley, 2020). As societies become more diverse, it is these horizontal relationships that weaken. This may result in lower levels of tax compliance in individuals, even when their vertical relationships with the state remain unchanged.

In addition to that, our analysis shows how the effect of societal diversity on tax morale extends beyond the typical divisions related to ethnicity, language or religion.[1] As our results highlight, differences in cultural values, although less visible, play an

[1] Such divisions have been the focus of an extensive literature in economics and other social sciences. See Alesina and Ferrara (2005), Stichnoth and Straeten (2013), and Ashraf and Galor (2018) for surveys.

important role in this context. This suggests that social identification operates in more complex ways than is commonly understood in the literature.[2] Greater dispersion of cultural values can also lead to the loosening of social ties and undermine the sense of civic duty of individuals to pay their taxes.

To establish these results we proceed as follows. In the next section we discuss the relationship between diversity, identity and tax morale and explain how our contribution relates to previous work in the literature. In Section 3 we describe our empirical strategy and discuss how we measure our key variables of interest. In Section 4 we present our main regression results. In Section 5 we report a series of robustness checks on our main results. In Section 6 we offer some concluding remarks on our findings and their implications. The paper's online appendix contains further information about the data and a series of additional regression results, which are not presented in the main part of the paper to conserve space.

2. Contribution and relation to the literature

As we are not the first to study the relationship between societal diversity and tax morale, it is worthwhile to provide some background regarding the nature of this relationship and how it has been approached by previous work in the literature. Some early empirical evidence regarding this relationship was provided by LaPorta et al. (1999) and Alesina et al. (2003). They document that tax morale across countries is negatively associated with diversity along ethnic, linguistic or religious lines. More recent work by Li (2010) and Alm et al. (2016) has studied this relationship more systematically based on regional and individual level data. In particular, Li (2010) shows that self-reported tax morale tends to be higher among individuals who belong to the majority groups in their countries in terms of their ethnicity, language or religion, while Alm et al. (2016) show that actual tax filing rates in the United States are lower in counties that are racially more diverse.

Although any empirical analysis about tax morale is subject to the caveat that morale is not something directly observable, it should be noted that similar findings have been obtained for income redistribution and public goods provision, which can be measured more accurately. Both across and within countries, there is extensive evidence that income redistribution and public goods provision are lower in places where societal diversity is higher (Stichnoth and Straeten, 2013). This pattern applies to different dimensions of diversity, but it is more clearly visible when diversity is driven by deep-rooted social cleavages (Desmet et al., 2009, 2017; Gershman and Rivera, 2018).

To explain the negative effect of societal diversity on tax morale, income redistribution and public goods provision, there are several common explanations that have been proposed by different authors in the literature. The first is that individuals from different social groups have different preferences regarding the role of the government (Alesina et al., 1999; Beach and Jones, 2017). Hence, an increasing number of social groups creates problems of preference aggregation, which results in lower levels of support for taxation and redistribution. A second yet related mechanism invokes income inequality across social groups, which gives rise to differential levels of support for income redistribution and other government programs (Banerjee et al., 2005; Baldwin and Huber, 2010). A third possibility is that in diverse societies segregation among social groups limits their access to the political process through which government decisions about redistribution and public goods provision are made (Habyarimana et al., 2007; Desmet et al., 2020). Finally, there is also the possibility that individuals' willingness to share resources is not the same across social groups as within these groups (Vigdor, 2004).

What these four explanations have in common is that they all indicate the aggregation of individual preferences as the main problem for income redistribution and public goods provision. Individuals are assumed to have clear preferences about the outcomes that they prefer. Yet, these preferences differ across individuals from different social groups. This creates a challenge for policy-makers who have to deal with a problem of preference aggregation in an environment of group heterogeneity. This is likely to lead to an inefficient outcome with which many individuals will not be satisfied and from which they will try to opt out by evading their own tax contribution.

In addition to these four explanations, as we discussed in the introduction, there is growing evidence that the willingness of individuals to pay their taxes depends on the extent to which they identify and sympathize with others (Christian and Alm, 2014; Dulleck et al., 2016; Dwenger et al., 2016; Hallsworth et al., 2017). These individual identity considerations imply that, beyond the preference aggregation problem, there is also a problem of preference interdependence (Luttmer, 2001). The tax morale of a given individual depends positively on how embedded they are in their social environment and how strong their concerns are about others around them.

So far Li (2010) is the only paper that has recognized and studied the relevance of individual identity considerations for tax morale. Using individual level data from the WVS-EVS, as we do, Li constructs different measures that reflect the social identity of individuals and relates those to self-expressed tax morale. She finds that tax morale tends to be higher for individuals who belong to the majority group in their country in terms of their ethnicity, language or religion. She also finds higher levels of tax morale among individuals who have stronger feelings of citizenship for their country. Based on these patterns she concludes that individual identity considerations are the main reason for the negative relationship between societal diversity and tax morale, observed at the aggregate level.

Our analysis goes in the same direction as that of Li, exploring how tax morale is affected by the social identities of individuals. It differs from Li, though, by taking a broader view of the social identity of individuals which transcends beyond their ethnic, linguistic or religious background. Instead it assumes that an individual's identity can be derived from participation in a wide range of social

[2] See Akerlof and Kranton (2000), Shayo (2009), or Benjamin et al. (2010) for common ways of modeling social identity.

groups and involvement in a variety of activities.[3] Moreover, there are countries where individuals tend to identify and associate themselves strongly with the social groups where they belong and countries where such associations are weaker (Gelfand et al., 2011). To capture the varying degree of social identification of individuals, we follow a recent line of research by Beugelsdijk and Klasing (2016), Desmet et al. (2017), and Beugelsdijk et al. (2019) that considers diversity in cultural values and quantifies that in a similar way as ethnolinguistic diversity. As these authors have shown, diversity in cultural values is a distinct dimension of diversity, which only partially overlaps with ethnic, linguistic or religious diversity.

Furthermore, when quantifying the social identity of an individual we need to do that based on an appropriate societal reference point. This could be a geographic entity or a social group that provides the natural reference point for individuals to make interpersonal comparisons. While a common choice is to take the country of residence as this reference point, this may not be the best choice in our context. This is because individuals tend to interact more with people in their immediate environment rather than the entire country. Hence, their perception of society is likely to reflect the societal composition in their region of residence and their willingness to pay taxes will be strongly influenced by conditions in that region. Beyond that, when it comes to taxes, one has to keep in mind that not all of them are levied by the national government. There are also regional and local taxes, which are used to finance local public goods and redistribution programs, and which individuals may perceive differently compared to national taxes. Assuming that the willingness of individuals to pay each of these taxes depends on their ties with each of these geographic entities, this suggests that on average the reference point for general tax morale should be a geographical entity below the national level.

With these points in mind, our aim in this paper is to analyze how tax morale is affected by both the overall level of diversity in society as well as the own identity considerations of each individual taxpayer. For this analysis we consider measures of societal diversity and individual identity based on ethnic, linguistic or religious cleavages already considered by Li (2010), income differences emphasized by Baldwin and Huber (2010) and also differences in a wide range of cultural values. Doing so we can quantify how similar each individual is relative to others in society and compare that with the overall level of diversity in society along each of the aforementioned cleavages. This allows us to capture social identification broadly and also assess the relevance of all these social cleavages, which is something that has not been done in the literature so far. Moreover, we conduct our analysis at different levels of aggregation, comparing the relationship across countries, across regions within countries and across individuals within regions. This way we can explore the relevance of different geographic entities as a reference point for making interpersonal comparisons. In the next section we explain in more detail the approach that we follow to construct our explanatory variables and to estimate these effects.

3. Measurement and estimation approach

As our aim is to understand how tax morale is influenced by various social cleavages and the extent of societal diversity resulting from those, we study this relationship both at the individual and at the aggregate level. To analyze the importance of these cleavages we need to consider a reference point for our comparisons. We start by taking this reference point to be the country where individuals reside, as it is common in the literature. We then proceed beyond that to make comparisons across individuals within sub-national regions, which are typically based on first-level administrative divisions within countries. This is because even within countries there are important differences in the socio-economic environment across regions and part of the taxes that individuals pay are going to the sub-national rather than the central government. In addition to that, individuals are likely to have a more accurate perception of the socio-economic environment in their region rather than in their country as a whole. Hence, taking a geographic unit of reference below the country level may allow for more meaningful comparisons.[4]

As a first test for our hypothesis, we estimate Eq. (1) that links the average level of tax morale in a given country c measured at a given point in time t, $TaxMorale_{c,t}$, to measures of similarity in the population of that country in terms of a given attribute, $Similarity_{c,t}$. We then change the reference point and estimate Eq. (2) that links the average level of tax morale in a given sub-national region r of country c, $TaxMorale_{r,c,t}$, to measures of similarity in the population of that region in terms of a given attribute, $Similarity_{r,c,t}$. Then we move to individual-level data and estimate Eq. (3) to see how the tax morale of individual i, $TaxMorale_{i,r,c,t}$, depends on the degree of similarity of that individual with others residing in the same region, $Similarity_{i,r,c,t}$.

$$TaxMorale_{c,t} = \alpha_t + \beta \cdot Similarity_{c,t} + \gamma' X_{c,t} + \varepsilon_{c,t}. \tag{1}$$

$$TaxMorale_{r,c,t} = \alpha_{c,t} + \beta \cdot Similarity_{r,c,t} + \gamma' X_{r,c,t} + \varepsilon_{r,c,t}. \tag{2}$$

$$TaxMorale_{i,r,c,t} = \alpha_{r,c,t} + \beta \cdot Similarity_{i,r,c,t} + \gamma' X_{i,r,c,t} + \varepsilon_{i,r,c,t}. \tag{3}$$

In all equations the matrix X denotes other determinants of tax morale and the vector α a set of included fixed effects that vary across specifications. This is because as we shift our analysis from the country level to the region and the individual level, we also

[3] The notions of social identity and social identification have a long tradition in social psychology as important determinants of individual behavior, going back to the work of Tajfel and Turner (1979). Within this line of research there is an active debate about what degree of involvement in social groups is necessary in order to influence the social identity of individuals. See Hogg (2018) for a recent survey.

[4] As our analysis reveals though, our findings do not hinge on the choice of the geographic unit of reference. Table A3 in the appendix presents some key checks regarding that.

apply country-time and region-time fixed effects, and adjust our controls accordingly. Using a finer set of fixed effects allows us to filter out variation in tax morale across countries and regions, due to laws and institutional differences. This way we can focus on comparisons within countries and regions across individuals that operate largely within the same socio-economic environment.[5]

Beyond analyzing the relationship between tax morale and similarity at different levels based on the above specifications, we further consider alternative ones that include different measures of similarity to assess their relative importance. As part of this exploration, we estimate a series of interaction regressions to assess how the effect of similarity of individuals relates to the overall level of similarity in society. This is because the degree of similarity of a given individual with others may become more salient depending on the overall level of similarity across individuals in a given region, as discussed in Section 2. This is captured by Eq. (4) where the interaction effect of individual and regional similarity is estimated conditional on a set of region-time fixed effects.

$$TaxMorale_{i,r,c,t} = \alpha_{r,c,t} + \beta_1 \cdot Similarity_{i,r,c,t} + \beta_2 \cdot Similarity_{i,r,c,t} \cdot Similarity_{r,c,t} + \delta' X_{i,r,c,t} + \varepsilon_{i,r,c,t}. \quad (4)$$

We estimate the above regressions based on data derived from the integrated longitudinal data set of the World Values Survey and European Values Study (WVS/EVS), which covers all seven survey waves conducted between 1981 and 2020. Throughout our analysis we distinguish between data collected in different survey waves. Hence, the time index t in our regression specifications reflects the survey wave as part of which the data were collected. To capture any variation related to the timing of each survey wave, we always include wave-specific fixed effects. In the subsections below we describe how we measure tax morale and similarity at the individual and aggregate level. In the appendix we provide more information about the data and detailed descriptions of all variables used in our empirical analysis.

3.1. Measuring tax morale

The notion of tax morale refers to an individual's non-pecuniary motives as well as to factors outside the standard expected utility framework that drive tax compliance (Luttmer and Singhal, 2014). It is commonly measured from individual responses to survey questions on the moral justifiability of tax evasion, which have been shown to predict actual tax compliance behavior well (Alm and Torgler, 2006; Cummings et al., 2009; Dulleck et al., 2016; Mare et al., 2020).

The WVS/EVS questionnaire includes a question that asks individuals to indicate on a scale from 1 to 10 whether *"cheating on their taxes is justified"*. This question has been frequently used in empirical studies of tax morale in the literature (Slemrod, 2002; Alm and Torgler, 2006; Frey and Torgler, 2007; Hug and Spoerri, 2011; Besley, 2020). Given the way the question is phrased, an answer of 1 indicates that a respondent finds tax evasion never justified and an answer of 10 as always justified. Since our focus is on measuring tax morale, we invert the response scale, so that higher values indicate individuals with higher levels of tax morale. Averaging these values across individuals residing in the same country or region we obtain our corresponding aggregate measures of tax morale.[6]

3.2. Individual and aggregate measures of similarity

To measure how similar or dissimilar individuals are relative to others in society, we consider the extent of similarity across individual respondents from a given country or region interviewed as part of the same wave of the WVS/EVS survey in terms of five key social cleavages. These cleavages consist of their household income, their ethnicity, their main language, their religion and a range of cultural values. Using any of these cleavages a, we can partition the set of individuals from a given geographic entity r, which could be a country or a region, into different groups $1, \ldots, G^a$. For each of these groups g let s_r^g denote the share of the group in that entity. We can then quantify the aggregate level of similarity in terms of cleavage a based on the index:

$$Similarity_r^a = \sum_{g=1}^{G^a} (s_r^g)^2. \quad (5)$$

This index ranges from 0 to 1 and captures the probability that two randomly selected individuals from geographic entity r belong to the same group in terms of cleavage a. It corresponds to the Herfindahl–Hirschman index of concentration and is the opposite of the fractionalization index that has been widely used in the literature to measure ethno-liguistic diversity (Alesina et al., 2003; Desmet et al., 2009). In contrast to this line of research that focuses on aggregate measures, we also want to measure the extent to which a specific individual is similar in terms of cleavage a to other residents of entity r. This can be done following the same logic. Suppose that individual i belongs to group g whose relative size is given by the population share s_r^g. Since this share also reflects the probability of that individual coming across another person from the same entity falling into the same group along cleavage a, we can define the individual similarity index as:

$$Similarity_{i,r}^a = s_r^g. \quad (6)$$

[5] As the included set of fixed effects cannot account for all possible confounding factors that could influence the observed levels of tax morale, as part of our analysis we also consider an estimation strategy based on instrumental variables, which we describe in more detail in Section 5.1.

[6] The individual observations can also be weighted with the reported weights in the survey to make these averages more representative, which is something that we do as part of our robustness checks.

Using expressions (5) and (6) we compute measures of similarity at the country, region and individual level for each of the aforementioned social cleavages.[7] When doing so based on income we split individual respondents into 10 income groups using the country-specific income classes reported in the survey. When doing so based on ethnicity, language or religion we consider in each country all groups that have a population share of 5 percent or higher.

To measure similarity in terms of cultural values, we build on the approach of Beugelsdijk and Klasing (2016), Desmet et al. (2017) and Beugelsdijk et al. (2019), who group individuals based on the answers that they give to different value questions in the WVS/EVS survey. For this purpose we use the answers to 96 different questions listed in the appendix, which are in the core part of the WVS/EVS questionnaire. These are all questions that can be answered on a rating scale and indicate the importance that individuals attribute to a particular value. For each of these questions we measure similarity at the aggregate and individual level based on expressions (5) and (6), and then we average the resulting scores across all 96 questions to obtain broad measures of value similarity.

Throughout our analysis we focus on the notion of similarity between individuals and think of diversity as the opposite of that.[8] We measure similarity mostly using the above described similarity measures given their simplicity and straightforward interpretation. For each of these measures we also consider alternatives based on different metrics, different weighting of the data to reflect the actual population composition, and in the case of value similarity different sets of value questions. As none of these alternative measures make any difference for our results, we do not consider them as part of our main analysis, but we discuss them as part of our robustness checks.

4. Regressions results

4.1. Similarity and tax morale at the country level

We start our analysis at the country level in order to permit comparisons with previous work in the literature. Beyond our main explanatory variable, our regression specification, Eq. (1), includes a set of control variables which capture differences between countries in terms of their geographic characteristics and level of economic development. These variables, whose coefficient estimates are not reported for brevity, are nighttime luminosity per capita, population density, distance from the coast and centrality of its capital location. Our specification also includes a set of wave-specific continental fixed effects to capture additional unobserved variation in the level of tax morale across countries and time. To facilitate the interpretation of the estimated coefficients, all variables are standardized.

The estimation results are presented in Table 1, which is organized in three panels. Panel A presents the effects of each of our five measures of similarity in terms of income, ethnicity, language, religion and cultural values estimated separately based on all available observations. Panel B presents the same effects estimated based on a common sample. Finally, panel C presents the estimates from a horse-race regression where all five similarity measures are included together. Although the estimated coefficients are not always statistically significant, in general we see that tax morale tends to be higher in countries where the population is more similar in terms of these social cleavages.

This is broadly consistent with the findings of LaPorta et al. (1999), Alesina et al. (2003) and Li (2010), who have documented that ethnic, linguistic and religious diversity are all negatively related to tax morale. Looking at the magnitudes of the estimated effects, though, we see that average tax morale is most strongly related to similarity in terms of cultural values. For this measure we not only obtain a much larger estimated effect, but this effect also clearly dominates the others, as we see in the horse-race regression of panel C.[9] Even if we do not focus on the horse-race regression due to the fact that the different similarity measures are highly correlated, the estimates of panels A and B suggest that a one standard deviation increase in value similarity raises tax morale by half a standard deviation, which is a sizable effect.

4.2. Similarity and tax morale at the regional level

While the cross-country comparisons presented in Table 1 lend support to our hypotheses, they may also be driven by a variety of other factors that influence tax morale in each country. To insulate our analysis from these factors, we proceed to compare tax morale across regions within the same country. Our estimations in this case are based on Eq. (2), which includes country-wave fixed effects. This allows us to control for variation coming from the institutional framework of each country and other unobserved factors that may influence tax morale across countries and over time. The specification further includes a similar set of control variables (nighttime luminosity per capita, population density, distance from the coast and distance from the country's capital) as the regressions reported in Table 1, but now measured at the regional level.

[7] In a slight abuse of terminology we refer below to the latter index, which reflects the degree of similarity of a given individual, as individual similarity. This is to contrast it with the former index which reflects similarity at the aggregate level, be that a country, region or another reference unit.

[8] Other authors in the literature who have used such measures sometimes provide a different interpretation based on sympathy/antipathy (Desmet et al., 2017) or social connectedness (Stuart and Taylor, 2020). These varying interpretations are natural, as the notion of societal diversity can be understood in different ways, and they can also be applied to our analysis.

[9] In Table A1 in the appendix we present some additional regression results to document the robustness of this pattern.

Table 1
Similarity and tax morale at the country level.

	(1)	(2)	(3)	(4)	(5)
Dep. variable:	Country tax morale				
Country Sim. wrt	Income	Ethnicity	Language	Religion	Cult. values
Panel A - Effects Estimated from Separate Regressions using all Available Observations					
Coefficient	0.007	0.133	0.196**	0.026	0.467***
Std. Error	(0.057)	(0.102)	(0.078)	(0.053)	(0.063)
Observations	345	217	236	399	404
Countries	110	96	99	110	111
Panel B - Effects Estimated from Separate Regressions based on Common Sample					
Coefficient	−0.049	0.068	0.107	0.102	0.519***
Std. Error	(0.098)	(0.109)	(0.088)	(0.093)	(0.092)
Observations	188	188	188	188	188
Countries	92	92	92	92	92
Panel C - Effects Estimated Jointly in the Same Regression					
Coefficient	−0.055	0.178	−0.078	−0.154	0.659***
Std. Error	(0.092)	(0.120)	(0.091)	(0.102)	(0.114)
Observations			188		
Countries			92		

Notes: This table compares the effect of various measures of similarity on tax morale at the country level. All estimates are based on ordinary least squares (OLS) with fixed effects imposed at the continent-wave level. The estimations control without reporting the estimates for the following variables: luminosity per capita, population density, distance to sea, centrality of capital location. All variables are standardized to ensure comparability. Heteroskedasticity robust standard errors clustered at the continent-wave level are reported in parenthesis. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table 2
Similarity and tax morale at the region level.

	(1)	(2)	(3)	(4)	(5)
Dep. variable:	Regional tax morale				
Region Sim. wrt	Income	Ethnicity	Language	Religion	Cult. values
Panel A - Effects Estimated from Separate Regressions using all Available Observations					
Coefficient	0.020	0.018	0.047*	0.037	0.229***
Std. Error	(0.021)	(0.030)	(0.026)	(0.025)	(0.032)
Observations	4070	2770	3117	4832	4851
Regions	2069	1671	1853	2249	2255
Panel B - Effects Estimated from Separate Regressions based on Common Sample					
Coefficient	0.002	0.013	0.035	0.046	0.246***
Std. Error	(0.027)	(0.030)	(0.025)	(0.034)	(0.045)
Observations	2599	2599	2599	2599	2599
Regions	1661	1661	1661	1661	1661
Panel C - Effects Estimated Jointly in the Same Regression					
Coefficient	−0.137***	−0.027	0.015	−0.024	0.346***
Std. Error	(0.028)	(0.031)	(0.026)	(0.032)	(0.040)
Observations			2599		
Regions			1661		

Notes: This table compares the effect of various measures of similarity on tax morale at the region level. All estimates are based on ordinary least squares (OLS) with fixed effects imposed at the country-wave level. The estimations control without reporting the estimates for the following variables: luminosity per capita, population density, distance to sea, distance to capital. All variables are standardized to ensure comparability. Heteroskedasticity robust standard errors clustered at the country-wave level are reported in parenthesis. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

The reported estimates are presented in Table 2, which follows the same structure as Table 1, and are based on standardized versions of all explanatory variables. Panel A presents the effects of each of our five similarity measures estimated separately from all available observations, panel B presents the same effects estimated based on a common sample and panel C presents the effects estimated jointly in a horse-race regression. The general pattern that we see is very similar to what we saw across countries. All five measures of similarity are positively related with tax morale, although some effects are weak. The obtained coefficients in this case are smaller than those in Table 1. This is not surprising, as the country-wave fixed effects filter out a substantial part of the observed variation in tax morale. Nevertheless, we still see that similarity in terms of cultural values is the strongest predictor of

Table 3
Regional similarity and individual tax morale.

	(1)	(2)	(3)	(4)	(5)
Dep. variable:	Individual tax morale				
Region Sim. wrt	Income	Ethnicity	Language	Religion	Cult. values
Panel A - Effects Estimated from Separate Regressions using all Available Observations					
Coefficient	0.015**	0.009	0.029***	0.022*	0.164***
Std. Error	(0.007)	(0.012)	(0.011)	(0.012)	(0.016)
Observations	369731	270649	308037	464967	467147
Regions	2035	1665	1853	2228	2234
Panel B - Effects Estimated from Separate Regressions based on Common Sample					
Coefficient	0.013	0.007	0.032**	0.022	0.168***
Std. Error	(0.008)	(0.013)	(0.012)	(0.014)	(0.015)
Observations	258827	258827	258827	258827	258827
Regions	1656	1656	1656	1656	1656
Panel C - Effects Estimated Jointly in the Same Regression					
Coefficient	−0.021***	−0.009	0.015	−0.010	0.177***
Std. Error	(0.008)	(0.014)	(0.015)	(0.013)	(0.015)
Observations			258827		
Regions			1656		

Notes: This table compares the effect of various measures of regional similarity on individual tax morale. All estimates are based on ordinary least squares (OLS) with fixed effects imposed at the country-wave level. The estimations control without reporting the estimates for the following regional characteristics: luminosity per capita, population density, distance to capital, distance to sea. They also control for the following individual characteristics: age, gender, marital status, employment status, education level. All variables are standardized to ensure comparability. Heteroskedasticity robust standard errors clustered at the country-wave level are reported in parenthesis. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

differences in tax morale across regions.[10] As the estimates reveal, a one standard deviation increase in that measure of similarity raises tax morale by approximately one quarter of a standard deviation, which is about half of the cross-country effect.

4.3. Similarity and tax morale at the individual level

The effects reported in Tables 1 and 2 pertain to the average level of tax morale in a given country or region. To understand better what is driving these effects, in Tables 3 and 4 we shift our focus to the tax morale of individual survey respondents. In Table 3 we relate this to the same regional measures of similarity as in Table 2, retaining our set of regional controls and country-wave fixed effects. Then in Table 4 we switch to the corresponding measures of similarity at the individual level. Hence, we estimate Eq. (3), which links the level of tax morale of individual respondents with how similar these individuals are to other people in their region of residence in terms of the five cleavages. In this case we include region-wave instead of country-wave fixed effects. These filter out any unobserved differences between regions and waves, and absorb the regional controls. To capture other sources of variation across individuals we include in both tables a set of individual-level controls. These include age, gender, marital status, employment status and education level, all of which have been shown to affect individual tax morale.[11]

The structure of Tables 3 and 4 follows those of Tables 1 and 2. The reported estimates are based on standardized versions of all variables, while the estimates for the controls are not reported for brevity. Panel A presents the effects of each of our five similarity measures estimated separately from all available observations, panel B presents the same effects estimated based on a common sample and panel C presents the effects from a horse-race regression.[12] Looking first at panels A and B of Table 3 we see that all our regional similarity measures apart from ethnic similarity have a clear positive association with individual tax morale. This implies that individuals who reside in regions that are relatively homogeneous in terms of income, language, religion and cultural values, exhibit higher levels of tax morale relative to individuals residing in more diverse regions within the same country. Comparing the magnitudes of the effects and looking at the horse-race regression in panel C, we see again that similarity in terms of cultural values has the strongest effect on individual tax morale with a one standard deviation increase raising tax morale by about 0.17 of a standard deviation.

In Table 4 we obtain similar effects on tax morale with our individual similarity measures. In this case we find a clear positive effect for all five of our measures. Hence, no matter whether we measure similarity in terms of income, ethnicity, language, religion or cultural values, we see that individuals who are more similar to others in their region of residence tend to exhibit higher levels of

[10] In Table A2 in the appendix we document the robustness of this pattern with some additional regressions.

[11] There is an extensive literature on how particular individual characteristics predict the likelihood of tax compliance going back to Clotfelter (1983) and Slemrod (1985). See Alm (2019) for a survey of this literature.

[12] In Tables A4 and A5 in the appendix we provide a series of robustness checks on the specifications of Tables 3 and 4 respectively.

Table 4
Individual similarity and individual tax morale.

	(1)	(2)	(3)	(4)	(5)
Dep. variable:	Individual tax morale				
Ind. Sim. wrt	Income	Ethnicity	Language	Religion	Cult. values
Panel A - Effects Estimated from Separate Regressions using all Available Observations					
Coefficient	0.010***	0.009*	0.008*	0.011***	0.349***
Std. Error	(0.002)	(0.005)	(0.005)	(0.004)	(0.007)
Observations	334061	259433	290612	433580	467147
Regions	2035	1665	1853	2228	2234
Panel B - Effects Estimated from Separate Regressions based on Common Sample					
Coefficient	0.010***	0.010*	0.008	0.019***	0.389***
Std. Error	(0.003)	(0.005)	(0.005)	(0.005)	(0.009)
Observations	229204	229204	229204	229204	229204
Regions	1656	1656	1656	1656	1656
Panel C - Effects Estimated Jointly in the Same Regression					
Coefficient	−0.003	−0.001	−0.002	−0.014***	0.392***
Std. Error	(0.003)	(0.005)	(0.005)	(0.004)	(0.009)
Observations			229204		
Regions			1656		

Notes: This table compares the effect of various measures of individual similarity on individual tax morale. All estimates are based on ordinary least squares (OLS) with fixed effects imposed at the region-wave level. The estimations control without reporting the estimates for the following individual characteristics: age, gender, marital status, employment status, education level. All variables are standardized to ensure comparability. Heteroskedasticity robust standard errors clustered at the region-wave level are reported in parenthesis. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

tax morale.[13] When comparing the magnitudes of the estimated effects, however, we see clear differences across the five measures of similarity. In particular, we find the effect of similarity in terms of values to dwarf the effects of the others in all three panels of Table 4. As the level of similarity in the values of a given individual to others in the region increases by one standard deviation, this raises the tax morale of the individual by 0.37 of a standard deviation. This is more than twice the effect of regional value similarity that we obtained in Table 3 and much larger than that of most other individual-level control variables such as the respondents' gender or education, which have been emphasized in the literature.[14]

Taken together the results shown in Table 4 indicate that individual measures of similarity relate positively to individual tax morale and this is particularly the case for similarity in terms of cultural values. These results are consistent with those in the previous tables, but they also highlight that interpersonal variation is highly relevant. The level of tax morale exhibited by a given individual is not only influenced by how much diversity there is overall in the region. It also hinges on how similar that individual is to others. Individuals who are more similar to others within their region, particularly when it comes to the values that they espouse, tend to be characterized by higher tax morale. This suggests that identity considerations matter and is consistent with the view of social identity advanced by Turner et al. (1987), according to which the extent of altruistic behavior of individuals towards other members of a social group depends on their level of identification with the group.[15]

4.4. Interdependence and interactions across similarity measures

The evidence that we have presented so far highlights how similarity across individuals, particularly in terms of cultural values, is an important determinant of tax morale. Yet, similarity in cultural values is not entirely distinct from similarity in attributes such as ethnicity, language and religion. Given that, in Table 5 we explore their potential interdependence in the following way. In addition to measuring value similarity for individuals relative to others residing in the same region, we take as a reference point their ethnic, linguistic or religious group in the country and measure value similarity within these groups. Using these additional measures of value similarity, we compare the importance of value similarity of individuals with people who share the same ethnicity, language or religion relative to their similarity with people in their region in terms of these three cleavages and relative to their value similarity with people in their region. Doing this three-way comparison, we can assess the importance of similarity along each of these three cleavages relative to values and their interdependence.[16]

[13] In Table A6 in the appendix we demonstrate that this effect is robust when replacing our main individual similarity measure with a simple majority group dummy as in Li (2010).

[14] See, for example, Torgler and Valev (2010), and Rodriguez-Justicia and Theilen (2018).

[15] The view of social identity by Turner et al. (1987) is often referred to in the literature as self-categorization theory in order to distinguish it from the original version of social identity theory proposed by Tajfel and Turner (1979).

[16] Desmet et al. (2017) provide a similar three-way comparison at the country level using an alternative measure of overlap between culture and ethnicity based on the χ^2 index. In Table A8 in the appendix we use the same index to show that the effects of value similarity as well as those of ethnic and linguistic similarity are stronger in cases where the overlap is greater.

Table 5
Interdependence of similarity across different cleavages.

	(1)	(2)	(3)	(4)	(5)
Dep. variable:	Individual tax morale				
Panel A - Value Similarity versus Ethnic Similarity					
Ind. Value Sim. (Ethn. Ref. Point)	0.408***	0.271***	0.408***		0.273***
	(0.009)	(0.020)	(0.009)		(0.020)
Ind. Value Sim. (Reg. Ref. Point)		0.149***		0.390***	0.147***
		(0.020)		(0.009)	(0.020)
Ind. Ethnic Sim. (Reg. Ref. Point)			0.014***	−0.004	0.007
			(0.005)	(0.005)	(0.005)
Observations	260228	260228	260228	260228	260228
Panel B - Value Similarity versus Linguistic Similarity					
Ind. Value Sim. (Ling. Ref. Point)	0.407***	0.273***	0.407***		0.275***
	(0.008)	(0.021)	(0.008)		(0.021)
Ind. Value Sim. (Reg. Ref. Point)		0.144***		0.385***	0.142***
		(0.020)		(0.008)	(0.020)
Ind. Linguistic Sim. (Reg. Ref. Point)			0.010**	−0.005	0.005
			(0.005)	(0.004)	(0.004)
Observations	291407	291407	291407	291407	291407
Panel C - Value Similarity versus Religious Similarity					
Ind. Value Sim. (Relig. Ref. Point)	0.336***	0.245***	0.335***		0.243***
	(0.006)	(0.012)	(0.006)		(0.012)
Ind. Value Sim. (Reg. Ref. Point)		0.115***		0.351***	0.117***
		(0.014)		(0.007)	(0.014)
Ind. Religious Sim. (Reg. Ref. Point)			0.004	−0.020***	−0.004
			(0.003)	(0.003)	(0.003)
Observations	434374	434374	434374	434374	434374

Notes: This table explores the effect on individual tax morale of similarity in terms of cultural values versus similarity in terms of ethnicity, language and religion. All estimates are based on ordinary least squares (OLS) with fixed effects imposed at the region-wave level. The estimations control without reporting the estimates for the following individual characteristics: age, gender, marital status, employment status, education level. All variables are standardized to ensure comparability. Heteroskedasticity robust standard errors clustered at the region-wave level, in line with the fixed effects, are reported in parenthesis. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Panel A of Table 5 presents the results for this comparison between cultural values and ethnicity, while panels B and C present the results of the same comparison for language and religion respectively. In each of the three panels column 1 shows just the effect of value similarity relative to others who share the same ethnicity, language or religion. In all cases this effect is positive and sizable. Column 2 contrasts this effect with that of value similarity when the reference point is the region of residence and finds that both effects matter. This is not surprising given the partial overlap between the two value similarity measures. Columns 3, 4 and 5 in each panel include in the specification our main measures of similarity in terms of ethnicity, language and religion respectively. This is first compared with value similarity relative to others that share this attribute in column 3 and with value similarity relative to others residing in the same region. Then in column 5 we present the estimates for the full three-way comparisons. As we can see in all three panels, the effect of ethnic, linguistic and religious similarity that we saw in Table 4 can to a large extent be accounted for by similarity in cultural values. Hence, it disappears when controlling for value similarity relative to others in their ethnic, linguistic and religious group or their region of residence.

In Tables A7 and A8 in the appendix we provide some additional evidence regarding the overlap of cultural values with ethnicity, language and religion. In particular, Table A7 documents a similar set of results when analyzing jointly the interdependence of culture with ethnicity, language and religion. The results also do not hinge on taking sub-national regions as the geographic unit of reference. In all cases the resulting estimates are consistent with the conclusion emerging from Table 5 that similarity in cultural values can largely account for the effects of ethnic, linguistic and religious similarity. This does not mean that these other notions of similarity do not matter. As shown in Table A8, these other notions of similarity become more relevant when they overlap with value similarity. It does mean, however, that the relevant notions of societal diversity influencing tax morale extend beyond ethnic, linguistic or religious divisions.

In the regressions presented in Tables 4 and 5 we have explored how the tax morale of individuals is affected by their degree of similarity to others. Yet in the regressions of Table 3 we have documented that tax morale is also influenced by the overall level of similarity in their region of residence. This raises the question of whether these two effects are related and if so in what way. With this question in mind, we combine the explanatory variables from these two tables and estimate Eq. (4) which allows for an interaction effect between our measures of individual and regional similarity. From this specification we can assess whether the effect of individual similarity on tax morale is sensitive to the overall level of similarity across individuals in their region of residence.

Table 6 presents the estimates for these interaction regressions. In the first five columns we take our individual similarity measure in terms of values, which we found to have the strongest effect on individual tax morale, and estimate interaction effects with each

Table 6
Interaction regressions with individual and regional similarity.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dep. variable:	Individual tax morale								
Ind. Sim. wrt	Cult. Values	Cult. Values	Cult. Values	Cult. Values	Cult. Values	Income	Ethnicity	Language	Religion
Reg. Sim. wrt	Income	Ethnicity	Language	Religion	Cult. Values	Cult. Values	Cult. Values	Cult. Values	Cult. Values
Panel A - Effects Estimated from Separate Regressions using all Available Observations									
Ind. Similarity	0.360***	0.374***	0.361***	0.349***	0.356***	0.011***	0.008	0.007	0.011***
	(0.007)	(0.008)	(0.008)	(0.007)	(0.007)	(0.003)	(0.005)	(0.005)	(0.003)
Ind. Similarity x	−0.025***	−0.035***	−0.054***	−0.034***	−0.070***	−0.000	−0.007*	−0.006	−0.005
x Reg. Similarity	(0.006)	(0.008)	(0.009)	(0.008)	(0.007)	(0.002)	(0.004)	(0.004)	(0.003)
Observations	369731	270649	308037	464967	467147	334061	259433	290612	433580
Panel B - Effects Estimated from Separate Regressions based on Common Sample									
Ind. Similarity	0.394***	0.387***	0.383***	0.387***	0.396***	0.010***	0.008	0.006	0.019***
	(0.009)	(0.009)	(0.009)	(0.010)	(0.009)	(0.003)	(0.005)	(0.005)	(0.005)
Ind. Similarity x	−0.032***	−0.022**	−0.048***	−0.032***	−0.074***	0.000	−0.008*	−0.008*	−0.007
x Reg. Similarity	(0.009)	(0.009)	(0.010)	(0.010)	(0.009)	(0.002)	(0.004)	(0.005)	(0.004)
Observations	229204	229204	229204	229204	229204	229204	229204	229204	229204

Notes: This table explores how the effect of individual similarity on individual tax morale varies with the level of regional similarity. The estimated interaction effects are based on different combinations of individual and regional similarity measures indicated at the top of each column. All estimates are based on ordinary least squares (OLS) with fixed effects imposed at the region-wave level, which absorb the effect of regional similarity. The estimations control without reporting the estimates for the following individual characteristics: age, gender, marital status, employment status, education level. All variables are standardized to ensure comparability. Heteroskedasticity robust standard errors clustered at the region-wave level, in line with the fixed effects, are reported in parenthesis. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

of our five regional similarity measures, indicated at the top of each column. In the last four columns we take each of the other individual similarity measures in terms of income, ethnicity, language and religion and estimate interaction effects with our regional similarity measure in terms of values, which was also found to have the strongest effect on individual tax morale. Panel A of the table presents these interaction regressions based on all available observations and panel B based on a common sample. All regressions are estimated with region-wave fixed effects, which absorb all region-specific variables, including the regional similarity measures.[17]

Looking at the estimates in columns 1 to 5, we see in all cases a clear negative interaction effect between individual value similarity and all our regional similarity measures. This implies that as regional similarity in terms of any of the five cleavages increases, the positive relationship between individual value similarity and tax morale weakens. The tax morale of a given individual is less sensitive to how similar that individual is to others in the region, if that individual resides in a region where people are more similar and diversity is low. When turning to the other measures of individual similarity in columns 6 to 9 we also see a similar pattern. However, in this case the magnitude of the interaction effect is much smaller and in several cases statistically insignificant. This is consistent with the weaker effects that we found for these other measures of individual similarity in Table 4, yet qualitatively it suggests a similar pattern.

To assess how the link between tax morale and individual similarity varies as regional similarity changes, we plot in Fig. 1 the marginal effect of individual value similarity on individual tax morale for different levels of regional value similarity. We focus on this combination, for which we found the strongest interaction effect. The plot is thus based on the estimates of column 5 in panel A of Table 6. In the figure we also present the 95% confidence interval as well as a histogram of the distribution of regional value similarity, which is measured on the horizontal axis.

As the plot indicates, there is a tight negative relationship between the marginal effect of individual value similarity on individual tax morale and regional value similarity for the typical values of regional value similarity. When we move beyond that range, as the histogram in the background reveals, the relationship weakens and estimation precision drops. This is because there are only a few regions where individuals are so similar in terms of their values to fall in this category. Even for that range, however, individual value similarity has a positive effect on tax morale. Hence greater regional similarity moderates, but does not overturn the base result.

Taken together these findings lend further support to the conclusions about social identity that we advanced earlier. The fact that the effect of individual similarity is lower for individuals residing in more homogeneous regions is consistent with our interpretation of this variable as reflecting identity considerations of each individual. In a region where people are very similar to each other, each individual's sense of identification with members of their own social groups is expected to be relatively weak. As societal diversity increases and the overall level of similarity between individuals decreases, though, this is likely to trigger closer identification of individuals with their own group. This will have an asymmetric effect on the tax morale of individuals whose own social group is large and their level of individual similarity is high relative to those whose own social group is small and their individual similarity is lower. For the former individuals, closer identification with members of their own group raises their concerns for most people in society and consequently increases their willingness to pay taxes. For the latter individuals, on the other hand, their concerns about the majority of people in society will decrease and their tax morale will decline. Thus, greater societal diversity will increase the

[17] In Table A9 in the appendix we explore some alternative ways of estimating these interaction effects.

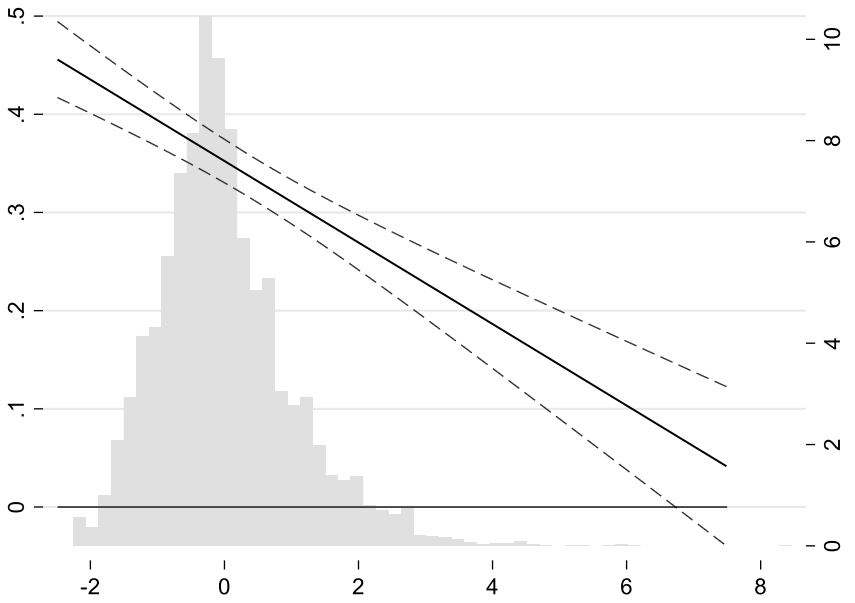

Fig. 1. Interaction plot between individual and regional value similarity.
Notes: The figure plots how the effect of individual value similarity on tax morale varies with the level of regional value similarity based on the regression estimates of column 5 in panel A of Table 6. The solid line corresponds to the point estimate and the dotted lines mark the upper and lower bounds of the 95% confidence interval. The lines are superimposed above the histogram of the distribution of regional value similarity across all regions in our sample.

differences in tax morale between individuals characterized by high and low similarity to others and should make tax morale more sensitive to variation in individual similarity.

This interpretation is also consistent with the broader view of social identity advanced by Turner et al. (1987). In particular it echoes the conclusions from Turner and Oakes (1989), who argue that the comparison with other social groups is more evident in a more diverse environment and diversity tends to strengthen an individual's own categorization as a member of a specific group. It is also consistent with a variety of evidence indicating that a more diverse social environment triggers stronger identification of individuals with the social group where they belong (Charness et al., 2007; Fong and Luttmer, 2009; Benjamin et al., 2010; Finseraas et al., 2019). Societal diversity, hence, makes group identity more salient.

5. Robustness checks

In this section we present some further regressions in order to assess the robustness of our main results presented above. In particular, these regressions explore the robustness of our main results with respect to different estimation methods, to alternative measures of similarity, to using different sets of value questions and to the inclusion of additional control variables.

5.1. Different estimation methods

Throughout our main analysis we have estimated the relationship between tax morale and the different similarity measures with ordinary least squares (OLS). We opted for this baseline estimation approach for simplicity and in order to facilitate comparisons across different specifications with different controls and fixed effects. However, our individual tax morale measure is not a continuous variable. It reflects individual responses on an ordered scale of integers from 1 to 10. Given that, in this subsection we estimate our individual level specification using probit and logit models. The estimation results are reported in Table 7 below.

In column 1 of the table we first repeat for comparison the OLS estimates for the effect of individual value similarity from column 5 in panel B of Table 4. We study this regression specification in particular, as individual value similarity is the main measure we focus on and, as already shown, it has the largest effect on individual tax morale. In columns 2 and 3 we estimate the same specification using ordered probit and ordered logit models respectively. For these models the table reports both the estimated coefficients and the marginal effects, estimated at the mean value of individual similarity. In both cases we obtain a clear positive relationship between individual tax morale and value similarity. Also the marginal effects from the two models are very similar.

In columns 4, 5 and 6 we follow an alternative approach and estimate the effect of individual value similarity on a binary measure of individual tax morale. This is based on a re-coding of our individual tax morale measure with 1 indicating individuals who find tax evasion completely unjustified and have a score of 10 on the original ordered scale, and 0 indicating individuals who find it at least partially justified and have scores from 1 to 9 on the original ordered scale. This is a common approach in the literature on tax morale, as the distribution of individual responses in the ordered scale is quite skewed with a large share

Table 7
Robustness checks with different estimation methods.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. variable:				Individual tax morale				
Estimation Method	OLS	O. Probit	O. Logit	OLS	Probit	Logit	OLS	2SLS
Ind. Value Similarity	0.349***	0.440***	0.789***	0.188***	0.483***	0.813***	0.011***	0.214***
	(0.007)	(0.003)	(0.005)	(0.003)	(0.003)	(0.006)	(0.003)	(0.057)
Marginal Effect		0.165***	0.180***		0.182***	0.188***		
		(0.002)	(0.002)		(0.002)	(0.002)		
Observations	467942	467942	467942	466694	466694	466694	467942	467942
Regions	2235	2235	2235	2233	2233	2233	2235	2235

Notes: This table explores the robustness of the effect of individual value similarity on individual tax morale when considering different estimation methods. Column 1 reports our baseline OLS estimates from Table 4. Columns 2 and 3 report the estimates respectively for the ordered probit and logit model, estimated based on the same dependent variable. Columns 4, 5 and 6 report the estimates respectively for a simple OLS, probit and logit, estimated based on the modified binary version of the dependent variable. Column 7 reports the OLS estimates when we use the predicted measure of individual value similarity instead of the actual one. Column 8 reports the 2SLS estimates when we use the predicted measure of individual value similarity as an instrument for the actual one. The estimates in all columns include fixed effects at the region-wave level and the following individual-level controls: age, gender, marital status, employment status and education level. All variables are standardized to ensure comparability. Heteroskedasticity robust standard errors clustered at the level of the fixed effects are reported in parentheses. For the estimates of columns 2–5 we also report the marginal effects estimated at the mean. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

of respondents answering that tax evasion is never justified.[18] Using this binary measure as the dependent variable, we estimate in column 4 our main specification with OLS, in column 5 with a binary probit and in column 6 with a binary logit. Doing so yields similar conclusions as before. Individual value similarity is found in all cases to have a positive effect on tax morale and the estimated marginal effects are similar across the different models. A one standard deviation increase in value similarity is found to increase the probability of an individual considering tax evasion to be unjustified by about 18.5%.

In the remaining two columns of Table 7 we use an instrumental variable estimation approach in order to account for the potential endogeneity of individual value similarity. We construct predicted individual value similarity scores for each individual and use that as an instrument for their actual value similarity scores. The predicted value similarity scores are obtained by first estimating an ordered probit regression where the responses to each of the 96 employed value questions are regressed on our core set of individual controls and region-wave fixed effects. Based on the estimated coefficients, we then compute for each individual a predicted response to a given value question. This corresponds to the response with the highest probability score according to the ordered probit regression. We then assign to each individual respondent a similarity score for that question that reflects the share of other people in the region whose actual response to that question is equal to the individual's predicted response.[19]

After computing predicted value similarity scores for each respondent and question, we average the scores across questions and use the resulting average score as our measure of predicted value similarity. In the regression of column 7 we estimate the same regression as in column 1 using this measure instead of the actual value similarity score. In column 8 we estimate a 2SLS regression with predicted value similarity serving as an instrument for actual value similarity. In both cases, we observe a positive relation with individual tax morale, while the obtained first-stage F-statistic of 354.11 confirms the strength of our instrument. This suggests that our main results are not driven by potential biases due to endogeneity. The reduced-form estimates are lower than the OLS estimates, as predicted value similarity only reflects part of the observed variation in actual value similarity. Yet, the 2SLS estimates are quite close to the marginal effects from the probit and logit models which is reassuring.

5.2. Alternative measures of value similarity

The measures of individual and regional similarity that we use for our main analysis all reflect the average probability of an individual or a group of individuals coming across another individual from the same region sharing a given attribute, such as the same income class, ethnicity, main language, religion or cultural values. Using such measures of similarity that are based on the Herfindahl–Hirschman index of concentration has a long tradition in economics and other social sciences. These measures are useful in our context because they facilitate comparisons regarding the importance of similarity across social cleavages. However, these measures reflect only the distribution of individuals within a given region across the different groups along each cleavage. They do not reflect how different these groups are or other aspects of differentiation across individuals.

With this in mind, in this section we explore the robustness of our main results to alternative measures of similarity. We only focus on similarity in terms of cultural values, given its relative importance and the fact that there are several alternative ways to measure that. These robustness checks are reported in Table 8. In all cases we introduce these alternative measures in the specification of

[18] See, for example, Alm and Torgler (2006) and Torgler and Valev (2010).

[19] Suppose, for example, that we have a value question with three possible answers for which the actual shares of respondents in a given region associated with answers 1, 2 and 3 are 0.2, 0.3 and 0.5 respectively. Consider an individual whose actual response to this question is 1 and his actual individual value similarity score based on that question is 0.2. If the probit estimation gives a predicted answer to this question for that individual of 2 based on the employed regressors, then the individual's predicted value similarity score would be equal to 0.3.

Table 8
Robustness checks with alternative measures of individual value similarity.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. variable:	Individual tax morale							
Value Similarity	0.349***							
	(0.007)							
Weight. Value Similarity		0.197***						
		(0.006)						
Value Sim. Above Med.			0.179***					
			(0.004)					
Weigh. Value Sim. Above Med.				0.142***				
				(0.004)				
Values in Majority					0.210***			
					(0.005)			
Values in Minority						−0.093***		
						(0.003)		
Abs. Deviation of Values							−0.127***	
							(0.005)	
Std. Deviation of Values								−0.123***
								(0.004)
Observations	467942	467942	467942	467942	467942	467942	467942	467942
Regions	2235	2235	2235	2235	2235	2235	2235	2235
Countries	111	111	111	111	111	111	111	111

Notes: This table explores the robustness of the effect of individual value similarity on individual tax morale when considering different measures of similarity. Column 1 presents the effect for our preferred measure of value similarity, from Table 4. Column 2 presents the effect of the corresponding weighted measure of value similarity where the weights reflect distances on the response scales. In columns 3 and 4 the employed measures reflect the extent to which the responses of individuals place them on average above or below the median in their region in terms of the unweighted and weighted value similarity measures respectively. In columns 5 and 6 the employed measures reflect the extent to which the responses of individuals place them on average in the majority or the minority in their region respectively. In columns 7 and 8 value similarity is measured based on the average absolute deviation and standard deviation respectively of individual responses relative to other respondents from the same region. The estimates in all columns are based on ordinary least squares (OLS) with fixed effects imposed at the region-wave level. All regressions control without reporting the estimates for the following individual characteristics: age, gender, marital status, employment status, education level. All variables are standardized to ensure comparability. Heteroskedasticity robust standard errors clustered at the level of the fixed effects are reported in brackets. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

column 5 in panel B of Table 4. Column 1 of the table repeats for comparison the estimates for our main measure of individual value similarity. In column 2 we replace that with a weighted measure of value similarity where the probability of a given individual coming across another person with a certain type of values is weighted with the extent of difference in values between them.[20] Our weighted measure of value similarity is again computed separately for each question and then averaged across all 96 value questions from the WVS/EVS questionnaire. As we can see in column 2, this weighted measure of value similarity is also strongly positively related to individual tax morale.

In column 3 we employ an alternative measure of value similarity that indicates the share, out of the 96 value questions, for which an individual's value similarity score is above or equal to the corresponding median score in their region. In column 4 we employ a similar measure that indicates the share of questions for which an individual's weighted value similarity score is above or equal to the corresponding median score in their region. Using both these measures we see that individuals whose level of value similarity on average is more similar to the median in their region tend to be characterized by higher tax morale, confirming our earlier conclusion.

In column 5 we consider a simpler measure of value similarity for individuals that reflects the share of questions, out of the 96, for which an individual's values are equal to the mode in their region, placing them in the majority. In column 6 we consider a measure that reflects the opposite, namely the share of questions for which an individual's values are the least common in their region, placing them in the minority. As we can see from the estimation results, these two measures have unsurprisingly opposite effects on individual tax morale. Individuals whose values conform more closely to the majority view in their region tend to display higher tax morale, while individuals whose values conform more closely to the minority view in their region tend to display lower tax morale.

Finally in columns 7 and 8 we employ two alternative measures that reflect the extent of the deviation in values between a particular individual and other people from the same region. Specifically, in column 7 we consider the absolute deviation of the responses of a given individual to each value question from the mean response in their region averaged across all 96 questions.

[20] More formally, let s_r^g denote again the share of individuals in region r belonging to group g, which in the context of cultural values are defined based on the answer to a given value question. Consider a given individual i who belongs to group g_i. Suppose the difference in values between that individual with people belonging to group g is proportional to the distance in the response scale $|g - g_i|$. Then a weighted measure of individual value similarity that accounts for this distance would correspond to the index:

$$\widetilde{Similarity}^{a}_{i,r} = 1 - \sum_{g=1}^{G^a} s_r^g \cdot |g - g_i|.$$

Table 9
Robustness checks with alternative sets of value questions I.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dep. variable:	Individual tax morale						
Ind. Value Sim. (Sect. A)	0.027***						
	(0.003)						
Ind. Value Sim. (Sect. B)		0.016***					
		(0.003)					
Ind. Value Sim. (Sect. C)			0.016***				
			(0.003)				
Ind. Value Sim. (Sect. D)				0.030***			
				(0.003)			
Ind. Value Sim. (Sect. E)					0.091***		
					(0.004)		
Ind. Value Sim. (Sect. F)						0.302***	
						(0.007)	
Ind. Value Sim. (Sect. G)							0.034***
							(0.003)
Ind. Value Sim. (Residual)	0.365***	0.369***	0.338***	0.339***	0.294***	0.140***	0.338***
	(0.007)	(0.008)	(0.007)	(0.007)	(0.007)	(0.005)	(0.007)
Observations	467942	383313	467331	467454	467840	467930	448213
Regions	2235	2144	2235	2235	2235	2235	2235

Notes: This table explores the robustness of the effect of individual value similarity on individual tax morale when value similarity is measured based on alternative sets of values questions. The value questions are grouped based on the section (A to G) in which they appear in the EVS/WVS questionnaire. Columns 1 to 7 compare the effect of value similarity measured based on questions from a particular section of the survey with that of value similarity measured based on all remaining questions. The estimates in all columns are based on ordinary least squares (OLS) with fixed effects imposed at the region-wave level. All regressions control without reporting the estimates for the following individual characteristics: age, gender, marital status, employment status, education level. All variables are standardized to ensure comparability. Heteroskedasticity robust standard errors clustered at the level of the fixed effects are reported in parentheses. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

In column 8 we consider the standard deviation of the responses of a given individual from the mean response in their region across all 96 questions. As expected, both measures are found to have a strong negative relationship with individual tax morale. Taken together, the estimation results in Table 8 imply that our earlier conclusions about the strong positive relationship between individual tax morale and similarity in cultural values do not hinge on the exact way in which we measure similarity.

5.3. Alternative sets of value questions

Our main measure of value similarity, as well as the alternatives presented in Table 8, are constructed as averages of similarity measures across a broad set of 96 cultural values that are included in the core part of the WVS/EVS questionnaire. A natural question is whether the conclusions that we have reached using this measure are sensitive to the exact set of questions based on which we measure value similarity. To assess whether this is the case, in this subsection we construct alternative versions of our individual value similarity measure that are based on the responses of individuals to only a subset of these 96 questions.

In Table 9 we focus on subsets of these questions following the thematic structure of the WVS/EVS questionnaire which represents seven different themes. These are perceptions of life (Section A), environment (Section B), work (Section C), family (Section D), politics and society (Section E), religion and morality (Section F), and national identity (Section G). Given that the questions falling under the same section tend to reflect more closely related values, we construct seven alternative versions of our main individual value similarity measure that are equal to the average of the value similarity scores for the questions belonging to the same section. For each of these sections we also construct a corresponding residual value similarity measure that takes the average of the value similarity scores for the questions not belonging to this section.

Using these section-specific value similarity measures, together with the corresponding residual one, we re-estimate our key specification from column 5 in panel B of Table 4. Doing so we can assess the relative importance of value similarity for the values that fall within each of the seven sections of the WVS/EVS questionnaire. The estimates are reported in columns 1 to 7 of Table 9. As they reveal, tax morale is influenced by similarity in all types of cultural values. All section-specific individual value similarity measures are positively and significantly related to individual tax morale. Comparing the estimated effects across sections, though, we do see some differences. Specifically, we see that value similarity associated with religious and moral values, covered in Section F, appears to be relatively more important for tax morale, followed by value similarity in terms of values related to politics and society, covered in Section E. This suggests that sharing these types of values can play a crucial role in fostering a sense of identity and promoting pro-social behavior.

In Table 10 we further explore the robustness of our results to the types of value questions based on which we construct our measure of value similarity. This time, though, we focus on the response scales. For some value questions the response scale allows individuals to give only two possible responses. Others allow for 3, 4, 5 or 10 different responses. To account for these differences, we have from the beginning normalized the response scales for all questions. Nevertheless, questions that permit more responses may be better suited for measuring value similarity. To assess whether this matters, we construct five different measures of value

Table 10
Robustness checks with alternative sets of value questions II.

Dep. variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Individual tax morale								
	no 2pQ	no 3pQ	no 4pQ	no 5pQ	no 10pQ	uQ < 50%	uQ < 25%	N > 25	N > 50
Ind. Value Sim. (excl. 2pQ)	0.389***								
	(0.007)								
Ind. Value Sim. (excl. 3pQ)		0.333***							
		(0.007)							
Ind. Value Sim. (excl. 4pQ)			0.356***						
			(0.007)						
Ind. Value Sim. (excl. 5pQ)				0.342***					
				(0.007)					
Ind. Value Sim. (excl. 10pQ)					0.185***				
					(0.005)				
Ind. Value Similarity						0.351***	0.357***	0.351***	0.356***
						(0.007)	(0.012)	(0.007)	(0.008)
Observations	467942	467942	467942	467942	467942	453709	169448	455894	418568
Regions	2235	2235	2235	2235	2235	2229	1190	1873	1422

Notes: This table explores the robustness of the effect of individual value similarity on individual tax morale to variation in the response scale of the employed value questions and in the number of responses per respondent. In columns 1 to 5 our measure of value similarity is constructed excluding questions with two, three, four, five and ten point Likert-scales respectively. In columns 6 and 7 our sample excludes respondents who answered fewer than 50% and 75% of the 96 value questions respectively. In columns 8 and 9 our sample excludes regions where the number of respondents is fewer than 25 and 50. The estimates in all columns are based on ordinary least squares (OLS) with fixed effects imposed at the region-wave level. All regressions control without reporting the estimates for the following individual characteristics: age, gender, marital status, employment status, education level. All variables are standardized to ensure comparability. Heteroskedasticity robust standard errors clustered at the level of the fixed effects are reported in brackets. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

similarity that exclude questions with a particular response scale from the analysis and take the average of the value similarity scores across the remaining questions.

Columns 1 to 5 of the table report the estimation results when we consider each of these response-scale-specific variants. As the estimates reveal, dropping questions with 2-point response scales from the value similarity measure, as we do in column 1, increases the magnitude of the effect. This is because these questions capture less variation in cultural values across individuals. Dropping questions with 3-point, 4-point and 5-point response scales from the value similarity measure, as we do in columns 2, 3 and 4, does not change the estimated effect much. On the other hand, dropping from the value similarity measure the questions with the 10-point response scales, which capture better variation in cultural values, decreases the magnitude of the effect in column 5.

In columns 6 and 7 of the table we go back to include all 96 value questions in our measure of individual value similarity, but we impose cutoffs on the number of survey questions that an individual should have answered in order to be included in the sample. We do so because individuals who answered only few of the 96 questions may have individual value similarity scores which are less informative. Specifically, in column 6 we retain in our sample only individuals who have answered at least 50% of the 96 questions. In column 7 we raise that cutoff further and retain in the sample only individuals who have answered at least 75% of the 96 questions. In both cases we see that our results are robust to imposing these cutoffs.

In columns 8 and 9 we impose cutoffs on the number of respondents per region and exclude from our sample regions for which the number of individual respondents is low. This is because our value similarity scores may be less accurate for regions where we have only few survey participants. Specifically, in column 8 we only retain regions with at least 25 survey respondents and in column 9 we only retain regions with at least 50 respondents. While this approach reduces our sample size, in both cases we see that it does not affect our main results.

5.4. Additional individual controls

The regressions presented in our main tables include only a small set of individual controls, namely age, gender, marital status, employment status and education level. This is largely for consistency reasons, as these variables are reported for most survey respondents and, hence, we can use them for our instrumentation strategy, described in a previous subsection. Beyond these variables, though, the literature has identified several other individual characteristics that are systematically related to individual tax morale. With this in mind, in Table 11 we explore the robustness of our main regression results to the inclusion of additional individual characteristics as controls. Here we report for brevity the effect of these controls when introduced in the specification of column 5 in panel A of Table 6, which includes individual value similarity interacted with regional value similarity.

In column 1 of Table 11 we control for differences in the respondents' household income. In column 2 we condition on the level of religiosity of the respondents, emphasized by Torgler (2006). In column 3 we account for the respondents' level of generalized morality, as defined by Tabellini (2010). In column 4 we control for the respondents' level of generalized trust and in column 5 for the level of trust that respondents exhibit towards their neighbors. In column 6 we account for the respondents' perception of tax evasion by other people in their country, highlighted by Frey and Torgler (2007). In column 7 we control for the respondents' attitudes towards foreigners, proxied by their stated preferences about having a foreign neighbor. In columns 8 and 9 we follow Besley

Table 11
Robustness checks with additional individual characteristics.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Dep. variable:	Individual tax morale										
Ind. Value Similarity	0.371***	0.350***	0.355***	0.358***	0.352***	0.335***	0.356***	0.353***	0.364***	0.359***	0.359***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.010)	(0.026)	(0.007)	(0.007)	(0.010)	(0.007)	(0.007)
Ind. Value Similarity ×	−0.071***	−0.070***	−0.069***	−0.068***	−0.074***	−0.117***	−0.071***	−0.069***	−0.063***	−0.069***	−0.070***
x Reg. Value Similarity	(0.007)	(0.007)	(0.007)	(0.007)	(0.010)	(0.022)	(0.007)	(0.007)	(0.010)	(0.007)	(0.007)
Income	−0.015***										
	(0.003)										
Religiosity		0.032***									
		(0.003)									
Morality			0.009***								
			(0.002)								
Trust				0.017***							
				(0.002)							
Trust in Neighbors					0.027***						
					(0.003)						
Compatriots Tax Honesty						0.164***					
						(0.012)					
Dislike Foreign Neighbors							−0.019***				
							(0.003)				
Conf. in Civil Service								0.032***			
								(0.003)			
Satisf. with Pol. System									0.013***		
									(0.003)		
Preference for Equality										0.008***	
										(0.002)	
Preference for Active Gov.											0.006**
											(0.003)
Observations	385961	453889	467922	451920	263332	33087	448705	445304	284184	447357	460307
Regions	2145	2226	2235	2235	1859	366	2210	2211	1925	2234	2234

Notes: This table explores the robustness of the effect of individual value similarity and its interaction with regional value similarity on individual tax morale when controlling for additional individual characteristics. The estimates in all columns are based on ordinary least squares (OLS) with fixed effects imposed at the region-wave level. All regressions control without reporting the estimates for the following individual characteristics: age, gender, marital status, employment status, education level. All variables are standardized to ensure comparability. Heteroskedasticity robust standard errors clustered at the region-wave level are reported in brackets. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

(2020) and consider the respondents' perception of institutions, measured by their confidence in the country's civil service and their satisfaction with their country's political system. In columns 10 and 11 we account for the respondents' preferences towards equality and government intervention, emphasized by Alesina and Angeletos (2005). All these variables are also standardized and measured based on specific questions from the survey described in the appendix.[21]

Looking at the estimated effects for the additional control variables, first of all, we can confirm previous conclusions reached by the literature regarding the effects of these variables. Individuals who are richer and less sympathetic to foreigners tend to have lower tax morale. Individuals who are more religious, exhibit higher levels of generalized morality and trust others more tend to have higher tax morale. This is also the case for individuals who are more confident in other people paying their taxes, believe more strongly in government institutions, and are more supportive of income redistribution and an active government. Even when we condition on these effects in our regressions, though, we still observe the main pattern uncovered in our baseline regressions. Individuals who are more similar to others in terms of their values are characterized by higher levels of tax morale and this effect becomes weaker if they reside in a region where people are generally more similar.[22]

5.5. Additional interactions with regional characteristics

In the regression results reported in Table 6 we explored interaction effects of individual similarity with different regional similarity measures. The regional similarity measures that we consider, however, may also reflect other differences across regions that influence tax morale. Hence, to ensure the correct interpretation of our main results, it is important that we assess whether the interaction effects that we have documented are driven by other regional characteristics. For this purpose, we have turned to the literature and identified several such characteristics that have been linked with tax morale. For each of these characteristics we estimate an interaction effect with our main individual value similarity measure. The estimation results are reported in Table 12. As

[21] As these questions are not included in all survey waves, there is naturally some variation in sample size across columns.

[22] Despite the fact that several of these additional control variables are significant predictors of individual tax morale, none of them appear to moderate the relationship between individual similarity and tax morale. This means our main result is not driven by specific values and attitudes such as social trust, tolerance towards migrants or preferences for equality.

Table 12
Robustness checks with additional regional characteristics.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dep. variable:	Individual tax morale								
Ind. Value Similarity	0.352***	0.356***	0.374***	0.335***	0.336***	0.227***	0.356***	0.346***	0.346***
	(0.007)	(0.007)	(0.011)	(0.007)	(0.007)	(0.014)	(0.007)	(0.009)	(0.011)
Ind. Value Sim. × Reg. Value Sim.	−0.079***	−0.069***	−0.078***	−0.102***	−0.102***	−0.094***	−0.070***	−0.068***	−0.068***
	(0.007)	(0.007)	(0.012)	(0.007)	(0.007)	(0.017)	(0.007)	(0.010)	(0.012)
Ind. Value Sim. × Luminosity per capita	−0.029***								
	(0.006)								
Ind. Value Sim. × Population Density		0.008							
		(0.005)							
Ind. Value Sim. × Immigrant Pop. Share			0.051***						
			(0.014)						
Ind. Value Sim. × Nat. Qual. of Governance				−0.083***					
				(0.008)					
Ind. Value Sim. × Nat. Gov. Effectiveness					−0.077***				
					(0.007)				
Ind. Value Sim. × Reg. Qual. of Governance						−0.044***			
						(0.012)			
Ind. Value Sim. × Reg. Conf. in Civil Service							−0.001		
							(0.006)		
Ind. Value Sim. × Reg. Sat. with Pol. System								−0.025**	
								(0.010)	
Ind. Value Sim. × Reg. Election Participation									−0.026***
									(0.003)
Observations	467942	467942	150766	423541	450393	96567	466692	304993	203397
Regions	2235	2235	1441	2006	2104	496	2221	1925	1705

Notes: This table explores the robustness of the interaction effect between individual value similarity and regional value similarity on individual tax morale to the inclusion of interactions with other regional characteristics. The estimates in all columns are based on ordinary least squares (OLS) with fixed effects imposed at the region-wave level. All regressions control without reporting the estimates for the following individual characteristics: age, gender, marital status, employment status, education level. All variables are standardized to ensure comparability. Heteroskedasticity robust standard errors clustered at the region-wave level are reported in brackets. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

we did in Table 11, we introduce these interactions in the specification of column 5 in panel A of Table 6. As this specification includes region-wave fixed effects, for each of these regional characteristics we can only estimate the interaction effect with individual value similarity, but not the level effect.

In column 1 of Table 12 we introduce an interaction effect of individual value similarity with the level of economic development of the region, proxied by the observed luminosity of the region at night relative to its population.[23] In column 2 we include an interaction with population density which reflects each region's broader economic and social organization. In column 3 we do the same with the share of respondents who are immigrants, which reflects the exposure of the region's population to migrants. In columns 4 and 5 we introduce interaction terms of individual value similarity with two widely used measures of institutional quality at the country level. These are respectively the government effectiveness index from the World Bank's Worldwide Governance Indicators and the ICRG quality of governance index. In column 6 we do the same with a regional quality of governance indicator by the Quality of Governance Institute which is available for European Union regions. In columns 7 and 8 we consider interactions with the regional average levels of confidence in the civil service and satisfaction with democracy as indicators of the perceived quality of local institutions. Finally, in column 9 we look at average levels of participation in elections by individuals in each region to capture the local level of civic engagement. All these variables are also standardized and measured as explained in the appendix.

Looking at the estimates of the interaction terms between individual value similarity and each of these variables, we see a clear pattern. The positive relationship between individual value similarity and tax morale generally appears to be less pronounced in regions that are more developed economically, have better institutions and where the population is more civically engaged. This pattern is consistent with previous findings in the literature, as economic and institutional development can foster tax morale (Torgler, 2005; Hug and Spoerri, 2011; Ivanyna et al., 2016; Besley, 2020). It is also consistent with the interaction effects reported in Table 6, as regions that are more homogeneous tend to be characterized by higher levels of economic development and institutional quality. The only variables for which we find a positive interaction effect with individual value similarity is the share of immigrants and population density. This is not surprising as populous urban areas tend to be more diverse, more congested and also attract more migrants compared to rural areas.

What is striking, however, is that in all cases the negative interaction effect between individual and regional value similarity is still visible in the data even when we condition on these additional interaction effects. Moreover, in all cases we see that the

[23] We use this variable as a measure of economic development because it is hard to obtain comparable measures of regional income levels for the large number of sub-national regions that we consider in our analysis. For more details on why per capita luminosity is a good proxy for economic development see Henderson et al. (2012).

additional interaction effects are always weaker than our main interaction effect with regional value similarity. This suggests that our main results are unlikely to be driven by other regional determinants of tax evasion, even though these variables are definitely relevant for individual tax morale.

6. Conclusion

In this paper we explore how societal diversity in terms of various social cleavages that include income, ethnicity, language, religion and cultural values can influence tax morale. Using data from the World Values Survey, we measure and compare tax morale across countries, across regions within countries and across individuals within regions. We consistently find greater diversity along these cleavages to be associated with lower tax morale at the country, region and individual level. Moreover, our regression results demonstrate that within a given country and region individuals who are more similar to others in terms of these cleavages tend to display higher levels of tax morale.

Furthermore, we compare how the tax morale of individuals is influenced by their similarity to other respondents in their region of residence and the overall level of similarity across respondents in the region. We find that the two effects complement each other. As regional diversity increases and the overall level of similarity across individuals decreases, individual tax morale becomes more sensitive to how similar that individual is to others in the region. These patterns are shown to hold conditional on various individual determinants of tax morale and on a full set of region fixed effects that filter out all sources of variation across regions.

Assessing the relative importance of these different social cleavages, we show that, although these patterns apply in all cases, similarity in terms of cultural values appears to be particularly important. The strong effect of similarity in values on tax morale is shown to apply also when comparing individuals that share the same ethnicity, language or religion. Taken together these results suggest that there is a strong individual component to tax morale, which is visible even when comparing individuals that live in the same region and operate within the same local economic and institutional environment. This individual component appears to reflect people's own social identity, which depends on how strongly they identify themselves with the social groups to which they belong.

Although the role of social identity in the context of tax evasion has already been discussed by some authors in the literature, we are the first to assess its effect by quantifying broadly the degree of social identification across individuals. When doing so it is important to look at similarity in terms of a range of cultural values and not only particular social cleavages such as ethnicity, language or religion, that may not fully reflect the social identity of individuals. While our results confirm the relevance of ethnic, linguistic and religious cleavages in this context, these cleavages appear to matter mostly to the extent that they overlap with differences in cultural values. By looking at individual values we can get a better sense of how individuals identify with others around them, and it is this element of identification that can motivate individuals to deviate from their own narrow interests and behave in a more pro-social way when it comes to tax compliance and related decisions.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jce.2023.04.005.

References

Akerlof, G.A., Kranton, R.E., 2000. Economics and identity. Q. J. Econ. 115 (3), 715–753.
Alesina, A., Angeletos, G.M., 2005. Fairness and redistribution. Amer. Econ. Rev. 95 (4), 960–980.
Alesina, A., Baqir, R., Easterly, W., 1999. Public goods and ethnic divisions. Q. J. Econ. 114 (4), 1243–1284.
Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., Wacziarg, R., 2003. Fractionalization. J. Econ. Growth 8 (2), 155–194.
Alesina, A., Ferrara, E.La., 2005. Ethnic diversity and economic performance. J. Econ. Lit. 43 (3), 762–800.
Allingham, M.G., Sandmo, A., 1972. Income tax evasion: a theoretical analysis. J. Public Economics 1 (3–4), 323–338.
Alm, J., 2019. What motivates tax compliance? J. Econ. Surv. 33 (2), 353–388.
Alm, J., Clark, J., Leibel, K., 2016. Enforcement, socioeconomic diversity, and tax filing compliance in the United States. South. Econ. J. 82 (3), 725–747.
Alm, J., Torgler, B., 2006. Culture differences and tax morale in the United States and in Europe. J. Econ. Psychol. 27 (2), 224–246.
Ashraf, Q.H., Galor, O., 2018. The macrogenoeconomics of comparative development. J. Econ. Lit. 56 (3), 1119–1155.
Baldwin, K., Huber, J.D., 2010. Economic versus cultural differences: Forms of ethnic diversity and public goods provision. Am. Polit. Sci. Rev. 104 (4), 644–662.
Banerjee, A., Iyer, L., Somanathan, R., 2005. History, social divisions, and public goods in rural India. J. Eur. Econom. Assoc. 3 (2–3), 639–647.
Beach, B., Jones, D.B., 2017. Gridlock: Ethnic diversity in government and the provision of public goods. American Econ. J.: Economic Policy 9 (1), 112–136.
Benjamin, D.J., Choi, J.J., Strickland, A.J., 2010. Social identity and preferences. Amer. Econ. Rev. 100 (4), 1913–1928.
Besley, T., 2020. State capacity, reciprocity, and the social contract. Econometrica 88 (4), 1307–1335.
Beugelsdijk, S., Klasing, M.J., 2016. Diversity and trust: The role of shared values. J. Comp. Econ. 44 (3), 522–540.
Beugelsdijk, S., Klasing, M.J., Milionis, P., 2019. Value diversity and regional economic development. Scand. J. Econ. 121, 153–181.
Charness, G., Rigotti, L., Rustichini, A., 2007. Individual behavior and group membership. Amer. Econ. Rev. 97 (4), 1340–1352.
Christian, R.C., Alm, J., 2014. Empathy, sympathy, and tax compliance. J. Econ. Psychol. 40, 62–82.
Clotfelter, C.T., 1983. Tax evasion and tax rates: An analysis of individual returns. Rev. Econ. Stat. 65 (3), 363.
Cowell, F.A., Gordon, J.P., 1988. Unwillingness to pay: Tax evasion and public good provision. J. Public Economics 36 (3), 305–321.

Cummings, R.G., Martinez-Vazquez, J., McKee, M., Torgler, B., 2009. Tax morale affects tax compliance: Evidence from surveys and an artefactual field experiment. J. Econ. Behav. Organ. 70 (3), 447–457.
Dahlberg, M., Edmark, K., Lundqvist, H., 2012. Ethnic diversity and preferences for redistribution. J. Polit. Econ. 120 (1), 41–76.
Desmet, K., Gomes, J.F., Ortuno-Ortin, I., 2020. The geography of linguistic diversity and the provision of public goods. J. Dev. Econ. 143, 1023–1084.
Desmet, K., Ortuno-Ortin, I., Wacziarg, R., 2017. Culture, ethnicity, and diversity. Amer. Econ. Rev. 107 (9), 2479–2513.
Desmet, K., Ortuno-Ortin, I., Weber, S., 2009. Linguistic diversity and redistribution. J. Eur. Econom. Assoc. 7 (6), 1291–1318.
Dulleck, U., Fooken, J., Newton, C., Ristl, A., Schaffner, M., Torgler, B., 2016. Tax compliance and psychic costs: Behavioral experimental evidence using a physiological marker. J. Public Economics 134, 9–18.
Dwenger, N., Kleven, H., Rasul, I., Rincke, J., 2016. Extrinsic and intrinsic motivations for tax compliance: Evidence from a field experiment in Germany. American Econ. J.: Economic Policy 8 (3), 203–232.
Feld, L.P., Frey, B.S., 2007. Tax compliance as the result of a psychological tax contract: The role of incentives and responsive regulation. Law and Policy 29 (1), 102–120.
Filippin, A., Fiorio, C.V., Viviano, E., 2013. The effect of tax enforcement on tax morale. Eur. J. Political Econ. 32, 320–331.
Finseraas, H., Hanson, T., Johnsen, A.A., Kotsadam, A., Torsvik, G., 2019. Trust, ethnic diversity, and personal contact: A field experiment. J. Public Economics 173, 72–84.
Fong, C.M., Luttmer, E.F.P., 2009. What determines giving to hurricane katrina victims? experimental evidence on racial group loyalty. American Econ. J.: Applied Economics 1 (2), 64–87.
Fortin, B., Lacroix, G., Villeval, M.-C., 2007. Tax evasion and social interactions. J. Public Economics 91 (11–12), 2089–2112.
Frey, B.S., Torgler, B., 2007. Tax morale and conditional cooperation. J. Comp. Econ. 35 (1), 136–159.
Gelfand, M.J., Raver, J.L., Nishii, L., Leslie, L.M., Lun, J., Lim, B.C., et al., 2011. Differences between tight and loose cultures: A 33-nation study. Science 332 (6033), 1100–1104.
Gershman, B., Rivera, D., 2018. Subnational diversity in sub-Saharan Africa: Insights from a new dataset. J. Dev. Econ. 133, 231–263.
Gruendler, K., Koellner, S., 2020. Culture, diversity, and the welfare state. J. Comp. Econ. 48 (4), 913–932.
Habyarimana, J., Humphreys, M., Posner, D.N., Weinstein, J.M., 2007. Why does ethnic diversity undermine public goods provision? American Polit. Sci. Rev. 101 (4), 709–725.
Hallsworth, M., List, J.A., Metcalfe, R.D., Vlaev, I., 2017. The behavioralist as tax collector: Using natural field experiments to enhance tax compliance. J. Public Economics 148, 14–31.
Henderson, J.V., Storeygard, A., Weil, D.N., 2012. Measuring economic growth from outer space. Amer. Econ. Rev. 102 (2), 994–1028.
Hogg, M.A., 2018. Social identity theory. In: Burke, P.J. (Ed.), Contemporary Social Psychological Theories. Ch. 5, Stanford University Press, pp. 112–138.
Hug, S., Spoerri, F., 2011. Referendums, trust, and tax evasion. Eur. J. Political Econ. 27 (1), 120–131.
Ivanyna, M., Moumouras, A., Rangazas, P., 2016. The culture of corruption, tax evasion, and economic growth. Econ. Inq. 54 (1), 520–542.
LaPorta, R., Silanes, F.Lopez-de., Shleifer, A., Vishny, R.W., 1999. The quality of government. J. Law, Econ. Org. 15 (1), 222–279.
Li, S.X., 2010. Social identities, ethnic diversity, and tax morale. Public Finance Rev. 38 (2), 146–177.
Litina, A., Palivos, T., 2016. Corruption, tax evasion and social values. J. Econ. Behav. Organ. 124, 164–177.
Luttmer, E.F.P., 2001. Group loyalty and the taste for redistribution. J. Polit. Econ. 109 (3), 500–528.
Luttmer, E.F.P., Singhal, M., 2014. Tax morale. J. Econ. Perspect. 28 (4), 149–168.
Mare, M., Motroni, A., Porcelli, F., 2020. How family ties affect trust, tax morale and underground economy. J. Econ. Behav. Organ. 174, 235–252.
Myles, G.D., Naylor, R.A., 1996. A model of tax evasion with group conformity and social customs. Eur. J. Political Econ. 12 (1), 49–66.
Rodriguez-Justicia, D., Theilen, B., 2018. Education and tax morale. J. Econ. Psychol. 64, 18–48.
Shayo, M., 2009. A model of social identity with an application to political economy: Nation, class, and redistribution. American Polit. Sci. Rev. 103 (2), 147–174.
Slemrod, J., 1985. An empirical test for tax evasion. Rev. Econ. Stat. 67 (2), 232.
Slemrod, J., 2002. Trust in public finance. NBER Working Paper No. 9187.
Slemrod, J., Yitzhaki, S., 2002. Tax avoidance, evasion, and administration. In: Auerbach, M. (Ed.), Handbook of Public Economics. vol. 3, Elsevier, pp. 1423–1470, 22.
Srinivasan, T., 1973. Tax evasion: A model. J. Public Economics 2 (4), 339–346.
Stichnoth, H., Straeten, K.van.der., 2013. Ethnic diversity, public spending, and individual support for the welfare state: a review of the empirical literature. J. Econ. Surv. 26 (2), 364–389.
Stuart, B.A., Taylor, E.J., 2020. The effect of social connectedness on crime: Evidence from the great migration. Rev. Econ. Stat. 1–46.
Tabellini, G., 2010. Culture and institutions: Economic development in the regions of europe. J. Eur. Econom. Assoc. 8 (4), 677–716.
Tajfel, H., Turner, J.C., 1979. An integrative theory of intergroup conflict. In: Austin, S. (Ed.), The Social Psychology of Intergroup Relations. (Ch 3), Brooks-Cole, Monterey, CA, pp. 33–37.
Torgler, B., 2005. Tax morale and direct democracy. Eur. J. Political Econ. 21 (2), 525–531.
Torgler, B., 2006. The importance of faith: Tax morale and religiosity. J. Econ. Behav. Organ. 61 (1), 81–109.
Torgler, B., Valev, N.T., 2010. Gender and public attitudes toward corruption and tax evasion. Contemp. Econ. Policy 28 (4), 554–568.
Traxler, C., 2010. Social norms and conditional cooperative taxpayers. Eur. J. Political Econ. 26 (1), 89–103.
Turner, J.C., Hogg, M.A., Oakes, P.J., Reicher, S.D., Wetherell, M.S., 1987. Rediscovering the Social Group : A Self-Categorization Theory. Basil Blackwell, Oxford.
Turner, J.C., Oakes, P.J., 1989. Self-categorization theory and social influence. In: Paulus, P. (Ed.), The Psychology of Group Influence. vol. 23, Erlbaum, Hillsdale, NJ, pp. 3–275.
Vigdor, J.L., 2004. Community composition and collective action: Analyzing initial mail response to the 2000 census. Rev. Econ. Stat. 86 (1), 303–312.