Cultural Identity and Norms of Cooperation and Trust in Italy*

Emanuele Bracco[1], Federica Liberini[2],

Ben Lockwood[3], Francesco Porcelli[4,6],

Michela Redoano[3], and Daniel Sgroi[3,5,6*]

[1]Università di Verona

[2]University of Bath

[3]University of Warwick

[4]Università di Bari

[5]IZA

[6]ESRC CAGE Centre

*E-mail (corresponding author): daniel.sgroi@warwick.ac.uk

April 2022

*Funding for this project was provided by the ESRC CAGE Centre, ESRC Grant Ref RES-626-28-0001. Ethical approval granted by the University of Warwick HSSREC (Ref HSS 64/18-19). Registered on the AEA RCT registry (https://doi.org/10.1257/rct.3816-1.0). The authors declare no conflicts of interest.

Abstract

In a survey involving 1,547 respondents across three Italian cities we exploit regional variation in background, language and diet to investigate the relationship between cultural identity, trust and cooperation. Respondents with relatives who originate in the north of Italy, and who share common cultural characteristics, contributed 15% more in a public goods game, displayed greater trust in government and greater willingness to pay taxes, than respondents whose language and diet identified them as being southern. However, self-reported identity, a mainstay of the survey literature, had no predictive power. This highlights the importance of identity, but only if measured appropriately.

JEL Classification: Z13, C83, H41.

Keywords: identity, culture, cooperation, trust, social capital, public goods game, language, diet, survey.

1 Introduction

In this paper our focus will be on the links between identity and three important social norms: cooperation, trust and honesty. Our key question concerns the extent to which cooperation, trust and honesty travel with individuals when they move from region to region, or whether individuals adopt the norms of the region within which they reside. Cooperation, trust and honesty are important components of what is commonly called "social capital". Social capital is an umbrella term covering many aspects of social interaction that together allow societies to work effectively, including interpersonal relationships, a shared sense of cultural heritage, a shared understanding, shared norms and shared values. The importance of social capital is hard to overstate: it has been found to correlate with health (Wilkinson, 1996), longevity (Putnam, 2000), income equality (Wilkinson, 1996; Kawachi et al., 1997), economic growth (Helliwell and Putnam, 1995; Knack and Keefer, 1997; Zak and Knack, 2001), trade (Guiso et al., 2008a), well-functioning institutions (Putnam, 2000; Knack, 2002), child welfare (Côté and Healy, 2001), public services outcomes such as educational achievement (Coleman, 1988), financial markets (Guiso et al., 2004), financial development (Guiso et al., 2004), corruption and crime (Putnam, 2000; Buonanno et al., 2009) and has also been shown to be persistent, with short-run shocks potentially lasting for many years (Guiso et al., 2008b, 2016). If we can understand how cooperation, trust and honesty change over time with migration, and how they reflect identity, we will be better placed to understand the evolution of social capital over time.

To make our contribution clear, consider a hypothetical nation with low norms of cooperation and trust in the south and high norms in the north. An individual may move from south to north within their own lifetime, or perhaps their parents or grandparents migrated to the north. After living in the north, will such an individual take on the social norms of the north, or retain the social norms of the south? To answer this question we first need to measure identity. The standard method is through self-reporting, typically in survey or census data (for example Todd (2007) or Ha (2007)), but identity may operate at a subconscious level, exerting an influence that may be unknown to the individual, which might make a self-reported measure misleading. Our results show a powerful link between identity, cooperation,

trust and honesty, but *only* when we use a novel way to measure identity that takes full account of an individual's cultural background as distinct from a self-reported measure of identity.

Italy is a country that experienced huge internal migration from south to north in the aftermath of World War 2, and which has significant internal differences in language (dialect), diet and the distribution of social capital across the country (Putnam et al., 1993; Bigoni et al., 2019). It is this regional variation in background which allows us to generate a measure of cultural identity formed from differences in language and diet, and contrast this with a more conventional self-reported measure. Our primary method is a survey involving 1,547 (with some incentivized experimental features) spread across Milan, Turin and Rome which involves incentivized games designed to reveal attitudes towards trust and cooperation, as well as novel methods of investigating linguistic and dietary preferences and more conventional survey questions. Crucially for us, our respondents have very different family backgrounds which allows us to identify the extent to which their current home matters more or less than the birth-place of their parents or grandparents. We find that respondents with relatives who originate in the north of Italy, and who share common linguistic and dietary preferences, contributed significantly more in a public goods game, displayed greater trust in the government and had greater willingness to pay higher taxes than did those whose language and diet identified them as being from the south. Self-reported identity, on the other hand, had no predictive power at all. This suggests that self-identification may mask the prominent role played by identity in establishing norms of trust and cooperation. We also find that respondents with a family background that extends beyond the city where they live and were born, often behave in ways that are consistent with the place of birth of their maternal grandmothers, rather than other residents of their home town. Our results suggest that migrating to a different region may not be enough to initiate changes in social norms which may in fact take multiple generations to develop.

For our self-reported measure, we take a weighted average from a set of questions that are of the form: "Where do you think of yourself as coming from?" or "Which which place of origin of your relatives do you identify?" and combine this with a question about football team preference. We weight their answers via (polychoric) principal component analysis, normalizing to a number between 0 and 1, where 1 indicates a more southern identity. For our more subtle language and dietary based measure

which we refer to as an index of cultural heritage, we ask them to name a picture, identify a certain local idiom, or translate words in dialect as well as asking about their dietary preferences. We again weight their answers according to principal component analysis, creating a number between 0 and 1, where 1 indicates a more southern identity. We use a dummy to capture those who tend to use neutral answers that are not linked to a particular region which allows us to remove them from the identity measure.[1] We find a remarkable disconnect between the two measures of identity, indicating a surprising lack of awareness of cultural identity among respondents, with the role and importance of grandparents and their origin especially prone to underestimation. This is surprising given the predictive power of our novel identity measure as opposed to the more conventional self-reported measure which has no predictive power.

In summary, our first contribution is to develop a new measure of cultural identity based on language and diet, and second to show that using this new measure we can identify a strong relationship between cultural identity and the tendency towards cooperation and trust which would not be visible if we relied only on self-reported measures of identity. Moreover we also argue that cultural identity can be slow to change with elements of the subconscious continuity of cultural identity spanning at least two generations.

2 Design

We use a survey with some incentivized experimental features that was administered by Qualtrics Research Services through their online panel in April 2019. In total 1,547 respondents took part, just over 500 from each of three major Italian cities: Milan, Turin and Rome. The respondent population was pre-screened to admit only respondents with both sets of grandparents of Italian origin, but otherwise was pre-selected by Qualtrics Research Services to ensure a demographic spread that resembles the wider Italian population apart from immigration status which was filtered to give us sufficient variation. Qualtrics survey software was used to perform the study which took approximately 25 minutes.[2] Further

[1] The full method is described in much more detail in the design section below.

[2] The average length of 25 minutes per respondent was based on prior testing by Qualtrics Research Services.

respondent details can be found in table 1. While the survey is *not* a randomized control trial we did pre-register some elements of the design in the AEA RCT Registry ((Bracco et al., 2019)), for example the planned size of the survey and some of the key features, which provide an element of prior commitment. Qualtrics Research Services also operate a screening process that removes respondents who used detectable random answers and also removed any respondents who provided incomplete sets of answers. This was done prior to our receipt of data and so the 1,547 respondents whose data we received did not include any with incomplete answers or with detectable random answers.

The base earnings for each respondent was 3 euros (based on the standard payment set by Qualtrics Research Services), though in addition respondents were informed about potential additional earnings linked to two games and a random draw took place which selected a number of respondents who would receive a bonus equal to the amount accumulated at the end of the given game. Respondents were informed that the bonus could result in payments of upto 50 euros for each of the games. The full survey can be found in the appendix at the end of the paper but we describe the key elements of the survey here.[3]

First, respondents were asked a set of questions about their personal characteristics, their origins and their family's origins. Respondents were asked for their current city of residence (which was restricted to Milan, Turin or Rome) as well as their place of birth. They were asked their age when they moved to their current city of residence and various questions about when and why they moved, as well as a series of questions about the origin of their parents and grandparents. They were also asked a number of subjective questions such as where they believe they are from, which football team they support, and how close they felt to their parents and grandparents. The answers to this set of questions were used to form measures of self-reported identity. The questions are presented in pages 4-6 of the English translation of

[3]An English translation of the survey is presented in the appendix. This is made as compact as possible by removing duplicated questions from the original: for instance if identical questions were asked about a respondent's mother, father and all four grandparents, only the versions pertaining to the mother are in the translation. The translation also does not include standard demographic questions. The full original Italian version is presented at the end of the appendix and this does include all questions. Note that while the survey appears long, respondents typically only saw a subset determined by their own answers and in some cases randomization which explains why the average duration is around 25 minutes. The survey company ran pilots in advance to determine the average completion time in order to assist us in setting a payment level that follows convention for an interactive incentivized survey of this length.

the survey in the Appendix.[4]

Second, respondents were asked a series of language and diet questions designed to tease out their underlying cultural identity. Figure 1 in the Appendix provides a graphical indication of the wide variation in Italian dialect which is based on Pellegrini (1977)'s seminal division of Italy into linguistic subdivisions and is described in detail in section 3.1 below. Words in Italian dialect can be very different: take for instance the word for towel (*asciugamano*) which in dialect could be *tuvagghia*, *sciugaman* or *macrame* depending upon the region of Italy. Each respondent was asked to listen to four recorded sentences in the regional dialect of his or her grandparents. They were asked to translate the sentence and comment on their understanding. They were then asked to decide which of a series of words in dialect were most used by their own family. Next they were asked to translate various regional sayings (assigned based on their parents' place of birth), select which word they would use to describe their loved ones, and to state which word they might use to describe a water melon (having seen a photo), a fruit which has many different names across different dialects. This set of questions was used to form measures of identity derived from cultural heritage. The questions themselves were carefully selected to be linked to particular regions within the full range of Italy, and we also included several neutral responses to both the language and dietary questions which allow us to distinguish between those with a genuine preference for regional (local or otherwise) language and food as opposed to more neutral choices. We then control for neutral language and food choices though the use of dummies in our regressions to follow. The questions are presented in pages 6-13 of the English translation of the survey in the Appendix and we also provide the original survey in Italian immediately after the English translation in the Appendix.[5]

Third, respondents played two incentivized games. The first was a public goods game: a traditional

[4]Where questions are identical the version relating to the mother is presented in the translation though they were repeated for the father and all four grandparents. The full text is presented in the original Italian version of the survey.

[5]Note that many of these questions have an objectively correct answer which is a useful feature of this part of the survey when we construct our various measures, but also note that these linguistic questions were not incentivized. Both of these features are helpful in downplaying any potential priming effect or the temptation for respondents to expend energy in an attempt to find out the correct answer. For those questions that did not have a correct answer we typically offered a large range of possible choices and so it would be difficult for respondents to feel guided towards any particular response.

game used to measure an individual's propensity to cooperate with others. Respondents were put into pairs and each was allocated 20 euros. They were then asked how much they wished to put into a communal pot. Each respondent then received 20 - [their contribution to the pot] + 3/4 [the sum of the pot] in the game, and stood a chance of earning this as a real bonus. Before being asked to make their decision each respondent was shown a worked example and took a test designed to aid and check their understanding. A highly cooperative pair of partners might put all of their 20 euros into the pot and would then perform better than a pair of individually rational individuals. The best possible payoff is to put nothing in the pot while your partner places their full endowment into the pot (in which case they would earn 35 euros, while their partner would earn 15 euros). In this way a two-player public goods game is effectively the same as a two-player Prisoner's Dilemma but with a finer action set (a contribution can be selected rather than just cooperate or not) which is helpful in measuring the degree of cooperation. Respondents were also asked to guess the contributions of their partners. The second game was a simple test of honesty: respondents were asked to flip 10 coins and declare how many heads they flipped, potentially earning a bonus based on the number of heads.[6] While the coin flips were entirely private, we can say with some confidence that the higher the number of declared heads the more likely is the respondent to have lied especially when this is averaged over large numbers of players. This is especially clear when we compare the probability of flipping at least 8 heads (5.5%) with the percentage of times respondents declared they had flipped at least 8 heads (32.1%). These two games act as our core incentivized measures of cooperation, trust and honesty. The questions are presented in pages 13-15 of the English translation of the survey in the Appendix.

Finally, respondents were asked a series of questions designed to measure their self-reported level of social capital across several dimensions including cooperation and trust. For example, they were asked to list what behaviors they felt were socially acceptable (for instance not voting, evading taxation or using public transport without payment) and asked to indicate the level of trust they placed in groups such as the police, neighbours, or the state. They were asked whether they paid their taxes, or engaged in charity or other civic duties, as well as indicating how they would act when placed in a moral dilemma.

[6]For instance those who declared 8 or more heads had a potential bonus of at least 20 euros.

They were also asked how actively they engaged in the political process, for instance whether they voted regularly or kept upto date with political affairs. At the end of this part of the survey they were asked some factual questions about the current political situation including for instance the name of the mayor of their town of residence. The questions are presented in pages 15-25 of the English translation of the survey in the Appendix.

3 Results

In the results that follow we will differentiate between self-reported measures of identity and measures that are derived from responses to questions designed to tease out underlying cultural heritage. Before beginning our analysis we first go into some detail on our subdivision of Italy into 3 or 6 regions, and provide definitions of "local origin" and our various different measures of identity. We also note here the respondent characteristics obtained as part of the survey which are summarized in table 1 suggest that our sample was balanced across the Milan, Rome and Turin sub-samples.

3.1 Definitions

Regional Categorization. In what follows we will refer to the "north", "centre" and "south" of Italy which forms our trichotomy of Italy. Figure 1 provides a graphical exposition of the three regions broken down into sub-regions. The north (the darkest two shades of grey) will typically be allocated a value of 1, the central region (the middle two shades of grey) a value of 2 and the southern region (the lightest two shades of grey) a value of 3. Occasionally we will subdivide further into 6 regions indicated in the figure by the six different shades of grey in situations where we do not consider diet, in which case the north-west region (which includes Turin and Milan) is allocated a value of 1, the north-east region a value of 2, Tuscany is allocated a value of 3, the central region (which includes Rome) a value of 4, the southern region a value of 5 and Calabria/Sicily a value of 6. Our division of Italy into these compact 3 or 6 macro-regions accords with the dialect-based classification following Pellegrini (1977).

We note that while it is possible to classify the dialects of Italy into hundreds of different local variants, Pellegrini identifies 7 main Italo-Romance subdivisions. We make use of 6 of his 7 core linguistic regions except for the small "Trentini Centrali" region that is embedded within the "Veneti" region since it does not feature prominently within our data-set. When we restrict attention to 3 regions (especially when we consider non-language features where the breakdown into 6 becomes more subjective) the zones roughly correspond to "North" which includes the "Gallo-Italico" and "Veneti" regions, "Centre" which includes the "Toscani e Corsi" and "Mediano" regions, and "South" which includes the "Meridionali" and "Meridionali Estremi" regions.

Local Origins. We will often categorize respondents as being of "local origin". In order to be classified as having local origins, the respondent and both parents had to be born in the region that contains their current city of residence. For Milan, Turin and Rome the relevant region is Lombardy, Piedmont and Lazio respectively.[7]

We next move on to consider how we measure identity. There are two important dichotomies. First, we distinguish between self-reported identity and identity derived from our cultural indicators based on language and dietary choice which will give us two distinct scores. Second, we need to differentiate between two concepts of identity: the regional measures that will form the main part of our study and a separate set of"closeness" measures that link individuals to specific relatives. We begin by considering regional identity measures.

Regional Identity Measures. When considering regional identity measures, we distinguish between what we call the "regional self-reported index" and the "regional cultural heritage index". The former considers respondent i's responses to a direct question about their own perceived identity and two questions related to birth town and resident town football support. These are weighted using a principal component analysis with factor loadings provided in table 2 with further details in the notes below the

[7]We might go further still and define respondents as having "super local origins" where we also add a further requirement that the maternal grandmother, a relative which we will see has special importance, is also born in the local region. We show in the appendix in various variations on our main regressions that this refinement does not generate a significant change in the results. We do include this measure in our comparison of respondent characteristics by town of residence in table 1.

table. The result is a number that provides an individual-specific measure from 1 to 3 where a lower number indicates a more northern identity. The latter more novel measure instead considers language and dietary preferences and attempts to capture the more subtle influence of cultural heritage on identity. The regional cultural heritage index is computed as a weighted average of a number of questions described in the design section which measure respondent i's average food and dessert preferences, the average dialect spoken and the average performance in the various language tests. The weights are determined by a principal component analysis with factor loadings provided in table 3 with further details in the notes below the table. Scores are coded as 1, 2 or 3 for a northern, central or southern dietary or linguistic choice in each question which generates an overall measure that is once again between 1 and 3, with a lower number again indicating a more northern identity. In effect both regional identity measures provide an indication of the "southernness" of respondents. Any choices that are neutral in nature are controlled for through the use of a dummy (and so do not feature in the identity measure). Neutral answers are of course important in enabling respondents to signal that they have no particular regional preference. For example, consider the question "If someone asked where are you from what would you answer?" The options allowed were: (i) I am European, (ii) I am Italian, (iii) I live in Rome/Milan/Turin but I am from ..., (IV) I am from Rome/Milan/Turin, (V) I am from Rome/Milan/Turin but my family comes from ..., (VI) other. Answers that indicated "European" or "other" would be classified as neutral and are categorized through the use of a dummy (and so do not feature in the identity measure). The other answers would generate a numeral response from 1 to 3 determined by whether the answers was in the north, centre or south of Italy and the score would then be added to the overall measure weighted by the principal component analysis.

Relative-specific Identity Measures. We also consider the "closeness" of a respondent to a particular relative. For example a respondent might have a maternal grandmother from a region in the south of Italy, or a father who was born in a region in the north. In that sense our measures are linked not just to the respondent i, but is also specific to the relative j. This will enable us to determine which relatives are more influential in the development of cultural identity. We refer to the closeness measures as the "score based on self-reported identity" and the "score based on cultural heritage" and will always link

this to a particular relative drawn from the set mother, father, maternal grandmother, maternal grandfather, paternal grandmother, paternal grandfather. For each relative j the measure is a weighted average of a number of different components, with the weights determined by principal component analysis. For the score based on self-reported identity of respondent i with respect to relative j we include the self-reported closeness to relative j, the self-reported closeness to relative j's place of origin, the direct report by the respondent of whether they consider themselves as coming from relative j's place of origin and self-reported support for a football team that is based in the father's place of birth. For the score based on the cultural heritage of respondent i with respect to relative j we include property ownership in relative j's place of origin, together with answers to our questions that seek to test understanding of relative j's regional dialect, closeness to relative j's traditional food and dessert and closeness to relative j's dialect. We control for whether respondent i has neutral preferences in food and dessert and uses neutral language through the use of a dummy which allows us to remove neutral preferences from our measure.

3.2 Identity

Tables 4 and 5 regress our measures of identity (one for each respondent) against the physical distance between the town of residence (Rome, Milan or Turin) of each respondent and the birth place of all family members. Distance from own-birth place is calculated as: code of the region of birthplace (1-6 as shown in figure 1) minus code of region of current residence (on the same scale). Distance from birthplace of any relative is calculated in the same way. For example a positive coefficient in the regression indicates that the birthplace of the relative or respondent is to the south of the current town of residence. In general, our coefficients are positive since our cities were chosen in the centre and north to reflect the flow of immigrants from south to north over Italy's recent history.

When regressing the cultural heritage index on the distance measures, in table 4, we find a highly significant correlation between the regional index and being born in a region further to the south than the town of residence (p-values typically between 0.001 and 0.006) and when looking at the birth-place of

the respondent's parents and grandparents (p-values typically between 0.003 and 0.007). In other words, even controlling for own-birthplace, the birthplace of family members seem to be important determinants of our cultural heritage identity measure.

In contrast, we find no correlation between self-reported identity and the distance from the town of residence of any relative of the respondent in table 5 with own-birthplace and town of residence instead providing the main explanatory variables. It is interesting to note that the distance from own-birthplace is not only significant (typically with p-values between 0.008 and 0.016) but also provides a negative relationship with self-reported identity. In other words, an immigrant who was born further south and then moved to Milan, Turin or Rome, identifies more closely with their new town of residence than someone who was born further north.

What is clear from the data is that self-reported identity fails to pick up what seems to be an important relationship that only becomes apparent when we instead shift attention to our alternative measure of cultural identity based on language and diet: family origin matters for our cultural heritage based measure of identity but not for self-reported identity. In the light of the findings in tables 4 and 5 it seems apparent that self-reported closeness and our alternative index using language and diet provide quite different measures. We can investigate this more directly by examining the importance of each type of relative using our relative-specific identify measures which instead focus on closeness to a specific relative. Figure 2 displays these values with higher bars indicating higher levels of closeness.[8]

Let us first focus on relative-specific identity that is derived from self-reported identity and which is detailed in the top two sets of bar graphs in figure 2. On average parents seem very important for all respondents and also when we restrict our attention to only those with local origins. This importance remains evident across all three towns of residence as shown in the top right set of bar graphs, and in every case parents seem more important than grandparents. In fact the pattern is remarkably consistent as in every single one of the four sets of bar graphs the mother is the family member to which respondents self-reported that they were closest, followed by the father, next paternal grandparents and finally maternal

[8]Note that we have removed any null or invalid responses. Neutral responses are identified by two dummies: one that is 1 for respondents with neutral food preferences and zero otherwise, and another that is 1 for respondents with neutral dessert preferences and zero otherwise. This explains why $N = 719$ in the figure.

grandparents.

The bottom part of figure 2 switches focus to relative-specific identity based instead on cultural heritage. Despite what seems a clear message from the top part of figure 2, glancing at the bottom part presents a very different picture. Our alternative measure based on language and dietary preferences changes the ranking for both the full sample and those with local origins as seen in the bottom left part of figure 2 and for the three different Italian cities in the bottom right part of figure 2: both sets of grandparents are now seen as the closest to respondents with a slight (insignificant) lead by maternal grandparents, mothers follow and fathers are ranked last. This significant shift in the importance of different family members may help to shed some light on why the results in tables 4 and 5 are so different, and also helps to explain how our cultural heritage measure (which preserves the importance of grandparents relative to parents) performs well in the analysis to follow.

3.3 Diet

Before we move on to look at behavior we will first take a closer look at some patterns in the raw data relating to dietary preferences and language. We will also examine the role of the maternal grandmother in more detail in what is to follow as a preface to our behavioral results where we will see that the maternal grandmother appears to have a special importance.

We will start with the consumption of regional food for dessert and at Christmas. We use the trichotomy with "North" (equal to 1), "Centre" (equal to 2) and "South" (equal to 3), which generates a metric that increases in value the further south we move, and so a higher score is indicative of a more southern diet. We also remove "neutral" answers which are not regional before forming our average.

As we might expect since Turin and Milan are located in the north of Italy, there is a higher prevalence of northern food eaten by the residents of those cities at Christmas than in Rome, which is located in the centre of Italy. The distribution of dessert does not show a particular regional trend. This is shown in figure 3 which displays the distribution of dessert and Christmas food consumed by the respondents,

grouped by place of residence. [9]

Figure 4 shows the distribution of the respondents' preferred desserts and Christmas food where we differentiate by the origin of the maternal grandmother of the respondent once again following our trichotomy of northern, central or southern (indicated in dark grey, a mid-level grey and the lightest grey respectively). Recall again that the higher the bars the more southern is the food preference. Respondents are grouped into the full sample (the top two panels) and those with local origins (the bottom two panels) with confidence intervals shown at the 95% level. Looking at the full sample in the top two panels there is a clear pattern of more southern food preferences for both Christmas food and desserts the more southern are the origins of maternal grandmothers. For instance, those with maternal grandmothers from the south of Italy display very statistically significant preferences for southern food easily satisfying the 95% confidence intervals. The bottom two panels switch attention to those with local origins. Recall that having local origins means that the respondent and their parents were born in the region surrounding their current city of residence which makes it all the more remarkable for a maternal grandmother to have a strong effect as they do for instance when considering Christmas food choices for those living in Turin. Nevertheless the effect is much reduced compared to the top two panels as we might expect when parents are also local and with the reduced number of observations when we restrict to only those with local origins.[10]

3.4 Language

We will also take a slightly closer look at language using the same regional trichotomy as for diet. As with diet we remove "neutral" answers which are not regional before forming our average, awarding "1" for regional answers from the North, "2" for answers from the Centre and "3" for answers from the "South", which allows us to once again say that the higher is the score for each respondent, the more

[9] Note that the number of observations is restricted to remove respondents who display neutral food preferences in order to only consider those with particular regional food preferences, this gives us $N = 682$ in the left panel (which considers Christmas food) and $N = 591$ in the right panel which considers dessert preferences.

[10] As shown in figure 4 we have $N = 455$ and $N = 470$ for Christmas food and desserts respectively for those with local origins as opposed to $N = 984$ and $N = 1086$ for Christmas food and desserts respectively for the full sample).

southern is their linguistic preference or understanding.

Figure 5 summarizes the core results. Once again we either consider the full sample of survey respondents or only those with local origins. We see a pattern for language that appears similar to the pattern for diet. Having a maternal grandmother from a southern part of Italy greatly pushes up the scores especially for the full sample where our 95% confidence intervals are easily satisfied. This is especially evident when comparing those with maternal grandmothers from the north of Italy as opposed to those with maternal grandmothers from the central or southern regions. Across the full sample the pattern is very similar when dividing our sample by town of residence in the bottom two panels, especially for the full sample. For the sample restricted to local origins the pattern remains similar for the two northern cities, with a flatter distribution for Rome which is of course categorized as being central itself.[11]

Since language data can be estimated at the level of six different regions rather than three we also present figure 6 which provides an alternative to figure 5.[12] We first note a pattern of increases in the bar charts towards being more southern as we move towards cultural identity that is tilted towards the south of Italy both in terms of own birth region and the birth region of the maternal grandmother.[13] We also note a complete absence of any pattern when we instead consider self-reported identity. In other words, our novel measure of cultural identity provides an unambiguously clearer pattern even when we consider the case of six regions.

3.5 Opinions about Others

While dietary and linguistic differences may be interesting, and in some cases it might be surprising that the maternal grandmother exercises a significant influence, we also need to be sure that we have consistency with more general attitudes and behavior. Our data does provide clear evidence that cultural

[11]Note again that the sample with local origins is lower ($N = 719$) than for the full sample ($N = 1546$) as shown in figure 5.

[12]The alternative six region characterization is discussed in section 3.1 and can be seen in figure 1.

[13]Note that the first pair of regions (shown as columns 1 and 2) are hard to interpret as more northern or southern than each other in 1. However, column 3 (which is to the north of the central region in the trichotomy) vs column 4 (a more southern central region) or column 5 (the northern part of the south of Italy) vs column 6 (the deep south) seem more clearly ordered.

identity (as opposed to self-reported identity) has a significant effect on behavior.

The clearest way to support this is provided in a Figure 7 which displays the proportion of respondents who believe that people are overall not honest, not helpful and not trustworthy, derived from the answers to "most people would try to take advantage of others", "most people think mostly about themselves" and "one has to be very careful because you cannot trust people" respectively. The answers are broken down by the area of origin of the maternal grandmother, and we also consider the full sample and those with local origins separately.

Starting with the full sample and with honesty, we note that respondents with maternal grandmothers from the north of Italy seem to have a higher opinion of the honesty of others as compared to those with maternal grandmothers from further south, and this is supported by the displayed 95% confidence intervals. There is a similar pattern for the helpfulness of others and trust in others though the scale of the effect is diminished to the extent that in some cases we do not have 95% confidence in the results. Moving to the sample with local origins we see a very clear level of differentiation between those with maternal grandmothers from the north, central or southern area of Italy in terms of the honesty of others. There is a distinction between those with southern maternal grandmothers and the rest in terms of their belief in the helpfulness and trustworthiness of others, but with no distinction between those from central or northern Italy. Perhaps the most stark comparison occurs when comparing those with maternal grandmothers from the north and those with maternal grandmothers from the south: in every panel we have a clear pattern of lower levels of opinions from those with southern maternal grandmothers achieving 95% confidence even for those who were born in their city of residence and whose parents were also born in the local region.

3.6 Behavior, Cooperation and Honesty

So far we have detected important effects coming from family origins and feeding into opinions about others. These opinions, especially concerning trust, are in turn likely to have a knock-on effect on behavior, but the scale of this effect can only be determined empirically. We next attempt to quantify the

behavioral ramifications by examining behavior in our two different incentivized games.

We can start by providing some suggestive graphical analysis designed to give us an early indication of correlations between identity (measures in terms of cultural heritage or self-reported identity) and behavior. We note a negative correlation between our cultural heritage measure of identity and average contributions in the public goods (in the top left panel of figure 8) and another negative relationship between cultural heritage and the respondents' guesses about the contributions of their partners which we label as the "public good contribution of others" (in the centre left panel of figure 8). Using the same scale, we see a largely flat relationship when we instead consider self-reported identity in the respective right hand side panels of figure 8. We also note here that both measures of identity show no distinct pattern between southern or northern identity and honesty, as indicated by the two bottom panels. This suggests that while there is regional variation in attitudes towards cooperation this does not extend to honesty. This is merely suggestive but we will see a similar pattern when we conduct a regression analysis of the data later in this section.

Figure 9 provides an indication of the importance of regional identity: in the top left we see that own region has a significant effect on contributions particularly when we compare those who were born in the north to those born in the south. This is even clearer when we consider the mother or maternal grandmother's region of birth as shown in the top middle and top right panels respectively. There is a similar marked difference in beliefs about the average contributions of others between those with northern identities and southern identities, which appears similar whether we consider own region of birth, or the birth region of the mother or maternal grandmother (middle left, middle centre and middle right panels respectively). Once again we see no difference in terms of behavior in the lying game: own birthplace, mother's birthplace and maternal grandmother's birthplace seems to be irrelevant (as seen in the bottom three panels). This matches closely the pattern in figure 8: those who are from the south expect to see lower contributions from others and themselves contribute less in the public goods game which suggests that levels of trust and cooperation are higher for those who are from the north or whose family comes from the north, but this is quite distinct from honesty levels which are similar across Italy. We might conjecture that it is trust levels that are driving down cooperation levels, but that this is independent of

base levels of honesty.

Our final figure, figure 10, examines the data at the level of the three cities and allows us to consider the transmission of cultural identity over time by directly comparing the behavior of locals with third generation immigrants. Our definition of local here is a strong one: we require the respondent together with all of their parents and grandparents to have been born in the region surrounding the current city of residence. Our definition of third generation immigrant requires that recipients and their parents were born in the region surrounding the current city of residence but that their grandparents were born outside the local region. The difference between these groups is therefore entirely determined by the birthplace of grandparents. We see that despite sharing the same birthplace and having parents who were also born in the same region, behavior is quite different when we consider contributions in the public goods game (the top two panels) and beliefs about the contributions of others (the middle two panels). This is true for the full sample (the top left and middle left panels) where we can see the differences between the bar charts satisfy 95% confidence. In the top right and middle right panels we compare the same sample of local respondents with a sub-sample of third generation immigrants with southern roots and we see a similar pattern albeit with reduced significance from the Turin sub-sample likely because of reduced power.[14]

While figures 8, 9 and 10 provide a clear and consistent story they are devoid of controls. Table 6 presents a regression analysis that provides a fuller indication of the extent to which our cultural heritage index is indeed a better predictor of behavior than self-reported identity. The first two columns of table 6 present regressions of the amount contributed in the public goods game, against our cultural heritage index derived from the linguistic and dietary measures (column 1) and the self-reported identity index (column 2). What is clear from the results is that the measure based on language and diet is significant when considering contributions (p-value $= 0.008$), while the self-reported measure is not (p-value $= 0.986$). The effect is also quite large: an average contribution of 0.79 units lower for those with cultural ties to the south of Italy, which represents a contribution size of around 7.5% lower than the average

[14]The sample size for non-immigrants in figure 10 is $N = 372$. This falls to $N = 197$ when we consider third generation immigrants from the full sample, and down to $N = 81$ when we restrict attention to third generation immigrants with southern roots.

contribution. Column 3 performs a "horse race", pitting the cultural heritage index against self-reported identity with the cultural heritage index performing well p-value < 0.001) while self-reported identity performs poorly (p-value $= 0.876$). The overall message across all of these specifications is that self-reported identity is incapable of predicting behavior in the public goods game while the cultural heritage index derived from linguistic and dietary preferences is a highly significant predictor of behavior.

Columns 4, 5 and 6 in table 6 switch attention to a raw measure of dishonesty derived from reported numbers of heads in the coin-flipping game. We rewarded those who reported a number of heads of eight or more from ten flips. While we cannot know for sure that anyone reporting eight or more heads was lying it is much more likely that they are lying relative to those who reported seven or fewer heads and did not receive a bonus payment.[15] What is apparent from columns 4, 5 and 6 is that neither measure of identity predicts honesty: we cannot see any relationship between cultural or self-reported identity and the propensity to lie in our game which is entirely consistent with our earlier findings: since identity (northern, central or southern) does not correlate with honesty we would not expect either measure to have predictive power.[16,17]

Table 7 conducts a different exercise, considering the relationship between a respondent's public good contribution and the weighted average contribution of those born in the birth region of their grandparents (with weights derived from the relative-specific identity measures for cultural heritage to give us an idea of which grandparents matter more for cultural identity, as indicated by the term "CH weights" in the table). The first column indicates that there is some link between respondent's contributions and the average from their grandparents' place of birth (p-value < 0.05). This remains true when controlling for the respondent's own birth region (column 2). The relationship weakens somewhat when we also include the average contributions of those born in the birth region of their grandparents with weights

[15] As noted earlier we observed approximately six times as many reports of 8 or more heads as would be likely to occur with honest reporting.

[16] See figure 9 and the surrounding discussion in section 3.6.

[17] We also provide several variations of table 6 in the appendix which explore the role of having local or super-local origins as well as considering alternative ways to measure cultural heritage and self-reported identity. In each case the qualitative story remains the same: cultural heritage performs very well as a predictor of behavior in the public goods game while self-reported identity does not, and there is no indication that southern or northern identity is linked to differences in underlying honesty.

instead derived from self-reported closeness, though note that this self-reported measure is itself highly insignificant (and we can consider this another "horse race" between cultural heritage and self-reported identity albeit in a different context). Note that columns 1 to 3 use the 6 macro-regions while columns 4 to 6 repeat the same exercise using our standard trichotomy of Italy.[18] Finally, also note that we excluded the respondent's own contribution when considering the average from those born in the grand-parental birth region to avoid endogeneity. Table 8 instead considers the relative importance of different family member birth regions with columns 1-6 considering the mother, father, maternal grandmother, paternal grandmother, maternal grandfather and paternal grandfather respectively. We see here a very clear story indicating that the most powerful predictor of behavior in the public goods game from among the various family members is the maternal grandmother (p-value < 0.01 in column 3) which supports the notion that the maternal grandmother plays an important role in cultural transmission among our respondents.[19]

4 Discussion

Cooperation and trust are key components of social capital and are vital for a well-functioning society. Social capital in turn has a strong positive relationship with health, income equality, economic growth, trade, well-functioning institutions, child welfare, education and financial development, and a strong negative relationship with corruption and crime. It seems likely that culture and identity are linked to cooperation and trust, but conventional methods of measuring identity are too subjective to shed much light on the underlying relationship. This is something we confirm in our own study: a traditional form

[18] Recall that our standard trichotomy has a more general focus on food and language, while the 6 macro-region variation is more directly focused on language.

[19] We also note the correlation between regional blood donation data, a common measure of social capital taken from Nannicini et al. (2013) with (a) the respondents' public goods game contributions and with (b) our survey measures of trust (taken from the average response to the questions on the honesty of other people, the helpfulness of other people and the trustworthiness of other people. We see a positive and significant correlation with regional blood donation data taken if we consider the regional birthplace of the respondent's maternal grandmother as the relevant region of comparison. For (a) the pairwise correlation coefficient is 0.5363 (p-value = 0.0393) and for (b) the pairwise correlation is 0.4987 (p-value = 0.0585). We find no meaningful correlation when we instead consider the region of the birthplace of the respondent in our comparison (we instead find pairwise correlations of -0.0868 (p-value = 0.7584) and -0.0463 (p-value = 0.8699) respectively).

of self-reported identity displays no statistical relationship with our measures of cooperation and trust. This might be because there is no underlying relationship between identity and cooperation and trust, or it could be because the measures for any of these three concepts are suspect. The measures we use for cooperation and trust have been successfully employed elsewhere and utilize features such as incentives and ease of measurement (public goods games in particular have been heavily used at least partly for this reason). We argue that it is the measure of identity that is suspect: partly because the concept of identity may be so ingrained that it may lie in the subconscious, making a self-reported subjective measure inappropriate. In response, we develop a new way to measure identity through linguistic and dietary preferences among internal migrants. The techniques we employ are more objective than a self-reported measure since they test knowledge directly, asking respondents whether they understand forms of dialect which reveals the extent to which they may have been influenced by their family history. We supplement this with subjective elements, particularly from diet though there is a degree of masking even there: it is not obvious that food preferences reveal identity though again, they may reveal a subconscious predilection towards certain regional types of food. What matters from our perspective is of course whether our measure has predictive power and this also doubles as a test of the veracity of the measure. We find that there is indeed a strong relationship between cooperation and trust and our new measure of identity and using this new measure, we reveal and explain the important role played by family background, and especially grandparents, and the slow rate at which social norms change over time.

Our results also show the importance of grandparents when considering the cultural identity and behavior of the respondents in our study. When thinking about how grandparents influence the behavior of their grandchildren, one thing that might come to mind is genetics. While genetics may play an important role (Mann, 1994), parents typically have more genetic material in common with their children than do grandparents and so the *independent* influence that grandparents (especially the maternal grandmother) have on identity, attitudes and behavior is likely to reflect more than genetics. This is especially true in our results which showcase the important role that maternal grandmothers play in building linguistic and dietary identity and even in shaping social capital and behavior. We are not alone in highlighting the importance of grandparents. Using data from the British Household Panel Survey and the UK House-

hold Longitudinal Survey, Zhang and Li (2019) show that grandparents have a significant impact on occupational aspirations, educational attainment and class that is independent of the effect that comes through parents. Work of this sort is part of an older literature that stresses the importance of early life experiences on behavior (Hebert-Myers et al., 2006) which may of course provide a way for grandparents to exercise influence.The importance of grandparents highlighted in this literature may explain why the effects we see are strong even when parents are born near the city of residence, with grandparents providing the link to a cultural past that might otherwise be lost.

Notwithstanding the role of maternal grandparents in Italy, it is perhaps not surprising that those with a cultural heritage that hails from one part of a country might still retain knowledge of dialects or enjoy food that provides them with some cultural identity. What is surprising is that social norms can also transfer in this way: the single most surprising and perhaps important result in our work is to show that the regional identity of our grandparents (especially maternal grandmothers) might still play an important role in determining our own attitudes towards others, and this role is important whether we realize it or not. Self-reported identity does not play such a role, and using self-reports as a measure of identity fails to reveal the importance of identity as a predictive variable. The ramifications are significant: given the extraordinary importance of cooperation and trust as key elements of social capital which in turn has a strong effect on everything from crime rates and longevity to economic growth and trade, the fact that behavior may take several generations to change, even in the face of migration to a very different environment, tells us that bad experiences in any form of social or state interaction can have very long-lasting repercussions on future behavior.

Like many researchers before us, we also wish to highlight Italy as a useful laboratory in which to examine concepts of national and regional identity. Italy has only developed as a nation state quite recently in Western history, and has a strong tradition of regional independence, significant linguistic differences across the country, high levels of internal migration (especially in the aftermath of World War II) and extreme differences in levels of social capital between the north and south (Putnam et al., 1993; Bigoni et al., 2019). The seminal work on social capital in economics by Putnam (1993) makes good use of many of these features but aims to draw general conclusions from the Italian experience:

of crucial importance is that all countries have regional variation in social capital and in culture but the quality of data and simple nature of migration in Italy (largely from south to north) make it relatively easy to use Italian data.

References

BIGONI, M., S. BORTOLOTTI, M. CASARI, AND D. GAMBETTA (2019): "At the Root of the North–South Cooperation Gap in Italy: Preferences or Beliefs?" *Economic Journal*, 129, 1139–1152.

BRACCO, E., F. LIBERINI, B. LOCKWOOD, F. PORCELLI, M. REDOANO, AND D. SGROI (2019): "Social Capital, Electoral Behaviour and Government Performance in Italy: the Role of Culture and Family Origin," *AEA RCT Registry*.

BUONANNO, P., D. MONTOLIO, AND P. VANIN (2009): "Does Social Capital Reduce Crime?" *The Journal of Law & Economics*, 52, 145–170.

COLEMAN, J. S. (1988): "Social Capital in the Creation of Human Capital," *American Journal of Sociology*, 94, S95–S120.

CÔTÉ, S. AND T. HEALY (2001): *The Well-Being of Nations: The Role of Human and Social Capital. Education and Skills.*, ERIC.

GUISO, L., P. SAPIENZA, AND L. ZINGALES (2004): "The Role of Social Capital in Financial Development," *American Economic Review*, 94, 526–556.

——— (2008a): "Long Term Persistence," NBER Working Papers 14278, National Bureau of Economic Research, Inc.

——— (2008b): "Social Capital as Good Culture," *The Journal of the European Economic Association*, 6, 295–320.

——— (2016): "Long Term Persistence," *Journal of European Economic Association*, 14, 1401–1436.

HA, SHANG E.AND JANG, S.-J. (2007): "National Identity, National Pride, and Happiness: The Case of South Korea," *Social Indicators Research*, 121, 471–482.

HEBERT-MYERS, H., C. L. GUTTENTAG, P. R. SWANK, K. E. SMITH, AND S. H. LANDRY (2006): "The Importance of Language, Social, and Behavioral Skills Across Early and Later Childhood as Predictors of Social Competence With Peers," *Applied Developmental Science*, 10, 174–187.

HELLIWELL, J. AND R. PUTNAM (1995): "Economic Growth and Social Capital in Italy," *Eastern Economic Journal*, 21, 295–307.

KAWACHI, I., B. P. KENNEDY, K. LOCHNER, AND D. PROTHROW-STITH (1997): "Social capital, income inequality, and mortality." *American Journal of Public Health*, 87, 1491–1498, pMID: 9314802.

KNACK, S. (2002): "Social Capital and the Quality of Government: Evidence from the States," *American Journal of Political Science*, 46, 772–785.

KNACK, S. AND P. KEEFER (1997): "Does Social Capital Have an Economic Payoff? A Cross-Country Investigation," *The Quarterly Journal of Economics*, 112, 1251–1288.

MANN, C. C. (1994): "Behavioral genetics in transition: a mass of evidence - animal and human - shows that genes influence behavior. But the attempt to pin down which genes influence behaviors has proved frustratingly difficult," *Science*, 264, 1686.

NANNICINI, T., A. STELLA, G. TABELLINI, AND U. TROIANO (2013): "Social capital and political accountability," *American Economic Journal: Economic Policy*, 5, 222–50.

PELLEGRINI, G. B. (1977): *Map of the dialects of Italy*, Pacini.

PUTNAM, R. D. (2000): *Bowling Alone: The Collapse and Revival of American Community*, New York: Simon & Schuster.

PUTNAM, R. D., R. LEONARDI, AND R. NANETTI (1993): *Making Democracy Work: Civic Traditions in Modern Italy*, Princeton: Princeton University Press.

TODD, J. (2007): "National identity in transition? Moving out of conflict in (Northern) Ireland," *Nations and Nationalism*, 13, 565–571.

WILKINSON, R. (1996): *Unhealthy Societies: The Afflictions of Inequality*, Studies in culture and communication, Routledge.

ZAK, P. J. AND S. KNACK (2001): "Trust and Growth," *The Economic Journal*, 111, 295–321.

ZHANG, M. AND Y. LI (2019): "Family fortunes: The persisting grandparents' effects in contemporary British society," *Social Science Research*, 77, 179–192.

Acknowledgments

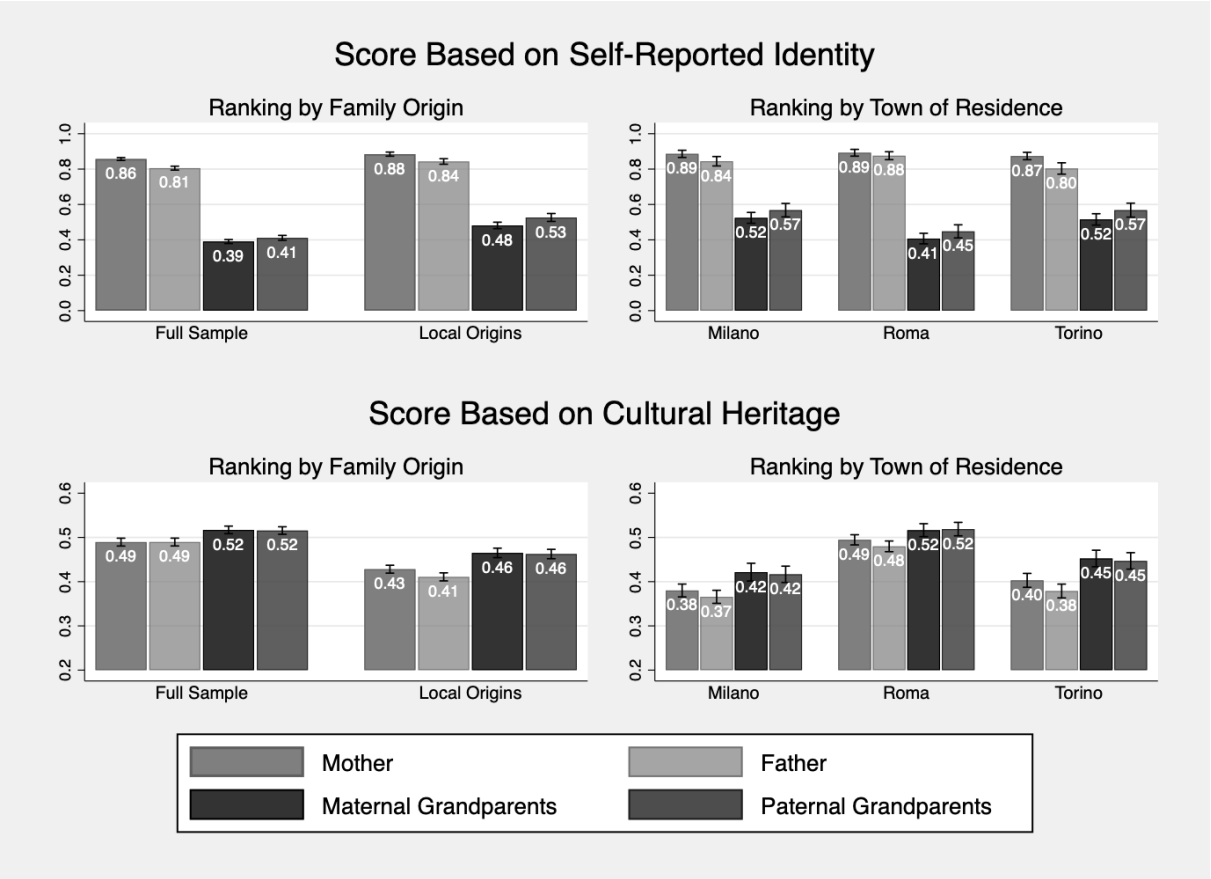
Funding for this project was provided by the ESRC CAGE Centre, ESRC Grant Ref RES-626-28-0001. Ethical approval granted by the University of Warwick HSSREC (Ref HSS 64/18-19). Registered on the AEA RCT registry (https://doi.org/10.1257/rct.3816-1.0). The authors declare no conflicts of interest.

Figures and Tables

Fig. 1: Map of Italian Regions

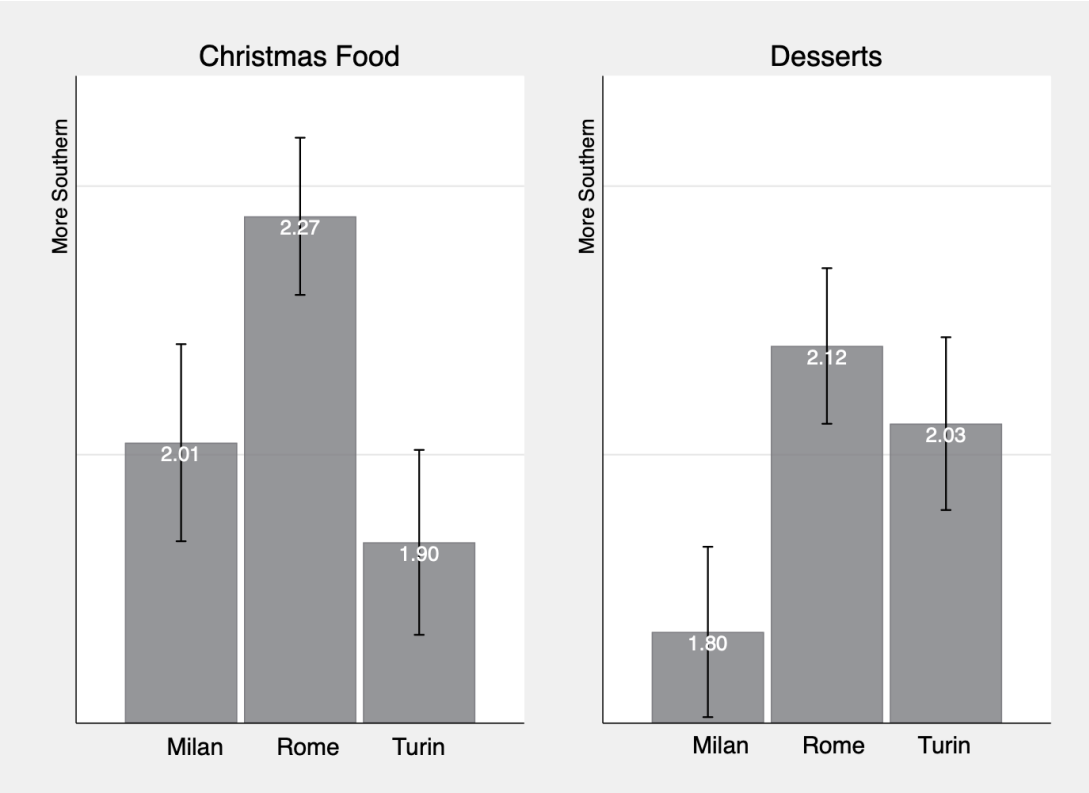
Notes: The figure illustrates our trichotomy of Italy into three regions: the "north" (the darkest grey) which is assigned a value of 1 in the main text, "centre" (a mid-level grey) which is assigned a value of 2, and "south" (the lightest grey) which is assigned a value of 3. The three can be further divided into six as indicated by the darker and lighter shades. This six region variation is used as indicated in the main text in situations where diet is not considered.

Fig. 2: Relative-specific Identify Scores based on Self-Reported Identity and Cultural Heritage



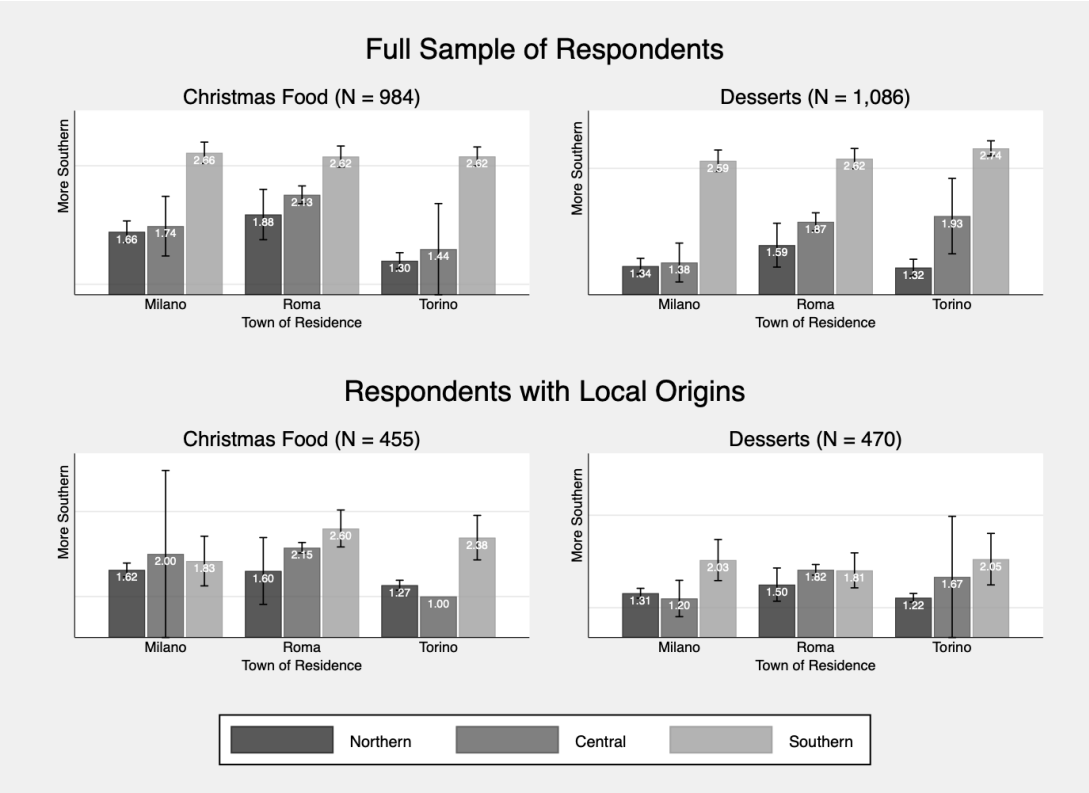
Notes: "Ranking by Town of Residence" includes only respondents defined to have a local origin. A respondent is classified as having local origins if the respondent and both parents were born in the region where the respondent currently resides, $N = 719$. A full definition of the "Self-Reported Identity" and "Cultural Heritage" measures is provided in Section 3.1. Confidence intervals are shown at the 95% level.

Fig. 3: Food Preferences over Town of Residence



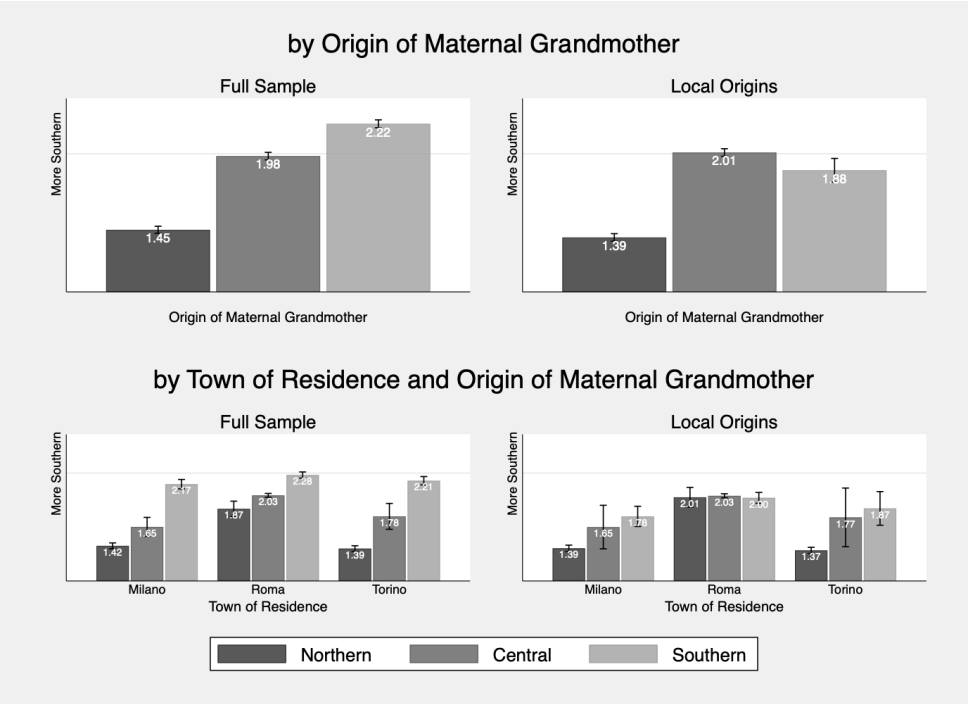
Notes: In this figure we include all respondents who reported that they did not consume neutral (non-regional) food for Christmas (left chart, $N = 682$) or dessert (right chart, $N = 591$) and then group these respondents by place of residence. We see a higher prevalence of northern food eaten at Christmas by the residents of the two northern cities, Milan and Turin, than in Rome, which is located in the centre of Italy. The distribution of dessert does not show a particular regional trend. Confidence intervals are shown at the 95% level.

Fig. 4: Food Preference by the Origin of the Maternal Grandmother



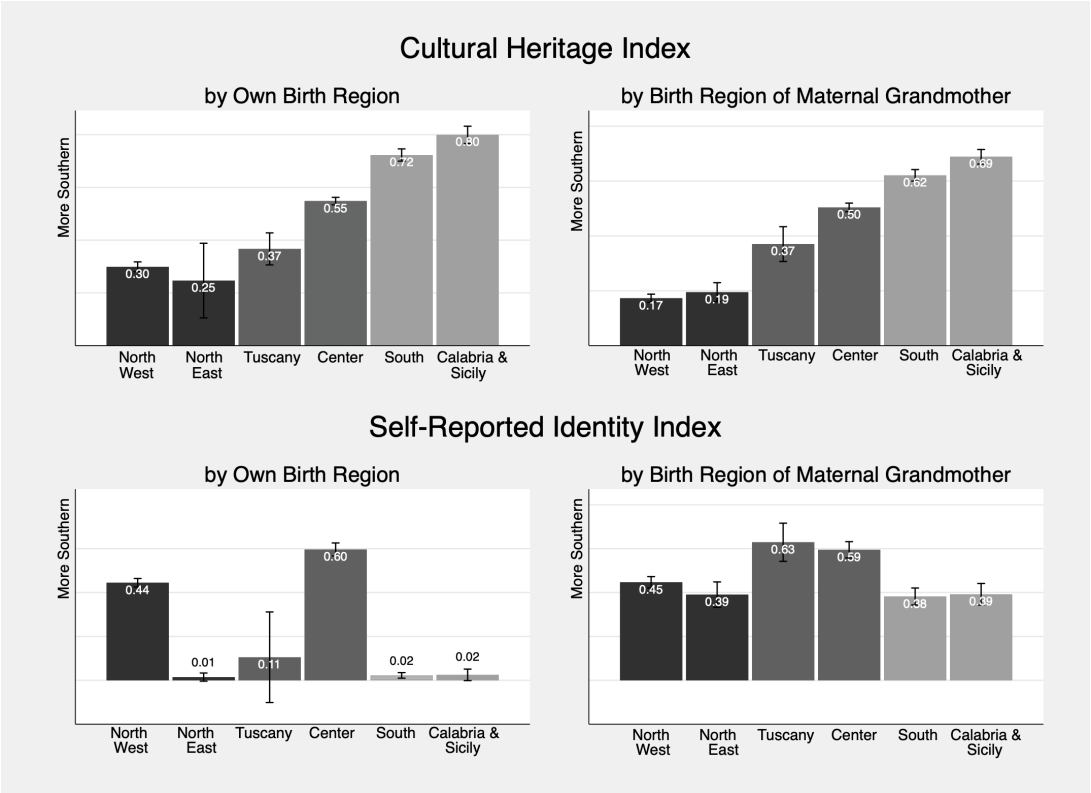
Notes: A respondent is classified as having local origins if the respondent and both parents were born in the region where the respondent currently resides. This results in a reduced number of observations: for Christmas food and desserts $N = 455$ and $N = 470$ respectively, as compared with the full sample where for Christmas food and desserts $N = 984$ and $N = 1086$ respectively. Even for the full sample our number of observations is below the total surveyed since we exclude those who have neutral food preferences. Confidence intervals are shown at the 95% level.

Fig. 5: Spoken Dialect by Town of Residence and Origin of Maternal Grandmother



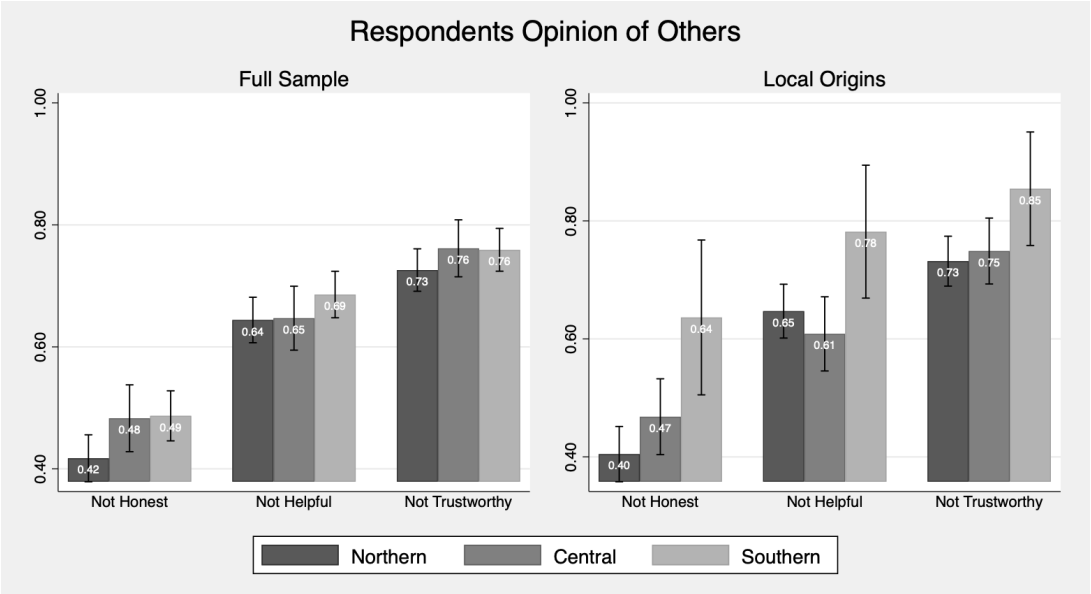
Notes: A respondent is classified as having local origins if the respondent and both parents were born in the region where the respondent currently resides. The dialect spoken by the respondent is classified as being Northern, Central or Southern based on the methodology described in Section 3.1. Sample size is $N = 1546$ for the full sample and $N = 719$ for the sample with local origins. Confidence intervals are shown at the 95% level.

Fig. 6: Identity at the Macro-Regional Level by Town of Residence and Origin of Maternal Grandmother



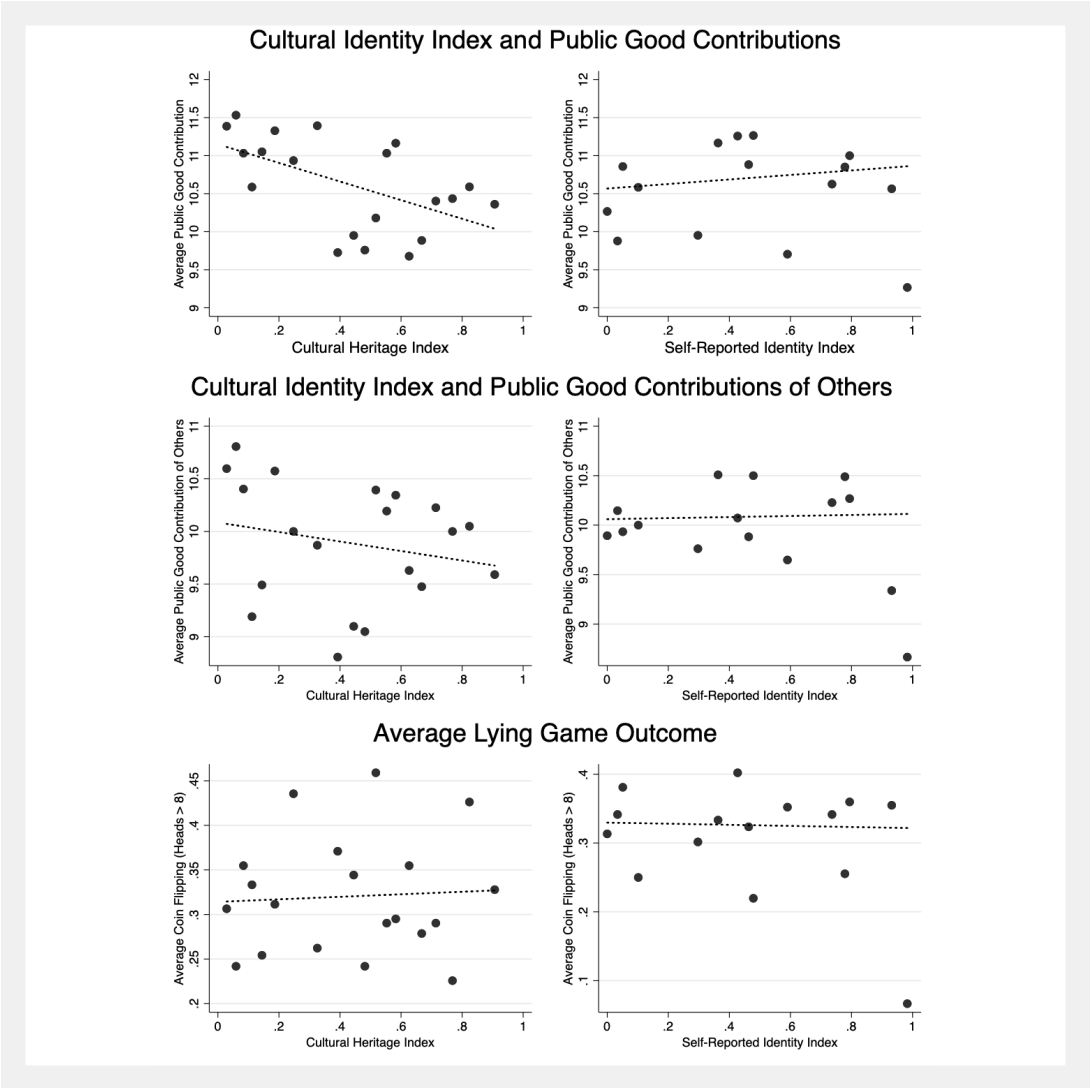
Notes: Respondents are grouped on the basis on their own birth macro-region (left panels) and of the birth macro-region of their maternal grandmother (right panels). For each group, the bins in the top panels report the average cultural heritage index by group and the bins in the bottom panels report the average self-reported identity index. These vary between 0 and 1, with values closer to 1 indicating the identity of the respondent is culturally closer to the south. Sample size is $N = 1230$. Confidence intervals are shown at the 95% level.

Fig. 7: Opinions over Attributes



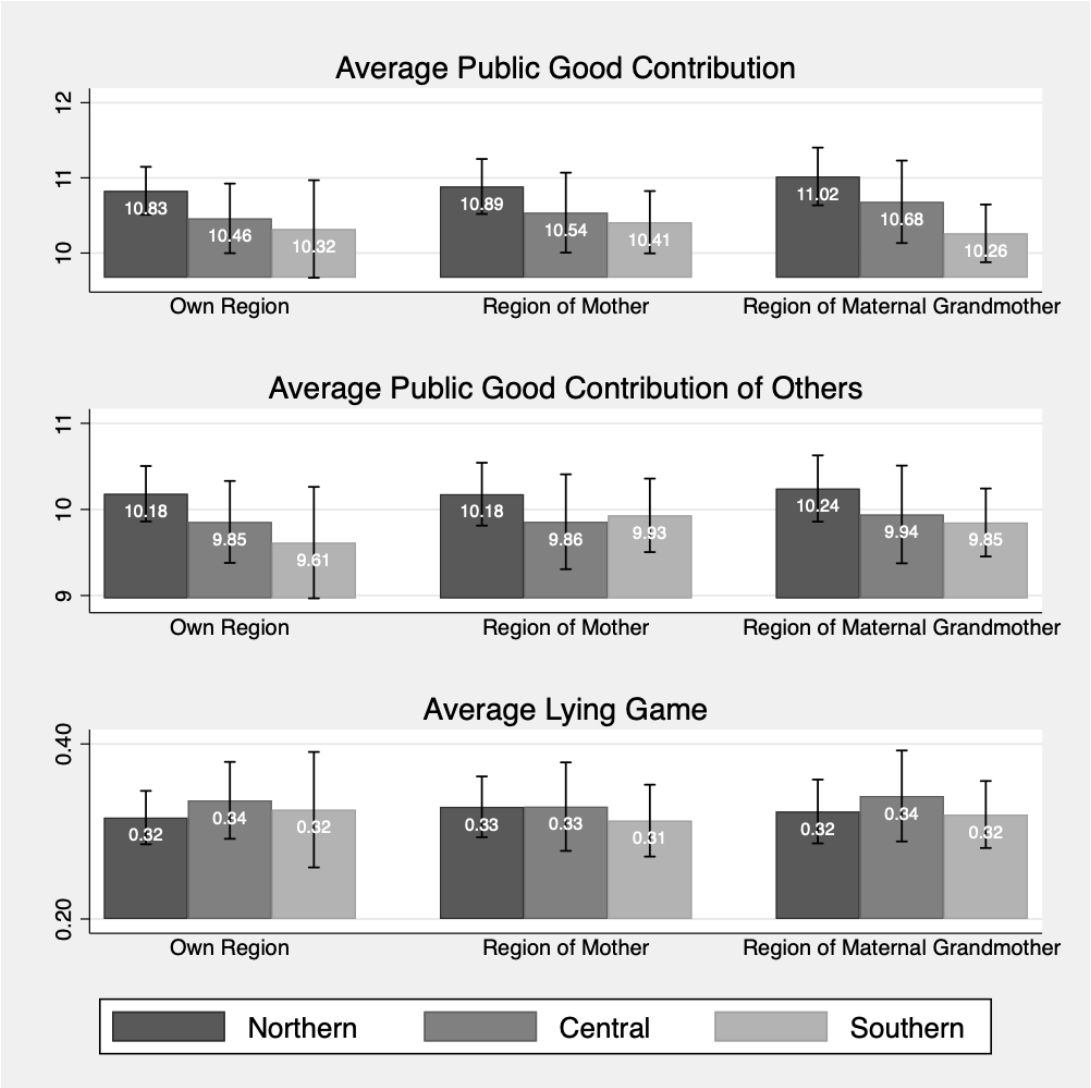
Notes: The bars represent the share of respondents believing others are not honest ("most people would try to take advantage of others"), not helpful ("most people think mostly about themselves") or not trustworthy ("one has to be very careful because you cannot trust people"). Sample size is $N = 1547$ for the full sample and $N = 719$ for the sample with local origins. Respondents are grouped according to the region of birth of their maternal grandmother. Confidence intervals are shown at the 95% level.

Fig. 8: Correlations between Identity and Behavior



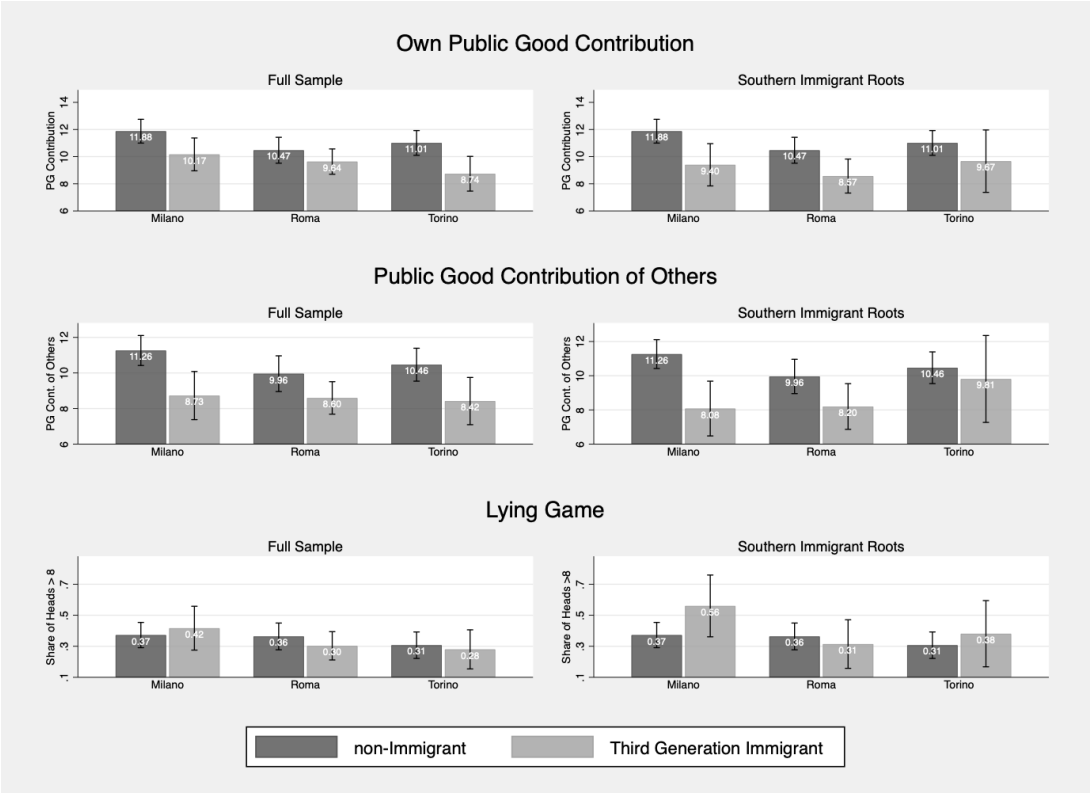
Notes: The binned scatter-plots provide non-parametric plotting of the average game outcome for each value of the cultural heritage index (left panel) and of the self-reported identity index (right panel).

Fig. 9: Behavior by Geographical Region



Notes: The figure shows average game outcomes across respondents: "Average Public Good Contribution" is the public good contribution of the respondent in the incentivized public goods game, "Average Public Good Contribution of Others" is the guess made by the respondent of the contribution that their partner will make and "Average Lying Game" is the share of respondents reporting a number of heads bigger than 8 in the coin-flipping game. Results are grouped by region of origin: own-birth region, the birth region of the respondent's mother and the birth region of the respondent's maternal grandmother. Sample size is $N = 1547$. Confidence intervals are shown at the 95% level.

Fig. 10: Behavior of Non-immigrants and Third Generation Immigrants



Notes: In the figure "Average Public Good Contribution" is the public good contribution of the respondent in the incentivized public goods game, "Average Public Good Contribution of Others" is the guess made by the respondent of the contribution that their partner will make and "Average Lying Game" is the share of respondents reporting a number of heads bigger than 8 in the coin-flipping game. The figure shows average game outcomes for two groups of respondents: those defined as non-immigrants and those defined as third generation immigrants. The former are all the respondents who were born in the same region where they reside and whose parents and grandparents were all born in the same region. The latter are, instead, all the respondents who were born in the same region where they reside, with parents born in the same region, but with at least one of the maternal grandparents born in a different region. The right panel restricts third generation immigrants to those whose migrant maternal grandparent has southern origins. Sample size for the non-immigrants is $N = 372$. Sample size for third generation immigrants is $N = 197$ for the left panel and $N = 81$ for the right panel when we further restrict the sample to only those with southern roots. Confidence intervals are shown at the 95% level.

Table 1: Respondent Characteristics

	MILAN (N = 514)		ROME (N=517)		TURIN (N=516)	
	Mean	SD	Mean	SD	Mean	SD
Local origins*	0.4903		0.4836		0.4205	
Super Local origins* !!!	0.4125		0.4236		0.3469	
Female* =1 if female	0.5175		0.5455		0.5601	
Married* =1 if married	0.4553		0.4101		0.4147	
Children* =1 if R has children	0.4883		0.4410		0.4244	
Degree* = 1 if R has University degree	0.3443		0.3927		0.3334	
Income* = 1 if R's income is 30K Euros +	0.5038		0.4333		0.4070	
Age	41.0564	18.8234	39.2447	17.1760	41.5039	16.0603
Age !!!	27.0195	18.84904	25.205	17.20388	27.4942	16.07576
Geographical identity:						
Cultural heritage index	0.3275	0.2671	0.5820	0.1565	0.3889	0.2904
Self-reported index	0.4039	0.2895	0.5441	0.3150	0.4328	0.2917
Public good own-contribution	11.0448	4.6912	10.4197	5.0309	10.5039	4.8498
Public good expected partner contribution	10.2023	4.8044	9.8201	5.0850	10.0446	4.9424
Coin-flipping (8 or more heads)*	0.3288	0.4702	0.3289	0.4702	0.3140	0.4645
Trust in state*	0.3145	0.4648	0.2964	0.4571	0.2691	0.4439
Trust in police*	0.7080	0.4551	0.6186	0.4862	0.6383	0.4810
Trust in church*	0.3505	0.4776	0.3113	0.4635	0.2556	0.4367
People not helpful*	0.6634	0.4730	0.6712	0.4702	0.6492	0.4777
People not honest*	0.4202	0.4941	0.4913	0.5004	0.4612	0.4990
People not trustworthy*	0.7335	0.4426	0.7698	0.4214	0.7306	0.4441
Acceptable benefits*	0.8735	0.3327	0.8859	0.3183	0.9244	0.2646
Acceptable no tax*	0.8988	0.3018	0.8801	0.3252	0.8915	0.3113
Acceptable corruption*	0.9066	0.2913	0.8956	0.3061	0.9264	0.2614
Acceptable no TV licence*	0.8249	0.3804	0.7911	0.4069	0.8062	0.3957

Notes: * Denote dummy variables. Local origins is set equal to one when the respondent and both parents were born in the same region as the city of residence and zero otherwise. Public good own contribution is the amount in euros the respondent contributes in the public goods game. Public good expected partner contribution is the amount in euros the respondent reports they believe their partner will contribute. Coin-flipping (8 or more heads) is set equal to one if the respondent declares 8 or more heads in the coin-flipping game. "People not helpful" is set equal to one if the respondent replies that: "People are usually selfish" to the question: "Do you think people usually want to help each other or do you think they are usually selfish?" and zero otherwise. "People not honest" is set equal to one if the respondent replies that: "People would take advantage of me" to the question: "Do you think that people, given the chance, would take advantage of you or would they act honestly?" and zero otherwise. "People not trustworthy" is set equal to one if the respondent replies that: "You must be very cautious" to the question: "Do you think you should trust the majority of people, or do you think you should be cautious with people?" and zero otherwise. Acceptable Benefits, no Tax, Corruption, no TV licence are dummies set equal to one if the respondent states that claiming non deserved benefits, evading taxes, receiving bribes or not paying TV licence are not acceptable behaviors. The variables are coded equal to one if the respondent replied 4 or 5 to the questions that these behavior are 1=always acceptable to 5=never acceptable.

Table 2: Polychoric Factor Loadings for the Self-reported Index of Geographical Identity

Variable	Mean	SD	Factor1	Uniqueness
Self-Reported Region	0.189	0.230	0.3585	0.8715
Birth Town Football Team Supporter	0.434	0.496	0.3942	0.8446
Resident Town Football Team Supporter	0.527	0.499	0.4968	0.7532

Notes: The table presents means, standard deviations and factor loadings for the variables used to build the self-reported index of geographical identity. All variables range between 0 and 1. This variable was built by taking into account how close the respondent felt to the region of each of their family members. The other two variables are binary indicators with value 1 if the respondent supports either the football team of his birth town or the football team of their town of residence.

Table 3: Factor Loadings for the Cultural Heritage Index of Geographical Identity

Variable	Mean	SD	Factor1	Uniqueness
Standardized Mean Food Preference	2.17e-08	1	0.8072	0.3146
Standardized Mean Dessert Preference	1.05e-09	1	0.9107	0.1687
Standardized Mean Dialect Used	9.50e-10	1	0.8821	0.2188
Dummy for Neutral Food	.4395604	.4964941	-0.0106	0.3904
Dummy for Neutral Dessert	.2979961	.4575257	-0.0371	0.4918
Count of Words in Neutral Language	1.321267	1.213691	-0.0943	0.9782

Notes: The table presents means, standard deviations and factor loadings for the variables used to build the cultural heritage index of geographical identity.

Table 4: Regression of Cultural Heritage Index of Geographical Identity (identity drawn from language and diet data)

Dependent Variable: Cultural Heritage Index	(1)	(2)	(3)	(4)	(5)	(6)
Town of Residence:						
Rome	0.402***	0.388***	0.426***	0.412***	0.395***	0.397***
	(0.00932)	(0.00286)	(0.00717)	(0.00779)	(0.00263)	(0.00321)
Turin	0.0136*	0.00491	0.0105	0.00816	0.00331	0.0104
	(0.00430)	(0.00204)	(0.00672)	(0.00637)	(0.00511)	(0.00413)
Distance from:						
Own-birthplace	0.0727***	0.0630***	0.0659***	0.0686***	0.0682***	0.0689***
	(0.00167)	(0.00587)	(0.00336)	(0.00137)	(0.00219)	(0.00362)
Mother birthplace	0.0646***					
	(0.00618)					
Father birthplace		0.0727***				
		(0.00676)				
Maternal GM birthplace			0.0755***			
			(0.00488)			
Maternal GF birthplace				0.0704***		
				(0.00453)		
Paternal GM birthplace					0.0696***	
					(0.00393)	
Paternal GF birthplace						0.0673***
						(0.00447)
Family Heterogenity	0.0815**	0.0588*	0.0524**	0.0559***	0.0541*	0.0520*
	(0.00849)	(0.0190)	(0.00697)	(0.00466)	(0.0129)	(0.0155)
Observations	1,183	1,181	1,179	1,180	1,179	1,179
R-squared	0.696	0.704	0.733	0.711	0.693	0.682

Notes: The table presents OLS estimations of the cultural heritage index, which is calculated through principal component analysis as detailed in table 3. Distance from own-birth place is calculated as: code of the region of birthplace (1-6, see figure 1) minus code of region of current residence (1-6). Distance from birthplace of any relative is calculated in the same way. GM/GF stands for grandmother/grandfather respectively. "Family Heterogeneity" is the standard deviation of the regional composition of the respondent's family (= 0 if all family members were born in the same macro-region). All regressions control for gender, age, marital status, income, children and education level. Standard errors are presented in parentheses under the coefficients and are clustered at the level of the town of residence. Milan is the baseline town of residence. The stars indicate statistical significance (* $<$ 1%, ** $<$ 5%, *** $<$ 10%).

Table 5: Regression of Self-Reported Index of Geographical Identity Questions (identity drawn from self-reported data)

Dependent Variable: Self-Reported Index	(1)	(2)	(3)	(4)	(5)	(6)
Town of Residence						
Rome	0.118***	0.118***	0.118***	0.120***	0.120***	0.126***
	(0.00884)	(0.00957)	(0.00784)	(0.00856)	(0.00960)	(0.00812)
Turin	0.0253*	0.0253*	0.0278*	0.0272*	0.0291*	0.0271**
	(0.00735)	(0.00773)	(0.00662)	(0.00677)	(0.00684)	(0.00587)
Distance from						
Own-birthplace	0.0830**	0.0797**	0.0849**	0.0826**	0.0778**	0.0838***
	(0.0124)	(0.0104)	(0.0131)	(0.0158)	(0.0125)	(0.00809)
Mother birthplace	-0.00196					
	(0.00706)					
Father birthplace		-0.00752				
		(0.0116)				
Maternal GM birthplace			-0.000308			
			(0.00480)			
Maternal GF birthplace				0.000343		
				(0.00621)		
Paternal GM birthplace					-0.00518	
					(0.0100)	
Paternal GF birthplace						0.000595
						(0.00851)
Family Heterogenity	-0.000728	0.00850	-0.00150	-0.000340	0.00684	-0.000308
	(0.00875)	(0.0151)	(0.00751)	(0.00676)	(0.0165)	(0.0184)
Observations	1,183	1,183	1,180	1,183	1,182	1,181
R-squared	0.130	0.135	0.130	0.127	0.129	0.134

Notes: the table presents OLS estimations of the self-reported identity index, calculated through principal component analysis as detailed in table 2. Distance from own-birth place, distance from birthplace of relatives and family heterogeneity are measured as in table 4. GM/GF stands for grandmother/grandfather. All regressions control for gender, age, marital status, income, children and education level. Standard errors are clustered at the level of the town of residence. Milan is the baseline town of residence. The stars indicate statistical significance (* $<$ 1%, ** $<$ 5%, *** $<$ 10%) and standard errors are given in parentheses below coefficient values.

Table 6: Regressions of Game Outcomes on Geographical Identity

Dependent Variable:	Public Good (own contribution)			Coin Flipping (Heads>7)		
	(1)	(2)	(3)	(4)	(5)	(6)
Respondent's Geographical Identity:						
Cultural Heritage Index	-0.790***		-0.639***	-0.0105		-0.00444
	(0.266)		(0.147)	(0.0560)		(0.0960)
Self-reported Index		-0.00559	-0.0542		-0.00873	0.000426
		(0.319)	(0.343)		(0.0417)	(0.0667)
PGG (expected partner contributions)	0.732***	0.749***	0.745***			
	(0.0170)	(0.00778)	(0.0108)			
Town of residence						
Rome	-0.258	-0.409***	-0.310*	-0.0221	-0.0130	-0.0163
	(0.227)	(0.119)	(0.171)	(0.0217)	(0.0189)	(0.0293)
Turin	-0.424***	-0.696***	-0.597***	-0.0360*	-0.0569**	-0.0507**
	(0.141)	(0.104)	(0.123)	(0.0176)	(0.0200)	(0.0233)
Wrong answer	0.0124	0.111	0.0200	0.0349	0.0520**	0.0618**
	(0.116)	(0.0920)	(0.142)	(0.0210)	(0.0204)	(0.0220)
Link to maternal grandmother's birth region	0.0630	0.0224	0.0498	0.0171	0.00957	0.0105
	(0.0534)	(0.0692)	(0.0661)	(0.0126)	(0.0138)	(0.0141)
Observations	1,181	1,007	930	1,181	1,007	930
R-squared	0.611	0.624	0.625	0.025	0.030	0.030

Notes: the table presents OLS estimations of contributions in the public good game (PGG) and coin-flipping game. The two identity measures are the same as those used in tables 4 and 5. "Wrong answer" is a dummy indicating whether a respondent failed a simple understanding check (they were asked for the payoff in the public goods game when they were presented with a simple example of behavior, see survey question 595 in the original Italian survey). 1 indicates a failure of understanding. "Link to maternal grandmother's birth region" is a dummy that is equal to 1 if the respondent owns a house in the region of birth of the maternal grandmother and this house is used regularly. All regressions control for gender, age, marital status, income, children and education level. Standard errors are clustered at birth region level. Milan is the baseline town of residence. The stars indicate statistical significance (* $<$ 1%, ** $<$ 5%, *** $<$ 10%) and standard errors are given in parentheses below coefficient values.

Table 7: Regressions of Public Goods Game Outcomes on the Average Contribution in Grandparents' Birth Regions Weighted by Relative-specific Regional Identity Measures for Cultural Heritage or Self-reported Identity Respectively

Dependent Variable:	Six Macro Regions			Three Macro Regions		
Public Good (own contributions)	(1)	(2)	(3)	(4)	(5)	(5)
Average PG contributions:						
GP Regions (CH Weights)	0.00594**	0.00600**	0.00619*	0.00746*	0.00760*	0.00785*
	(0.00197)	(0.00198)	(0.00197)	(0.00316)	(0.00321)	(0.00320)
R. Birth Region		-0.477	-0.450		-0.0493	-0.0460
		(0.549)	(0.475)		(0.00321)	(0.00320)
GP Regions (SR Weights)			-0.0493			-0.0280
			(0.0279)			(0.0151)
Town of Residence						
Rome	-0.329**	-0.472	-0.479	-0.330**	-0.475	-0.471
	(0.117)	(0.281)	(0.243)	(0.117)	(0.300)	(0.299)
Turin	-0.527***	-0.506***	-0.505*	-0.528***	-0.525***	-0.516***
	(0.0820)	(0.101)	(0.121)	(0.0817)	(0.0865)	(0.0877)
PGG (expected partner contributions)	0.752***	0.751***	0.751***	0.752***	0.751***	0.752***
	(0.00765)	(0.00663)	(0.00262)	(0.00764)	(0.00690)	(0.00692)
Wrong answer	0.0974	0.0994	0.0990	0.0972	0.0982	0.0964
	(0.0837)	(0.0809)	(0.0744)	(0.0840)	(0.0818)	(0.0819)
Link to maternal grandmother's birth region	0.0335	0.0286	0.0311	0.0328	0.0295	0.0322
	(0.0506)	(0.0487)	(0.0588)	(0.0509)	(0.0499)	(0.0490)
Observations	1,166	1,166	1,166	1,166	1,166	1,166
R-squared	0.622	0.623	0.623	0.622	0.623	0.623

Notes: The table presents OLS estimations of contributions in the public good game. "Average PG contributions" are the average contributions of those born in the same region as the respondent's grandparents weighted by closeness to their grandparents with weights derived from our relative-specific regional identity measures for cultural heritage (indicated as "GP Regions (CH Weights)") or self-reported identity (indicated as "GP Regions (GP Regions (SR Weights)") excluding the respondent's own contribution. "R. Birth Region" is the average public goods game contribution of those born in the respondent's region, excluding the respondent's own contribution. Regions are either the six macro regions or the trichotomy of Italy as displayed in fig. 1. "Wrong answer" is a dummy indicating whether a respondent failed a simple understanding check (they were asked for the payoff in the public goods game when they were presented with a simple example of behavior, see survey question 595 in the original Italian survey). 1 indicates a failure of understanding. "Link to maternal grandmother's birth region" is a dummy that is equal to 1 if the respondent owns a house in the region of birth of the maternal grandmother and this house is used regularly. All regressions control for gender, age, marital status, income, children and education level. Standard errors are clustered by birth region. Milan is the baseline town of residence. The stars indicate statistical significance (* $<$ 1%, ** $<$ 5%, *** $<$ 10%) and standard errors are given in parentheses below coefficient values.

Table 8: Regressions of Public Goods Game Outcomes on Family Members' Birth Regions

Dependent Variable: Public Good (own contribution)	(1)	(2)	(3)	(4)	(5)	(6)
Average PG contribution by:						
Own-birthplace	-0.591	-0.494	-0.836	-0.887	-0.912	-0.634
	(0.655)	(0.608)	(0.660)	(0.633)	(0.669)	(0.600)
Mother birthplace	0.156					
	(0.213)					
Father birthplace		0.000289				
		(0.141)				
Maternal GM birthplace			0.385***			
			(0.0909)			
Paternal GM birthplace				0.422*		
				(0.197)		
Maternal GF birthplace					0.414**	
					(0.140)	
Paternal GF birthplace						0.200
Town of Residence						(0.213)
Rome	-0.471	-0.471	-0.516	-0.527	-0.457	-0.473
	(0.303)	(0.306)	(0.312)	(0.304)	(0.298)	(0.302)
Turin	-0.521***	-0.525***	-0.502***	-0.497***	-0.498***	-0.516***
	(0.0982)	(0.102)	(0.0902)	(0.104)	(0.0992)	(0.105)
PGG (expected partner contributions)	0.751***	0.751***	0.751***	0.751***	0.750***	0.751***
	(0.00677)	(0.00699)	(0.00692)	(0.00681)	(0.00664)	(0.00671)
Wrong answer	0.103	0.0990	0.109	0.100	0.116	0.101
	(0.0775)	(0.0798)	(0.0732)	(0.0764)	(0.0738)	(0.0808)
Link to maternal grandmother's birth region	0.0254	0.0264	0.0170	0.00591	0.0172	0.0208
	(0.0379)	(0.0415)	(0.0337)	(0.0424)	(0.0357)	(0.0429)
Observations	1,166	1,166	1,166	1,166	1,166	1,166
R-squared	0.622	0.622	0.623	0.623	0.623	0.623

Notes: The table presents OLS estimations of contributions in the public good game. "Mean outcome in grandparents and mother regions" are the average contributions of those born in the same region as the respondent's grandparents weighted by closeness to their grandparents with weights derived from our cultural identity measure or through self-reports excluding the respondent's own contribution. "Wrong answer" is a dummy indicating whether a respondent failed a simple understanding check (they were asked for the payoff in the public goods game when they were presented with a simple example of behavior, see survey question 595 in the original Italian survey). 1 indicates a failure of understanding. "Link to maternal grandmother's birth region" is a dummy that is equal to 1 if the respondent owns a house in the region of birth of the maternal grandmother and this house is used regularly. All regressions control for gender, age, marital status, income, children and education level. Standard errors are clustered at the level of the town of residence. Milan is the baseline town of residence. The stars indicate statistical significance (* $<$ 1%, ** $<$ 5%, *** $<$ 10%) and standard errors are given in parentheses below coefficient values.

Supplementary Online Appendix

The supplementary appendix first provides some robustness checks for the key regression analysis in the main paper, before providing a summary of the key questions in the survey translated into English and a full version of the original Italian survey.

Variations on the Main Tables

Table A1 provides an alternative to table 6 where we switch from the full sample to a narrower sample of respondents who have super-local origins (i.e where the respondent, both parents and the parental grandmother were born in the local region). We also interact super local origins with a dummy for each of our three cities. We see that our cultural heritage region remains highly significant (as do the three interactions) as opposed to self-reported identity. Table A2 provides another alternative specification where we use the respondent's mother's birth region as a dummy. Once again our cultural heritage measure remains significant while the self-reported identity measure does not.

Table A3 instead regressed game outcomes on geographical identity using an alternative measure of identity: here we use a dummy to measure geographical identity where the cultural heritage dummy = 1 if the respondent's maternal grandmother was born in the south and the self-reported identity dummy = 1 if the respondent was born in the south. Table A4 returns to our standard measures for cultural heritage and self-reported identity when regressing game outcomes on identity but limits the sample to those of local origin (where the respondent and both parents were born in the local region). For column 1 (public goods game contributions) we once again see that the cultural heritage measure is highly significant while the self-reported measure is not. Neither are significant in columns 2 and 3 (the expected contributions of partners and the coin-tossing game respectively). Columns 4-6 of table A4 conducts the same exercise but using the dummy method from table A3 and generates only marginal significance (at the 10% level) in columns 4 and 6. Finally, table A5 replicates the analysis in tableA4 but restricting the sample further still to only those with super-local origins (adding the further requirement that the respondent's maternal grandmother was also born in the local region). Given the severe restrictions, our sample falls dramatically to 374 and we see only marginal significance for the cultural heritage measure (with the self-reported measure remaining insignificant). Note that we cannot use the dummy methodology in this table as this is incompatible with the super-local requirement.

Table A6 presents a regression of the respondent's partner's expected contribution in the public goods game on the respondent's geographical identity. We see no significance for either measure of identity.

Table A1: Regressions of Game Outcomes on Geographical Identity and Super-Local/Town of Residence Dummies

Dependent Variable:	Public Good (own contribution)			Coin Flipping (Heads>7)		
	(1)	(2)	(3)	(4)	(5)	(6))
Respondent's Geographical Identity:						
Cultural Heritage Index	-0.761**		-0.636**	0.0121		0.0344
	(0.311)		(0.226)	(0.0818)		(0.120)
Self-reported Index		0.0259	0.0148		-0.0158	-0.00271
		(0.325)	(0.356)		(0.0438)	(0.0695)
PGG (expected partner contributions)	0.731***	0.749***	0.745***			
	(0.0168)	(0.00783)	(0.0109)			
Town of residence						
Milan x Super Local	0.152	0.274**	0.104	0.0182	0.0238	0.0344
	(0.111)	(0.123)	(0.139)	(0.0370)	(0.0276)	(0.0412)
Rome	-0.0788	-0.0901	0.00264	-0.0268	-0.0206	-0.0310
	(0.351)	(0.270)	(0.321)	(0.0314)	(0.0371)	(0.0346)
Rome x Super Local	-0.409**	-0.510***	-0.593***	-0.0115	0.0174	0.00991
	(0.151)	(0.102)	(0.120)	(0.0291)	(0.0295)	(0.0336)
Turin	-0.393**	-0.618***	-0.595***	-0.0382	-0.0531*	-0.0495
	(0.177)	(0.145)	(0.178)	(0.0262)	(0.0296)	(0.0300)
Turin x Super Local	-0.319**	-0.526***	-0.492***	-0.0108	-0.0331	-0.0137
	(0.146)	(0.110)	(0.124)	(0.0464)	(0.0279)	(0.0244)
Wrong answer	0.0164	0.113	0.0169	0.0350	0.0529**	0.0624**
	(0.115)	(0.0926)	(0.145)	(0.0211)	(0.0207)	(0.0222)
Link to maternal grandmother birth region	0.0810	0.0470	0.0824	0.0159	0.00727	0.00795
	(0.0504)	(0.0732)	(0.0590)	(0.0127)	(0.0137)	(0.0137)
Observations	1,181	1,007	930	1,181	1,007	930
R-squared	0.611	0.625	0.626	0.025	0.030	0.031

Notes: The table presents OLS estimations of contributions in the public good game (PGG) and coin-flipping game. The two identity measures are the same as those used in tables 4 and 5. "Wrong answer" is a dummy indicating whether a respondent failed a simple understanding check (they were asked for the payoff in the public goods game when they were presented with a simple example of behavior, see survey question 595 in the original Italian survey). 1 indicates a failure of understanding. "Link to maternal grandmother" is a dummy that is equal to 1 if the respondent owns a house in the region of birth of the maternal grandmother and this house is used regularly. "Wrong answer" is a dummy indicating whether a respondent failed a simple understanding check (they were asked for the payoff in the public goods game when they were presented with a simple example of behavior, see survey question 595 in the original Italian survey). 1 indicates a failure of understanding. "Link to maternal grandmother's birth region" is a dummy that is equal to 1 if the respondent owns a house in the region of birth of the maternal grandmother and this house is used regularly. All regressions control for gender, age, marital status, income, children and education level. Standard errors are clustered at birth region level. Milan is the baseline town of residence. The stars indicate statistical significance (* $<$ 1%, ** $<$ 5%, *** $<$ 10%) and standard errors are given in parentheses below coefficient values.

Table A2: Regressions of Game Outcomes on Geographical Identity and Mother Birth Region Dummies

Dependent Variable:	Public Good (own contribution)			Coin Flipping (Heads>7)		
	(1)	(2)	(3)	(4)	(5)	(6)
Respondent's Geographical Identity:						
Cultural Heritage Index	-0.716***		-0.541**	0.0557		0.0740
	(0.243)		(0.253)	(0.104)		(0.113)
Self-reported Index		-0.0441	-0.0458		-0.0187	-0.00387
		(0.345)	(0.418)		(0.0461)	(0.0676)
PGG (expected partner contributions)	0.731***	0.7493***	0.743***			
	(0.0165)	(0.00747)	(0.0102)			
Town of residence						
Rome	-0.174	-0.208	-0.142	-0.0315	-0.0320	-0.0463
	(0.408)	(0.370)	(0.365)	(0.0351)	(0.0438)	(0.0346)
Turin	-0.434***	-0.689***	-0.603***	-0.0368*	-0.0567**	-0.0513**
	(0.140)	(0.113)	(0.130)	(0.0187)	(0.0206)	(0.0232)
Mother's Birth Region:						
North East	-0.109	-0.149	-0.281	0.0988	0.0197	0.0516
	(0.496)	(0.588)	(0.612)	(0.0707)	(0.0549)	(0.0424)
Tuscany	0.702	-0.348	-0.394	0.0392	0.0144	0.00777
	(0.718)	(0.587)	(0.601)	(0.0746)	(0.110)	(0.114)
Center	-0.236	-0.341	-0.354	-0.00618	0.0322	0.0189
	(0.457)	(0.523)	(0.542)	(0.0455)	(0.0441)	(0.0562)
South	-0.268	-0.309	-0.190	-0.0635	-0.0522	-0.0881
	(0.279)	(0.344)	(0.384)	(0.0661)	(0.0372)	(0.0544)
Calabria and Sicily	0.103	-0.311	-0.145	-0.00591	0.0199	-0.0110
	(0.300)	(0.256)	(0.242)	(0.0709)	(0.0654)	(0.0794)
PGwrongans	0.0295	0.130	0.0281	0.0360*	0.0535**	0.0630**
	(0.118)	(0.0931)	(0.156)	(0.0190)	(0.0218)	(0.0219)
Link to maternal grandmother's birth region	0.0760	0.0297	0.0617	0.0151	0.00833	0.00836
	(0.0526)	(0.0791)	(0.0671)	(0.0141)	(0.0133)	(0.0145)
Observations	1,178	1,005	928	1,178	1,005	928
R-squared	0.612	0.624	0.625	0.030	0.033	0.034

Notes: The table presents OLS estimations of contributions in the public good game (PGG) and coin-flipping game. The two identity measures are the same as those used in tables 4 and 5. "Wrong answer" is a dummy indicating whether a respondent failed a simple understanding check (they were asked for the payoff in the public goods game when they were presented with a simple example of behavior, see survey question 595 in the original Italian survey). 1 indicates a failure of understanding. "Link to maternal grandmother" is a dummy that is equal to 1 if the respondent owns a house in the region of birth of the maternal grandmother and this house is used regularly. All regressions control for gender, age, marital status, income, children and education level. Standard errors are clustered at birth region level. Milan is the baseline town of residence. The stars indicate statistical significance (* $<$ 1%, ** $<$ 5%, *** $<$ 10%) and standard errors are given in parentheses below coefficient values.

Table A3: Regressions of Game Outcomes on Geographical Identity, Dummy South Version

Dependent Variable:	Public Good (own contribution)			Coin Flipping (Heads>7)		
	(1)	(2)	(3)	(4)	(5)	(6)
Dummy South based on:						
Cultural Heritage	-0.377***		-0.370***	-0.0186		-0.0186
	(0.113)		(0.0970)	(0.0329)		(0.0329)
Self-reported Identity		-0.259	-0.0205	0.00686	-0.00513	0.00686
		(0.244)	(0.258)	(0.0362)	(0.0319)	(0.0362)
PGG (expected partner contributions)	0.739***	0.738***	0.739***			
	(0.0160)	(0.0161)	(0.0163)			
Town of Residence						
Rome	-0.383**	-0.382**	-0.383**	-0.0176	-0.0175	-0.0176
	(0.159)	(0.160)	(0.159)	(0.0179)	(0.0178)	(0.0179)
Turin	-0.494***	-0.528***	-0.494***	-0.0361**	-0.0378**	-0.0361**
	(0.111)	(0.116)	(0.110)	(0.0153)	(0.0156)	(0.0153)
Wrong answer	0.0766	0.0648	0.0763	0.0271	0.0266	0.0271
	(0.0948)	(0.0953)	(0.0931)	(0.0200)	(0.0201)	(0.0200)
Link to maternal grandmother's birth region	0.0263	0.0476	0.0269	0.0136	0.0146	0.0136
	(0.0650)	(0.0609)	(0.0645)	(0.0140)	(0.0128)	(0.0140)
Observations	1,277	1,277	1,277	1,277	1,277	1,277
R-squared	0.612	0.611	0.612	0.023	0.023	0.023

Notes: The table presents OLS estimations of contributions in the public good game (PGG) and coin-flipping game. The two identity measures are the same as those used in tables 4 and 5. "South Dummies" are a simple indicator of (a) having a maternal grandmother from the south (defined as the two lightest grey regions in figure 1) for the cultural heritage version or (b) being born in the south (again, defined as defined as the two lightest grey regions in figure 1) for the self-reported identity version. "Wrong answer" is a dummy indicating whether a respondent failed a simple understanding check (they were asked for the payoff in the public goods game when they were presented with a simple example of behavior, see survey question 595 in the original Italian survey). 1 indicates a failure of understanding. "Link to maternal grandmother's birth region" is a dummy that is equal to 1 if the respondent owns a house in the region of birth of the maternal grandmother and this house is used regularly. All regressions control for gender, age, marital status, income, children and education level. Standard errors are clustered at birth region level. Milan is the baseline town of residence. The stars indicate statistical significance (* $<$ 1%, ** $<$ 5%, *** $<$ 10%) and standard errors are given in parentheses below coefficient values.

Table A4: Regressions of Game Outcomes on Geographical Identity, Local Origin Sample

Dependent Variable:	Indexes			South Dummies		
	PGG	Exp PGG	Coins (H>7)	PGG	Exp PGG	Coins (H>7)
	(1)	(2)	(3)	(4)	(5)	(6)
Respondent's Geographical Identity:						
Cultural Heritage Index	-2.012***	2.658	-0.00722			
	(0.566)	(1.843)	(0.326)			
Self-Reported Index	0.205	-0.411	0.0235			
	(0.428)	(0.881)	(0.168)			
South Dummy (based on CHI)				-0.752*	0.0995	0.0702*
				(0.362)	(0.698)	(0.0333)
South Dummy (based on SRI)				0.707	3.135	0.322
				(1.588)	(3.489)	(0.278)
PGG (expected partner contributions)	0.763***			0.765***		
	(0.0173)			(0.0155)		
Town of Residence						
Rome	-0.199	-2.079**	-0.0269	-0.680***	-1.045***	-0.0236**
	(0.293)	(0.863)	(0.130)	(0.0862)	(0.157)	(0.00905)
Turin	-0.682***	-0.261	-0.0594*	-0.771***	-0.304	-0.0574**
	(0.0583)	(0.197)	(0.0288)	(0.0877)	(0.193)	(0.0242)
Wrong answer	0.226	-0.0134	0.0534	0.304**	-0.0832	-1.17e-05
	(0.271)	(0.526)	(0.0542)	(0.130)	(0.614)	(0.0234)
Link to maternal grandmother's birth region	0.0647	0.294	0.0165	0.0480	0.129	0.0120*
	(0.0409)	(0.259)	(0.0105)	(0.109)	(0.174)	(0.00636)
Observations	450	450	450	575	575	575
R-squared	0.668	0.070	0.043	0.664	0.071	0.052

Notes: the table presents OLS estimations of contributions in the public good game (PGG) and coin-flipping game. The two identity measures are the same as those used in tables 4 and 5. The sample is restricted to those with "local origins" which is defined as the set of respondents born in the town of residence and whose parents were born in the region of respondent's residence, i.e. Lombardia (Milan), Piedmont (Turin) or Lazio (Rome). "Wrong answer" is a dummy indicating whether a respondent failed a simple understanding check (they were asked for the payoff in the public goods game when they were presented with a simple example of behavior, see survey question 595 in the original Italian survey). 1 indicates a failure of understanding. "Link to maternal grandmother's birth region" is a dummy that is equal to 1 if the respondent owns a house in the region of birth of the maternal grandmother and this house is used regularly. All regressions control for gender, age, marital status, income, children and education level. Standard errors are clustered at birth region level. Milan is the baseline town of residence. The stars indicate statistical significance (* $<$ 1%, ** $<$ 5%, *** $<$ 10%) and standard errors are given in parentheses below coefficient values.

Table A5: Regressions of Game Outcomes on Geographical Identity, Super-Local Origin Sample

Dependent Variable:	PGG	Exp PGG	Coins (H>7)
	(1)	(2)	(3)
Respondent's Geographical Identity:			
Cultural Heritage Index	-1.546*	1.243	-0.144
	(0.748)	(3.435)	(0.353)
Self-Reported Index	-0.0749	-0.136	0.0688
	(0.491)	(0.485)	(0.157)
PGG (expected partner contributions)	0.756***		
	(0.0156)		
Town of Residence			
Rome	-0.473	-1.152	0.0302
	(0.448)	(1.387)	(0.134)
Turin	-0.779***	0.0318	-0.0666**
	(0.166)	(0.351)	(0.0266)
Wrong answer	0.437	0.144	0.0685
	(0.256)	(0.349)	(0.0713)
Link to maternal grandmother's birth region	0.144	0.163	-0.000830
	(0.159)	(0.188)	(0.0270)
Observations	374	374	374
R-squared	0.649	0.091	0.074

Notes: the table presents OLS estimations of contributions in the public good game (PGG) and coin-flipping game. The two identity measures are the same as those used in tables 4 and 5. The sample is restricted to those with "super-local origins" which is defined as the set of respondents born in the town of residence and whose parents and at least the maternal grandmother were born in the region of respondent's residence, i.e. Lombardia (Milan), Piedmont (Turin) or Lazio (Rome). This definition implies that south dummies are not applicable since the maternal grandmother must be born in the local region which cannot be from the south. "Wrong answer" is a dummy indicating whether a respondent failed a simple understanding check (they were asked for the payoff in the public goods game when they were presented with a simple example of behavior, see survey question 595 in the original Italian survey). 1 indicates a failure of understanding. "Link to maternal grandmother's birth region" is a dummy that is equal to 1 if the respondent owns a house in the region of birth of the maternal grandmother and this house is used regularly. All regressions control for gender, age, marital status, income, children and education level. Standard errors are clustered at birth region level. Milan is the baseline town of residence. The stars indicate statistical significance (* $<$ 1%, ** $<$ 5%, *** $<$ 10%) and standard errors are given in parentheses below coefficient values.

Table A6: Regressions of Partner's Expected Contribution in the Public Goods Game on Geographical Identity

Dependent Variable: PGG (expected partner contributions)	(1)	(2)	(3)
Respondent's Geographical Identity:			
Cultural Heritage Index	-0.131		0.0635
	(0.344)		(0.417)
Self-reported Index		-0.0565	-0.109
		(0.290)	(0.333)
Town of Residence			
Rome	-0.393*	-0.593***	-0.645***
	(0.203)	(0.148)	(0.216)
Turin	-0.225	-0.587***	-0.512***
	(0.151)	(0.188)	(0.174)
Wrong answer	0.0362	0.320	0.166
	(0.288)	(0.435)	(0.390)
Link to maternal grandmother's birth region	0.115	0.0972	0.149*
	(0.0808)	(0.0632)	(0.0737)
Observations	1,181	1,007	930
R-squared	0.030	0.034	0.034

Notes: the table presents OLS estimations of contributions in the public good game (PGG) and coin-flipping game. The two identity measures are the same as those used in tables 4 and 5. "Wrong answer" is a dummy indicating whether a respondent failed a simple understanding check (they were asked for the payoff in the public goods game when they were presented with a simple example of behavior, see survey question 595 in the original Italian survey). 1 indicates a failure of understanding. "Link to maternal grandmother's birth region" is a dummy that is equal to 1 if the respondent owns a house in the region of birth of the maternal grandmother and this house is used regularly. All regressions control for gender, age, marital status, income, children and education level. Standard errors are clustered at birth region level. Milan is the baseline town of residence. The stars indicate statistical significance (* $<$ 1%, ** $<$ 5%, *** $<$ 10%) and standard errors are given in parentheses below coefficient values.