



ISSN 2282-6483

Alma Mater Studiorum - Università di Bologna
DEPARTMENT OF ECONOMICS

**From Personal Values
to Social Norms**

Francesca Barigozzi
Natalia Montinari

Quaderni - Working Paper DSE N°1182



From Personal Values to Social Norms

Francesca Barigozzi* and Natalia Montinari†

March 2023

Abstract

In Experimental Economics, coordination games are used to elicit social norms as incentivized *beliefs about others' beliefs*. Conversely, representative surveys like the World Values Survey elicit social norms as personal attitudes and values that are *independent of others' beliefs*. Using a representative survey of the Italian population ($N = 1,501$), we compare the two ways of measuring social norms with gender roles as a working example and find the following results. At the aggregated level, appropriateness ratings obtained under the two elicitation methods follow the same pattern but differ significantly in magnitude, with the incentivized social norm elicitation depicting a more conservative view on gender roles than the unincentivized one. The analysis carried out at the individual level allows us to explain the previous result. Most respondents report personal values as more progressive than the *perceived* norm, which may be consistent with a desirability and/or a self-image bias. This occurs irrespectively of whether respondents correctly perceive the social norm or not. We conclude that analyses based on personal values lead to a proxy of gender norms significantly more progressive than the norms elicited in coordination games.

Keywords: elicitation of social norms, representative surveys, coordination games, personal values.

JEL classification: A13; C90; D01; J16.

*University of Bologna, Italy, Email: francesca.barigozzi@unibo.it.

†University of Bologna, Italy, Email: natalia.montinari@unibo.it.

Non-technical abstract

In this study, the authors compare two methods of measuring social norms, using social norms on gender roles as an example. The first method involves eliciting personal values through representative surveys, such as the World Values Survey, and using these values as a proxy for social norms. The second method uses coordination games to elicit individuals' beliefs about others' beliefs.

To clarify this, consider the following statement in wave 7 (2017-2020) of the World Values Survey: "A preschool child is likely to suffer if his or her mother works full-time." In the World Values Survey, respondents are asked if they agree or disagree with the claim. The reply to the previous sentence captures the respondents' personal values. Conversely, in a coordination game, respondents are asked to guess whether most people in their reference group agree or disagree with the claim. Correct guesses are incentivized: participants who match the modal answer given by other participants in the same reference group receive a monetary reward.

Social norms are inherently collective in nature and require joint recognition and approval by members of a population. Thus, using personal values as a proxy for social norms may only partially capture the social element of norms. Moreover, when answering, respondents may want to appear more progressive than they are. Conversely, measuring social norms with coordination games allows one to focus on individuals' expectations of others' values or what others think is appropriate. Additionally, people tend to be more inclined to view others as conservative, relative to their own belief.

The study answers the following questions. Does eliciting social norms as personal values generate systematic differences from eliciting social norms as expectations of others' personal values? If yes, why? Is this difference related to personal agreement/disagreement with perceived social norms? Or is it related to whether social norms are correctly perceived or misperceived?

The authors find significant and systematic differences between the two measures of social norms. Personal values expressed by group members are significantly more progressive than the social norm elicited at the group level in the coordination game. Hence, one can conclude that when personal gender values are used as a proxy for social norms, gender norms appear more progressive than they are. The analysis carried out at the individual level allows us to explain the previous result. The authors show that most respondents disagree with the perceived gender norm. Specifically, respondents tend to report personal values as more progressive than the perceived norm, especially when they hold a university degree. This is consistent with a desirability and a self-image bias. However, the authors are not able to disentangle the two. Finally, the probability of being more progressive than the perceived norm is similar when considering only respondents who perceived the norms correctly and the whole sample. As a result, unlike disagreement with the perceived social norm, misperception of the social norm is just a mistake without any systematic regularity.

1 Introduction

Coordination games elicit social norms as people’s *beliefs about others’ beliefs*; representative surveys elicit personal attitudes instead and values that are *independent of others’ beliefs*. In this study, using values on gender roles, we compare the two ways of measuring social norms and discuss the associated methodological implications.

In many empirical studies, replies to the World Values Survey (WVS) or other representative surveys are used to infer information about the social norms prevailing in the countries under study (as for gender norms, and among many others, see Fortin, 2005; Kleven et al., 2019a; Kleven et al., 2019b; Bertrand 2020).¹ To clarify the objective of our paper, let us consider the following claim appearing in wave 7 (2017-2020) of the WVS: “A preschool child is likely to suffer if his or her mother works full-time.” Respondents are asked if they strongly agree, agree, disagree, or strongly disagree with the claim. Results for Italy are reported in Table 1 and show that the 54.5% of women and the 50.9% of men “agree” or “strongly agree” with the claim. By eliciting personal values in a representative sample of the population of a given country, one obtains information on the personal values and judgements *that are prevalent in that country*, and the latter are used as a proxy for social norms. Specifically, given the answers to the claim mentioned above, one infers that most Italians hold traditional (and negative) attitudes towards full-time working mothers with young children. As we explain below, however, using personal values as a proxy for social norms, researchers and policymakers are disregarding an important dimension of norms.

Collectively approving or disapproving given judgments or behaviors is at the very heart of the definition of social norms. Indeed, Ostrom (2000, pages 143-144) defines social norms as “shared understandings about actions that are obligatory, permitted, or forbidden.” In different words, the *social* element of norms requires that they be jointly recognized, or mutually held, by members of a population and that people’s approval and disapproval sustain them. Hence, to elicit the collective agreement with some judgments/behavior and its perceived appropriateness in a society, one should not focus on personal values but on the *expectation of others’ personal values*. The latter corresponds to *people’s beliefs about “what others think is an appropriated judgment or behavior.”* Recent experimental literature uses coordination games carried out among groups in the field or in the lab to elicit social norms (see Krupka and Weber, 2013, and reference within). Participants in

¹For example, Kleven et al. (2019, page 125) “provide evidence on the relationship between child penalties and elicited gender norms. The norm variable is taken from the International Social Survey Program (ISSP), focusing on whether women with children under school age or in school should work outside the home (full-time or part-time) or stay at home.” They plot their estimated long-run child penalties in earnings against the fraction of respondents who think women should stay at home and conclude that: “The countries that feature larger child penalties are also characterized by much more gender-conservative views. This evidence, while not necessarily causal, is consistent with a potentially important role in gender norms.”

	Total	Male	Female
Agree Strongly	12.6	11.4	13.8
Agree	40.1	39.5	40.7
Disagree	37.6	38.1	37.0
Strongly Disagree	7.0	7.2	6.8
Don't Know	2.6	3.5	1.7
No answer	0.2	0.2	0.1
N	2,282	1,183	1,098

Table 1: World Value Survey (wave 7, 2017-2020). Percentage of Italian people agreeing and disagreeing with the statement "Pre-school child suffers with working mother."

experiments are provided monetary incentives to match the responses of others. Thus, participants play a pure matching coordination game whose goal is to anticipate the extent to which others will rate behavior as socially appropriate or inappropriate (instead of revealing their personal values on the proposed behavior). Social norms elicited using this methodology have been found to predict individuals' behavior in a variety of situations, such as prosocial behavior, bribing, discrimination, and saving behavior (e.g. Gächter et al. 2013; Burks and Krupka, 2012; Barr et al. 2018, and Fromell et al. 2021).²

Before proceeding further, a caveat is worth mentioning. The WVS elicits personal values and judgments and many of its claims are, by their own nature, descriptive (precisely as the claim "A preschool child is likely to suffer if his or her mother works full-time"), whereas the experimental literature on social norms focuses on prescriptive claims on behaviors (i.e. on actions that are obligatory, permitted, or forbidden). To compare the two types of social norms' elicitation, we apply the experimental methodology to both prescriptive and descriptive claims. Note, however, that moving from a descriptive claim to a prescriptive norm on behaviors is pretty natural. For example, if an individual agrees with the statement "A preschool child is likely to suffer if his or her mother works full-time," he/she will tend to judge inappropriate that the mother of a young kid works full-time. In turn, by observing that most people agree with this claim, one expects that Italian people tend to think that mothers *should not* work full-time when their children are below six.

In the experimental literature, in line with the definition of Ostrom (2000), social norms are interpreted as behaviors that are *jointly recognized, or collectively perceived, by members of the group as appropriate*. By eliciting actions that are considered to be socially appropriate for the group of participants in the experiment, this methodology identifies social norms as second-order

²For an excellent survey on the tools for the measurement of social norms in Experimental Economics, see Gorges and Nosenzo (2020).

beliefs. However, a major limitation of the experimental literature is that the group of participants in the experiment is necessarily small and, being mostly composed by university students, may not be representative of the entire society, especially for matters such as gender norms.

In this study, we compare the two ways of measuring social norms and values: via personal values and via incentivized expectations of others' personal values.³ To do so, we conducted an incentivized survey on a representative sample of the Italian population ($N = 1,501$) and elicited social norms as (unincentivized) personal values and as (incentivized) modal responses in a coordination game. In the coordination game, we ask respondents to match the choice of a group of people similar to them for gender, age group, and residence area. In such a way, we create homogeneous subgroups in which respondents guess modal responses. In addition, we control for respondents' personal characteristics (e.g. civil status, education, employment status, presence of children, etc., etc.) and personality traits.

We analyzed the five claims below. Claims 1-4 belong to the WVS, while we personally phrased the fifth one.

1. Men make better political leaders than women do;
2. When jobs are scarce, men should have more rights to a job than women;
3. University education is more important for a boy than for a girl;
4. A preschool kid is likely to suffer if his or her mother works full time;
5. A woman should be ready to reduce the time devoted to her job for family reasons.

Our goal is answering to the following questions using social norms on gender elicited from the five claims above.

Does eliciting social norms as personal values generate systematic differences from eliciting social norms as expectations of others' personal values? If yes, why? Is this difference related to personal agreement/disagreement with perceived social norms? Or is it related to whether social norms are correctly perceived or misperceived?

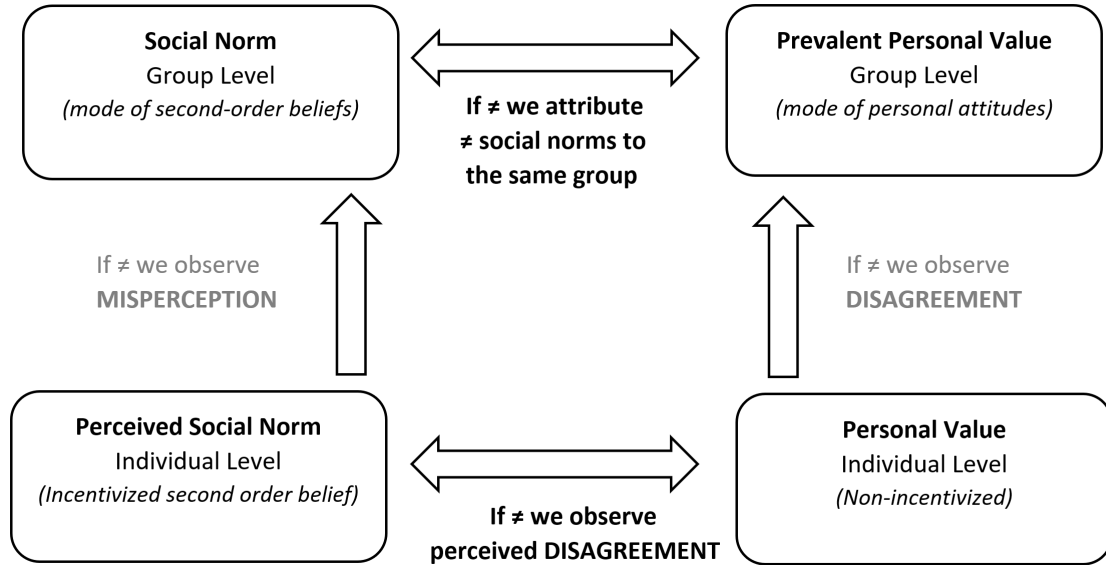
We first show that the five claims differ in the consensus reached by society on the specific matter expressed in the claim. For the first three claims, which contain an implicit comparison between men's and women's abilities and rights, we observe that all groups in the sample share a progressive attitude, and the corresponding social norm appears to be "strong." For the other two claims, which

³To the best of our knowledge, Bursztyn et al. (2020) is the only existing study where personal values and social norm elicitation are considered. The authors provide evidence of misperception of gender norms in Saudi Arabia. The objective and the methodology used in their and in our paper are different; we refer the reader to Subsection 4.3.1 for comparing the two approaches and results.

refer to the role of women as the main caregiver within the family, there is a more fragmented view, and the corresponding social norm is “weak.”

Next, we carry on a set of comparisons between perceived social norms and personal values, as described in Figure 1.

Figure 1: Relationship between Social Norms^a and Personal Values



^a Perceived social norms are elicited at the individual level using the incentivized method introduced by Krupka Weber (2013): participants are incentivized to match the modal response provided by other subjects similar to them (in terms of gender, age group and geographical area of living). Such elicitation results define the incentivized measure of the social norm at the group level. Personal values are elicited at the individual level (unincentivized) by answering the question, "How much do you agree with the following statement?". Such elicitation results define the non-incentivized measure of prevalent personal value at the group level. Moving along the vertical dimension of the figure, we compare individual vs group, while moving along the horizontal dimension, we compare the same group (top row) or the same individual (bottom row).

We start by comparing social norms elicited as the mode of second-order beliefs and the prevalent personal value, i.e. the personal attitude shared by most respondents; see the upper part of Figure 1. While they mirror each other, we show that average appropriateness scores obtained under the two methods are significantly different. Specifically, personal values expressed by group members are more open than the social norm elicited as the group’s modal reply in the coordination game. In addition, the existence of systematic differences between the two elicitation methods is confirmed for the claims that identify “strong norms” and the ones corresponding to “weak norms.” While for the claims that identify “strong norms,” the difference mainly affects the degree of support for the

norm, in the case of claims identifying "weak norms," the two elicitation methods provide different results (a change in the resulting social norm).

Second, to understand the reason why the two elicitation methods provide different results, we move to the individual level and the bottom part of Figure 1. We compare personal values and perceived social norms (correctly perceived or not) to investigate personal agreement/disagreement with the perceived norm and the direction of the possible disagreement (i.e. the fact of holding personal values more or less conservative than the perceived social norm). We find that many respondents report a personal attitude different from the perceived social norm. Most of them report personal values as more progressive than the perceived norm. Such behavior is positively associated with holding a university degree and possibly consistent with a desirability or self-image bias.

Finally, our results hold irrespectively from the fact that individuals correctly perceive the social norm. Misperception (identified in the left-hand side of Figure 1) is widespread also for strong norms. No clear patterns about the direction of misperception are observed, and situations in which respondents perceive a more progressive and a less progressive norm than the actual one are both observed.

From a methodological perspective, our results are important to understand the relationship between the currently used proxy for social norms elicited from representative surveys such as the WVS and the measure of social norms based on normative expectations obtained through incentivized experimental methods. Acknowledging the existing difference between the elicitation of normative expectations and of personal values is crucial to improving the design and interpretation of surveys, especially on matters that are sensitive to social desirability (i.e. a tendency to respond in ways that are thought to be appropriate) and a self-image bias; see Section 4.3 for more on this point. In addition, the appropriate design of policies aimed at reducing gender gaps in the labor market requires a correct interpretation and measure of gender norms.

As a second methodological contribution, we show how social expectations can be easily measured by incorporating an incentivized social norm elicitation in a representative survey with relatively small associated monetary costs.

2 The representative survey

We collected data on a representative sample of the Italian population ($N = 1,501$)⁴. Representativeness holds with respect to the following characteristics: gender (male: 41.57%; female: 58.43%), age range (25 – 34 (19.85%); 35 – 59 (52.43%); 50 – 64 (27.71%)), residence area (North (47.90%),

⁴The size of our sample is in between the two most recent waves of the WVS for Italy, i.e. wave 5 ($N = 1,012$) and wave 7 ($N = 2,282$).

Center (18.92%) and South of Italy (33.18%)) and, education (percentage of people holding a tertiary degree: 35.38%); see Table 2. Descriptive statistics of the variable used are provided in table A1 in the Appendix.

The data were collected by the professional company Scenari S.r.l. in June 2020 from a panel of 10.000 participants using the Computer-assisted web interviewing (CAWI) methodology, an Internet surveying technique whose main advantage is to have a lower cost compared to other methods, basically because there is no need for interviewers to hold the survey. On average, participants spent 23.4 minutes completing the survey (standard deviation: 10.5 min.).

Age group	North		Center		South and Islands	
	Male	Female	Male	Female	Male	Female
Age 25-34	63	67	20	26	58	64
Age 35-49	133	244	68	92	105	145
Age 50-64	91	121	32	46	54	72
Total	287	432	120	164	217	281
N (M+F)	719		284		498	

Table 2: Representative Sample N=1,501^a

^aThe sample was collected in June 2020 with a CAWI methodology from Scenari s.r.l., a private company owning a panel of 10,000 individuals. The final sample (N=1,501) is representative with respect to gender (male, female), age range (25-34; 35-49; 50-64), and residence area (North, Center, and South of Italy). The table displays the eighteen groups relevant to our social norm elicitation.

The survey was composed of three parts. In the first part, participants answered questions relating to their demographic information and household composition. In the second part, we elicited gender norms following the methodology introduced by Krupka and Weber (2013). We proposed four vignettes and a question composed of the five claims we listed in the Introduction.⁵ The four vignettes were presented randomly but always before the question containing the five claims. Within this question, the five claims were presented in random order. Specifically, participants had to guess, for each of the five claims, the answer assigned by the majority of people similar to them (with respect to gender, age group, and residence area), i.e. their second-order beliefs. In the third part, participants answered questions about i) their employment and the employment of other members of their household; ii) the allocation of the chores within their household (before, during and after the lockdown associated with the first wave of the COVID-19 emergency); iii) their (unincentivized) personal values on the same questions encountered in part 2 (i.e. the vignettes and the question with the five claims); iv) their political preferences, their religious beliefs, the relative

⁵The four vignettes refer to the allocation of chores within a household and were presented under two different framings. They are not included in this study but are analyzed in Barigozzi, Gaggini and Montinari (2022), however in all the analyses presented here, we control for the framing effects.

importance of different spheres of life (e.g. family, work, friends); v) some personality traits (TIPI, cognitive reflection tests).

While answering the survey, participants could not go back to the previous questions they had already answered, and they were unaware of the content of the different parts of the survey. The company offers incentives to motivate members of the panel to take part in surveys adopting a point-based system. Participants receive points for each survey they complete, depending on the survey length. Every 50 points they can get a 10 Euros Amazon gift card. For our survey, the company offered 20 points, and in part 2, we provided additional incentives: participants who correctly guessed the answer to the five claims given by most individuals in their reference group (i.e. of the same gender, age group, and residence area) were rewarded with 3 Euros per correct guess paid for with an Amazon gift card. Participants were informed that after the completion of the data collection, one of the questions presented in part 2 as well as 10% of participants (i.e. $N = 150$) would be randomly selected and receive the earnings associated with their correct guesses through the company.⁶ If the question containing the five claims is selected, the participants who are among the 10% chosen to receive payment for responses can earn up to 15 Euros if they make correct guesses on all 5 claims, in addition to the points assigned to all subjects for their participation in the survey. Specifically, of the 150 participants randomly selected, 39% provided 2 correct answers out of 3 in the vignette that had been randomly selected, earning on average 5 Euros, for a total cost of 745 Euros, paid for incentives.

We now move to the five claims that we use to measure gender norms and personal values that are listed in the Introduction. The participants encountered each claim twice, first in part 2 and then in part 3 of the survey. In part 2, for each claim, they are asked to guess the answer given by most people in their reference group on a four-point scale. The possible answers are: "strongly disagree," "disagree," "agree," or "strongly agree." Correct guesses are incentivized: participants who match the modal answer given by other participants in the same reference group receive additional 3 euros for each claim, while they earn nothing more otherwise. As shown by Krupka and Weber (2003), this incentive structure transforms the task into a pure matching coordination game where respondents are incentivized to reveal their normative expectations (not their personal values). After the elicitation of second-order beliefs, at the end of part 3, participants are also asked their personal opinion on each of the five claims stating how much they agree/disagree on the same four-point scale: "strongly disagree;" "disagree;" "agree;" "strongly agree." This allows for eliciting personal values.

We only implement one sequence of elicitation, collecting the incentivized measures first and then the unincentivized ones. Robustness of Krupka and Weber's (2013) method with respect to the

⁶Charness et al. (2016) provide evidence that paying for only a subset of individuals or for a subset of decisions is as effective as the "pay all" approach. See also Burks and Krupka (2012) who run a social norm elicitation and randomly selected 25% of participants for the payment of the social norm elicitation task.

order of elicitation of first and second-order beliefs is reported by Koenig-Kersting (2021), along with more general evidence of the robustness of this methodology to several variations: i.e. to the timing of play of the game with respect to the elicitation (D’Adda et al., 2016) and to the interests at the stake of the respondent (i.e. stakeholder or spectator, Erkut et al. 2015).

3 The conceptual framework and hypothesis

The following definitions are useful, together with Figure 1, to present and interpret our results.

Definition 1 *A **Reference Group** is a set of people characterized by the same gender (male, female), age range (25-34; 35-59; 50-64), and residence area (North, Center and, South and Islands of Italy).*

In this study, we consider 18 different reference groups whose size is depicted in Table 2. Social norms are elicited for each different reference group because different groups of individuals may hold different norms, as evidenced, for example, in Burks and Krupka (2012). When analyzing the social appropriateness score, we pool some of the reference groups by using the sample weights to account for the different sizes and characteristics of the groups while still preserving representativeness.

Definition 2 *The **Personal Value** of a respondent is his/her personal agreement/disagreement with each of the five claims.*

Our definition of Personal Value resembles the definition of "first-order normative beliefs" very closely (see e.g. Bicchieri et al. 2014) or simply "first-order beliefs" (see e.g. Koenig-Kersting 2021 and references within) used in the experimental literature to identify personal beliefs about what should be done in a specific context.⁷ Personal Values allow for disentangling the individual from the collective dimension of the appropriateness of specific behaviors and, in our case, the personal perception of a given claim from the social one.

Definition 3 *A **Perceived Social Norm** is the (incentivized) second-order belief reported by a respondent; i.e. the reply that, according to the respondent, is chosen by most people in his/her reference group.*⁸

While both Personal Values and Perceived Social Norms are measured at the individual level, by aggregating the individuals’ replies, we obtain two measures at the group level:

⁷Note that this definition is also different than the one used in the game theory literature where first-order beliefs are "beliefs on other players’ actions;" see, e.g., Fudenberg and Tirole (1991).

⁸Note that this definition of perceived social expectations includes empirical expectations (beliefs about others’ behavior) and normative expectations (beliefs about what others think should be done); see Bicchieri (2006), and Bicchieri et al. (2014) and (2022).

Definition 4 *Following Krupka and Weber (2013), a **Social Norm** is the mode of the distribution of second-order beliefs reported by members of a group on a specific claim.*

Definition 5 *A **Prevalent Personal Value** identifies the personal value (overall agreeing/overall disagreeing) shared by most individuals in a referent group on a specific claim.*

As an example, in Table 1, the Prevalent Personal Value is "agree" with the statement because the majority of male and female respondents reported that they "agree" or "strongly agree" with the claim.

The following remark is important at this stage. In empirical papers that use personal values as a proxy for social norms, no specific indication about how to derive a social norm from personal values is provided because only relationships between a given market outcome and aggregated replies to survey questions are documented (see Footnote 1 again for an example). Similarly, reports prepared by international organizations use the sum of the percentage of people agreeing and strongly agreeing with a claim in a given country to express how conservative gender values are in that country.⁹ To account for this, we will compare the Social Norm and the Prevalent Personal Value by using both the methodology introduced in the experimental literature by Krupka and Weber (2013) and the percentage of individuals agreeing/disagreeing used in the empirical literature.

To analyze the relationship between the individual and the collective dimension of Social Norms and Prevalent Personal Values, two last definitions are necessary:

Definition 6 ***Disagreement** with the Perceived Social Norm occurs when a respondent's Personal Value differs from the Perceived Social Norm. Agreement occurs when the two coincide.*

Disagreement is therefore measured at the individual level by comparing the answers given by the same individual revealing his/her beliefs about how much others will agree with a specific claim (the Perceived Social Norm) and when asked how much s/he agrees with the same claim (the Personal Value, see the bottom part of Figure 1). Personal Values and Perceived Social Norms do not need to be the same, especially on issues such as gender equality. On the one side, individuals may be concerned about appearing open and progressive; on the other, they may feel that norms are changing fast. In addition, consider that Disagreement can occur irrespective of whether the social norm is correctly perceived or not.

Our last definition refers to the left-hand side of Figure 1.

⁹For example, Special Eurobarometer 465, Report on Gender Equality 2017, on page 14, writes that: "[In Europe,] More than four in ten (44%) believe the most important role of a woman is to take care of her home and family, with 17% in total agreement. The majority, however, disagree (54%). There is no notable difference between the opinions of men and women."

Definition 7 *Misperception* of a social norm occurs when a respondent’s Perceived Social Norm differs from the Social Norm. The norm is correctly perceived in the opposite case.

Figure 1 in the Introduction summarizes our conceptual framework and represents the road map for the analysis of results. Before proceeding, we formulate our main hypothesis.

Hypothesis 1 *Social Norms (resulting from the Krupka and Weber, 2013, incentivized elicitation method) are less progressive compared to Prevalent Personal Values. At the individual level, Perceived Social Norms are less progressive than Personal Values.*

Hypothesis 1 contains a two-level comparison: at the group level, we compare Social Norms (Definition 4) and Prevalent Personal Values (Definition 5), resulting in a given group from the two elicitation methods (top horizontal dimension of Figure 1). At the individual level, we compare Personal Values (Definition 2) with Perceived Social Norms (Definition 3), highlighting Disagreement (bottom horizontal dimension of Figure 1).

Our hypothesis is based on the observation that Krupka and Weber’s (2013) elicitation method can limit the individual impact of a self-image and a desirability bias compared to Prevalent Personal Values in two ways. First, on topics publicly debated, such as gender equality, respondents tend to think of themselves as individuals more progressive than the average person in the reference group (out of social desirability motivations). Second, the social desirability bias may have a lower impact or no impact at all when eliciting social expectations because respondents may find less or no difficulties in revealing that “others” hold conservative views or agree with gender inequalities.¹⁰ Finally, note that, at the individual level, Hypothesis 1 holds irrespectively from the fact that respondents can identify the modal rating of answers given in their group, that is, the Social Norm (i.e. moving along the vertical dimension of Figure 1).

4 Results and discussion

Our results are presented in two steps, first at the aggregated and then at the individual level. In Section 4.2, we compare Social Norms and Prevalent Personal Values resulting from the two elicitation methods. In Section 4.3, we study respondents’ agreement/disagreement with the Perceived Social Norm. In Section 4.1, as an intermediate result, we derive “strong” and “weak” Social Norms.

¹⁰Note that the accuracy of answers is maximized when the elicitation of empirical and normative expectations is incentivized; see Bicchieri et al. (2014). Incentives motivate respondents to exert effort to formulate accurate guesses (Osband, 1989; see also Goetz et al., 1984). Moreover, by inducing effortful thinking, incentives can limit some of the biases to which “automatic” or System 1 thinking is subject (Epley and Gilovich, 2005). Therefore, these incentives for accuracy provide extra motivation to overcome social desirability and answer honestly (Osband, 1989).

When conducting the analysis, we use two methods for aggregating individual answers.

1. As in the experimental literature on social norms, we use the *appropriateness norm rating*, obtained by converting subjects' ratings in the task to numerical values. Specifically, one assigns evenly spaced values of +1 to the rating "Very socially appropriate", +0.33 to "Somewhat socially appropriate", -0.33 to "Somewhat socially inappropriate", and -1 to "Very socially inappropriate". By aggregating the individual answers, we can assign a value of social appropriateness to each claim. While this procedure is typically used in the experimental literature on social norms, for the sake of comparison, we also apply it to individual replies about Personal Values. The rating takes positive values for claims that, on average, are evaluated as socially appropriate and negative values for inappropriate ones. Such ratings are reported in Table 3 for Social Norms (column 7 of all the panels) and Personal Values (column 13 of all the panels).
2. As in the empirical literature using representative surveys on personal values, we measure the percentage of people replying "strongly agree" or "agree" to a given claim; see also Table 1 and the discussion in the introduction of the paper. This measure is reported in the panels of Table 3; see columns 6 and 12 for Social Norms and Personal Values, respectively.

In all our analyses, we use the sample weights, which guarantee the representativeness of our sample for within/between-group comparison and regression analysis.¹¹ Note that this procedure allows us to study the eighteen reference groups displayed in Table 2 and their aggregations (e.g., men versus women, North versus South and Islands, etc. etc.). Comparisons reported in the tables are based on a set of paired t-tests and proportion tests. For both types of tests, we followed the Benjamini-Hochberg False Discovery Rate method (Benjamini and Hochberg, 1995) for multiple tests adjustment: we sorted all the p-values in ascending rank and multiplied each by the number of separate tests being performed (five in our case, one for each possible claim) before dividing each by its rank- thus greater adjustments are made to smaller p-values.

Before comparing Social Norms and Prevalent Personal Values, we study the five claims presenting different aggregations of the groups and disentangling "strong" and "weak" Social Norms.

4.1 Weak and strong norms

In this section, we characterize the Social Norm for each claim and the consensus it obtains in each group and in different groups aggregation.

¹¹Specifically, we implement the command "svy" in Stata.

For each claim, Table A2 in the Appendix displays the distribution of evaluations obtained from the social norm (SN) elicitation (columns 2-5 on the left) and the Personal Value (PV) elicitation (columns 8-11 on the right). In each group, for both elicitation methods, the modal response is reported in bold. In addition, the tables display: i) the mean of the appropriateness rating (columns 7 and 13); ii) the percentage of individuals (weakly or strongly) agreeing with the claim (columns 6 and 12); iii) a set of pr-tests and paired t-tests (columns 14-15) on which we will comment in the next subsection.

Analyzing panels a-c of Tables A2, one can note that for claims 1-3, the social norm is always "strongly disagree" with the claim in all the analyzed groups. This means that, when one considers women's participation in politics and the labor market and women's access to education, a progressive attitude prevails in Italy and is equally shared between men and women of different ages and areas of residence. We observe that, for claims 1-3, the social norm can be defined as "strong" because the distribution of responses is strongly unimodal, i.e. the percentage of people choosing the modal response of their group is extremely high. For example, take claim 3 reported in Panel c: "University education is more important for a boy than for a girl:" the mode is characterized by a percentage of responses ranging from 49% to 85% and being over 60% in 16 of the 18 reference groups. Similar patterns emerge for claims 1 and 2, but with slightly smaller magnitudes.

Interestingly, the same trend emerges when considering Prevalent Personal Values: for claims 1-3, the modal rating remains "very socially inappropriate" in all groups. However, the percentage of respondents choosing "very socially inappropriate" is systematically larger when eliciting Personal Values than when eliciting second-order beliefs.

Panels d-e of Table A2 illustrate the answers to claims 4 and 5, which are more controversial and display a "weak" norm because two or even three responses are characterized by very close frequencies in many groups. As for claim 4, "A preschool kid is likely to suffer if his or her mother works full time" (see panel d of Table A2), in most groups, the norm is "agree" with the claim and, in some groups, the norm is "strongly agree". Variability is much higher when we look at claim 5, "A woman should be ready to reduce the time devoted to her job for family reasons" (see panel e of Table A2). Here the norm goes from "strongly disagree" (e.g. for males in the age range 25-34), to "agree."¹² The mode is characterized by a percentage of responses strictly below 48% in all groups for claim 4 and below 40% for claim 5.

A similar trend emerges when considering Prevalent Personal Values in claims 4 and 5. The vari-

¹²Notably, sometimes women perceive the norm as more conservative than men (see young women in the South and North). The fact that the norm sometimes happens to be more conservative for women than for men, both in the case of claims 4 and 5, suggests that being the main caregiver in the family is associated with the perception of a strong social expectation, while this is not the case when considering Personal Values (see panel d in Table A.2).

ability between groups in modal responses is higher, and the percentages associated with replies are lower with respect to claims 1-3. Finally, for claim 5, a larger difference between Social Norms and Prevalent Personal Values can be observed.¹³ This comparison is the focus of our following analysis.

The next section contains a rigorous analysis of the difference between the two elicitation methods at the aggregated level (with different groups aggregations), accounting for individual heterogeneity.

4.2 Social Norms and Prevalent Personal Values

In this section, we compare Social Norms (Definition 4) and Prevalent Personal Values (Definition 5) resulting from the two elicitation methods, highlighting systematic differences at the aggregated level. Our findings are coherent with the first part of Hypothesis 1 and summarized in Result 1. The second part of Hypothesis 1 is analyzed in the next section.

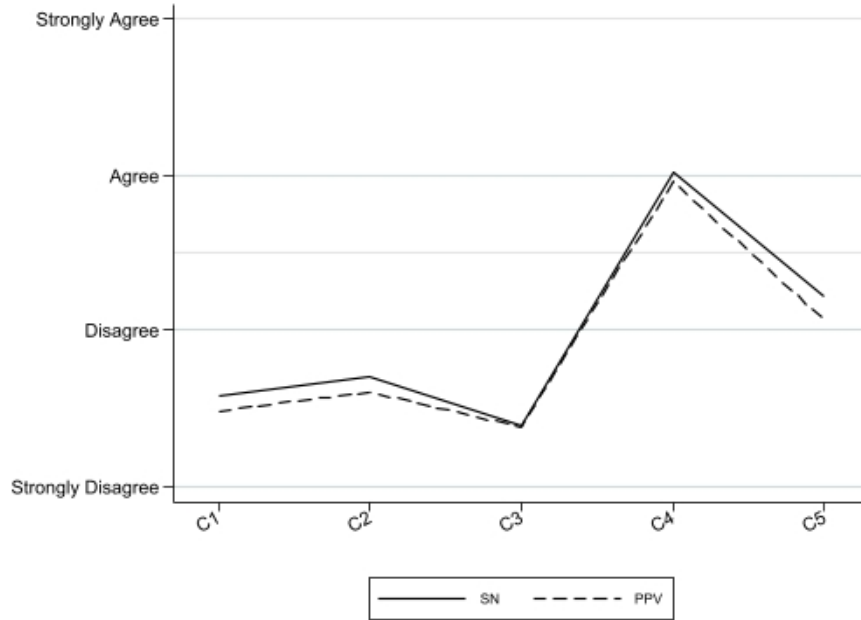
Result 1 *Social Norms and Prevalent Personal Values are highly correlated, but systematic differences emerge. At the aggregated level, the results based on the Social Norms elicitation method are more conservative compared to the ones obtained from Personal Values.*

Result 1 holds for both aggregation methods listed in section 4: support can be found in Table 3, and in Figures 2 and 3 in the main text, and in Tables A2 contained in the Appendix. Starting from the Figures, they plot the appropriateness function obtained by converting subjects' ratings in the task to numerical values. Figure 2 plots the appropriateness rating for Social Norms and Prevalent Personal Values both for the whole sample (a) overall and (b) by gender, while Figure 3 splits the sample by age groups (25-34; 35-49; and 50-64) listed in the left panel, and by geographical areas (North, Center; South and Islands) listed in the right panel, respectively.

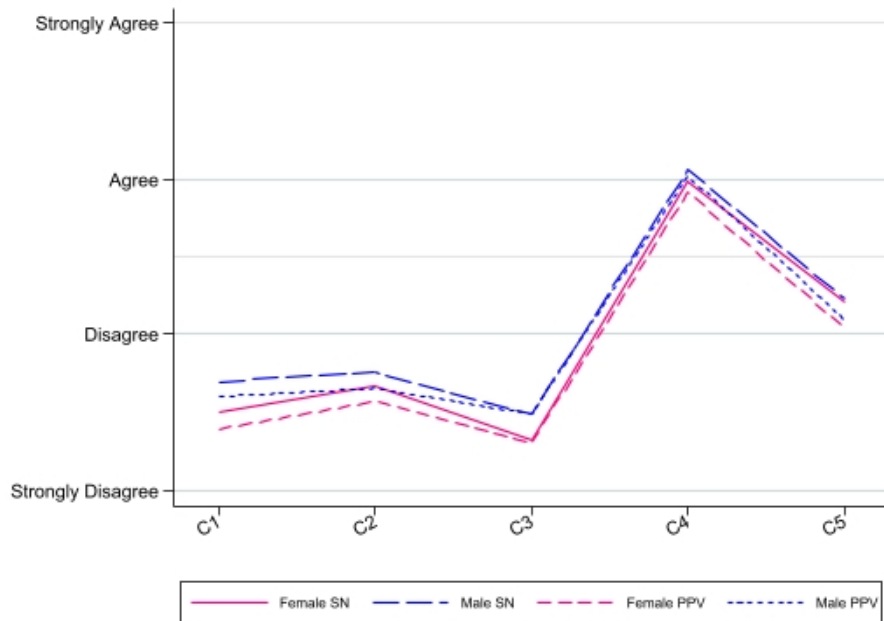
Figures 2-3 and Table 3 show some evidence and some regularities. First, in all figures, it can be noted how the average score of claims 1-3 and 5 is located in the negative domain of judgment, between -0.33 and -1. Claim 4, instead, obtains an average social appropriateness score located in the positive domain. Second, in each group analyzed, the two appropriateness scores appear to be very close to each other. Still, the scores associated with the Social Norm function lie above those associated with the Prevalent Personal Values function, evidencing a more progressive attitude than the one that emerges from the social norm elicitation. This trend can be better noted in Table 3 when comparing the two elicitation methods focusing on columns 7 and 13 (for the mean appropriateness ratings) and columns 6 and 12 (referring to the percentage of individuals agreeing with the claim).

¹³Consider female responses in claim 5. In many groups, the modal response for Personal Values is "strongly disagree" but the modal response for second-order beliefs is "agree" (e.g. see women in the South).

Figure 2: Social Norms (SN) and Prevalent Personal Values (PPV)



(a) Overall



(b) By gender

[Insert Table 3 here.]

As mentioned before, Table 3 displays the distribution of evaluations for each claim obtained both from the social norm (SN) elicitation (columns 2-5) and from the personal Values (PV) elicitation (columns 8-11). The modal response represents each group's Social Norm (Prevalent Personal Value), which is reported in bold. In addition, the table reports for both elicitation methods: i) the mean of the appropriateness rating (columns 7 and 13); ii) the percentage of individuals (weakly or strongly) agreeing with the claim (columns 6 and 12); iii) a set of pr-test and paired t-tests (columns 14 and 15) performed on the appropriateness mean ratings of the SN vs PA¹⁴ where the null hypothesis is that both means (proportions) are the same. In all tests, we followed the Benjamini-Hochberg False Discovery Rate method (Benjamini and Hochberg, 1995) for multiple tests adjustment. The distribution of the evaluations is provided for the overall sample and then on the age groups (25-34; 35-49, 50-65) and the geographical areas (North, Center, South, and Islands) and repeats the estimates by gender in each subgroup.

Inspection of Table 3 shows that for claims 1-3, the Social Norm is always "strongly disagree" with the claim in all groups. Results are the same when considering the Prevalent Personal Value. However, when comparing the differences in the mean appropriateness rating (see column 15), we find that for claims 1, 2, and 5, statistical significance is achieved for most groups considered, while for claim 4, differences are only significant for the selected group aggregation (i.e. Age group 35-49, North and South both overall and for women). When considering the percentage of individuals (weakly or strongly) agreeing with a specific claim rather than the mean appropriateness rating, we find that the same trends found for the appropriateness rating hold (see column 14) and that differences between the two elicitation methods are significant for most of the sub-samples for which the difference of the mean appropriateness rating was also statistically significant.¹⁵

Results are qualitatively similar when we consider the last two columns of Tables A1-A5 in the Appendix, where the same analysis is presented for each claim in the 18 reference groups.

Taken together, the findings from this section document the existence of systematic differences between the two elicitation methods, both for the claims that identify strong norms and the ones that identify weak norms. Moreover, such differences are quite robust to group aggregation and the

¹⁴According to our hypothesis that Personal Values are more open than Social Norms, all p-values are from one side paired t-tests (proportion tests), unless differently specified.

¹⁵In Table A.3, we report results from OLS regressions to control for individual heterogeneity. In all models, the dependent variable is obtained at the individual level for each claim as the difference between the answer given in the incentivized social norm elicitation and the personal value. As independent variables, we include socio-demographic information about the respondent, the job status, education, and the relative importance assigned to different dimensions in life as well as the political orientation. We also control for municipal characteristics. We find that in all claims except Claim 4, holding a university degree is significantly associated with a greater difference between the two answers, while other controls have less robust effects across the different claims.

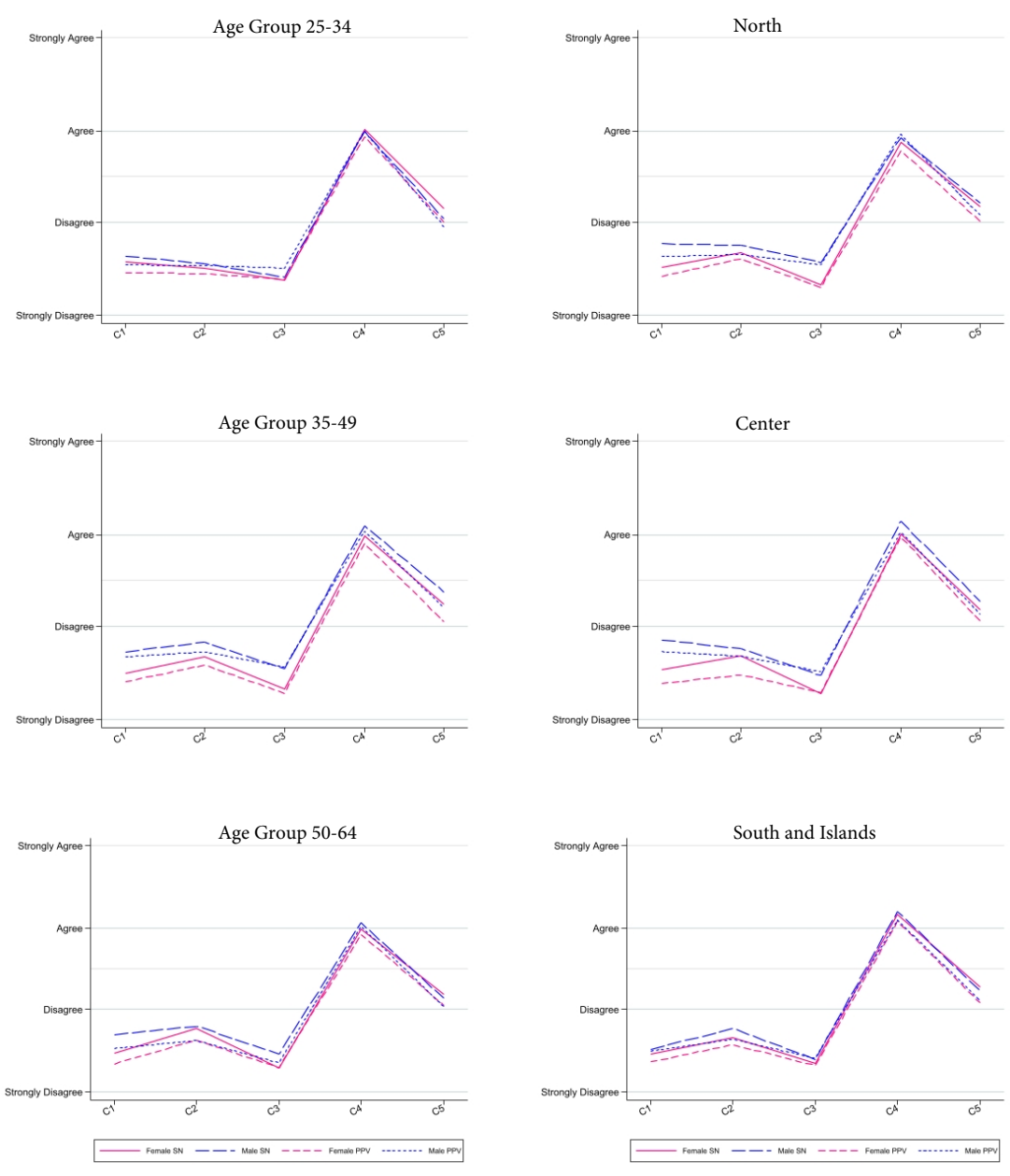


Figure 3: Social Norm (SN) and Prevalent Personal Values (PPV).
 Note: Age groups (25-34; 35-49; 50-64) are displayed in the panels on the left, while geographical areas (North; Center; South, and Islands) are displayed in the panels on the right.

methods for aggregating individual replies. At the aggregated level, social norms derived from the answers to Personal Values are more progressive than those derived from the incentivized Krupka-Weber’s (2013) methodology. While, for strong Social Norms, such differences mainly affect the degree of support for the norm, in the case of weak Social Norms, we also observe a change in the resulting Social Norms.¹⁶ Overall, the observed differences can be consistent with both a self-image bias and a social desirability bias. However, it is difficult to disentangle the two sources of bias.

4.3 (Dis)agreement with perceived norms

In this section, the analysis is carried out at the individual level. We compare Personal Values (Definition 2) and Perceived Social Norms (Definition 3); see the bottom part of Figure 1. Our main findings are reported in Result 2.

Result 2 *A substantial percentage of respondents report a Personal Value different from the Perceived Social Norm. Most of them report Personal Values as more progressive than the perceived norm. Such behavior is positively associated with holding a university degree and possibly consistent with a desirability or self-image bias.*

Support for Result 2 can be found in Tables 4 and 5. In Table 4, we define a dummy variable at the individual level for each claim. The variable takes value 1 if the individual submitted a different answer in the norm elicitation question (part 2 of the incentivized survey) and in the personal value question (part 3 of the non-incentivized survey); while a value of 0 is assigned if the two answers are the same. Hence, the first column in Table 3 shows the percentage of individuals in the sample who, for each claim, reported a Personal Value different from the Perceived Social Norm. Considering the five claims, this percentage ranges from 20.17% in claim 3 to 32.75% in claim 5. As expected, disagreement with the norm is relatively lower in the case of strong norms (i.e. claims 1-3), for which consensus is higher.

Going more in detail and looking at the second and third columns of Table 3, we observe that people reporting a Personal Value more progressive than the Perceived Social Norm are a large majority for all claims. Specifically, considering the pool of respondents who disagree with Perceived Social Norms, those indicating Personal Values more progressive than the Perceived Social Norm are in the range of [52.09%, 65.02%] and percentages are now similar for claims with weak and strong norms with differences in these proportions achieving statistical significance according to a set of one-sample tests of proportions (all p-values < 0.000, except for claim 3, where $p = 0.260$, see Table 3 for details). Two interpretations are possible here: either respondents who disagree with

¹⁶Specifically, an inspection of Table 3 indicates the following. In claim 4, one group out of eighteen changes its norm with the elicitation methods. In claim 5 instead, fifteen groups out of the eighteen change the social norm.

Claim	Disagreement ^a		
	Overall	MORE	LESS
		progressive than the (perceived) social norm	
1. Men make better political leaders than women do	22.76% (.0122)	65.02% (.0294)	34.98% (.0294)
2. When Jobs are scarce men should have more right to a job than women	24.64% (.0126)	62.15% (.0288)	37.85% (.0288)
3. University education is more important for a boy than for a girl	20.17% (.0117)	52.09% (.0324)	47.91% (.0324)
4. A preschool kid is likely to suffer if his or her mother works full time	24.86% (.0125)	61.62% (.0282)	38.38% (.0282)
5. A woman should be ready to reduce the time devoted to her job for family reasons	32.75% (.0136)	64.13% (.0242)	35.87% (.0242)

Table 4: Disagreement with the Perceived Social Norm (Std. Error in parenthesis)

^a In each panel, the percentages reported are obtained with the Stata command “.svy” accounting for the sample weights in order to ensure representativeness within each group considered. Disagreement measures the % of individuals who answer differently in the incentivized social norm elicitation and the personal value question (non-incentivized). For each claim, we define a dummy variable at the individual level which takes value 1 if the individual submitted a different answer in the norm elicitation question and in the personal value question; a value of 0 is assigned if the two answers are the same. We further distinguish cases in which Personal Values are more (less) progressive than the Social Norm. Note that in our case, being more progressive implies assigning “disagree” with the statement to a greater extent. According to a set of one-sample tests of proportions, all values reported in the table are significantly different from 50% (p-values < 0.0000) except from the claim “University education is more important for a boy than for a girl”, specifically: 52.09% and 47.91% are not significantly different from 50% (p= 0.260). All p-values are corrected using the Benjamini-Hochberg False Discovery Rate method (Benjamini and Hochberg, 1995).

the Perceived Social Norm are truthfully holding more progressive views, or they are affected by a form of (desirability or self-image) bias and would like to appear more open than other individuals in their reference group.

To explain the evidence reported in Table 3, the following remark is relevant. When filling out the survey, individuals could not go back to the previous questions to check replies, and the survey content was not known in advance. Therefore it was not possible to anticipate that the questions used in part 2 will be proposed again in part 3 to elicit Personal Values. Hence, respondents had no reason to misreport Perceived Social Norms and Personal Values strategically. In addition, if the disagreement we observed was the result of the imperfect recall of previous answers in the incentivized part, then we should have observed that individuals are equally likely to recall being more and less progressive than what they stated in the norm elicitation.

To better investigate the determinants of this behavior, we run a set of Logit regressions, whose marginal effects are reported in Table 4. The table reports five models, one for each claim. In each model, our dependent variable is defined at the individual level and takes a value of 1 if the Personal Value reported in part 3 of the survey is more progressive than the Perceived Social Norm provided in part 2 of the survey, and 0 otherwise (i.e. if the two are not different, or if Personal Values are more conservative than Perceived Social Norms). As independent variables we include socio-demographic information about the respondent, such as gender (through the dummy *female* which takes value 1 if the respondent is a female and 0 otherwise); age group (through a set of dummies, age 25 – 34 and age 50 – 64, leaving as omitted category the age group 35 – 49), presence of children (through the dummy *children* that takes value 1 if the respondent reports having at least one child and 0 otherwise); area of residence (with a set of dummies: *Center* -which takes value 1 if the respondent is resident in the Center of Italy and 0 otherwise-; and a dummy *South and Islands* which takes value 1 if the respondent is resident in the South of Italy or in one of the Islands, and 0 otherwise leaving the North as the omitted category). We also account for the civil status (through a dummy *Married or Cohabitant* that takes value 1 if the respondent reports to be married or cohabitant, and 0 if the respondent is single, widow, or separated/divorced); the job status (using the dummy *working* equal to 1 if the respondent has a job and 0 otherwise) and education (including the dummy *university degree* that is equal to 1 if the respondent holds a university degree and 0 otherwise). We control for the relative importance assigned to different dimensions in life with a set of dummies (e.g. *free time*, *community engagement*, *work* and *family* which take value 1 if the respondent assigns to each of them the highest relative importance and 0 otherwise, leaving as omitted variable respondents who report assigning equal importance to all dimensions) and for the political orientation (through the dummy *center-right* that takes the value 1 if the respondent self-reports being in line with the ideas of a political party from the center-right, and 0 otherwise). Notably, we also include the dummy *correctly Perceived Social Norm* that takes the value 1 if the individual correctly perceived the social norm in the incentivized task and 0 otherwise, and a set of dummies that account for the size of the municipality where the individual lives by using data on the population at the municipality level provided by the Urban Index-Indicatori per le politiche urbane, Presidenza del Consiglio dei ministri (<https://www.urbanindex.it>).¹⁷ Table 4 reports the marginal effects measuring the amount of change in the dependent variable that will be produced by a 1-unit change in the explanatory variable, when all other explanatory variables in the model are considered at their mean value.

Consider Models 1-3 first, which refer to claims related to strong norms: the only variable that appears to be systematically able to predict this specific form of mismatch between Personal Values

¹⁷We use the postal code provided by the respondents to match the information on municipalities contained in the Urban Index Dataset. However, since postal codes are not unique for each municipality, we lose data on 102 individuals in Model 1 and Model 2; 58 individuals in Model 3; and 203 individuals in Models 4 and 5.

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
	Claim 1	Claim 2	Claim 3	Claim 4	Claim 5
Estimation Method	Logit Regression, Marginal Effects (dy/dx)				
Dependent Variable	1 if personal values are more open than the perceived social norm elicitation in the claim, 0 otherwise ^a				
Independent Variables					
Age Range (Baseline: Age range 35-49)					
Age range 25-34	.0040 (.0653)	-.0070 (.0654)	.0056 (.0791)	-.0483 (.0324)	-.0155 (.0337)
Age range 50-64	.0353 (.0558)	.0799 (.0498)	-.0119 (.0683)	-.0373 (.0244)	-.0097 (.0279)
Female	.0733 (.0506)	.0212 (.0485)	.0910 (.0623)	.0122 (.0245)	.0249 (.0268)
Geographical Area of residence (Baseline: North)					
Center	.0145 (.0625)	.1019* (.0557)	.0392 (.0733)	.0005 (.0304)	.0035 (.0329)
South and Islands	-.0465 (.0587)	.0430 (.0557)	.0505 (.0751)	.0414 (.0264)	-.0194 (.0309)
Civil Status (Baseline: Single, Widower, Separated-Divorced)					
Married or Cohabitant	.0227 (.0561)	-.0141 (.0588)	.1832*** (.0700)	.0072 (.0274)	-.0165 (.0327)
Having Children	-.0248 (.0558)	.0297 (.0550)	-.0209 (.0678)	-.0076 (.0259)	.0236 (.0303)
University degree	.0916** (.0461)	.1402*** (.0437)	.1648*** (.0581)	-.0295 (.0225)	.0441* (.0242)
Working	.0823 (.0552)	.1327 (.0520)	.0297 (.0737)	.0700** (.0276)	.0306 (.0295)
Most Important dimension in life (baseline: all dimensions equally important)					
Free time	.0395 (.0751)	-.0318 (.0741)	-.0310 (.0872)	.0011 (.0325)	-.0531 (.0394)
Community	.0240 (.1417)	.0260 (.1396)	-.1625 (.1609)	.1203* (.0638)	.1026 (.0802)
Work	.0086 (.0596)	-.0282 (.0602)	-.0638 (.0738)	-.0193 (.0287)	.0179 (.0340)
Family	-.0142 (.0622)	-.0257 (.0596)	-.0352 (.0713)	-.0809*** (.0276)	-.0650* (.0341)
Political view: Center-Right	-.0553 (.0574)	-.0307 (.0514)	.0438 (.0646)	-.0406* (.0271)	.0200 (.0284)
Correctly Perceived Social Norm	-	-	-	.0296 (.0224)	-.0052 (.0263)
Municipality Dimensions	Yes	Yes	Yes	Yes	Yes
N Observations	504	561	329	1,298	1,298
Population size	10,837.73	12,356.36	7,244.05	27,884.28	27,884.28
F	(17, 487)=0.87	(17, 544)=1.68	(17, 312)=1.32	(18, 1280)=2.21	(18, 1280)=1.04
Prob > F	0.6139	0.0421	0.1782	0.0025	0.4127

Table 5: Disagreement with the Perceived Social Norm: stating more open Personal Values than the Perceived Social Norm, Marginal effects on Logit regressions (Standard Errors in parenthesis)^a

^aNote: Table 5 presents the results from a set of Logit estimations. Marginal effects are reported following the dy/dx method: the results measure the amount of change in the dependent variable that will be produced by a 1-unit change in the explanatory variable when all other explanatory variables in the model are considered at their mean value. The dependent variable corresponds to reporting a personal value more open than the guess expressed in the incentivized social norm elicitation. In models 1-3 the variable Correctly Perceived Social Norm is omitted since Correctly Perceived Social Norm=1 perfectly predicts y=0: the reason is that for claims 1-3, in all reference groups, the Social Norm is “strongly disagree”, and therefore, it is not possible, to have personal values more open than the norm, once it has been correctly identified. In all models, we include controls for the size of the municipality of residence using data provided by the Urban Index (<https://www.urbanindex.it>), but we lose data on 102 individuals in Model 1 and Model 2; 58 individuals in Model 3; 203 individuals in Models 4 and 5, since we are not able to match all municipalities in our dataset. ***, ** and * indicate significance at the 1% level, 5% level and 10% level, respectively.

and the Perceived Social Norm is the education level: respondents holding a university degree are significantly more likely to report more open Personal Values compared to the Perceived Social Norm as reported in the incentivized task. The magnitude of the effect varies between the different models. In model 1, it is equal to about 9 percentage points, while it corresponds to 14 percentage points in model 2 and 16 percentage points in model 3.

Note that, for these three models, the dummy *correctly perceived social norm* is omitted in the estimation since it perfectly predicts a value of 0 of our dependent variable. The reason is that in claims 1-3, in all reference groups, the social norm is “strongly disagree” and therefore only misperceivers hold a personal value more progressive than the perceived norm.

Consider now Models 4 and 5, which refer to weak Social Norms. In model 4, we find that a university degree does not achieve statistical significance. However, respondents who report having a job are about seven percentage points more likely to displaying this form of disagreement, as well as individuals who report the highest relative importance of being involved in community life, with an effect almost twice as big in terms of magnitude (+ 12%). Interestingly, individuals who assign the highest relative importance to family appear to be about 8 percentage points less likely to exhibit Personal Values more progressive than the Perceived Social Norm as well as for those individuals who feel in line with parties from the center-right, which typically have a more conservative political agenda (for them, there is a 4 percent points reduction in the probability of observing the mismatch). We focus now on model 5, which refers to the claim “A woman should be ready to reduce the time devoted to her job for family reasons:” holding a university degree is again positively affecting the probability of displaying this form of disagreement with the perceived social norm (the magnitude of the effect is equal to 4%). As evidenced by model 4, we find a negative and significant effect for those individuals who assign the highest relative importance to family. The education level is now only partially significant, and just for claim 5. This, again, is in line with the fact that claims 4-5 express weak norms: being more progressive than a weak norm is quite an elusive attitude.

Other independent variables do not achieve statistical significance. Specifically, “correctly perceiving the norm” makes sense for claims 4-5 but is not significant in those models. We thus conclude that the probability of being more progressive than the perceived norm is not different when considering only respondents who perceived the norms correctly and the whole sample. This, in turn, suggests that being more progressive than the perceived norm is a general tendency of respondents.

4.3.1 Misperception

Misperception of a social norm is the difference between a Perceived Social Norm (Definition 3) and the Social Norm (Definition 4), as depicted in the left-hand side of Figure 1. Inspection of the third column of Table 6 reveals that the percentage of individuals who misperceive the social norm is sizeable. First, for claims 1-3, where we identify strong norms, there is a considerable fraction

of misperception: from 26.5%, for claim 3, to 45.29%, for claim 2, of total respondents, even if the fraction of individuals who perceive the norm correctly is significantly higher than the one of respondents who misperceive it. Second, for claims 4-5, for which we identified weak social norms, misperception refers to the 52.99% and 57.59% of respondents, respectively, and it is significantly higher than the fraction of individuals who correctly perceive it.

In what follows, we shortly discuss how (mis)perception intersects the agreement/disagreement with Perceived Social Norms. For each claim, Table 6 displays the percentage of individuals who misperceive the social norm in their reference group and of those who disagree with the perceived norm. The table also reports the p-values from a set of proportion tests comparing the percentages displayed within the cells and the p-values of a Design-based Pearson association test which tests the overall association between the two dimensions.

[Insert Table 6 here.]

Inspection of Table 6 reveals that, for claims 1-3, there is an association between agreeing with the Perceived Social Norm and being able to correctly identify the norm, while the opposite is true for individuals who disagree with the norm, whose majority is misperceiving the social norm. When considering claims 4 and 5, we note that the association between agreement and the ability to correctly perceive the social norm is weaker. In addition, in claim 5, the Pearson association test does not achieve statistical significance.

Taken together, our results suggest that (mis)perception is associated with the (strong or weak) nature of the norm. According to intuition, agreement with the norm and the ability to correctly perceive it are more likely for strong norms than weak ones. For strong norms, most respondents who disagree with the perceived norm are indeed misperceiving it. For weak norms, instead, no clear pattern is observed. Overall, this result appears to be driven by those individuals who correctly perceive the norm, while it seems weaker in magnitude for those individuals who disagree with the Perceived Social Norm.

Let us compare our results on misperception to the ones of Bursztyn et al. (2020). They study the gender norm prevailing among Saudi Arabian men about women working outside the home. Specifically, the authors ask a sample of Saudi Arabian men whether they agree or disagree with the claim: “In my opinion, women should be allowed to work outside of the home.” Participants are then asked to estimate the percentage of other participants they expect to agree with the claim, thus deriving a measure of Misperception of the Social Norm. Their and our paper have different focuses. However, they have in common the comparison of personal values and incentivized second-order beliefs finding a large misperception of Social Norms. Bursztyn et al. (2020) find that about 75% of

respondents substantially underestimate the level of support for women working outside the home by other similar men. The authors thus conclude that most men in Saudi Arabia perceive this gender norm as being much less progressive than it actually is. In this respect, Table A.4 in Appendix shows that misperception of gender norms can go in opposite directions. Specifically, in the case of claim 5, like in Bursztyn et al. (2020), we observe that most misperceivers perceive the norm as less progressive than it actually is. But, as for claim 4, most misperceivers perceive the norm as more progressive than it actually is. Notably, policy implications are very different in the two instances. Providing information on misperception allows increasing acceptance of working women in the case of Bursztyn et al. (2020), but it would discourage full-time work for mothers in the case of our claim 5.

To conclude, differently than in the case of disagreement, where the tendency was to perceive a norm as more conservative than own personal value, we find that misperception of gender norms can go in both directions. We observe four situations where misperceivers perceive the social norm as more progressive than it actually is (like in Bursztyn et al., 2020) and one situation in which the opposite holds. Finally, at the individual level, conscientiousness is the only factor that is found to have a positive robust effect on the probability of identifying the social norm correctly (see Table A.5 in the Appendix).

5 Conclusion

Using a representative sample of the Italian population, we analyze gender norms' elicitation via coordination games and personal values. First, we compare Prevalent Personal Values and Social Norms at the group level. Then, we compare Personal Values and *Perceived* Social Norms at the individual level. Our analysis delivers the following results. Values expressed by group members are significantly more progressive than the social norm elicited by the group's modal reply in the coordination game. Hence we conclude that when personal gender values are used as a proxy for social norms, gender norms appear more progressive than they actually are. The analysis carried out at the individual level allows us to explain the previous result. We show that most of the population disagrees with the perceived gender norm. Specifically, respondents tend to report personal values as more progressive than the perceived norm and especially so when they hold a university degree. This is consistent with a desirability and/or a self-image bias. However, we are not able to disentangle the two.

Finally, the probability of being more progressive than the perceived norm is not different when considering only respondents who perceived the norms correctly and the whole sample. Hence we find that, unlike Disagreement, Misperception of the social norm is just a mistake without any systematic regularity.

Eliciting personal values is obviously cheaper and simpler than applying Krupka and Weber’s methodology. In addition, the WVS allows for important comparisons among countries and for studying trends and cultural changes via the construction of panel datasets. Hence, we expect surveys on personal values to remain the standard source of information about social norms on gender for international organizations (and empirical economists) in the years to come. However, we showed that social expectations could be easily measured by incorporating an incentivized social norm elicitation in a representative survey with relatively small associated monetary costs.

The methodological contribution of our study in terms of social norms elicitation is clear. What about its policy relevance?

Surveys eliciting personal values tend to generate a measure of gender norms significantly more progressive than the one based on incentivized social norms. This implies that policy interventions built on such biased proxies for social norms may be inappropriate. For example, policymakers might disregard the persistence of a social stigma against full-time work for mothers with small children; see our claim 4. As a result, policymakers could interpret the persistence of low employment rates for mothers with young children mainly as a matter of preferences, thus ignoring social pressure. This could lead to under-investment in cash and in-kind transfers devoted to families with small children, and in turn, the sub-optimal provision of those transfers could contribute to the persistence of gender gaps in the labor market. Our study informs policymakers about this risk.

References

- [1] Barigozzi F., Gaggin C. and Montinari N., 2022, “Elicitation of gender norms in a representative sample,” mimeo, University of Bologna.
- [2] Barr, A., Lane, T., & Nosenzo, D. (2018). ”On the social inappropriateness of discrimination.” *Journal of Public Economics*, 164, 153-164.
- [3] Benjamini, Y., Hochberg, Y. (1995). “Controlling the false discovery rate: a practical and powerful approach to multiple testing.” *Journal of the Royal statistical society: series B (Methodological)*, 57(1), 289-300.
- [4] Bertrand M., 2020, “Gender in the Twenty-First Century,” *AEA Papers and Proceedings*, 110: 1–24.
- [5] Bicchieri C., 2006. “The Grammar of Society: The Nature and Dynamics of Social Norms,” Cambridge: Cambridge University Press.

- [6] Bicchieri, C., Lindemans, J. W., Jiang, T. (2014). "A structured approach to a diagnostic of collective practices," *Frontiers in psychology*, 5, 1418.
- [7] Bursztyjn L., A.L. González, and D. Yanagizawa-Drott, 2020, "Misperceived Social Norms: Women Working Outside the Home in Saudi Arabia." *American Economic Review*, 110 (10): 2997-3029.
- [8] Burks, S.V. and Krupka, E.L., 2012. "A Multimethod Approach to Identifying Norms and Normative Expectations within a Corporate Hierarchy: Evidence from the Financial Services Industry." *Management Science*, 58(1), 203-217.
- [9] Charness, G., U. Gneezy, & B. Halladay, 2016, "Experimental methods: Pay one or pay all." *Journal of Economic Behavior & Organization*, 131, 141-150.
- [10] D'Adda, G., Drouvelis, M., Nosenzo, D. (2016). "Norm elicitation in within-subject designs: Testing for order effects." *Journal of Behavioral and Experimental Economics*, 62, 1-7.
- [11] Dohmen T., Falk A., Huffman D., Sunde U., Schupp J. and Wagner G.G., 2011, "Individual risk attitudes: measurement, determinants, and behavioral consequences," *Journal of the European Economic Association*, 9: 522-550.
- [12] Epley, N., and Gilovich, T. (2005). "When effortful thinking influences judgmental anchoring: differential effects of forewarning and incentives on self-generated and externally provided anchors," *Journal of Behavioral Decision Making* 18, 199–212. doi: 10.1002/bdm.495
- [13] Erkut, H., Nosenzo, D., Sefton, M. (2015). "Identifying social norms using coordination games: Spectators vs. stakeholders," *Economics Letters*, 130, 28-31.
- [14] Fortin N., 2005, "Gender Role Attitudes and the Labour-Market Outcomes of Women Across OECD Countries." *Oxford Review of Economic Policy* 21(3), 416–438.
- [15] Fromell, H., Nosenzo, D., Owens, T., & Tufano, F. (2021). "One size does not fit all: Plurality of social norms and saving behavior in Kenya." *Journal of Economic Behavior & Organization*, 192, 73-91.
- [16] Gächter, S., Nosenzo, D., Sefton, M. (2013), "Peer effects in pro-social behavior: social norms or social preferences?" *J. Eur. Econ. Assoc.*, 11 (3), 548-573
- [17] Görges L. and D. Nosenzo, 2020, "Measuring Social Norms in Economics: Why It Is Important and How It Is Done," *Analyse & Kritik*; 42(2):285–311.
- [18] Gosling, S. D., Rentfrow, P. J., & Swann, W. B., Jr. 2003. "A Very Brief Measure of the Big Five Personality Domains." *Journal of Research in Personality*, 37, 504-528.

- [19] Goetz, E. G., Tyler, T. R., and Cook, F. L. (1984). “Promised incentives in media research: a look at data quality, sample representativeness, and response rate,” *J. Marketing Res.* 21, 148–154. doi: 10.2307/3151697
- [20] Haerpfer, C., Inglehart, R., Moreno, A., Welzel, C., Kizilova, K., Diez-Medrano J., M. Lagos, P. Norris, E. Ponarin Puranen (eds.). 2022. World Values Survey: Round Seven – Country -Pooled Datafile Version 4.0. Madrid, Spain Vienna, Austria: JD Systems Institute WVSA Secretariat. doi:10.14281/18241.18.
- [21] Kleven H., Landais C. , Posch J., Steinhauer A., Zweimuller J., 2019a. “Child Penalties Across Countries: Evidence and Explanations.” *AEA Papers and Proceedings* 109, 122–126.
- [22] Kleven H., Landais C., Sogaard J.E., 2019b. “Children and Gender Inequality: Evidence from Denmark,” *American Economic Journal: Applied Economics*, 11(4), 181–209.
- [23] König-Kersting, Christian, (2021), “On the Robustness of Social Norm Elicitation, Working Papers, Faculty of Economics and Statistics,” University of Innsbruck, <https://EconPapers.repec.org/RePEc:inn:wpaper:2021-02>.
- [24] Krupka E.L., Weber R.A. 2013, “Identifying Social Norms Using Coordination Games: Why Does Dictator Game Sharing Vary?” *Journal of the European Economic Association* 11(3): 495-524.
- [25] Osband, K. (1989). “Optimal forecasting incentives,” *Journal of Political Economy* 97, 1091–1112. doi: 10.1086/261644
- [26] Ostrom E., 2000. “Collective Action and the Evolution of Social Norms.” *Journal of Economic Perspectives* 14, 137–158.

Table 3. Social appropriateness rating and general agreement

Panel a) Claim 1: “Men make better political leaders than women do”

	Social Norm						Personal Value						Benjamini-Hochberg adjusted p-values	
	--	-	+	++	SN>0	Mean rating	--	-	+	++	PV>0	Mean rating	pr-test	t-test
Overall	59.46	26.16	11.08	3.29	14.37	-6.115	66.59	20.31	10.40	2.70	13.10	-6.716	0.095	0.000
Females	63.34	25.26	9.12	2.28	11.40	-6.639	72.30	17.16	8.81	1.73	10.54	-7.332	0.211	0.000
Males	55.50	27.09	13.09	4.32	17.41	-5.580	60.76	23.54	12.03	3.68	15.71	-6.087	0.149	0.010
Age 25-34	58.35	27.08	10.89	3.68	14.57	-6.001	64.41	21.85	11.39	2.35	13.74	-6.551	0.407	0.049
Age 35-49	57.93	26.13	12.87	3.08	15.95	-5.922	62.86	22.51	11.13	3.51	14.64	-6.310	0.221	0.021
Age 50-64	61.52	25.75	9.43	3.30	12.73	-6.361	71.31	17.41	9.20	2.08	11.28	-7.194	0.252	0.000
Females 25-34	56.51	29.95	10.72	2.82	13.54	-6.004	65.93	20.79	10.67	2.61	13.28	-6.656	0.461	0.121
Females 35-49	64.81	23.01	9.98	2.19	12.17	-6.692	71.20	18.32	8.67	1.82	10.49	-7.526	0.134	0.004
Females 50-64	65.19	25.18	7.52	2.10	9.62	-6.892	76.38	14.33	8.06	1.23	9.29	-7.722	0.429	0.003
Males 25-34	60.14	24.28	11.05	4.53	15.58	-5.997	62.93	22.89	12.09	2.09	14.18	-6.444	0.375	0.183
Males 35-49	50.96	29.27	15.79	3.98	19.77	-5.144	54.43	26.74	13.61	5.22	18.83	-5.354	0.354	0.252
Males 50-64	57.64	26.35	11.44	4.56	16.00	-5.800	65.95	20.67	10.41	2.97	13.38	-6.636	0.315	0.048
North	56.92	27.25	12.67	3.16	15.83	-5.857	64.38	23.02	10.52	2.09	12.61	-6.641	0.019	0.000
Center	55.50	28.03	11.30	5.17	16.47	-5.585	66.32	19.63	8.33	5.71	14.04	-6.433	0.160	0.008
South	65.15	23.64	8.84	2.38	11.22	-6.766	69.68	17.12	11.45	1.75	13.20	-6.989	0.197	0.278
North females	61.85	27.26	8.78	2.12	10.90	-6.583	70.11	19.81	9.44	0.65	10.09	-7.288	0.266	0.000
North males	51.96	27.24	16.59	4.20	20.79	-5.127	58.61	26.24	11.61	3.54	15.15	-5.990	0.021	0.013
Center females	59.10	27.96	8.86	4.08	12.94	-6.132	72.63	16.36	7.65	3.37	11.02	-7.214	0.448	0.005
Center males	51.75	28.11	13.85	6.29	20.14	-5.017	59.75	23.04	9.05	8.15	17.20	-5.622	0.263	0.139
South females	67.78	21.07	9.72	1.43	11.15	-7.010	74.98	14.15	8.68	2.20	10.88	-7.458	0.440	0.067
South males	62.44	26.29	7.93	3.35	11.28	-6.515	64.22	20.18	14.31	1.29	15.60	-6.487	0.163	0.468

Panel b) Claim 2: “When Jobs are scarce men should have more right to a job than women”

	Social Norm						Personal Value						Benjamini-Hochberg adjusted p-values	
	--	-	+	++	SN>0	Mean rating	--	-	+	++	PV>0	Mean rating	pr-test	t-test
Overall	54.71	21.29	19.13	4.88	24.01	-5.054	59.81	19.54	16.15	4.49	20.64	-5.644	0.003	0.000
Females	56.44	20.71	19.19	3.66	22.85	-5.328	62.37	17.37	16.97	3.28	20.25	-5.922	0.036	0.000
Males	52.94	21.88	19.06	6.12	25.18	-4.775	57.20	21.75	15.32	5.73	21.05	-5.359	0.038	0.010
Age 25-34	63.04	19.82	12.71	4.42	17.13	-6.096	65.70	17.55	13.05	3.70	16.75	-6.349	0.423	0.174
Age 35-49	53.14	21.70	19.74	5.41	25.15	-4.838	57.88	19.80	17.02	5.30	22.32	-5.350	0.055	0.006
Age 50-64	52.11	21.61	21.71	4.58	26.29	-4.749	58.78	20.26	16.86	4.10	20.96	-5.580	0.028	0.001
Females 25-34	63.09	22.49	12.98	1.44	14.42	-6.479	66.21	18.93	14.14	0.72	14.86	-6.707	0.539	0.346
Females 35-49	56.64	21.85	17.21	4.30	21.51	-5.387	62.04	18.22	16.06	3.68	19.74	-5.907	0.244	0.011
Females 50-64	53.09	18.77	24.04	4.11	28.15	-4.723	60.86	15.83	19.19	4.12	23.31	-5.563	0.108	0.004
Males 25-34	62.99	17.23	12.45	7.33	19.78	-5.724	65.21	16.21	11.99	6.59	18.58	-6.001	0.349	0.290
Males 35-49	49.60	21.56	22.30	6.54	28.84	-4.282	53.68	21.41	17.98	6.94	24.92	-4.787	0.105	0.088
Males 50-64	51.07	24.61	19.25	5.07	24.32	-4.777	56.58	24.95	14.38	4.09	18.47	-5.598	0.238	0.063
North	56.23	19.49	19.35	4.92	24.27	-5.136	59.04	20.82	16.14	4.00	20.14	-5.657	0.008	0.005
Center	52.72	24.61	17.81	4.87	22.68	-5.009	62.71	19.24	11.99	6.06	18.05	-5.905	0.090	0.008
South	53.84	21.74	19.59	4.83	24.42	-4.973	59.16	18.01	18.59	4.24	22.83	-5.474	0.223	0.295
North females	57.44	19.50	19.33	3.73	23.06	-5.377	60.10	18.63	17.65	3.62	21.27	-5.680	0.273	0.130
North males	55.01	19.49	19.38	6.12	25.50	-4.893	57.97	23.02	14.62	4.39	19.01	-5.635	0.018	0.015
Center females	53.67	24.30	17.51	4.52	22.03	-5.139	66.11	18.19	13.80	1.90	15.70	-6.566	0.110	0.003
Center males	51.72	24.93	18.12	5.23	23.35	-4.874	59.18	20.32	10.12	10.38	20.50	-5.217	0.179	0.208
South females	56.75	20.20	19.98	3.07	23.05	-5.375	63.18	15.24	17.94	3.64	21.58	-5.865	0.412	0.077
South males	50.85	23.33	19.18	6.63	25.81	-4.559	55.03	20.86	19.26	4.85	24.11	-5.070	0.319	0.212

Panel c) Claim 3 “University education is more important for a boy than for a girl”

	Social Norm						Personal Value						Benjamini-Hochberg adjusted p-values	
	--	-	+	++	SN>0	Mean rating	--	-	+	++	PV>0	Mean rating	pr-test	t-test
Overall	73.65	16.54	6.40	3.40	9.80	-.7360	73.80	14.76	8.01	3.43	11.44	-.7259	0.057	0.389
Females	79.48	12.50	5.03	2.99	8.02	-.7896	80.70	9.86	6.71	2.73	9.44	-.7901	0.133	0.812
Males	67.70	20.68	7.80	3.82	11.62	-.6813	66.75	19.76	9.34	4.15	13.49	-.6603	0.149	0.174
Age 25-34	70.78	20.33	6.62	2.26	8.88	-.7305	68.35	18.21	10.35	3.10	13.45	-.6784	0.058	0.087
Age 35-49	71.89	15.89	8.25	3.96	12.21	-.7045	71.45	15.16	9.32	4.08	13.40	-.6930	0.205	0.288
Age 50-64	76.80	15.30	4.48	3.42	7.90	-.7696	78.80	12.65	5.57	2.98	8.55	-.7816	0.343	0.300
Females 25-34	72.02	18.11	6.70	3.17	9.87	-.7262	73.66	12.59	11.93	1.83	13.76	-.7205	0.585	0.449
Females 35-49	79.45	11.95	5.51	3.09	8.60	-.7848	82.66	8.48	6.17	2.69	8.86	-.8073	0.425	0.147
Females 50-64	83.07	10.35	3.77	2.81	6.58	-.8243	82.19	9.88	4.73	3.20	7.93	-.8069	0.286	0.267
Males 25-34	69.58	22.50	6.54	1.38	7.92	-.7346	63.18	23.67	8.81	4.33	13.14	-.6375	0.088	0.085
Males 35-49	64.24	19.88	11.03	4.85	15.88	-.8073	60.11	21.91	12.50	5.48	17.98	-.5774	0.336	0.120
Males 50-64	70.18	20.53	5.23	4.06	9.29	-.7117	75.22	15.58	6.45	2.74	9.19	-.7549	0.486	0.151
North	70.68	17.87	8.54	2.91	11.45	-.7085	71.67	16.48	9.19	2.66	11.85	-.7142	0.478	0.474
Center	73.82	19.09	4.34	2.75	7.09	-.7594	73.06	16.77	5.63	4.55	10.18	-.7219	0.150	0.147
South	77.51	13.30	4.75	4.43	9.18	-.7590	77.06	11.30	7.83	3.82	11.65	-.7438	0.161	0.354
North females	78.85	13.49	5.57	2.09	7.66	-.7937	79.96	11.20	6.99	1.85	8.84	-.7951	0.248	0.579
North males	62.46	22.28	11.53	3.73	15.26	-.6228	63.32	21.80	11.40	3.48	14.88	-.6328	0.434	0.375
Center females	79.38	14.82	3.60	2.21	5.81	-.8087	79.70	11.13	7.50	1.68	9.18	-.7922	0.270	0.413
Center males	68.03	23.54	5.12	3.31	8.43	-.7080	66.15	22.63	3.68	7.54	11.22	-.6486	0.212	0.156
South females	80.38	9.84	5.15	4.62	9.77	-.7730	82.26	7.36	5.88	4.51	10.39	-.7824	0.493	0.394
South males	74.56	16.87	4.34	4.23	8.57	-.7446	71.70	15.35	9.84	3.11	12.95	-.7041	0.128	0.197

Panel d) Claim 4 “A preschool kid is likely to suffer if his or her mother works full time”

	Social Norm						Personal Value						Benjamini-Hochberg adjusted p-values	
	--	-	+	++	SN>0	Mean rating	--	-	+	++	PV>0	Mean rating	pr-test	t-test
Overall	6.50	16.92	42.07	34.51	76.58	.3631	8.86	17.11	40.51	33.51	74.02	.3237	0.013	0.003
Females	6.56	16.81	44.25	32.38	76.63	.3488	9.36	17.01	42.25	31.38	73.63	.3034	0.025	0.004
Males	6.44	17.04	39.84	36.68	76.52	.3776	8.35	17.22	38.74	35.69	74.43	.3444	0.183	0.060
Age 25-34	7.49	14.21	44.06	34.24	78.30	.3660	9.53	15.81	40.45	34.21	74.66	.3282	0.101	0.109
Age 35-49	5.38	20.12	37.89	36.60	74.49	.3708	8.55	18.79	38.84	33.82	72.66	.3189	0.185	0.005
Age 50-64	7.10	15.15	45.17	32.59	77.76	.3540	8.84	16.13	42.18	32.86	75.04	.3261	0.113	0.118
Females 25-34	6.81	17.05	41.33	34.81	76.14	.3601	10.35	14.71	43.33	31.61	74.94	.3071	0.596	0.199
Females 35-49	.69	19.12	39.88	34.31	74.19	.3447	10.57	17.69	38.08	33.66	71.74	.2982	0.194	0.010
Females 50-64	6.32	14.48	49.82	29.38	79.20	.3473	7.75	17.44	45.72	29.09	74.81	.3067	0.058	0.055
Males 25-34	8.15	11.44	46.72	33.69	80.41	.3718	8.73	16.87	37.65	36.74	74.39	.3487	0.091	0.247
Males 35-49	4.06	21.14	35.88	38.92	74.80	.3972	6.50	19.90	39.61	33.99	73.60	.3399	0.401	0.079
Males 50-64	7.92	15.85	40.25	35.98	76.23	.3612	10.00	14.74	38.43	36.84	75.27	.3467	0.465	0.343
North	7.37	22.25	42.54	27.84	70.38	.2717	9.78	20.15	42.25	27.82	70.07	.2534	0.418	0.252
Center	6.94	13.36	44.81	34.89	79.70	.3833	9.73	12.27	44.56	33.44	78.00	.3326	0.225	0.132
South	5.08	11.92	39.85	43.14	82.99	.4727	7.14	15.89	35.86	41.11	76.97	.4056	0.005	0.003
North females	8.51	18.33	48.09	25.07	73.16	.2638	12.06	18.90	44.10	24.94	69.04	.2119	0.024	0.013
North males	6.22	26.19	36.95	30.63	67.58	.2796	7.48	21.40	40.39	30.72	71.11	.2950	0.093	0.374
Center females	5.35	17.57	44.51	32.58	77.09	.3612	8.73	13.16	45.14	32.97	78.11	.3480	0.454	0.350
Center males	8.59	8.97	45.14	37.30	82.44	.4064	10.78	11.34	43.95	33.93	77.88	.3391	0.229	0.173
South females	4.70	14.36	39.06	41.88	80.94	.4532	6.20	16.76	38.14	38.90	77.04	.3976	0.130	0.061
South males	5.48	9.41	40.67	44.45	85.12	.4929	8.11	15.00	33.50	43.38	76.88	.4138	0.023	0.053

Panel e) Claim 5 “A woman should be ready to reduce the time devoted to her job for family reasons”

	Social Norm						Personal Value						Benjamini-Hochberg adjusted p-values	
	--	-	+	++	SN>0	Mean rating	--	-	+	++	PV>0	Mean rating	Pr-test	p-value
Overall	27.16	30.53	34.76	7.55	42.31	-.1821	33.31	29.86	30.42	6.41	36.83	-.2671	0.000	0.000
Females	28.37	28.48	35.41	7.74	43.15	-.1834	36.29	25.62	31.99	6.10	38.09	-.2809	0.005	0.000
Males	25.93	32.62	34.10	7.35	41.45	-.1809	30.26	34.19	28.82	6.74	35.56	-.2530	0.008	0.003
Age 25-34	31.64	29.07	33.17	6.12	39.29	-.2461	37.01	29.04	28.19	5.76	33.95	-.3153	0.085	0.040
Age 35-49	22.37	30.13	37.93	9.56	47.49	-.1023	29.70	28.84	33.58	7.87	41.45	-.2027	0.003	0.000
Age 50-64	29.62	31.64	32.46	6.29	38.75	-.2306	34.99	31.25	28.44	5.32	33.76	-.3061	0.043	0.003
Females 25-34	27.88	30.44	35.04	6.64	41.68	-.1972	35.84	26.54	31.91	5.71	37.62	-.2836	0.310	0.120
Females 35-49	27.49	27.08	35.75	9.69	45.44	-.1494	37.17	22.27	33.38	7.17	40.55	-.2633	0.035	0.000
Females 50-64	29.45	28.88	35.26	6.41	41.67	-.2093	35.67	28.38	30.70	5.26	35.96	-.2965	0.048	0.007
Males 25-34	35.30	27.74	31.34	5.62	36.96	-.2849	38.15	31.47	24.58	5.80	30.38	-.3462	0.170	0.198
Males 35-49	17.20	33.22	40.14	9.44	49.58	-.0547	22.14	35.49	33.79	8.58	42.37	-.1413	0.033	0.013
Males 50-64	29.80	34.55	29.49	6.16	35.65	-.2531	34.28	34.29	26.05	5.38	31.43	-.3162	0.305	0.083
North	29.05	30.11	33.65	7.19	40.84	-.2069	34.19	32.27	27.75	5.79	33.54	-.2989	0.000	0.000
Center	27.14	32.21	32.86	7.79	40.65	-.1913	33.92	28.32	31.26	6.50	37.76	-.2646	0.181	0.021
South	24.66	30.11	37.34	7.89	45.23	-.1439	31.79	27.54	33.48	7.20	40.68	-.2263	0.061	0.003
North females	29.93	28.23	35.40	6.44	41.84	-.2112	37.42	28.06	29.75	4.78	34.53	-.3208	0.005	0.000
North males	28.17	32.00	31.89	7.94	39.83	-.2026	30.93	36.51	25.75	6.81	32.56	-.2768	0.024	0.019
Center females	28.07	31.98	30.92	9.04	39.96	-.1938	34.87	24.40	35.60	5.14	40.74	-.2603	0.404	0.078
Center males	26.17	32.45	34.89	6.50	41.39	-.1887	32.94	32.41	26.74	7.91	34.65	-.2690	0.133	0.220
South females	26.49	26.76	38.05	8.70	46.75	-.1407	35.65	23.13	32.83	8.39	41.22	-.2406	0.160	0.008
South males	22.78	33.56	36.61	7.05	43.66	-.1472	27.81	32.09	34.15	5.96	40.11	-.2116	0.198	0.085

Note. “-” means “very socially inappropriate”; “-” means “somewhat socially inappropriate”; “+” means “somewhat socially appropriate”; “++” means “very socially appropriate”. For each group, we report in bold the modal rating (i.e. the social norm and prevalent personal value). “SN<0” displays the sum of positive answers (i.e. “+” and “++”), while the column “mean rating” displays the average rating obtained by converting subjects’ ratings in the task to numerical values: negative values indicate openness (i.e. disagreement with the claim) and positive values indicate more conservative views (i.e. agreement with the claim). In each panel, the percentages reported are obtained with the Stata command “svy” accounting for the sample weights in order to ensure representativeness within each group considered. The second last column reports results from pr-test while the last column report results from a set of t-tests. In both cases, our underlying hypothesis is that personal values are more progressive than social norms, for this reason all p-values reported are from one-sided tests and corrected using the Benjamini and Hochberg False Discovery Rate Method.

Table 6. (Mis)perception of social norm and (dis)agreement

Panel a) Claim 1: “Men make better political leaders than women do”				
	Agree with the (perceived) norm	Disagree with the (perceived) norm	Overall	p-value pr-test agreement
Correctly Perceived	55.92%	3.55%	59.46 %	p=0.000
Mis-perceived	21.32%	19.22%	40.54%	p=0.000
Overall	77.24%	22.76%		
p-value pr-test misperception	p= 0.000	p= 0.000	p=0000	Pearson: p= 0.000
Panel b) Claim 2: “When jobs are scarce men should have more right to a job than women”				
	Agree with the (perceived) norm	Disagree with the (perceived) norm	Overall	p-value pr-test agreement
Correctly Perceived	49.84%	4.87%	54.71%	p=0.000
Mis-perceived	25.52%	19.77%	45.29%	p=0.000
Overall	75.36%	24.64%		
p-value pr-test misperception	p=0.000	p= 0.000	p=0.000	Pearson: p= 0.000
Panel c) Claim 3: “University education is more important for a boy than for a girl”				
	Agree with the (perceived) norm	Disagree with the (perceived) norm	Overall	p-value pr-test agreement
Correctly Perceived	66.55%	7.10%	73.65%	p=0.000
Mis-perceived	13.28%	13.07%	26.35%	p= 0.889
Overall	79.83%	20.17%		
p-value pr-test misperception	p=0.000	p= 0.000		Pearson: p= 0.000
Panel d) Claim 4 “A preschool child suffers if his or her the mother works fulltime”				
	Agree with the (perceived) norm	Disagree with the (perceived) norm	Overall	p-value pr-test agreement
Correctly Perceived	36.62%	10.39%	47.01%	p=0.000
Mis-perceived	38.52%	14.47%	52.99%	p=0.004
Overall	75.14%	24.86%		
p-value pr-test misperception	p=0.455	p=0.004	p=0.000	Pearson: p= 0.0387
Panel e) Claim 5: “A woman should be ready to reduce the time devoted to her job for family reasons”				
	Agree with the (perceived) norm	Disagree with the (perceived) norm	Overall	p-value pr-test agreement
Correctly Perceived	28.95%	13.46%	42.41%	p=0.000
Mis-perceived	38.30%	19.29%	57.59%	p=0.000
Overall	67.25%	32.75%		
p-value pr-test misperception	p=0.000	p=0.000	p=0.000	Pearson: p= 0.5225

Note. In each panel, the percentages reported are obtained with the Stata command “.svy” accounting for the sample weights in order to ensure representativeness within each group considered. The variables are defined at the individual level. In order to measure perception, we identify individuals who correctly identified the social norm *within their reference group* (i.e. respondents of the same gender, age group (i.e. 25-34; 35-49; 50-64), and area of residence (i.e. North, Centre, South and Islands)). To measure disagreement, we identify individuals who submitted a different answer in the norm elicitation question (part 2 of the survey-incentivized) and in the personal attitude question (part 3 of the survey, non-incentivized). All p-values are corrected using the Benjamini-Hochberg False Discovery Rate method (Benjamini and Hochberg, 1995).

From Personal Values to Social Norms

Francesca Barigozzi and Natalia Montinari

Section A.1 Additional Tables

Table A1. Summary Statistics

Variable	Mean(SD)/ Proportion	N	Min	Max	Description
Gender					
Female	58.43	877	0	1	Dummy, 1 if female, 0 otherwise
Male	41.57	624	0	1	Dummy, 1 if male, 0 otherwise
Geographical Areas					
North	47.90	719	0	1	Dummy, 1 if the respondent is resident in the North, 0 otherwise
Center	18.92	284	0	1	Dummy, 1 if the respondent is resident in the Center, 0 otherwise
South	33.18	498	0	1	Dummy, 1 if the respondent is resident in the South or Islands, 0 otherwise
Age Group					
25-34	19.85	298	0	1	Dummy, 1 if the respondent belongs to the age group 25-34, 0 otherwise
35-49	52.43	787	0	1	Dummy, 1 if the respondent belongs to the age group 35-49, 0 otherwise
50-64	27.71	416	0	1	Dummy, 1 if the respondent belongs to the age group 50-64, 0 otherwise
Civil Status					
Single Widower, Separated Divorced	27.85	418	0	1	Dummy, 1 if the respondent is single, separated, widower or divorced, 0 otherwise
Married or Cohabitant	72.15	1083	0	1	Dummy, 1 if the respondent is married or cohabitant, 0 otherwise
Having children					
Yes	41.37	621	0	1	Dummy, 1 if the respondent has children, 0 otherwise
No	58.63	880	0	1	
University Degree					
Yes	64.62	970	0	1	Dummy, 1 if the respondent has a University degree, 0 otherwise
No	35.38	531	0	1	
Working					
Yes	63.82	958	0	1	Dummy, 1 if the respondent has a children, 0 otherwise
No	36.18	543	0	1	
Important dimensions in life					
All equally important	3.06	46	0	1	All dimensions have the same important life dimension
Free time	16.86	253	0	1	Free time is the most important life dimension
Community	2.27	34	0	1	Community is the most important life dimension
Work	22.78	342	0	1	Work is the most important life dimension
Religion	1.80	27	0	1	Religion is the most important life dimension
Family	70.55	1059	0	1	Family the most important life dimension
Political view; center right	24.38	366	0	1	Dummy, 1 if the respondent self-report to be right wings, 0 otherwise
Framing	54.1	689	0	1	Dummy which account for the framing used in the vignette
Municipality Dimensions*					
Less than 5,000	4.55%	59/12 98	0	1	Dummy, 1 if the respondent's city of residence has less than 5,000 inhabitants, 0 otherwise
Between 5,000 and 10,000	7.99%	120/1 298	0	1	Dummy, 1 if the respondent's city of residence has between 5,000 and 10,000 inhabitants, 0 otherwise
Between 10,000 and 50,000	53.90%	809	0	1	Dummy, 1 if the respondent's city of residence has between 10,000 and 50,000 inhabitants, 0 otherwise
Gini Index*	.2139 (.0211)	1293	.162	.264	Inequality "between-groups"; thus representing a "lower bound" of the overall population inequality. Source: Urban Index
Male/Female Employment Ratio*	1.5232 (.2889)	1298	1.18	1.32	Ratio (percent) between male employment ratio (employed males wrt resident male population of 15 years or more) and female employment ratio (employed females wrt resident female population of 15 years or more).

TIPI: Ten Item Personality Inventory				
Agreeableness	5.3078 (1.0856)	1501	1.5	7
Conscientiousness	5.5190 (1.1203)	1501	1.5	7
Emotional Stability	4.5410 (1.2425)	1501	1	7
Openness	4.2818 (1.0476)	1501	1	7
Extraversion	3.9957 (1.3708)	1501	1	7

* Source: urbanindex.it; Atlante PRIN Postmetropoli, elaborazioni su dati MEF - Ministero dell'Economia e della Finanza

Table A2. Social appropriateness rating and general agreement by the reference group

Panel a). Claim 1: “Men make better political leaders than women do”

	Social Norm					SN>0	Mean rating	Personal Value				PV>0	Mean rating	pr-test	t-test
	--	-	+	++	--			-	+	++					
North															
Females 25-34	44.98	38.73	13.42	2.88	16.30	-.5045	54.78	26.48	18.74	0.00	18.74	-.5733	1.175	0.435	
Females 35-49	62.47	23.95	12.05	1.54	13.59	-.6486	69.15	19.61	9.99	1.26	11.25	-.7106	0.241	0.019	
Females 50-64	68.84	25.26	3.56	2.33	5.89	-.7367	77.92	17.01	4.72	0.35	5.07	-.8162	0.349	0.016	
Males 25-34	60.35	21.06	16.49	2.10	18.59	-.5976	64.07	24.12	10.80	1.01	11.81	-.6745	0.437	0.225	
Males 35-49	41.06	36.12	20.52	2.29	22.81	-.4392	48.06	30.60	17.04	4.30	21.34	-.4823	0.355	0.303	
Males 50-64	58.75	21.39	12.75	7.10	19.85	-.5450	66.46	22.94	6.61	4.00	10.61	-.6785	0.034	0.028	
Center															
Females 25-34	54.79	34.25	8.56	2.39	10.95	-.6087	68.14	29.47	0.00	2.39	2.39	-.7547	0.244	0.054	
Females 35-49	69.22	19.40	7.56	3.83	11.39	-.6930	71.67	15.21	8.47	4.65	13.12	-.6925	0.694	0.495	
Females 50-64	51.26	33.40	10.24	5.10	15.34	-.5381	75.58	11.55	10.29	2.57	12.86	-.7342	0.583	0.020	
Males 25-34	55.30	29.54	12.88	2.27	15.15	-.5853	71.97	15.15	10.61	2.27	12.88	-.7119	0.868	0.335	
Males 35-49	40.60	27.82	17.08	14.50	31.58	-.2964	44.03	33.3	9.02	13.65	22.67	-.3840	0.115	0.200	
Males 50-64	61.24	27.69	11.07	0.00	11.07	-.6672	69.54	16.61	8.31	5.54	13.85	-.6675	0.716	0.499	
South and Islands															
Females 25-34	70.73	17.60	8.71	2.96	11.67	-.7070	77.70	9.76	6.80	5.75	12.55	-.7292	0.528	0.491	
Females 35-49	65.31	23.94	8.65	2.09	10.74	-.6826	73.69	18.45	7.01	0.85	7.86	-.7661	0.178	0.026	
Females 50-64	68.59	20.15	11.26	0.00	11.26	-.7152	74.77	12.36	11.26	1.61	12.87	-.7352	0.473	0.386	
Males 25-34	62.34	25.32	3.88	8.47	12.35	-.6095	57.03	25.40	14.32	3.24	17.56	-.5745	0.128	0.425	
Males 35-49	70.98	20.61	8.41	0.00	8.41	-.7501	69.54	17.43	11.59	1.44	13.03	-.7003	0.193	0.233	
Males 50-64	53.93	32.55	9.81	3.71	13.52	-.5773	63.09	19.88	17.03	0.00	17.03	-.6404	0.700	0.396	

Panel b). Claim 2: “When Jobs are scarce men should have more right to a job than women”

	Social Norm						Personal Value						pr-test	t-test
	--	-	+	++	SN>0	Mean rating	--	-	+	++	PV>0	Mean rating		
North														
Females 25-34	59.67	22.44	17.89	0.00	17.89	-.6117	57.65	23.29	19.06	0.00	19.06	-.5905	0.674	0.755
Females 35-49	57.61	22.32	15.26	4.80	20.06	-.5514	63.11	16.82	16.68	3.39	20.07	-.5977	0.768	0.402
Females 50-64	56.28	15.48	23.85	4.39	28.24	-.4913	58.32	18.26	17.95	5.47	23.42	-.5296	0.440	0.740
Males 25-34	69.94	17.37	8.64	4.05	12.69	-.6878	73.14	15.78	8.12	2.96	11.08	-.7270	0.615	0.992
Males 35-49	45.04	24.74	25.19	5.03	30.22	-.3986	46.14	26.77	22.30	4.80	27.10	-.4281	0.555	0.215
Males 50-64	57.77	15.29	18.75	8.18	26.93	-.4845	62.45	22.76	10.11	4.67	14.78	-.6196	0.580	0.157
Center														
Females 25-34	69.52	28.08	2.39	0.00	2.39	-.7800	76.70	20.90	2.39	0.00	2.39	-.8281	0.384	0.690
Females 35-49	54.10	27.78	14.75	3.37	18.12	-.5503	62.02	21.13	14.75	2.09	16.84	-.6203	0.168	0.295
Females 50-64	46.12	19.22	26.99	7.67	34.66	-.3588	65.29	14.13	18.01	2.57	20.58	-.6143	0.795	0.878
Males 25-34	57.58	15.15	8.33	18.94	27.27	-.4089	62.12	10.61	8.33	18.94	27.27	-.4394	0.487	0.502
Males 35-49	44.89	29.75	21.70	3.65	25.35	-.4390	52.95	23.30	12.79	10.96	23.75	-.4545	0.820	0.357
Males 50-64	55.70	24.92	19.38	0.00	19.38	-.5753	64.01	22.15	8.31	5.54	13.85	-.6304	0.328	0.771
South and Islands														
Females 25-34	63.76	19.69	12.72	3.83	16.55	-.6223	70.73	12.89	14.46	1.92	16.38	-.6829	0.397	0.859
Females 35-49	56.87	17.60	21.34	4.19	25.53	-.5145	60.61	18.33	16.02	5.04	21.06	-.5633	0.299	0.118
Females 50-64	52.92	22.95	22.52	1.61	24.13	-.5145	61.64	13.55	21.59	3.22	24.81	-.5577	0.741	0.605
Males 25-34	57.75	18.11	18.91	5.22	24.13	-.5226	57.67	19.55	18.28	4.51	22.79	-.5358	0.023	0.007
Males 35-49	58.78	12.20	18.65	10.37	28.93	-.4628	64.62	12.81	15.08	7.50	22.58	-.5638	0.578	0.892
Males 50-64	38.86	37.56	19.88	3.71	23.59	-.4099	43.87	29.69	24.04	2.40	26.44	-.4333	0.842	0.695

Panel c). Claim 3 “University education is more important for a boy than for a girl”

	Social Norm						Personal Value						pr-test	t-test
	--	-	+	++	SN>0	Mean rating	--	-	+	++	PV>0	Mean rating		
North														
Females 25-34	72.32	17.72	8.52	1.44	9.96	-.7391	69.87	18.15	11.98	0.00	11.98	-.7191	0.728	0.410
Females 35-49	76.59	14.78	6.27	2.36	8.63	-.7703	80.29	10.51	7.08	2.13	9.21	-.7929	0.301	0.096
Females 50-64	83.93	10.36	3.56	2.14	5.70	-.8404	84.19	8.74	4.66	2.41	7.07	-.8313	0.025	0.056
Males 25-34	68.46	18.71	10.74	2.10	12.84	-.6899	67.31	21.82	7.91	2.96	10.87	-.6894	0.283	0.086
Males 35-49	56.07	26.53	16.19	1.20	17.39	-.5829	52.23	28.20	15.86	3.70	19.56	-.5260	0.318	0.790
Males 50-64	65.93	19.77	7.29	7.02	14.31	-.6303	72.41	15.44	8.65	3.50	12.15	-.7115	0.498	0.779
Center														
Females 25-34	64.36	33.24	2.39	0.00	2.39	-.7454	76.70	14.73	8.56	0.00	8.56	-.7874	0.571	0.903
Females 35-49	84.88	9.65	2.55	2.91	5.46	-.8431	82.69	7.65	5.46	4.19	9.65	-.7923	0.783	0.979
Females 50-64	80.83	11.50	5.15	2.52	7.67	-.8040	78.15	12.87	8.98	0.00	8.98	-.7944	0.747	0.521
Males 25-34	49.24	48.48	2.27	0.00	2.27	-.6449	49.24	42.42	0.00	8.33	8.33	-.5491	0.354	0.308
Males 35-49	67.56	15.36	11.60	5.48	17.08	-.6332	60.14	21.59	9.13	9.13	18.26	-.5512	0.723	0.165
Males 50-64	77.85	19.38	0.00	2.77	2.77	-.8148	80.62	13.84	0.00	5.54	5.54	-.7965	0.328	0.665
South and Islands														
Females 25-34	75.61	10.80	6.80	6.80	13.60	-.7014	76.48	5.05	13.59	4.88	18.47	-.6878	0.337	0.283
Females 35-49	80.03	9.50	6.28	4.19	10.47	-.7690	85.86	6.22	5.37	2.55	7.92	-.8359	0.149	0.015
Females 50-64	83.24	9.65	3.22	3.90	7.12	-.8146	81.88	9.65	2.29	6.18	8.47	-.7813	0.529	0.634
Males 25-34	81.17	13.69	3.88	1.26	5.14	-.8315	65.50	16.30	14.32	3.88	18.20	-.6228	0.009	0.030
Males 35-49	73.64	13.33	3.49	9.54	13.03	-.6734	71.07	13.33	9.85	5.75	15.60	-.6646	0.078	0.016
Males 50-64	71.61	22.28	5.46	0.65	6.11	-.7651	75.96	16.83	7.21	0.00	7.21	-.7914	0.508	0.987

Panel d). Claim 4: “A preschool kid is likely to suffer if his or her mother works full time”

	Social Norm						Personal Value						pr-test	t-test
	--	-	+	++	SN>0	Mean rating	--	-	+	++	PV>0	Mean rating		
North														
Females 25-34	9.36	14.44	44.48	31.72	76.20	.3228	12.23	12.99	46.35	28.42	74.77	.2720	0.470	0.174
Females 35-49	10.38	22.20	41.35	26.07	67.42	.2201	14.13	20.88	39.50	25.49	64.99	.1751	0.193	0.015
Females 50-64	6.35	16.39	56.14	21.12	77.26	.2789	10.02	19.67	47.46	22.85	70.31	.2200	0.697	0.013
Males 25-34	5.99	15.98	48.77	29.26	78.03	.3409	3.90	14.10	45.86	36.15	82.01	.4274	0.189	0.180
Males 35-49	3.28	32.61	34.55	29.56	64.11	.2692	5.07	26.89	38.78	29.25	68.03	.2810	0.710	0.364
Males 50-64	9.25	24.71	33.71	32.34	66.05	.2606	11.58	19.45	39.37	29.59	68.96	.2458	0.027	0.011
Center														
Females 25-34	8.56	19.52	43.83	28.08	71.91	.2754	14.73	7.18	52.39	25.69	78.08	.2588	0.195	0.043
Females 35-49	5.46	20.22	41.80	32.52	74.32	.3417	8.38	16.30	40.71	34.61	75.32	.3429	0.555	0.990
Females 50-64	3.79	14.13	47.43	34.66	82.09	.4187	6.36	12.82	46.17	34.66	80.83	.3931	0.699	0.016
Males 25-34	2.27	4.55	46.97	46.21	93.18	.5793	2.27	12.88	40.91	43.93	84.84	.5091	0.347	0.142
Males 35-49	6.34	11.82	39.41	42.43	81.84	.4520	9.02	10.85	47.69	32.44	80.13	.3557	0.092	0.160
Males 50-64	14.01	8.31	50.00	27.69	77.69	.2744	16.78	11.07	41.69	30.46	72.15	.2378	0.573	0.997
South and Islands														
Females 25-34	2.96	18.81	36.41	41.82	78.23	.4466	5.92	20.56	35.19	38.33	73.52	.3723	0.845	0.785
Females 35-49	2.43	14.26	36.72	46.58	83.30	.5157	7.07	14.20	34.57	44.15	78.72	.4380	0.284	0.031
Females 50-64	7.79	12.11	42.68	37.42	80.10	.3972	5.50	17.19	43.10	34.21	77.31	.3725	0.567	0.618
Males 25-34	13.61	9.73	44.23	32.43	76.66	.3021	17.57	22.07	26.58	33.78	60.36	.1770	0.154	0.510
Males 35-49	3.79	10.76	35.60	49.85	85.45	.5425	6.97	15.60	35.91	41.52	77.43	.4125	0.232	0.372
Males 50-64	2.40	7.86	43.67	46.07	89.74	.5548	3.71	10.27	35.15	50.88	86.03	.5538	0.560	0.317

Panel e). Claim 5: “A woman should be ready to reduce the time devoted to her job for family reasons”

	Social Norm						Personal Value						pr-test	t-test
	--	-	+	++	SN>0	Mean rating	--	-	+	++	PV>0	Mean rating		
North														
Females 25-34	25.23	31.73	38.14	4.90	43.04	-.1822	29.71	30.11	34.69	5.49	40.18	-.2271	0.642	0.513
Females 35-49	32.75	26.55	32.43	8.28	40.71	-.2253	39.64	24.10	31.59	4.67	36.26	-.3250	0.095	0.003
Females 50-64	29.36	28.26	37.00	5.39	42.39	-.2108	38.76	30.90	25.77	4.57	30.34	-.3589	0.004	0.002
Males 25-34	36.45	27.80	27.43	8.32	35.75	-.2825	38.87	34.11	17.84	9.19	27.03	-.3505	0.164	0.320
Males 35-49	17.32	35.22	40.14	6.82	46.96	-.0938	20.78	37.18	34.42	7.62	42.04	-.1407	0.247	0.343
Males 50-64	34.47	30.81	25.86	8.87	34.76	-.2108	37.21	37.00	20.93	4.87	25.80	-.3765	0.117	0.063
Center														
Females 25-34	30.48	32.87	28.08	8.56	36.64	-.2349	46.22	28.08	23.30	2.39	25.69	-.4541	0.119	0.041
Females 35-49	25.50	34.15	32.33	8.01	40.34	-.1809	32.60	25.32	36.98	5.10	42.08	-.2366	0.707	0.306
Females 50-64	29.47	29.47	30.83	10.24	41.07	-.1877	31.94	21.85	39.81	6.41	46.22	-.1960	0.349	0.897
Males 25-34	28.03	23.49	37.88	10.61	48.49	-.1268	40.15	19.70	31.82	8.33	40.15	-.2782	0.354	0.263
Males 35-49	12.45	36.09	43.29	8.17	51.46	-.0191	23.30	40.60	28.78	7.31	36.09	-.1990	0.023	0.003
Males 50-64	39.09	33.22	24.92	2.77	27.69	-.3906	39.09	30.46	22.15	8.31	30.46	-.3352	0.328	0.450
South and Islands														
Females 25-34	29.61	27.70	35.02	7.67	42.69	-.1953	37.63	21.60	33.10	7.67	40.77	-.2616	0.704	0.317
Females 35-49	21.55	23.49	42.34	12.62	54.96	-.0271	36.60	17.94	33.62	11.83	45.45	-.1959	0.012	0.000
Females 50-64	29.55	29.38	35.56	5.50	41.06	-.2201	33.70	28.87	31.92	5.50	37.42	-.2719	0.543	0.396
Males 25-34	37.66	29.82	32.52	0.00	32.52	-.3677	36.31	34.41	28.64	0.63	29.27	-.3758	0.482	0.873
Males 35-49	19.17	28.71	38.26	13.86	52.12	-.0216	23.35	30.07	35.91	10.68	46.59	-.1075	0.255	0.076
Males 50-64	17.68	40.61	37.36	4.36	41.72	-.1439	27.29	32.75	35.60	4.36	39.96	-.2199	0.813	0.294

Note. “-” means “very socially inappropriate”; “-” means “somewhat socially inappropriate”; “+” means “somewhat socially appropriate”; “++” means “very socially appropriate”. For each group, we report in bold the modal rating (i.e. the social norm and prevalent personal value). “SN<0” displays the sum of positive answers (i.e. “+” and “++”), while the column “mean rating” displays the average rating obtained by converting subjects’ ratings in the task to numerical values: negative values indicate openness (i.e. disagreement with the claim) and positive values indicate more conservative views (i.e. agreement with the claim). In each panel, the percentages reported are obtained with the Stata command “.svy” accounting for the sample weights in order to ensure representativeness within each group considered. The second last column reports results from pr-test while the last column report results from a set of t-tests. In both cases, our underlying hypothesis is that personal values are more progressive than social norms, for this reason all p-values reported are from one-sided tests and corrected using the Benjamini and Hochberg False Discovery Rate Method. Note The percentages reported are obtained with the Stata command “.svy” accounting for the sample weights in order to ensure representativeness within the subgroups. According to our hypothesis that personal values are more progressive than social norms, all p-values are from one-sided paired pr-tests comparing the proportion of SN >0 vs the proportion of PPV>0 and corrected using the Benjamini-Hochberg False Discovery Rate Method.

Table A.3. Difference between Perceived Social Norms and Prevalent Personal Values

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
	Claim 1	Claim 2	Claim 3	Claim 4	Claim 5
Estimation Method	OLS Regression				
Dependent Variable	Difference between social norm and personal value				
Independent Variables					
Age Range (Baseline: Age range 35-49)					
Age range 25-34	-.0013 (.0359)	-.0206 (.0382)	-.0638 (.0446)	-.0215 (.0351)	-.0329 (.0416)
Age range 50-64	.0410 (.0310)	.0333 (.0343)	.0260 (.0436)	-.0205 (.0301)	-.0175 (.0336)
Female	.0239 (.0278)	-.0014 (.0303)	.0103 (.0301)	.0551 (.0274)	.0493 (.0323)
Geographical Area of residence (Baseline: North)					
Center	.0070 (.0362)	.0201 (.0374)	-.0357 (.0386)	.0013 (.0351)	-.0503 (.0395)
South and Islands	-.1044** (.0434)	.0938* (.0507)	-.07205 (.0522)	.0115 (.0404)	-.0683 (.0497)
Civil Status (Baseline: Single, Widower, Separated-Divorced)					
Married or Cohabitant	.0242 (.0318)	.0564* (.0338)	.0643* (.0373)	.0269 (.0326)	.0003 (.0399)
Having Children	-.0105 (.0332)	-.0228 (.0349)	-.0397 (.0348)	-.0150 (.0319)	.0123 (.0386)
University degree	.0632** (.0295)	.0568* (.0301)	.0940*** (.0307)	-.0401 (.0249)	.0820** (.0322)
Working	.0600* (.0302)	.0552* (.0332)	.0292 (.0337)	.0937*** (.0306)	.0561 (.0359)
Most Important dimension in life (baseline: all dimensions equally important)					
Free time	-.0066 (.0432)	-.0809* (.0446)	-.0355 (.0445)	.0041 (.0416)	.0099 (.0462)
Community	.1369 (.1248)	.0323 (.1092)	-.1193 (.1444)	.2684* (.1379)	.0455 (.1166)
Work	.0850** (.0380)	-.0190 (.0411)	-.0291 (.0429)	.0176 (.0356)	.0551 (.0404)
Family	.0790* (.0418)	-.0388 (.0454)	-.0127 (.0447)	-.0247 (.0371)	-.0104 (.0457)
Political view: Center-Right	-.0296 (.0309)	-.0046 (.0342)	.0397 (.0366)	-.0561* (.0330)	.0010 (.0354)
(Male /Female) Occupation	.0678 (.0659)	.2026** (.0919)	.1229 (.0921)	.9472 (.0577)	.1038 (.0727)
Gini Index	1.6111* (.8414)	2.6461*** (.9379)	-.0482 (.9345)	.5087 (.9125)	.1196 (.9054)
Framing	Yes	Yes	Yes	Yes	Yes
Municipality Dimensions	Yes	Yes	Yes	Yes	Yes
N Observations	1,293	1,293	1,293	1,293	1,293
Population size	27,780.19	27,780.19	27,780.19	27,780.19	27,780.19
F	1.63	1.70	1.51	1.50	1.33
Prob > F (20, 1273)	0.0395	0.0274	0.0684	0.0709	0.1515
R-squared	0.0383	0.0323	0.0321	0.0318	0.0233

Note. Table A3 presents the results from a set of OLS estimations. The dependent variable is obtained as the difference between the answer given in the social norm elicitation and the one given in the personal value elicitation. In all models, we control for the framing of the experiment. We include additional controls at the municipality level: the Gini-index, the ratio between male and female employment ratio and the size of the municipality where the respondent lives using data provided by the Urban Index-Indicatori per le politiche urbane, Presidenza del Consiglio dei Ministri (<https://www.urbanindex.it>), but we lose data on 203 individuals since we are not able to match all municipalities in our dataset. ***, ** and * indicate significance at the 1% level, 5% level and 10% level, respectively.

In Table A.3, to control for individual heterogeneity, we report results from a set of OLS regressions. In all models, the dependent variable is obtained at the individual level for each claim as the difference between the answer given in the incentivized social norm elicitation and the personal value. Each model corresponds to one of the claims. As independent variables, we include socio-demographic information about the respondent, such as gender (through the dummy female which takes value 1 if the respondent is a female and 0 otherwise); age group (through a set of dummies, age 25–34 and age 50 – 64, leaving as omitted category the age group 35 – 49); area of residence (with a set of dummies: Center -which takes value 1 if the respondent is resident in the Center of Italy and 0 otherwise-; and a dummy South and Islands which takes value 1 if the respondent is resident in the South of Italy or in one of the Islands, and 0 otherwise leaving the North as omitted category). We also account for the civil status (through a dummy Married or Cohabitant that takes value 1 if the respondent reports to be married or cohabitant, and 0 if the respondent is single, widow, or separated/divorced); the presence of children (through the dummy children that takes value 1 if the respondent reports having at least one child and 0 otherwise); the job status (using the dummy working equal to 1 if the respondent has a job and 0 otherwise) and education (including the dummy university degree that is equal to 1 if the respondent holds a university degree and 0 otherwise). We control for the relative importance assigned to different dimensions in life with a set of dummies (e.g. free time, community engagement, work, and family which take value 1 if the respondent assigns to each of them the highest relative importance and 0 otherwise, leaving as omitted variable respondents who report assigning equal importance to all dimensions) and for the political orientation (through the dummy center-right that takes value 1 if the respondent self-reports being in line with the ideas of a political party from the center-right, and 0 otherwise). Finally, we control for the framing of the questionnaire, and we include a set of dummies at the municipal level provided by the Urban Index-Indicatori per le politiche urbane, Presidenza del Consiglio dei ministri (<https://www.urbanindex.it>): the size of the municipality where the individual lives, the ratio between male and female occupation, and the GINI index.

Table A4. Misperception of the social norm (Std. Error in parenthesis)

Claim	Misperception: % of individuals who answer wrongly in the social norm elicitation (incentivized)		
	Overall	Perceive a MORE progressive norm than the actual one	Perceive a LESS progressive norm than the actual one
1. Men make better political leaders than women do	40.54% (.0143)	-	100%
2. When jobs are scarce, Men should have more rights to a job than women	45.29% (.0146)	-	100%
3. University education is more important for a boy than for a girl	26.35% (.0129)	-	100%
4. A preschool kid is likely to suffer if his or her mother works full time	52.99% (.0146)	35.67% (.0191)	64.33% (0.0191)
5. A woman should be ready to reduce the time devoted to her job for family reasons	57.59% (.0144)	59.60% (.0382)	40.40% (.0382)

Note. The percentages reported are obtained with the Stata command “.svy” accounting for the sample weights to ensure representativeness within each group considered. In column 2, misperception is identified at the individual level: for each claim, we define a dummy variable that takes value 1 if, in the incentivized task proposed in part 2 of the survey, the individual correctly identified the social norm *within his/her reference group* (i.e. respondents of the same gender, age group (i.e. 25-34; 35-49; 50-64), and area of residence (i.e. North, Centre, South, and Islands)); a value of 0 is assigned otherwise. In columns 3 and 4, we identify whether the perceived social norm is more (less) progressive than the actual social norm. Note that, in claims 1-3, the elicited social norm in all reference groups is “*Strongly disagree*”: this implies that none can perceive a social norm more progressive than the actual social norm. According to a set of tests of proportions, all values reported in the last two columns of the table are significantly different from each other (p-values < 0.0000). All p-values are corrected using the Benjamini-Hochberg False Discovery Rate method (Benjamini and Hochberg, 1995).

Table A.5: Correctly perceiving the social norm (in the reference group), Marginal effects on Logit regression (Standard Error in parenthesis).

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
	Claim 1	Claim 2	Claim 3	Claim 4	Claim 5
Estimation Method	Logit Regression, Marginal Effects (dy/dx)				
Dependent Variable	1 if personal values are more open than the perceived social norm elicitation in the claim, 0 otherwise.				
Independent Variables					
Age Range (Baseline: Age range 35-49)					
Age range 25-34	.0040 (.0653)	-.0070 (.0654)	.0056 (.0791)	-.0483 (.0324)	-.0155 (.0337)
Age range 50-64	.0353 (.0558)	.0799 (.0498)	-.0119 (.0683)	-.0373 (.0244)	-.0097 (.0279)
Female	.0733 (.0506)	.0212 (.0485)	.0910 (.0623)	.0122 (.0245)	.0249 (.0268)
Geographical Area of residence (Baseline: North)					
Center	.0145 (.0625)	.1019* (.0557)	.0392 (.0733)	.0005 (.0304)	.0035 (.0329)
South and Islands	-.0465 (.0587)	.0430 (.0557)	.0505 (.0751)	.0414 (.0264)	-.0194 (.0309)
Civil Status (Baseline: Single, Widower, Separated-Divorced)					
Married or Cohabitant	.0227 (.0561)	-.0141 (.0588)	.1832*** (.0700)	.0072 (.0274)	-.0165 (.0327)
Having Children	-.0248 (.0558)	.0297 (.0550)	-.0209 (.0678)	-.0076 (.0259)	.0236 (.0303)
University degree	.0916** (.0461)	.1402*** (.0437)	.1648*** (.0581)	-.0295 (.0225)	.0441* (.0242)
Working	.0823 (.0552)	.1327 (.0520)	.0297 (.0737)	.0700** (.0276)	.0306 (.0295)
Most Important dimension in life (baseline: all dimensions equally important)					
Free time	.0395 (.0751)	-.0318 (.0741)	-.0310 (.0872)	.0011 (.0325)	-.0531 (.0394)
Community	.0240 (.1417)	.0260 (.1396)	-.1625 (.1609)	.1203* (.0638)	.1026 (.0802)
Work	.0086 (.0596)	-.0282 (.0602)	-.0638 (.0738)	-.0193 (.0287)	.0179 (.0340)
Family	-.0142 (.0622)	-.0257 (.0596)	-.0352 (.0713)	-.0809*** (.0276)	-.0650* (.0341)
Political view: Center-Right	-.0553 (.0574)	-.0307 (.0514)	.0438 (.0646)	-.0406* (.0271)	.0200 (.0284)
Correctly Perceived Social Norm	-	-	-	.0296 (.0224)	-.0052 (.0263)
Municipality Dimensions	Yes	Yes	Yes	Yes	Yes
N Observations	504	561	329	1,298	1,298
Population size	10,837.73	12,356.36	7,244.05	27,884.28	27,884.28
F	(17, 487)=0.87	(17, 544)=1.68	(17, 312)=1.32	(18, 1280)=2.21	(18, 1280)=1.04
Prob > F	0.6139	0.0421	0.1782	0.0025	0.4127

Note For each claim, we define a dummy variable at the individual level (N=1,501). The variable takes value 1 if, in the incentivized task proposed in part 2 of the survey, the individual correctly identified the social norm within his/her reference group (i.e. respondents of the same gender, age group (i.e. 25-34; 35-49; 50-64), and area of residence (i.e. North, Centre, South, and Islands)); a value of 0 is assigned otherwise. Note that, in claims 1-3, the elicited social norm in all reference groups is “Strongly disagree”: this implies that none can have a personal value more progressive than the social norm



Alma Mater Studiorum - Università di Bologna
DEPARTMENT OF ECONOMICS

Strada Maggiore 45
40125 Bologna - Italy
Tel. +39 051 2092604
Fax +39 051 2092664
<http://www.dse.unibo.it>