Gender Norms and Labor-Supply Expectations: Experimental Evidence from Adolescents

Elisabeth Grewenig (ifo Institute)
Philipp Lergetporer (ifo Institute)
Katharina Werner (ifo Institute)

Discussion Paper No. 259

October 26, 2020

Collaborative Research Center Transregio 190 | www.rationality-and-competition.de
Ludwig-Maximilians-Universität München | Humboldt-Universität zu Berlin
Spokesperson: Prof. Dr. Klaus M. Schmidt, University of Munich, 80539 Munich, Germany
+49 (89) 2180 3405 | info@rationality-and-competition.de

Gender Norms and Labor-Supply Expectations: Experimental Evidence from Adolescents*

Elisabeth Grewenig† Philipp Lergetporer‡ Katharina Werner§

October 23, 2020

Abstract

Gender gaps in labor-market outcomes often emerge with the arrival of the first child. We investigate a causal link between gender norms and labor-supply expectations within a survey experiment among 2,000 German adolescents. Using a hypothetical scenario, we document that the majority of girls expects to work 20 hours or less per week when having a young child, and expects from their partner to work 30 hours or more. Randomized treatments that highlight the existing traditional norm towards mothers significantly reduce girls' self-expected labor supply and thereby increase the expected gender difference in labor supply between their partners and themselves (the expected within-family gender gap). Treatment effects persist in a follow-up survey two weeks later, and extend to incentivized outcomes. In a second experiment, we highlight another, more gender-egalitarian, norm towards shared household responsibilities and show that this attenuates the expected within-family gender gap. Our results suggest that social norms play an important role in shaping gender gaps in labor-market outcomes around child birth.

JEL-Classification: J16, J22, C93, D83
Keywords: gender norms, female labor supply, survey experiment

*For helpful comments and discussion, we would like to thank Sofia Amaral, Pietro Biroli, Leonardo Bursztyn, Davide Cantoni, Christina Gathmann, Eleonora Guarnieri, Anna Gumpert, Paul Hufe, Seema Jayachandran, Rainer Kotschy, Andreas Peichl, Brendan Shanks, Ricardo Perez-Truglia, Chris Roth, Dominik Sachs, Simeon Schudy, Peter Schwardmann, Michèle Tertilt, Roel van Veldhuizen, Gregory Veramendi, Joachim Winter, Johannes Wohlfart and participants at the Annual Congress of the German Economic Association in Leipzig, its economic of education group in Vienna, the PhD Workshop in Empirical Economics in Potsdam, the Annual Congress of the European Economic Association in Manchester, the MGSE Colloquium in Munich, and the CRC retreat in Schwanenwerder. We are most grateful to Ludger Woessmann for his constant support throughout the project, and to Sarah Kersten for her help in preparing the surveys. Financial support by Deutsche Forschungsgemeinschaft through CRC TRR 190 (project number 280092119) and by the Leibniz Competition (SAW-2014-ifo-2) is gratefully acknowledged. IRB approval was obtained from the ethics committee of the Ludwigs-Maximilians-Universitaet Munich (Projects 2019-15 and 2018-04). The experiments presented in this study were pre-registered in the AEA RCT Registry (AEARCTR-0004950 and AEARCTR-0003021).

†Grewenig (corresponding author): ifo Institute at the University of Munich; grewenig@ifo.de.
‡Lergetporer: ifo Institute at the University of Munich; CESifo; lergetporer@ifo.de.
§Werner: ifo Institute at the University of Munich; CESifo; werner.k@ifo.de.

1 Introduction

The birth of the first child has large and persistent negative effects on labor-market outcomes of women, but not of men. This finding holds for different countries and over time (Angelov et al., 2015; Kuziemko et al., 2018; Kleven et al., 2019b). Estimates of so-called *child penalties*—the impact of children on earnings of women relative to men—range from 20–25 percent in Scandinavia to 30–40 percent in Anglo-Saxon countries, and 40–60 percent in German-speaking countries (Kleven and Landais, 2017; Kleven et al., 2019a). In fact, the arrival of children is one of the primary reasons for persistent gender inequalities on the labor market (Kleven et al., 2019b). While the disadvantages in the labor market due to childbirth for women compared to men are well documented, only little is known about the underlying causes. In this paper, we argue that gender norms concerning parental labor supply can cause gender differences in outcomes relevant to the labor market.

From a theoretical perspective, such norms may encourage women and men to adjust their labor-market choices to what seems socially appropriate for mothers and fathers to do and in this way, produce gender gaps on the labor market (e.g., Akerlof and Kranton, 2000, 2010; Bertrand et al., 2015; Cortes and Pan, 2020). Indeed, empirical studies show that existing social norms towards maternal labor supply correlate strongly with child penalties across countries (Steinhauer, 2018; Kleven et al., 2019b). Yet, establishing a causal link between social norms and labor-market outcomes is challenging because exogenous variations in norms are rare. We therefore design and run large-scale online survey experiments to experimentally study the effects of social norms on labor-supply expectations.

Our sample consists of 2,000 German adolescents aged between 14 and 17 years. As in many other developed countries, social norms concerning women in general are relatively gender-equal in Germany, but those directly addressing how mothers should behave on the labor market are still very traditional. We focus on adolescents as they already face important, labor-market relevant decisions such as educational or occupational choices. These choices may be affected by labor-supply expectations even prior to labor-market entry. Moreover, understanding the role of social norms is particularly relevant for adolescents, who are in a key phase of gender-differential socialization (e.g., Hill and Lynch, 1983; Priess et al., 2009) that may lay the foundation for later gender inequalities.

We measure labor-supply expectations by presenting adolescents a hypothetical scenario, in which we ask them to imagine themselves at the age of 30 having a child. We then ask respondents about labor-supply expectations for themselves and for their hypothetical partner, allowing us not only to investigate respondents' self-expected labor supply, but also expected labor-supply differences between them and their partner (i.e., the within-family gender gap).

The fact that most adolescents are not yet on the labor market allows us to study expected labor-supply before any actual demand-side restrictions on the labor market are likely to become relevant.[1] We follow a long tradition in economics that studies subjective expectations and preferences concerning the labor market and other domains (see e.g., Manski, 2004; Delavande, 2014; Wiswall and Zafar, 2018). To understand the role of gender norms in labor-market decisions, preferences and expectations about future labor supply are an important and relevant outcome to study because realized labor-market choices can be a result of many different combinations of beliefs (e.g., about gender norms), preferences, and labor-market conditions. By obtaining direct measures for expectations, we isolate the effect of social norms on expected future labor supply.

To experimentally investigate the causal link between gender norms and labor-supply expectations, we first present an experiment that emphasizes the existing social norm prescribing how much mothers and fathers of young children should work on the labor market.[2] We implement a novel experimental design to explicitly study the complementarity between the norm's salience and information about its exact content. In total, we therefore consider three treatments that augment the norm's salience, information, and both: The first treatment *salience* asks respondents to guess the share of German adults who hold the opinion that mothers and fathers, respectively, should reduce their labor-market supply while their children are young. These belief-elicitation questions prime respondents to think of the gender roles for mothers and fathers. The second treatment *information* informs respondents about a representative study showing that 91 percent of Germans think that mothers should reduce their labor supply while their children are young, and 41 percent think that fathers should do so (Wippermann, 2015).[3] The third treatment, *salience and information*, combines both treatments by first posing the belief-elicitation question from treatment *salience* and afterwards providing correct information about the social norm towards mothers and fathers from treatment

[1] In the hypothetical scenario, we explicitly ask respondents how many hours per week they would like to work. This gives us a measure for the adolescents' supply-side intentions that abstracts from respondents' assumptions about the equilibrium mechanisms for the allocation of labor or preferences of the employees. We refer to this measure as "labor-supply expectation" throughout the paper, but we do not mean to use this term to imply that we have elicited probabilistic expectations or the like.

[2] We focus on *injunctive*, or *prescriptive*, social norms (i.e., what behavior is commonly approved of by society), but not on *descriptive* social norms (i.e., what most members of society actually do) (e.g., Cialdini and Trost, 1998; Benabou and Tirole, 2011). While injunctive social norms convey directly what an individual ought to do in a certain situation, descriptive social norms also reflect factors outside of an individual's control (e.g., labor-demand constraints), which renders the interpretation of descriptive-norms effects unclear (see Goerges and Nosenzo (2020) for a recent discussion).

[3] Data from the International Social Survey Program (ISSP) 2012 confirms that traditional views concerning maternal labor supply are not unique to Germany. In fact, in many developed countries—including more gender-egalitarian Scandinavian countries—, most residents think that women with children under school age should work at most part-time (see section 2.2 for details).

information. After treatment, all respondents answer the same questions about their labor-supply expectations as the untreated control group.

In the control group, we already find a large gender gap in self-expected labor supply: Most girls (59 percent) expect to work 20 hours or less per week while most boys (70 percent) expect to work 30 hours or more, which translates into a gender difference of 7.3 work hours per week. To study expected within-family gender gaps in parental labor supply, we exploit the fact that we also elicited respondents' labor-supply expectations for their partner. At baseline, girls expect to work 7.7 hours less than their partner, and, conversely, boys expect to work 10.9 hours more than their partner. Moreover, adolescents hold biased beliefs on the content of the existing gender norm (measured in treatments *salience* and *salience and information*): They largely underestimate the share of German adults thinking that mothers should reduce their labor supply (average belief: 66 percent; true value: 91 percent), and they also underestimate the share thinking that fathers should do so (average belief: 35 percent; true value: 41 percent). These misperceptions indicate leeway for correcting inaccurate beliefs through our information treatments.

Among girls, all three treatments significantly reduce labor-supply expectations. Treatment *salience* significantly reduces their self-expected labor supply by 2.0 hours per week. Treatment *information* reduces labor-supply expectations by 2.6, and treatment *salience and information* by 3.4 hours per week. We draw the following conclusions from this treatment-effect pattern: First, the strong effects of treatment *salience* suggest that priming adolescents to think about the existing gender norm already alters their labor-supply expectations. Second, the effect of the combined treatment *salience and information* is significantly larger ($p<0.1$) than the effect of treatment salience, suggesting that providing information about the norm's content has an additional effect beyond the treatment *salience*. Third, the effect of the combined treatment *salience and information* is smaller than the sum of the two separate effects of treatment *salience* and treatment *information*, suggesting that part of the information effect operates through increasing the norm's salience. Overall, reductions in self-expectations translate into expected within-family gender gaps in labor supply that are more gender-unequal: On average, the treatments induce girls to reduce their expectations about their own labor supply by 2.2 hours compared to their partners' labor supply, thereby significantly increasing expected within-family gender gaps.

Next, we investigate how the gender norm affects boys, a question that—despite its relevance—received much less attention in the public and scientific discourse than the effects on girls. Again, all three treatments reduce self-expected labor supply by 1.3 hours in treatment *salience*, 1.6 hours in treatment *information* and 2.5 hours in treatment *salience and information*. Contrary to the findings on girls, boys' expectations for their own family become more gender-

equal in response to the treatments:[4] On average, the treatments cause boys to reduce their labor-supply expectations by 1.4 hours more than those for their partner. This leads to a reduction of their expected within-family gender gap.

Leveraging our follow-up survey about two weeks after the main survey, we investigate whether information-treatment effects persist beyond the immediate survey horizon.[5] Focusing on treatments *information* and *salience and information*, we find that the information treatments persistently affect both labor-supply expectations and beliefs about the content of the norm. The persistence of treatment effects suggests that they are not due to experimenter-demand effects (Haaland et al., 2020).

We then turn to analyzing the mechanisms underlying our treatment effects. For this purpose, we estimate heterogeneous information-treatment effects by respondents' prior beliefs about the norm's content within the sample of adolescents who received the prior belief elicitation question (treatments *salience* and *salience and information*).[6] The heterogeneity analysis yields two findings: First, treatment effects are already prevalent among respondents with accurate priors, suggesting that information effects are at least partly driven by salience-based information updating (Bleemer and Zafar, 2018). Second, treatments effects are larger in absolute terms (albeit not significant at conventional levels) the more respondents underestimate the content of the social norm, as would be expected for information-based updating.[7]

We subject our main findings to the following robustness checks: First, we address a concern frequently expressed when it comes to unincentivized expectations measures, namely that respondents do not have any monetary incentives to provide meaningful and honest answers. To test whether treatment effects carry over to outcomes with immediate monetary consequences, we additionally elicited beliefs about the German public's views on a set of gender-related statements posed in the European Values Study, and pay respondents for correct answers. Reassuringly, information provision strongly and significantly affects these incentivized outcomes, which shows that treatment effects are not confined to survey answers without direct monetary consequences. Second, since the traditional gender norm prescribes mothers' and fathers' labor supply, we would not necessarily expect the treatments to affect labor-supply

[4]For both girls and boys, the norm treatments change self-expectations but hardly affect the expectations for their partner.

[5]Since priming effects, like the once induced by treatment *salience*, are by definition short-term, we focus our persistence analysis on treatments with informational content, which is standard in the information-provision literature (e.g., Haaland et al., 2020).

[6]Information provision may impact individuals because it makes the importance of gender-norm issues for labor-market participation more salient (e.g., Schwarz and Vaughn, 2002; Chetty et al., 2009; DellaVigna, 2009; Bleemer and Zafar, 2018), or because respondents were misinformed about the exact content of the social norm and update their beliefs accordingly (e.g., Rockoff et al., 2012; Bursztyn et al., 2020).

[7]The interpretation of information-based updating is also consistent with the finding that respondents who were informed about the exact content of the norm hold more accurate posterior beliefs about the share agreeing to the respective norm statement in the follow-up survey.

expectations without child. In fact, treatment effects on labor-supply expectations without child (elicited in an alternative hypothetical scenario) are small and insignificant, which mitigates concerns that our main findings are due to unintended effects such as experimenter-demand effects.

Finally, we investigate whether social norms can shift outcomes in both directions. We therefore conduct a second experiment on a more egalitarian norm in the follow-up survey and particularly study whether this norm could lead to less gender-unequal expectations. The corresponding treatments experimentally highlight a more gender-egalitarian norm towards sharing household responsibilities.[8] In particular, the treatments leverage the fact that 89 percent of Germans think that men should take as much responsibility for the home and children as women (European Values Study 2017). Adolescents hold downward-biased beliefs about this share (average belief: 66 percent), and correcting these false beliefs through randomized information provision significantly reduces expected within-family gender gaps in labor supply after child birth by 1.3 hours per week among girls, and by 1.4 hours per week among boys.

To our knowledge, ours is the first paper to study the causal effects of gender norms on labor-supply expectations after child birth (see section 2.1 for a in-depth discussion on how we contribute to the existing literature). Our findings suggest that social norms indeed play an important role in shaping gender gaps in labor-market outcomes after child arrival. The extent to which social norms shape labor-market outcomes may indeed depend on the degree of egalitarianism in the underlying social prescription. Child penalties for mothers compared to fathers may thus attenuate if norms towards mothers were to become more gender-equal over time.

The remainder of the paper is structured as follows: In section 2 we discuss how we contribute to the existing literature, and provide background information on female labor-market participation in Germany. Section 3 describes our dataset and the experimental design. Section 4 provides descriptive evidence on adolescents' labor-supply expectations. Section 5 shows the results of the first experiment that highlights social norms prescribing parental labor supply, and presents several robustness tests. Section 6 presents results of our second experiment on norms regarding shared household responsibilities. Section 7 concludes.

[8]While the traditional gender norm towards mothers' and fathers' labor supply after child birth studied in the first experiment directly prescribes mothers' and fathers' labor supply, it is not the only gender norm that might be relevant for labor-supply decisions. For instance, different norms prescribe relative income within households (Bertrand et al., 2015), within-household division of work at home and on the labor market, or educational decisions (European Values Study 2017).

2 Background

In this section, we first discuss how our paper relates to, and extends, different strands of the economic literature. Then, we provide institutional background information on female labor-market participation and gender norms in Germany.

2.1 Related Literature

Our paper contributes to several strands of economic research. First, it adds to the growing literature on child penalties that shows that gender gaps in labor-market outcomes often arise with the birth of the first child (e.g., Bertrand et al., 2010; Angelov et al., 2015; Kleven and Landais, 2017; Kuziemko et al., 2018; Kleven et al., 2019a,b). A small subset of this literature studies potential underlying causes of child penalties, and argues that they are not inherent in the biological relationship between mother and child (e.g., Andresen and Nix, 2019; Kleven et al., 2020).[9] Instead, factors related to socialization have been suggested as likely explanations for why the arrival of children has stronger negative labor-market impacts on women than men. Our paper works towards an understanding of whether strong social norms prescribing mothers' and fathers' labor supply–one specific but important societal factor—can explain gender gaps in parental labor supply.

Second, our paper contributes to a larger strand of literature that studies gender gaps unrelated to parenthood. This literature argues that social norms may promote gender gaps in the labor market (for a survey of this literature see Bertrand et al., 2010; Olivetti and Petrongolo, 2016; Blau and Kahn, 2017; Jayachandran, 2019; Altonji and Blank, n.d.). In particular, several studies suggest that women's labor-market outcomes have a cultural component, which is often interpreted as indirect evidence of the importance of gender norms (e.g., Fernández et al., 2004; Fernández, 2007; Fernández and Fogli, 2009; Alesina et al., 2013; Fernández, 2013; Charles et al., 2018; Giuliano, 2018; Boelmann et al., 2020). In this respect, a few studies (e.g., Fortin, 2005; Giavazzi et al., 2013; Fortin, 2015) demonstrate direct cross-country correlations between labor-market outcomes of women and injunctive gender norms measured in large-scale surveys like the World Values Survey.[10] To our knowledge, only two papers investigate such correlations in the context of labor-market outcomes of parents (Steinhauer, 2018; Kleven

[9] Kleven et al. (2020) compare child penalties among biological and adoptive families and find that in both types of families, men's and women's labor-market trajectories are very similar until the arrival of the first child, and diverge with child arrival due to an abrupt and persistent negative shock on females' labor-market outcomes. Similarly, Andresen and Nix (2019) investigate child penalties among female same-sex couples that include the biological mother of the child. They find no long-term differences in labor-market outcomes between the biological mother and the "co-mother".

[10] In this literature, commonly analyzed items usually focus on women's role as caregiver vs. breadwinner (e.g., agreement/disagreement to the statements "Being a housewife is just as fulfilling as working for pay." or "When a wife earns more than her husband, it is almost certain to cause problems.").

et al., 2019b). In contrast to these descriptive studies, our survey experiment allows us to identify the direct and causal link between gender norms and expected labor-market outcomes.

In this sense, our paper is closest to the study by Bursztyn et al. (2020) which finds that experimentally shifting perceived norms towards female labor-market participation in Saudi Arabia increases the willingness of married men to let their wives join the labor force. While Saudi Arabia provides an interesting case study for the role of gender norms in an extremely gender-unequal setting, we test the causal link between social norms and labor-market outcomes in Germany, a country with gender equality laws similar to those of most other developed countries.[11] Importantly, we focus on social norms and labor-market outcomes of mothers rather than women in general, which is particularly important in the context of developed countries: In these countries, overall gender gaps in labor-market outcomes have decreased over time and are now relatively small (e.g., Blau and Kahn, 2017), but gender gaps among parents are still large and persistent (Kleven et al., 2019b). Furthermore, our treatment addresses the social norm towards both mothers and fathers, which extends the existing experimental literature that has so far exclusively focused on gender norms relating to females' labor supply. Finally, we not only investigate how gender norms affect decisions of girls regarding their own and their partner's labor supply as parents, but also of boys.

Third, our paper uses elements from the literature that leverages subjective expectation-data to study decision-making under uncertainty (Manski, 2004). These studies often focus on university students, and investigate expectations or preferences about family life, labor-market relevant decisions such as educational choice or investment in children (e.g., Arcidiacono et al., 2012; Cunha et al., 2013; Stinebrickner and Stinebrickner, 2013; Delavande and Zafar, 2019; Wiswall and Zafar, 2019), and sometimes also gender differences in expectations (e.g., Goldin et al., 2006; Zafar, 2013; Reuben et al., 2017; Wiswall and Zafar, 2018). The rationale for using expectations-data rather than realized outcomes is that observed choices can be consistent with many different combinations of beliefs and preferences (Manski, 2004), which renders the investigation of subjective beliefs and expectations highly relevant.

Finally, the fact that we leverage the norm's salience to measure its effects on labor-market expectations is related to the literature that uses salience-treatments to prime subjects' social identities (e.g., Benjamin et al., 2010; Boschini et al., 2012) or increase salience of topics like immigration (e.g., Alesina et al., 2018; Bleemer and Zafar, 2018; Aloud et al., 2020). For example, Aloud et al. (2020) focus on female university students in Saudi Arabia and investigate the effects of (i) priming them to think about their parents and family, and (ii)

[11] Until mid–2019, Saudi Arabia had very strict "guardianship laws" that would not allow women any actions related to work, leisure, health, finances, and law without the permission or company of a close male relative (The Economist, July 20, 2019).

informing them about their peers' aspirations on labor-market expectations. They find that both priming and information increase expected labor-force participation. In contrast to this literature, our treatments directly address the injunctive social norm for parents' work hours.

2.2 Institutional Background

Although Germany has a comparatively high female labor-market-participation rate of about 56 percent, it is still around 11 percentage points below the rate of males (OECD, 2017). Large gender differences do not only exist at this extensive margin, but also at the intensive margin: Panel A of Appendix Figure A1 shows the share of male and female part-time employees across countries. In Germany, 37 percent (9 percent) of all employed women (men) work part-time, resulting in a gender gap in part-time employment of 28 percentage points, the largest in all observed countries. Recent research suggests that the arrival of children is one of the primary reasons for persistent gender inequalities on the labor market (Kleven et al., 2019a). Indeed, Germany exhibits the largest long-run child penalty of 61 percent among all countries observed (see panel B of Appendix Figure A1).

Appendix Figure A2 provides direct evidence for the existence of conservative injunctive gender norms regarding the labor supply of mothers in different countries. The figure shows that 90 percent of Germans think that women with children under school age should work at most part-time (International Social Survey Program (ISSP) 2012). Comparing this share across developed countries, it is striking that (i) the norm prescribes mothers to reduce their working hours in all observed countries—including more gender-egalitarian Nordic countries—and (ii) German gender norms are among the most traditional. Among the large set of existing gender norms (e.g., prescribing the division of responsibility for the home and children, or relative within-household income shares), our main experiment focuses on the norm that directly prescribes parental labor supply, our main outcome.

Germany offers a wide range of family-friendly policies intended to support female labor supply. Since 2013, every child from the age of one has a legal claim for a public childcare place. Childcare is heavily subsidized, which implies comparably low average costs for parents of between 0 Euros and 400 Euros per month (Geis-Thoene, 2018). Parents are entitled to 12 months of paid parental leave after child birth, which can be extended to 14 months if each parent takes at least 2 months of parental leave. Parents are also eligible for unpaid and job-protected parental leave of up to 3 years for each child. Given this policy environment, our main outcomes of interest are labor-market expectations when the child is between 1 year (i.e.,

when the legal claim for a childcare place becomes effective, and paid parental leave ends) and 6 years old (i.e., when compulsory schooling starts) (see section 3.2.1).[12]

3 Data and Experimental Design

In this section, we first describe the data-collection and sampling process, and then present the experimental design, sample characteristics, and the empirical strategy.

3.1 Data Collection and Sample

Our online survey was conducted between October and December 2019 and covers a sample of 2,000 German adolescents aged between 14 and 17 years.[13] The main survey comprises 11 questions related to educational, career, and labor-market decisions. In addition, we elicited a rich set of sociodemographic characteristics at the end of the survey. Median completion time was 12 minutes.

Sampling and polling were carried out by the German polling firm *konkret Mafo* (https://www.konkret-mafo.de/) who fielded the survey via online access panels.[14] The recruitment of adolescent respondents took place in two ways: First, adolescents who were registered in the online panels were recruited directly (60 percent of our analytic sample). The remaining 40 percent were recruited indirectly via their parents who were registered in the online panels. These parents were first asked for their permission to survey their child. If the parents agreed, they received a survey link to be shared with their child.[15] All respondents answered the questionnaire autonomously on their own digital devices.

To test the persistence of treatment effects, we implemented a follow-up survey about two weeks after the main survey. The follow-up survey re-elicited some outcomes without repeating any treatments from the main survey, and included the second experiment on the effects of a more gender-egalitarian norm towards the end (see section 3.2.2 for details).

[12] Childcare take-up after the child's first birthday is 33 percent for one-year olds, and 66 percent for two-year olds (Alt et al., 2017). Besides factors like childcare-slot shortages, social norms towards maternal labor supply have been discussed as a potential reason for non-take up (e.g., Jessen et al., 2020).

[13] Our experimental setup is based on a short pilot experiment that the we conducted within the scope of the ifo Education Survey 2018 (see Online Appendix O.1.1. for further details).

[14] Throughout the paper, we present unweighted analyses that assign equal weights to each respondent. It is reassuring that re-weighting observations to match official statistics with respect to gender, age, sate of residence and municipality size does not affect our qualitative results (results available upon request). In the context of adult samples, Grewenig et al. (2018) show that online surveys represent the overall population (online and offline) well.

[15] To ensure that the children and not their parents answered the survey, we incorporated several plausibility checks of age and birth date. In case of failure to provide consistent answers, respondents were exited from the survey. Importantly, treatment effects are prevalent among respondents recruited in both modes (results available upon request).

3.2 Experimental Design

3.2.1 Main Survey

The main survey consists of three stages. In stage one, respondents were randomly assigned to a control group or to one of three social-norm treatments (treatment *salience*, *information* or *salience and information*). Stage two elicited labor-supply expectations, our main outcome of interest. In stage three, we asked additional questions, e.g., incentivized outcome questions.[16]

Treatments: Before eliciting outcomes, respondents were randomly assigned to one of four experimental groups with equal probability. One group is the untreated control group. The other three groups receive different norm treatments that emphasize the existing social norm related to parents' labor supply. Unlike previous experimental studies, our treatments emphasize the social prescriptions towards mothers and fathers, not only mothers.[17]

Treatment *salience* used a belief-elicitation question to prime respondents with gender-norm considerations.[18] Before stating the outcomes of interest, treated respondents were asked: *"What do you think, how many adults in Germany hold the opinion that mothers and fathers, respectively, should reduce their labor supply while the children are young? We do not think of the first months after child birth, but the time thereafter."* Respondents were provided with two open answer fields, one for mothers, and one for fathers (see panel A of Appendix Figure A3 for a screenshot). This treatment was designed to make the social norm salient without providing information about the norm's content.

The second treatment *information* did not elicit beliefs, but instead provided respondents with information about the share of German adults who hold the opinion that mothers and fathers of young children should reduce their labor supply. We drew on results from a representative study by the *Federal Ministry for Family Affairs, Senior Citizens, Women and Youth* fielded in 2015 (Wippermann, 2015): *"Out of 100 adults in Germany, 91 hold the opinion that the mother should reduce her labor supply while the children are young. At the same time, out of 100 adults in Germany, 41 hold the opinion that the father should reduce his labor supply while the children are young."* Reassuringly, the norm concerning mothers' labor supply in Germany collected by the ISSP in the year 2012 are practically identical to the one of Wippermann (2015) that we use (see section 2.2), which indicates the robustness and

[16] Online Appendix O3 presents the question wordings.

[17] Our gender-bifocal treatment is in contrast to most of previous empirical studies on gender norms in the sense that this literature often exclusively studies norms concerning women (e.g., Fernández and Fogli, 2009; Alesina et al., 2013; Aloud et al., 2020; Bursztyn et al., 2020). To understand the wholistic impact of gender norms on parental labor-market participation and associated gender gaps, we find it crucial to examine norms that also prescribe the behavior of men.

[18] Other examples of studies that use belief-elicitation questions to increase salience are Alesina et al. (2018) and Aloud et al. (2020).

persistence of the norm. Along with the verbal statement, respondents were shown a graphical illustration of the provided information (see panel B of Appendix Figure A3 for a screenshot).

The treatment *salience and information* combines both treatments: Respondents were first asked the belief-elicitation questions as in treatment *salience*, and were then provided with the factual information about the norm as in treatment *information*. This treatment allows us to study the complementarity between salience and information provision.

Labor-supply expectations: Our main outcomes of interest are the respondents' labor-supply expectations that we elicited after treatment administration. In a hypothetical scenario, respondents were asked to imagine being 30 years old, living with their partner and having a child aged between 1 and 6 years. Our objective was to abstain as much as possible from adolescents' assumptions about the equilibrium mechanisms for the allocation of labor, or preferences of the employees to obtain a meaningful measure for the adolescents' supply-side intentions. After presenting the hypothetical scenario, we therefore elicit labor-supply expectations as follows:[19] We then elicit labor-supply expectations as follows: *"What do you think, how many hours per week on average would you like to work in order to earn money?"* and *"And how many hours per week on average would you like your partner to work in order to earn money?"* Answers to both questions were recorded on a 5-point scale ("0 hours, i.e. not at all"; "about 10 hours"; "about 20 hours"; "about 30 hours"; "about 40 hours, i.e. full-time").

For our main analyses, we combine responses to these two expectations questions to analyze the following two outcomes: (a) self-expected labor supply with child; and (b) the expected within-family gender gap in labor supply with child, calculated as the difference in (i) expectations regarding the male family member's labor supply (i.e., male respondents' self-expectations, and female respondents' expectations regarding their partner), and (ii) expectations regarding the female family member's labor supply (i.e., female respondents' self-expectations, and male respondents' expectations regarding their partner).[20]

Since expectations play an important role in any decision under uncertainty, we believe that our main outcomes of interest are well suited to study the impact of gender norms on labor market-relevant behavior. By obtaining direct measures for labor-market expectations, we can isolate the causal effect of gender norms on adolescents' future labor-supply expectations. Reassuringly, several studies show that (labor-market) expectations strongly predict actual

[19] Note that it is common in the literature to study expectation for events that occur several years in the future as we do (e.g., Goldin, 2014; Reuben et al., 2017; Wiswall and Zafar, 2018).

[20] In order to avoid implying mixed-sex relationships, we do not refer to the gender of partners when asking about expectations regarding labor supply. Nonetheless, for the sake of simplicity and readability, we refer to partners of female respondents as "male" or "father" and partners of male respondents as "female" or "mothers" throughout the paper. If anything, we expect this slight inaccuracy to cause a downward bias in the observed within-family gender gaps.

(labor-market) realizations several years later (e.g., Wiswall and Zafar, 2018, 2019; Kunz and Staub, 2020).

3.2.2 Follow-up Survey

We invited all respondents to participate in the follow-up survey two weeks after the main survey to (i) investigate the persistence of treatment effects, and (ii) implement a second experiment focusing on a more gender-egalitarian norm related to sharing household responsibility. Between 14 and 35 days after the main survey (median time lag: 17 days), we re-surveyed 1,319 respondents (66 percent of the sample).[21] Below, we introduce the individual stages of the follow-up survey:

Persistence of treatment effects: To study persistent treatment effects on labor-supply expectations, we first re-elicited labor-supply expectations as in the main survey. We then asked all respondents the belief-elicitation questions from the treatment *salience* (see section 3.2.1) to assess whether the treatments led to persistent changes in respondents' beliefs about existing norms regarding the labor supply of mothers and fathers with young children.

Treatments: Thereafter, we implemented our second experiment to test the impact of a different, more gender-egalitarian norm on labor-supply expectations. For this purpose, respondents were randomly assigned to one of three experimental groups—a control and two treatment groups. Randomization in the second experiment was independent from treatment conditions in the first experiment. As pre-specified, we aimed to randomly allocated respondents to the control group with 50 percent probability and to each treatment group with 25 percent probability.[22]

The first treatment *egalitarian information* provides the following information: *"Out of 100 adults in Germany, 89 hold the opinion that men should take as much responsibility for the home and the children as women."* (European Values Study 2017). As in the main survey experiment, we provided a graphical illustration alongside the verbal statement (see Appendix Figure A4 for screenshots).

[21] Re-contact rate and time lag between main survey and follow-up are at the upper end of other recent studies featuring large-scale survey data from adult samples: Kuziemko et al. (2015), Haaland and Roth (forthcoming), Alesina et al. (2018), and Lergetporer et al. (2020) have re-contact rates of 14 percent, 66 percent, 24 percent, and 64 percent, and time lags of one month, one week, one week, and two weeks respectively.

[22] We chose these probabilities to maximize statistical power in regressions where we pool both treatment groups. Due to a programming error in the follow-up survey, group sizes turned out to be 28, 20 and 52 percent for treatment *egalitarian information*, treatment *salience and egalitarian information*, and the control group, respectively. Reassuringly, sociodemographic characteristics are well balanced across the experimental groups (see Appendix Table A4 for details).

The second treatment group *salience and egalitarian information* received the same information, but had to guess the share of Germans holding that opinion beforehand (similar to treatment *salience and information* in the main experiment).

The control group received no additional questions or information. We re-elicited expectations in the same manner as in the beginning of the follow-up survey and in the main survey.

By focusing on a more gender-egalitarian norm towards sharing household responsibility in the second experiment as compared to the traditional norm concerning parental labor supply used in the first experiment, we take advantage of the coexistence of different gender norms related to labor market behavior. The gender norms that we investigate in both experiments vary along several key dimensions: First, the egalitarian norm focuses on responsibility for tasks performed at the home (that are only indirectly related to our main outcome of interest—labor-supply expectations), while the traditional norm focuses on labor supply. Second, the egalitarian norm targets women and men more generally, while the traditional norm explicitly focuses on parents. Finally, the traditional norm entails specific recommendations for behavior (i.e., reducing labor supply), whereas the egalitarian norm refers to the vaguer concept of equal responsibility.

Debriefing: At the end of the survey, we showed a debriefing screen that (i) informed about the research question of the study (i.e., how gender norms affect labor-supply expectations), (ii) provided the content and data sources of both social-norm information treatments, and (iii) provided correct answers to the additional belief-elicitation questions that we posed during the survey (see section 5.3.1). Furthermore, to counteract that the treatments made gender norms salient, the debriefing also stressed that decisions regarding future work hours depend on many different important factors, and not only on social norms.

3.3 Sample Characteristics

Our analytic sample closely resembles the German population aged 14 to 17 years. Appendix Table A1 compares the characteristics of our sample to the respective population statistics in the German Microcensus.[23] Reassuringly, in most categories our respondents' characteristics match official statistics well. While there are slight differences between both samples with respect to gender, educational track, and maternal employment, our sample covers a broad and diverse spectrum of adolescents in Germany.

Appendix Table A2 presents balancing tests to check whether the randomization successfully balanced respondents' observable characteristics across the experimental groups in the main

[23]Research Data Centres of the Federal Statistical Office and the statistical offices of the Laender, Microcensus, census year 2015.

survey. The first column presents the average characteristics of the control group, and the subsequent columns present characteristics of treatment groups *salience*, *information*, and *salience and information* along with the respective differences to the control group. Reassuringly, only one out of 69 pairwise comparisons is significant at the 5-percent level, and four at the 10-percent level, as would be expected by pure chance. Thus, random assignment worked as intended.

Appendix Table A3 investigates whether participation in the follow-up survey is related to treatment assignment in the first experiment that was implemented in the main survey. Regressing a dummy for follow-up-survey participation on treatment indicators and covariates shows insignificant coefficients on treatments *salience* and *information*, and only a marginally significant coefficient on treatment *salience and information*. Furthermore, males, younger respondents, those living in large cities, and those without a degree are more likely to participate in the follow-up survey. Importantly, among those who participated in the follow-up survey, covariates are well-balanced across treatments of the first experiment (see Appendix Table A4), implying that treatment-effect estimates of the first experiment on outcomes measured in the follow-up survey are unbiased. Finally, Appendix Table A5 confirms that the randomization in the second experiment implemented in the follow-up survey was also successful in balancing respondents' characteristics across experimental groups.

3.4 Empirical Strategy

We estimate the causal effects of the social-norm treatments using the following regression model:

$$y_i = \alpha_0 + \sum_j \alpha_{1j} T_i^j + \delta' X_i + \varepsilon_i, \; with \; j \in \{S, I, S\&I\} \tag{1}$$

where y_i is the outcome variable of interest, and T_i^S, T_i^I, $T_i^{S\&I}$ are treatment indicators equal to 1 if respondent i received treatment *salience*, *information*, or *salience and information*, and 0 otherwise. X_i is a vector of control variables, and ε_i is the error term. Average treatment effects α_{1S}, α_{1I}, and $\alpha_{1S\&I}$ are identified because of random assignment. In some selected analyses, we pool treatments to facilitate exposition.

To analyze whether treatment effects are heterogeneous across gender, we extend our basic regression model to:

$$y_i = \beta_0 + \sum_j \beta_{1j} T_i^j + \sum_j \beta_{2j} T_i^j * female_i + \beta_3 female_i + \delta' X_i + \varepsilon_i, \; with \; j \in \{S, I, S\&I\} \tag{2}$$

The treatment effect for boys is given by β_1, and β_2 gives the additional effect for girls.

Since we elicited labor-supply expectations from each respondent twice in the follow-up survey (once at the very beginning to assess treatment-effect persistence of the first experiment, and again after the second experiment), we can estimate treatment effects on stacked data and include individual fixed effects to increase statistical precision. In the corresponding analysis, we therefore estimate the following regression model:

$$y_{it} = \alpha_0 + \sum_j \alpha_{1j} T_{it}^j + \mu_i + \varepsilon_{it}, \; with j \in \{EI, S\&EI\} \tag{3}$$

where y_{it} is the outcome variable of interest of respondent i at time t (before or after eventual treatment administration), and T_{it}^{EI}, $T_{it}^{S\&EI}$ are indicators for treatment *egalitarian information* and *salience and egalitarian information*, respectively. μ_i are individual fixed effects and ε_{it} is the error term. Standard errors are clustered at the individual level.

Finally, to deal with the potential issue of multiple-hypothesis testing, our regression tables further present adjusted p-values following the procedure by List et al. (2019). We adjust for multiple treatments, multiple subgroups (girls and boys) and—where applicable—for multiple outcomes.

4 Descriptive Results: Labor-Supply Expectations

We start by describing labor-supply expectations in the control group and discuss how they relate to respondents' characteristics.

Figure 1 depicts self-expected labor supply of girls and boys in the control group. The gender difference in expected work hours is large: While the majority of girls (59 percent) expect to work 20 hours or less (panel A), the majority of boys (80 percent) expect to work 30 hours or more (panel B). This difference corresponds to a gender gap of 7.3 work hours per week (23.8 versus 31.1 hours). Furthermore, expectations are consistent with gender gaps in actual labor supply: In the German Microcensus (2015), 17 percent of mothers with children aged between 1 and 6 years work full-time, while 46 percent work part-time, and 38 percent do not work at all. On the contrary, most fathers (87 percent) work full-time, and only 7 percent do not work at all.

The fact that expected gender gaps in labor supply are large is also reflected in our second outcome variable of interest: The average expected within-family gender gap in labor supply is 9.1 work hours per week.

Figure 2 shows how the two measures of labor-supply expectations vary across different sociodemographic subgroups, and Table A6 presents the corresponding bivariate regressions. Respondents' gender matters beyond self-expected labor supply, since girls expect a significantly

smaller within-family gender gap than males (see panel B of Figure 1). East German respondents expect to work longer hours than West German respondents (panel A), and they expect a smaller within-family gender gap (panel B). These findings reflect the well-documented fact that labor-force participation of women and mothers is traditionally higher in East Germany than in West Germany (e.g., Boelmann et al., 2020). Finally, respondents whose mothers currently work full-time expect smaller within-family gaps, which is in line with the literature on intergenerational transmission of gender norms (e.g., Fernández et al., 2004).

5 Effects of the Norm Prescribing Parental Labor Supply

In this section, we first describe baseline beliefs about the existing traditional gender norm concerning parental labor supply in Germany. Next, we present the effects of the norm treatments on labor-supply expectations. Finally, we show treatment effects on additional outcomes that are not directly targeted by the norm.

5.1 Beliefs about the Norm

Figure 3 depicts respondents' prior beliefs about the share of German adults who hold the opinion that mothers and fathers, respectively, should reduce their labor supply while their children are young—elicited in the treatments *salience* and *salience and information*. On average, respondents believe that 66 percent of Germans think that mothers should reduce their labor supply, and the median belief is 70 percent (true value: 91 percent). Thus, most adolescents underestimate the actual share of Germans holding this opinion (see panel A). The same pattern applies to the social norm concerning fathers (see panel B): The mean (median) belief is that 35 (31) percent of Germans think that fathers should reduce their labor supply (true value: 41 percent). Interestingly, prior beliefs do not differ systematically by respondents' gender: Girls' mean (median) belief about the norm towards mothers is 66 percent (70 percent), and it is 35 percent (35 percent) towards fathers. The respective figures for boys are 65 percent (70 percent) and 35 percent (30 percent).

In a nutshell, respondents underestimate the difference between the social norms with respect to mothers and fathers. The stark misperceptions of the prevailing social norm regarding labor supply of mothers and fathers indicate potential leeway for correcting these beliefs through information provision in treatments *information and salience* and *information*. In the next section, we study norm-treatment effects on expected labor supply.

5.2 Treatment Effects on Labor-Supply Expectations

Since the gender norm regarding parental labor supply prescribes different labor-market behavior for mothers and fathers, we present treatment-effect estimates separately for girls and for boys.

5.2.1 Girls' Labor-Supply Expectations

The left part of Figure 4 displays self-expected labor supply in the control group and in the pooled treatment groups for girls. The treatments significantly reduce self-expected labor supply: The share of girls expecting to work 20 hours or less per week increases significantly from 59 percent to 67 percent (see panel A), and, conversely, the share expecting to work 30 hours or more significantly decreases from 41 percent to 33 percent (see panel B).

Turning to regression results, panel A of Table 1 shows that the pooled treatment effects correspond to a significant reduction of 2.6 hours in expected weekly work hours of girls (see column 1).[24] In Table 2 we present the effects of the three treatments separately. Each treatment has a highly significant and negative impact on girls' self-expected labor supply. The effect of the combined treatment *salience and information* is significantly ($p<0.1$) larger than the effect of treatment *salience*, and the effect size of treatment *information* is in-between the two. This pattern leads to three important insights: First, the strong effects of treatment *salience* suggests that priming adolescents to think about the existing gender norm already alters their labor-supply expectations. Second, the significant difference between treatment *salience* and the combined treatment *salience and information* suggests that providing information about the norm's content has an additional effect beyond the treatment *salience*. Third, the effect of treatment *salience and information* is smaller than the sum of the effects of treatment *salience* and treatment *information*, suggesting that information provision partly affects adolescents' expectations by making the norm salient. We discuss the role of the norms' salience in more detail in section 5.2.4.

Column 2 of Table 1 presents the pooled treatment effects on the expected within-family gender gap in labor supply, i.e., respondents' expected work hours of the male family member minus that of the female family member. The social-norm treatments significantly increase girls' expected within-family gender gap from 7.7 weekly work hours by 2.2 hours. Each treatment has a separate highly significant and positive impact on the expected within-family gender gap (see Table 2) which can be explained by the fact the treatments primarily decrease

[24] Online Appendix Table O1 presents treatment effects separately on each of the five answer categories.

self-expectations, but not their expectations for the partner (see Online Appendix Table O1 for separate treatment effects on labor-supply expectations for the partner).[25]

To deal with the potential issue of multiple-hypothesis testing, the main tables additionally display adjusted p-values following the methodology of List et al. (2019). We find that levels of significance do not change substantially when adjusting for multiple subgroups (Table 1) or multiple subgroups as well as multiple treatments (Table 2).

In sum, girls' labor-supply expectations react strongly to treatments that highlight the traditional social norm on of how much mothers of young children should work on the labor market. Girls' expectations become more gender-unequal in the sense that they expect to work fewer hours themselves and expect a larger within-family gender gap in labor-market hours. Thus, our results indicate that gender norms play an important role in explaining gender gaps in labor-market outcomes after child birth.

5.2.2 Boys' Labor-Supply Expectations

Next, we investigate how the gender norm affects expected labor supply of boys. The right part of Figure 4 reveals that the pooled social norm treatments also reduce self-expected labor supply for boys: The share of boys expecting to work at most 20 hours per week significantly increases from 20 percent to 28 percent in response to the treatments (see panel A). At the same time, the share of boys expecting to work 30 hours or more significantly decreases from 80 percent to 72 percent (see panel B), which is entirely driven by a decrease in the share of boys expecting to work 40 hours.

Panel B of Table 1 show that these treatment effects correspond to a significant reduction of 1.8 expected weekly work hours (column 1).[26] Table 2 depicts treatment effects separately for each of the three norm treatments. In line with the results for girls, each of the three treatments has a negative impact on boys' self-expected labor supply (see column 1 of Table 2). The effect of the combined treatment *salience and information* is the largest one and highly significant ($p<0.01$), the coefficient on treatment-indicator *information* is marginally significant ($p<0.1$), and the coefficient on treatment *salience* does not reach statistical significance.

Columns 2 of Table 1 and Table 2 present treatment effects on the expected within-family gender gap in labor supply. In contrast to our findings for girls, boys' expectations for their own family tend to become more gender-equal when being confronted with the existing social norm. The treatments decrease (albeit not significantly so) boys' expected within-family gender gap from 11 weekly work hours by 1.4 hours, which is again due to the fact the treatments

[25] In line with the fact that we hardly find treatment effects on partners' labor-supply expectations, we do not find any treatment effects on girls' preferences for a set of partner attributes, either (e.g., whether the partner helps with the household or raising children etc.) (results available upon request).

[26] Online Appendix Table O2 presents treatment effects separately for each of the five answer categories.

decrease self-expectations, but do not affect expectations for the partner (see Online Appendix Table O2).[27]

Overall, we find that the gender norm prescribing mothers' and fathers' labor supply strongly impact adolescents' expectations of their own labor supply. While the norm makes girls' expectations more gender-unequal, boys' expectations become more gender-equal. This latter result is particularly noteworthy given that the question how gender norms affect boys' labor supply has not yet been studied in the literature.

5.2.3 Persistence of Treatment

Next, we check whether treatment effects persist beyond the immediate survey horizon into the follow-up survey conducted about two weeks later. Given that priming effects, such as the ones induced by treatment *salience*, are by definition temporary and short-term (e.g., Forehand et al., 2002; Benjamin et al., 2010), we only expect persistent effects from treatments that entail information provision. Therefore, our persistence-analysis compares the pooled experimental groups that received and did not receive information on the norm's content (i.e., treatments *information and salience* and *information* versus treatment *salience* and the control group). As the previous section shows that the treatments do not affect respondents' labor supply expectations for their partner, we restrict our analyses of persistence on self-expected labor supply.

Table 3 combines data from the main survey and the follow-up survey and regresses self-expected labor supply on an information-treatment dummy, a follow-up-survey dummy, and the interaction of both indicators. For the overall sample, column 1 shows that information-treatment effects persist in the follow-up survey. As expected, the treatment effect in the follow-up survey tends to be somewhat smaller than the one in the main survey (likely due to imperfect recall), although the difference between treatment effects is not statistically significant (see coefficient on the interaction term). Columns 2 and 3 report persistent treatment effects separately for girls and boys. While treatment effects in the follow-up survey for these subsamples are remarkably similar in magnitude to the full sample, they do not reach statistical significance due to limited statistical power.[28]

[27] In addition, Appendix Table A7 depicts treatment-effect heterogeneities by gender on both labor-supply expectations. Column 1 confirms that the treatments equally affect self-expected labor supply among girls and boys. Column 2 shows treatment effect heterogeneities on the expected within-family gender gap, where we observe differences by gender for all treatments. While girls become more gender-unequal in their expectations for their own family, boys seem to become more gender-equal by expecting a smaller within-family gender gap. In online appendix table O3 we also present effect heterogeneities with respect to various other sociodemographic characteristics.

[28] In further analyses, we exploit variation in the time lag between main and follow-up survey, and find that treatment-effect persistence does not significantly differ for respondents who participated earlier vs. later in the follow-up (results available upon request). Appendix Table A8 analyses persistence for all three treatments

Turning to belief-updating, Appendix Table A9 investigates respondents' stated beliefs relative to accurate values. It shows that information provision persistently improves beliefs about the content of the norm, i.e., the share of Germans who hold the opinion that mothers and fathers, respectively, should reduce their labor supply. Interestingly, respondents seem to internalize in particular the existing norm for their own gender, which is consistent with the fact that the norm treatments mainly affect self-expected labor supply.[29]

In sum, the information treatments in the main survey lead to a persistent updating of self-expected labor supply and beliefs in the follow-up survey, which implies that (i) adolescents indeed understand and remember the provided information, and (ii) treatment effects are unlikely due to experimenter-demand effects (e.g., Cavallo et al., 2017; Bleemer and Zafar, 2018; Haaland et al., 2020).[30]

5.2.4 Heterogeneities by Prior Beliefs

The literature discusses two potential channels through which information provision may affect individuals' responses: salience-based versus information-based updating (Bleemer and Zafar, 2018). In our setting, information treatments may affect labor-supply expectations because information provision increases the salience of the gender norms (e.g., Schwarz and Vaughn, 2002; Chetty et al., 2009; DellaVigna, 2009; Bleemer and Zafar, 2018), or because it corrects respondents' misperceptions about the content of the norm (e.g., Rockoff et al., 2012; Bursztyn et al., 2020).[31] To investigate the relevance of these channels, we estimate heterogeneous information-treatment effects by respondents' prior beliefs (elicited in treatments *salience* and *salience and information*) using the following regression model:

$$y_i = \gamma_0 + \gamma_1 T_i^{S\&I} + \gamma_2 T_i^{S\&I} * Misperception_i + \gamma_3 Misperception_i + \delta' X_i + \varepsilon_i \qquad (4)$$

where $Misperception_i$ is the difference between the factual share of Germans thinking that respondent i's gender (i.e., mothers or fathers) should reduce his or her labor supply minus

separately. While statistical power is again limited, results suggest the effect of treatment *salience* does not persist, while the effects of treatments involving information provision have the expected sign and are partly significant.

[29] Appendix Figure A5 presents the entire distribution of beliefs.

[30] Quidt et al. (2018) and Mummolo and Peterson (2019) show that (survey) experiments are largely robust to experimenter-demand effects.

[31] The idea behind salience-based updating is that information about one specific norm increases the salience of that specific norm relative to other aspects that may affect labor-supply expectations. Labor-supply expectations could generally be influenced by a multitude of social-norm considerations, because (i) social identity is multidimensional (for instance, it can refer to gender, race, or social status (Akerlof and Kranton, 2000) and (ii) several social norms may coexist within each domain of social identity. In the context of gender norms, such norms may prescribe labor supply, relative income within households (Bertrand et al., 2015), or shared household responsibility, for instance.

respondent i's belief about this share. The coefficient γ_1 captures the average information effect for respondents with correct prior beliefs, and γ_2 captures the additional effect for respondents who initially misperceive the social norm. Information-based updating would imply that γ_2 is significant and negative, whereas salience-based updating would imply that γ_2 is close to zero.

Table 4 shows that the coefficients on the treatment *salience and information* are negative and, in the full sample, significantly different from zero. In addition, coefficients on the interaction term are negative as to be expected for information-based updating (but shy of statistical significance).[32]

Our results on the treatment effects for respondents with accurate priors suggest that providing information about the social norm's content at least partly affects labor-supply expectations by increasing the norm's salience. This interpretation is in line with the previous literature: Bleemer and Zafar (2018) find that the effects of college-returns information on intended college attendance do not vary by respondents' prior beliefs about college returns. Similarly, Alesina et al. (2018) argue that salience effects drive their negative findings of correcting natives' over-pessimistic beliefs about immigrants on natives' preferences for redistribution.[33]

5.3 Treatment Effects on Additional Outcomes

5.3.1 Incentivized Outcomes

A common critique against the expectations-literature is that the main outcomes of interest—survey-based expectations about future events or actions—have no immediate consequences for respondents, which raises concerns about the outcome variables' relevance.

To test whether treatment effects carry over to outcomes with direct monetary consequences, we next present results from a set of incentivized belief-questions that asked respondents to guess the shares of Germans who agree with the following gender-related statements (European Values Study 2017): (i) "A university education is more important for a boy than for a girl." (EVS: 16 percent); (ii) "When the mother works for pay, the children suffer." (EVS: 33 percent); (iii) "Being a housewife is just as fulfilling as working for pay." (EVS: 55 percent); and (iv) "If a woman earns more than her husband, it's almost certain to cause problems." (EVS: 20 percent). We introduce immediate monetary consequences for correct answers by paying each respondent two additional Euros for a roughly correct answer (defined as belonging to the better half of guesses) to the randomly chosen question. This experimental design allows us to test whether treatment effects carry over to outcomes with immediate monetary consequences.

[32] In line with the interpretation of information-based-updating, we also find significant treatment effects of the treatments that provide accurate information on beliefs elicited in the follow-up survey (see section 5.2.3).

[33] In Online Appendix O1, we study perceived peer pressure as a further potential mechanism driving our treatment effects, and find that the channel seems to be relevant for girls but not for boys.

Figure 5 provides a graphical depiction of the belief distributions for each item across experimental groups with and without information provision, and Table 5 reports the corresponding regression results.[34] Information provision significantly affects respondents' incentivized beliefs about the share of Germans agreeing to the different items. This finding, along with the fact that previous research has shown that unincentivized expectations are tightly linked to real outcomes, gives rise to our interpretation that treatment effects reflect effects of the gender norm on (labor-market) relevant outcomes.[35]

5.3.2 Labor-Supply Expectations Without Child

The traditional gender norm that we study prescribes labor supply of mothers and fathers. Consequently, the norm treatments should affect labor-supply expectations with child, but not necessarily without child. To perform this additional sanity check, we also elicited respondents' expected labor supply at the age of 30 in a hypothetical scenario without child.

Appendix Table A11 shows pooled treatment effects on labor-supply expectations without children. Treatment effects on both self-expected labor supply and the expected within-family gender gap are small and insignificant for both genders. The fact treatment effects are confined to those outcomes that are directly prescribed by the norm further raises confidence that our experimental results reflect genuine effects of the specific norm, as opposed to some unintended effects such as experimenter-demand effects.[36]

6 Effects of the Norm on Shared Household Responsibility

So far, we have shown that the prevailing traditional social norm prescribing mothers' and fathers' labor supply decreases labor-supply expectations, and thereby potentially promoting gender gaps in labor-market outcomes. We now investigate whether social norms can shift labor-market relevant outcomes in both directions. We therefore conduct a second experiment in

[34] We expect only the information content of treatments *information* as well as *information and salience* to spill over to the gender-related items as the incentivized outcome questions per se already induces all respondents to think about societal expectations and hence increase salience of the respective issues (similar to the questions posed in treatment *salience*). We therefore pool the two treatments *information and salience* and *information* and compare them to treatments *salience* and the control group. Appendix Table A10 reports effects of each treatment separately, and confirms that only those treatments that entail information provision affect the incentivized outcomes.

[35] Interestingly, respondents in the treatment group report more conservative beliefs, which undermines the accuracy of beliefs in all items but item (iii). In view of this result, it is particularly important to note that we provided accurate information about the different items in the debriefing stage at the very end of the survey.

[36] In Online Appendix O1, we study preferences for job attributes as additional indirect outcome variables, and find little evidence that these preferences are affected by the norm-treatments.

the follow-up survey to study whether a more egalitarian norm could lead to less gender-unequal outcomes. The corresponding treatments highlight a more gender-egalitarian norm towards sharing household responsibilities. In this section, we first describe baseline beliefs about the norm, and then present treatment effects on labor-supply expectations.

6.1 Beliefs about the Norm

Figure 6 depicts prior beliefs about the egalitarian gender norm elicited in treatment *salience and egalitarian information*. It shows that respondents underestimate the egalitarianism of the norm: The mean (median) guess is that 59 percent (60 percent) of Germans think that men should take as much responsibility for the household as women, whereas the true share in the German population is 89 percent. While both genders misperceive this norm, girls' beliefs tend to be more accurate than boys' beliefs (60 percent versus 55 percent median guess).

6.2 Treatment effects on Labor-Supply Expectations

Table 6 depicts pooled effects of treatments *egalitarian information and salience* and *egalitarian information* on labor-supply expectations. Focusing on the expected within-family gender gap in labor supply as the outcome of interest,[37] we indeed find that the treatments attenuates the expected gender gap in labor supply. In particular, the pooled treatments significantly decrease the expected gap by 1.3 hours per week (column 1). Girls expect a reduction of the gender gap by 1.3 hours (column 2), and boys by 1.4 hours (column 3). Appendix Table A12 shows that both treatments have statistically significant negative effects on the expected within-family gender gap. If anything, treatment effects tend to be stronger in the combined treatment *salience and egalitarian information* than in treatment egalitarian information, which resembles the patterns in the first experiment (see section 5).

In sum, these results show that the more egalitarian gender norm towards sharing household responsibility can lead to more gender-equal expectations regarding the within-family gender gap in labor supply after child birth. In the treatment groups both genders expect the mother to reduce her labor supply less relative to the father. Thus, the effects of gender norms on labor-market expectations can depend on the specific context—and the degree of gender-equality—of the respective norm.

[37] We focus on the expected within-family gap outcome of interest because this social norm explicitly addresses the household as a whole. Further analyses indeed reveal that treatment effects on the within-family gender gap are driven by changes in self-expectations as well as partners' expectations (results available upon request).

7 Conclusion

In many developed countries, gender differences in labor-market outcomes do not emerge until the arrival of the first child. We shed light on the causal relationship between labor-market outcomes and gender norms in large-scale experiments among 2,000 adolescents in Germany, a country with comparatively large child penalties in addition to a very traditional norm on how much mothers should work on the labor market. At baseline, most girls (59 percent) expect to work no more than 20 hours per week with a young child, and most boys (80 percent) expect to work at least 30 hours per week. Administering treatments that highlight the existence of a traditional gender norm in Germany—i.e., that 91 percent (41 percent) of Germans think that mothers (fathers) of young children should reduce labor supply—significantly reduces girls' labor-supply expectations by 2.6 hours per week, which increases the expected within-family gender gap in labor supply. While largely neglected by the literature so far, we also study how the gender norm affects boys' labor-supply expectations. Boys also expect to reduce their labor supply in response to the norm treatments, which translates into a reduced expected within-family gender gap. Finally, we show that an alternative treatment highlighting a more gender-egalitarian norm towards sharing household responsibility results in more gender-equal labor-supply expectations among both genders.

Our results indicate that social norms play an important role in shaping gender gaps in labor-market outcomes with young children.

While we discuss that effects of gender norms may indeed depend on the degree of egalitarianism in the underlying social prescription, the question whether and how social norms towards parents may change is beyond the scope of this study. Previous research has argued that the introduction of the birth control pill (Goldin and Katz, 2002), technological change (Alesina et al., 2013) or the implementation of parental leave reforms (Dahl et al., 2014) have changed societal views on gender issues over time. Accordingly, if social norms towards mothers were to become more egalitarian, child penalties for mothers compared to fathers may attenuate given our findings.

However, we view our results as "proof of concept" that norms regarding parental labor supply causally affect labor-market relevant outcomes in a country with gender-equal laws that resemble those in most other developed countries. We believe that expanding the scale and looking at how norms affect a broader set of outcomes, such as labor-market institutions, child care supply or employer behavior is an interesting avenue for future research.

References

Akerlof, George A. and Rachel E. Kranton (2000). "Economics and Identity". *The Quarterly Journal of Economics* 115(3), 715–753.

– (2010). *Identity Economics: How Our Identities Shape Our Work, Wages, and Well-Being*. Princeton University Press.

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

Alesina, Alberto, Armando Miano, and Stefanie Stantcheva (2018). "Immigration and Redistribution". *National Bureau of Economic Research Working Paper Series* No. 24733.

Aloud, Monira Essa, Sara Al-Rashood, Ina Ganguli, and Basit Zafar (2020). "Information and Social Norms: Experimental Evidence on the Labor Market Aspirations of Saudi Women". *National Bureau of Economic Research Working Paper Series* No. 26693.

Alt, Christian, Daniela Gesell, Sandra Hubert, Katrin Hüsken, Ralf Kuhnke, and Kerstin Lippert (2017). "DJI-Kinderbetreuungsreport 2017. Inanspruchnahme und Bedarfe aus Elternperspektive im Bundesländervergleich." *Deutsches Jugendinstitut (DJI), München*.

Altonji, Joseph G. and Rebecca M. Blank (n.d.). "Chapter 48 Race and Gender in the Labor Market". *Ashenfelter, Card (Hg.) 1999 – Handbook of Labor Economics*. Vol. 3, 3143–3259.

Andresen, Martin Eckhoff and Emily Nix (2019). "What Causes the Child Penalty and How Can It Be Reduced? Evidence from Same-Sex Couples and Policy Reforms". *Statistics Norway Discussion Paper* No. 902.

Angelov, Nikolay, Per Johansson, and Erica Lindahl (2015). "Parenthood and the Gender Gap in Pay". *Journal of Labor Economics* 34(3), 545–579.

Arcidiacono, Peter, V. Joseph Hotz, and Songman Kang (2012). "Modeling College Major Choices Using Elicited Measures of Expectations and Counterfactuals". *Journal of Econometrics* 166(1), 3–16.

Benabou, Roland and Jean Tirole (2011). "Laws and Norms". *National Bureau of Economic Research Working Paper Series* No. 17579.

Benjamin, Daniel J., James J. Choi, and A. Joshua Strickland (2010). "Social Identity and Preferences". *American Economic Review* 100(4), 1913–1928.

Bertrand, Marianne, Claudia Goldin, and Lawrence F. Katz (2010). "Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors". *American Economic Journal: Applied Economics* 2(3), 228–255.

Bertrand, Marianne, Emir Kamenica, and Jessica Pan (2015). "Gender Identity and Relative Income within Households". *The Quarterly Journal of Economics* 130(2), 571–614.

Blau, Francine D. and Lawrence M. Kahn (2017). "The Gender Wage Gap: Extent, Trends, and Explanations". *Journal of Economic Literature* 55(3), 789–865.

Bleemer, Zachary and Basit Zafar (2018). "Intended College Attendance: Evidence from an Experiment on College Returns and Costs". *Journal of Public Economics* 157, 184–211.

Boelmann, Barbara, Anna Raute, and Uta Schoenberg (2020). "Wind of Change? Cultural Determinants of Maternal Labor Supply". *Working Paper*.

Boschini, Anne, Astri Muren, and Mats Persson (2012). "Constructing Gender Differences in the Economics Lab". *Journal of Economic Behavior & Organization* 84(3), 741–752.

Bursztyn, Leonardo, Alessandra L. González, and David Yanagizawa-Drott (2020). "Misperceived Social Norms: Women Working Outside the Home in Saudi Arabia". *American Economic Review* 110(10), 2997–3029.

Cavallo, Alberto, Guillermo Cruces, and Ricardo Perez-Truglia (2017). "Inflation Expectations, Learning, and Supermarket Prices: Evidence from Survey Experiments". *American Economic Journal: Macroeconomics* 9(3), 1–35.

Charles, Kerwin Kofi, Jonathan Guryan, and Jessica Pan (2018). "The Effects of Sexism on American Women: The Role of Norms vs. Discrimination". *National Bureau of Economic Research Working Paper Series* No. 24904.

Chetty, Raj, Adam Looney, and Kory Kroft (2009). "Salience and Taxation: Theory and Evidence". *American Economic Review* 99(4), 1145–1177.

Cialdini, Robert B. and Noah J. Goldstein (2004). "Social Influence: Compliance and Conformity". *Annual Review of Psychology* 55, 591–621.

Cialdini, Robert B. and Melanie R. Trost (1998). "Social Influence: Social Norms, Conformity and Compliance". *The Handbook of Social Psychology*. New York, NY, US: McGraw-Hill, 151–192.

Cortes, Patricia and Jessica Pan (2020). "Chidren and the Remaining Gender Gaps in the Labor Market".

Cunha, Flávio, Irma Elo, and Jennifer Culhane (2013). "Eliciting Maternal Expectations about the Technology of Cognitive Skill Formation". *National Bureau of Economic Research Working Paper Series* No. 19144.

Dahl, Gordon B., Katrine V. Løken, and Magne Mogstad (2014). "Peer Effects in Program Participation". *American Economic Review* 104(7), 2049–2074.

Delavande, Adeline (2014). "Probabilistic Expectations in Developing Countries". *Annual Review of Economics* 6(1), 1–20.

Delavande, Adeline and Basit Zafar (2019). "University Choice: The Role of Expected Earnings, Nonpecuniary Outcomes, and Financial Constraints". *Journal of Political Economy* 127(5), 2343–2393.

DellaVigna, Stefano (2009). "Psychology and Economics: Evidence from the Field". *Journal of Economic Literature* 47(2), 315–372.

Fernández, Raquel (2007). "Women, Work, and Culture". *Journal of the European Economic Association* 5(2-3), 305–332.

– (2013). "Cultural Change as Learning: The Evolution of Female Labor Force Participation over a Century". *American Economic Review* 103(1), 472–500.

Fernández, Raquel and Alessandra Fogli (2009). "Culture: An Empirical Investigation of Beliefs, Work, and Fertility". *American Economic Journal: Macroeconomics* 1(1), 146–177.

Fernández, Raquel, Alessandra Fogli, and Claudia Olivetti (2004). "Mothers and Sons: Preference Formation and Female Labor Force Dynamics". *The Quarterly Journal of Economics* 119(4), 1249–1299.

Forehand, Mark R., Rohit Deshpandé, and Americus Reed (2002). "Identity Salience and the Influence of Differential Activation of the Social Self-Schema on Advertising Response". *The Journal of applied psychology* 87(6), 1086–1099.

Fortin, Nicole M. (2005). "Gender Role Attitudes and the Labour-market Outcomes of Women across OECD Countries". *Oxford Review of Economic Policy* 21(3), 416–438.

– (2015). "Gender Role Attitudes and Women's Labor Market Participation: Opting-Out, AIDS, and the Persistent Appeal of Housewifery". *Annals of Economics and Statistics* (117/118), 379–401.

Geis-Thoene, Wido (2018). "Familien müssen für die gleiche Betreuung in der Kita unterschiedlich viel zahlen - Ein Vergleich der Gebührenordnungen der größten Städte in Deutschland". *IW-Report* (50).

Giavazzi, Francesco, Fabio Schiantarelli, and Michel Serafinelli (2013). "Attitudes, Policies, and Work". *Journal of the European Economic Association* 11(6), 1256–1289.

Giuliano, Paola (2018). "Gender". *The Oxford Handbook of Women and the Economy*. Ed. by Susan Averett, Laura M. Argys, and Saul D. Hoffman. New York, NY: Oxford University Press, 644–672.

Goerges, Luise and Daniele Nosenzo (2020). "Social Norms and the Labor Market". *Handbook of Labor, Human Resources and Population Economics*. Ed. by Klaus F. Zimmermann. Cham: Springer International Publishing, 1–26.

Goldin, Claudia (2014). "A Grand Gender Convergence: Its Last Chapter". *American Economic Review* 104(4), 1091–1119.

Goldin, Claudia and Lawrence F. Katz (2002). "The Power of the Pill: Oral Contraceptives and Women's Career and Marriage Decisions". *Journal of Political Economy* 110(4), 730–770.

Goldin, Claudia, Lawrence F. Katz, and Ilyana Kuziemko (2006). "The Homecoming of American College Women: The Reversal of the College Gender Gap". *Journal of Economic Perspectives* 20(4), 133–156.

Grewenig, Elisabeth, Philipp Lergetporer, Katharina Werner, Ludger Woessmann, and Lisa Simon (2018). "Can Online Surveys Represent the Entire Population?" *Cesifo Working Papers* 7222.

Haaland, Ingar and Christopher Roth (forthcoming). "Labor Market Concerns and Support for Immigration". *Journal of Public Economics*.

Haaland, Ingar, Christopher Roth, and Johannes Wohlfart (2020). "Designing Information Provision Experiments". *Cesifo Working Papers* 8406.

Hill, John P. and Mary Ellen Lynch (1983). "The Intensification of Gender-Related Role Expectations during Early Adolescence". *Girls at Puberty: Biological and Psychosocial Perspectives*. Ed. by Jeanne Brooks-Gunn and Anne C. Petersen. Boston, MA: Springer US, 201–228.

Jayachandran, Seema (2019). "Social Norms as a Barrier to Women's Employment in Developing Countries". *Northwestern Working Paper*.

Jessen, Jonas, Sophia Schmitz, and Sevrin Waights (2020). "Understanding Day Care Enrolment Gaps". *Journal of Public Economics* 190, 104252.

Kleven, Henrik and Camille Landais (2017). "Gender Inequality and Economic Development: Fertility, Education and Norms". *Economica* 84(334), 180–209.

Kleven, Henrik, Camille Landais, Johanna Posch, Andreas Steinhauer, and Josef Zweimüller (2019a). "Child Penalties across Countries: Evidence and Explanations". *AEA Papers and Proceedings* 109.

Kleven, Henrik, Camille Landais, and Jakob Egholt Søgaard (2019b). "Children and Gender Inequality: Evidence from Denmark". *American Economic Journal: Applied Economics* 11(4), 181–209.

– (2020). "Does Biology Drive Child Penalties? Evidence from Biological and Adoptive Families". *National Bureau of Economic Research Working Paper Series* No. 27130.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz (2007). "Experimental Analysis of Neighborhood Effects". *Econometrica* 75(1), 83–119.

Kunz, Johannes S. and Kevin E. Staub (2020). "Early subjective completion beliefs and the demand for post-secondary education". *Journal of Economic Behavior & Organization* 177, 34–55.

Kuziemko, Ilyana, Michael I. Norton, Emmanuel Saez, and Stefanie Stantcheva (2015). "How Elastic Are Preferences for Redistribution? Evidence from Randomized Survey Experiments". *American Economic Review* 105(4), 1478–1508.

Kuziemko, Ilyana, Jessica Pan, Jenny Shen, and Ebonya Washington (2018). "The Mommy Effect: Do Women Anticipate the Employment Effects of Motherhood?" *National Bureau of Economic Research Working Paper Series* No. 24740.

Lergetporer, Philipp, Katharina Werner, and Ludger Woessmann (2020). "Educational Inequality and Public Policy Preferences: Evidence from Representative Survey Experiments". *Journal of Public Economics* 188, 104226.

List, John A., Azeem M. Shaikh, and Yang Xu (2019). "Multiple Hypothesis Testing in Experimental Economics". *Experimental Economics* 22(4), 773–793.

Maniadis, Zacharias, Fabio Tufano, and John A. List (2014). "One Swallow Doesn't Make a Summer: New Evidence on Anchoring Effects". *American Economic Review* 104(1), 277–290.

Manski, Charles F. (2004). "Measuring Expectations". *Econometrica* 72(5), 1329–1376.

Mummolo, Jonathan and Eric Peterson (2019). "Demand Effects in Survey Experiments: An Empirical Assessment". *American Political Science Review* 113(2), 517–529.

Olivetti, Claudia and Barbara Petrongolo (2016). "The Evolution of Gender Gaps in Industrialized Countries". *Annual Review of Economics* 8(1), 405–434.

Priess, Heather A., Sara M. Lindberg, and Janet Shibley Hyde (2009). "Adolescent Gender-Role Identity and Mental Health: Gender Intensification Revisited". *Child Development* 80(5), 1531–1544.

Quidt, Jonathan de, Johannes Haushofer, and Christopher Roth (2018). "Measuring and Bounding Experimenter Demand". *American Economic Review* 108(11), 3266–3302.

Reuben, Ernesto, Matthew Wiswall, and Basit Zafar (2017). "Preferences and Biases in Educational Choices and Labour Market Expectations: Shrinking the Black Box of Gender". *The Economic Journal* 127(604), 2153–2186.

Rockoff, Jonah E., Douglas O. Staiger, Thomas J. Kane, and Eric S. Taylor (2012). "Information and Employee Evaluation: Evidence from a Randomized Intervention in Public Schools". *American Economic Review* 102(7), 3184–3213.

Schwarz, Norbert and Leigh Ann Vaughn (2002). "The Availability Heuristic Revisited: Ease of Recall and Content of Recall as Distinct Sources of Information". *Heuristics and Biases: The Psychology of Intuitive Judgment*. Ed. by Thomas Gilovich, Dale Griffin, and Tversky Kahnemann. New York: Cambridge University Press, 103–119.

Steinhauer, Andreas (2018). "Working Moms, Childlessness, and Female Identity". *LIEPP Working Paper* No. 79.

Stinebrickner, Ralph and Todd R. Stinebrickner (2013). "A Major in Science? Initial Beliefs and Final Outcomes for College Major and Dropout". *The Review of Economic Studies* 81(1), 426–472.

The Economist (July 20, 2019). "Changing the Guard: Saudi Arabia Weighs Loosening more Controls on Women". *The Economist* 2019.

Wippermann, Carsten (2015). "Transparenz für mehr Entgeltgleichheit: Einflüsse auf den Gender Pay Gap (Berufswahl, Arbeitsmarkt, Partnerschaft, Rollenstereotype) und Perspektiven der Bevölkerung für Lohngerechtigkeit zwischen Frauen und Männern". *Bundesministeriums für Familie, Senioren, Frauen und Jugend*.

Wiswall, Matthew and Basit Zafar (2018). "Preference for the Workplace, Investment in Human Capital, and Gender". *The Quarterly Journal of Economics* 133(1), 457–507.

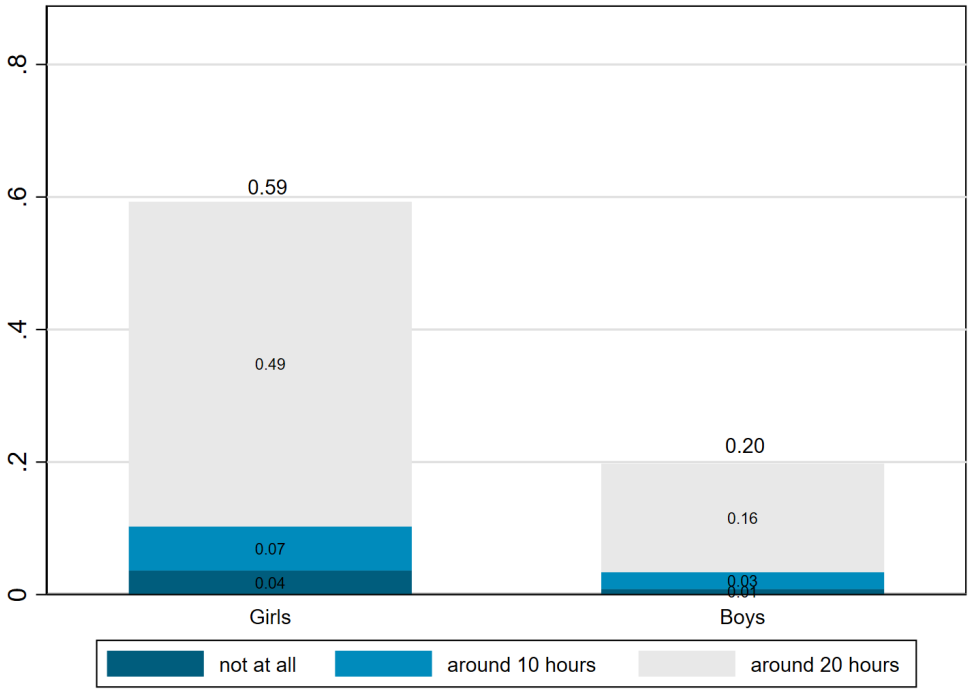
– (2019). "Human Capital Investments and Expectations about Career and Family". *Working Paper*.

Woessmann, Ludger, Philipp Lergetporer, Elisabeth Grewenig, Sarah Kersten, and Katharina Werner (2018). "Denken Jugendliche anders über Bildungspolitik als Erwachsene?" *fo Schnelldienst 71* (17), 31–45.

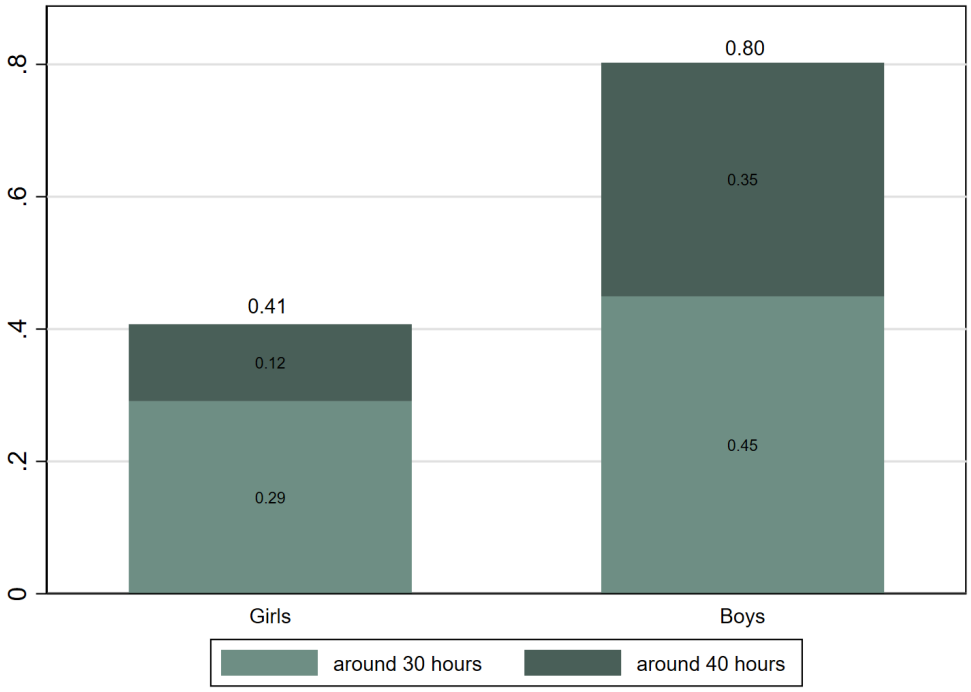
Zafar, Basit (2013). "College Major Choice and the Gender Gap". *Journal of Human Resources* 48(3), 545–595.

Figure 1: Labor-Supply Expectations

Panel A: Expectations to Work at Most 20 Hours per Week



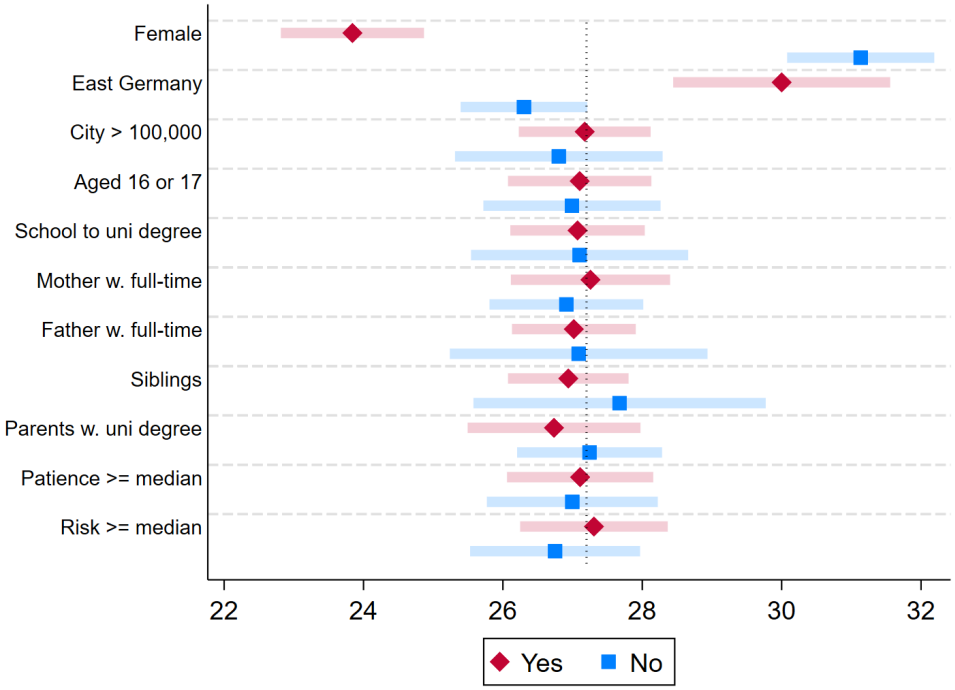
Panel B: Expectations to Work at Least 30 Hours per Week



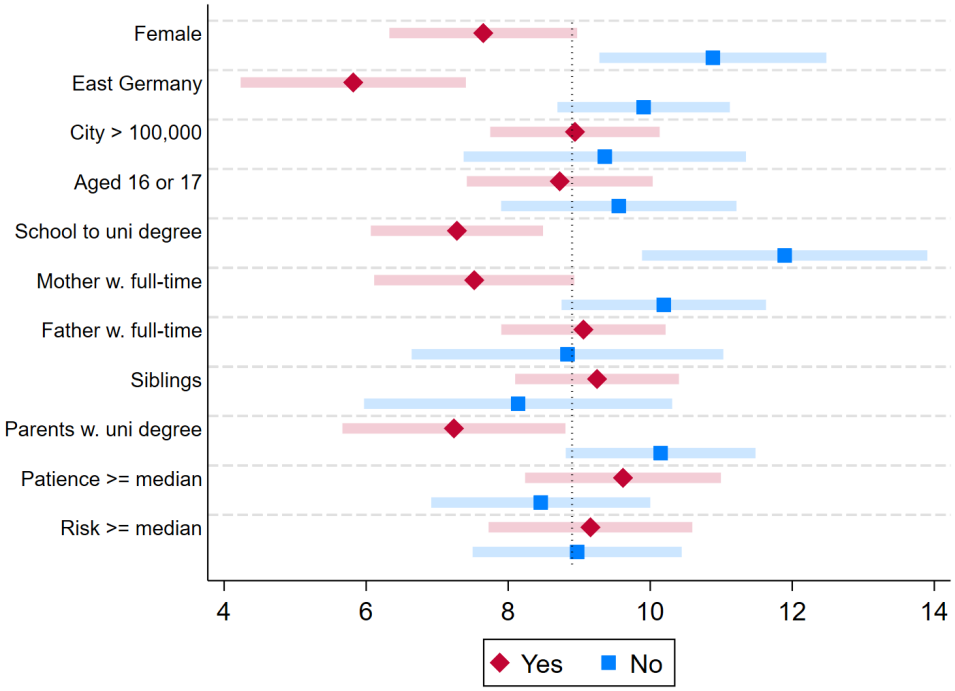
Notes: Responses to the question "Imagine you are 30 years old and you have a child aged between 1 and 6 years with your partner. What do you think, how many hours per week on average would you like to work in order to earn money?". Sample: Respondents in the control group.

Figure 2: Labor-Supply Expectations across Sociodemographic Characteristics

Panel A: Self-Expected Labor Supply



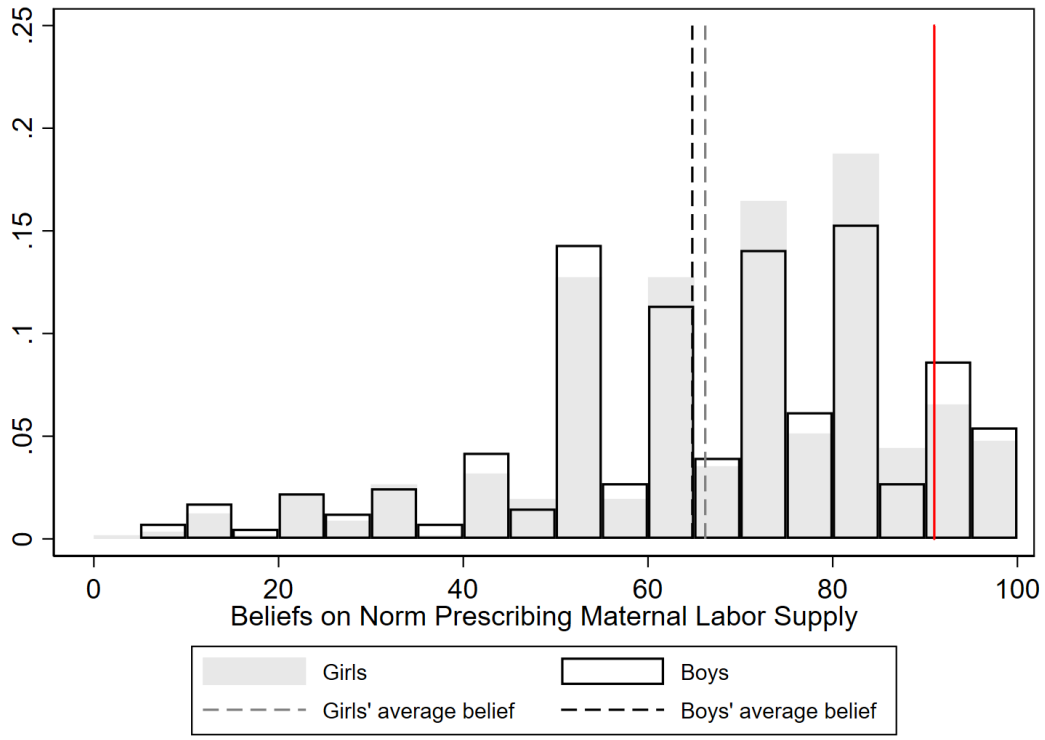
Panel B: Expected Within-Family Gender Gap



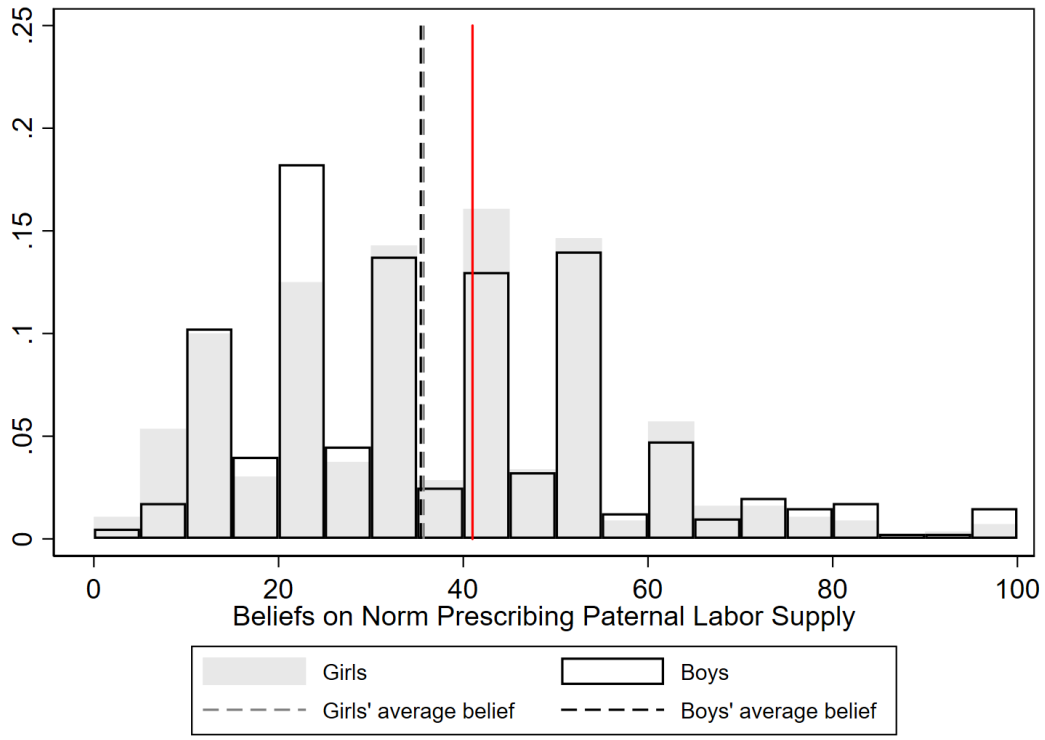
Notes: The figure shows the control group answers to the expected labor supply for different groups of respondents. The shaded areas are 95 percent confidence intervals around the average response. Panel A: Hours expected to work per week with child 1-6 (0=0 hours, i.e. not at all; 10= 10 hours; 20= 20 hours, i.e. part-time; 30=30 hours, 40=40 hours, i.e. full-time). Panel B: Responses to labor supply for both parents (self and partner) with higher values indicating higher labor market supply of men relative to women. Average hours (full sample) indicated by vertical, dotted line.

Figure 3: Distribution of Beliefs about the Norm Prescribing Parental Labor Supply

Panel A: Norm that Mothers Should Reduce Their Labor-Supply when Children Are Young



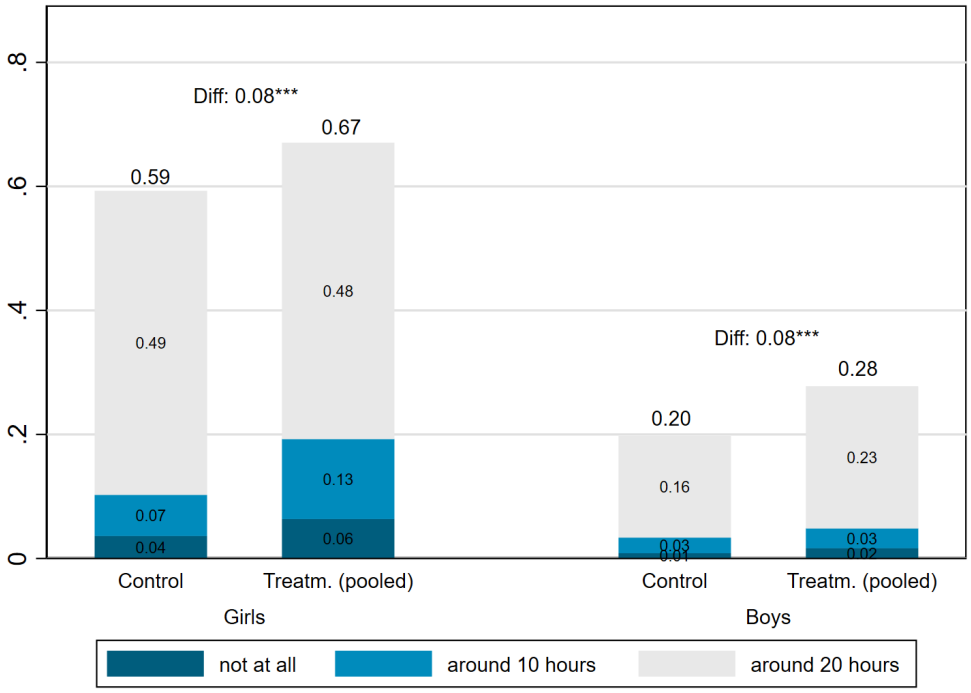
Panel B: Norm that Fathers Should Reduce Their Labor-Supply when Children Are Young



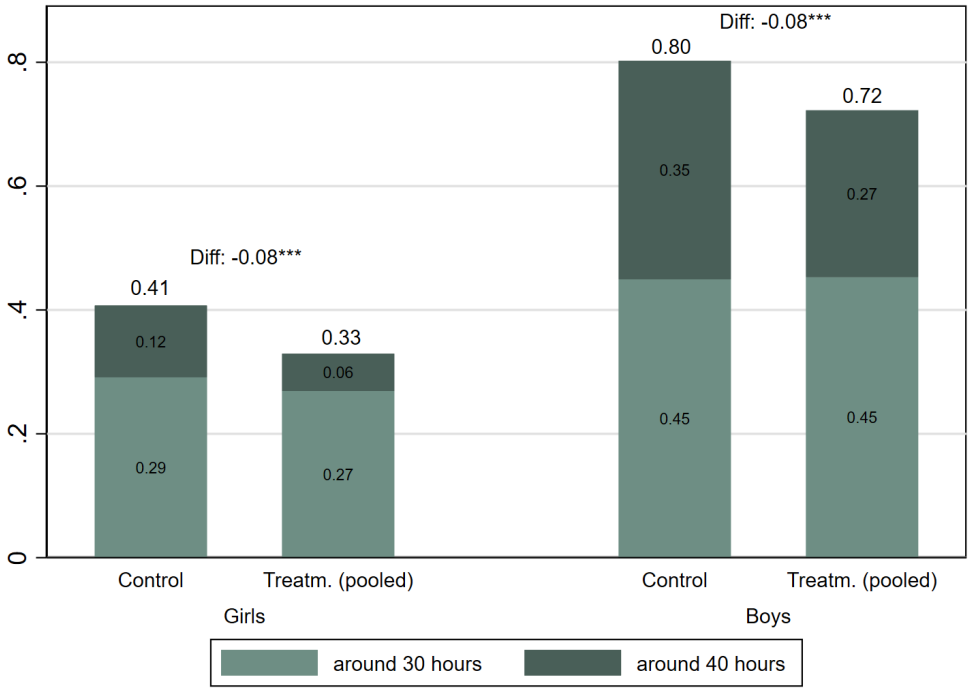
Notes: Beliefs about the extent of the norm that parents should reduce their labor market supply when children are young. Correct values indicated by vertical lines. Sample: Respondents in experimental groups *salience* or *salience and information*.

Figure 4: Self-expected Labor Supply across Treatments

Panel A: Expectations to Work at Most 20 Hours per Week

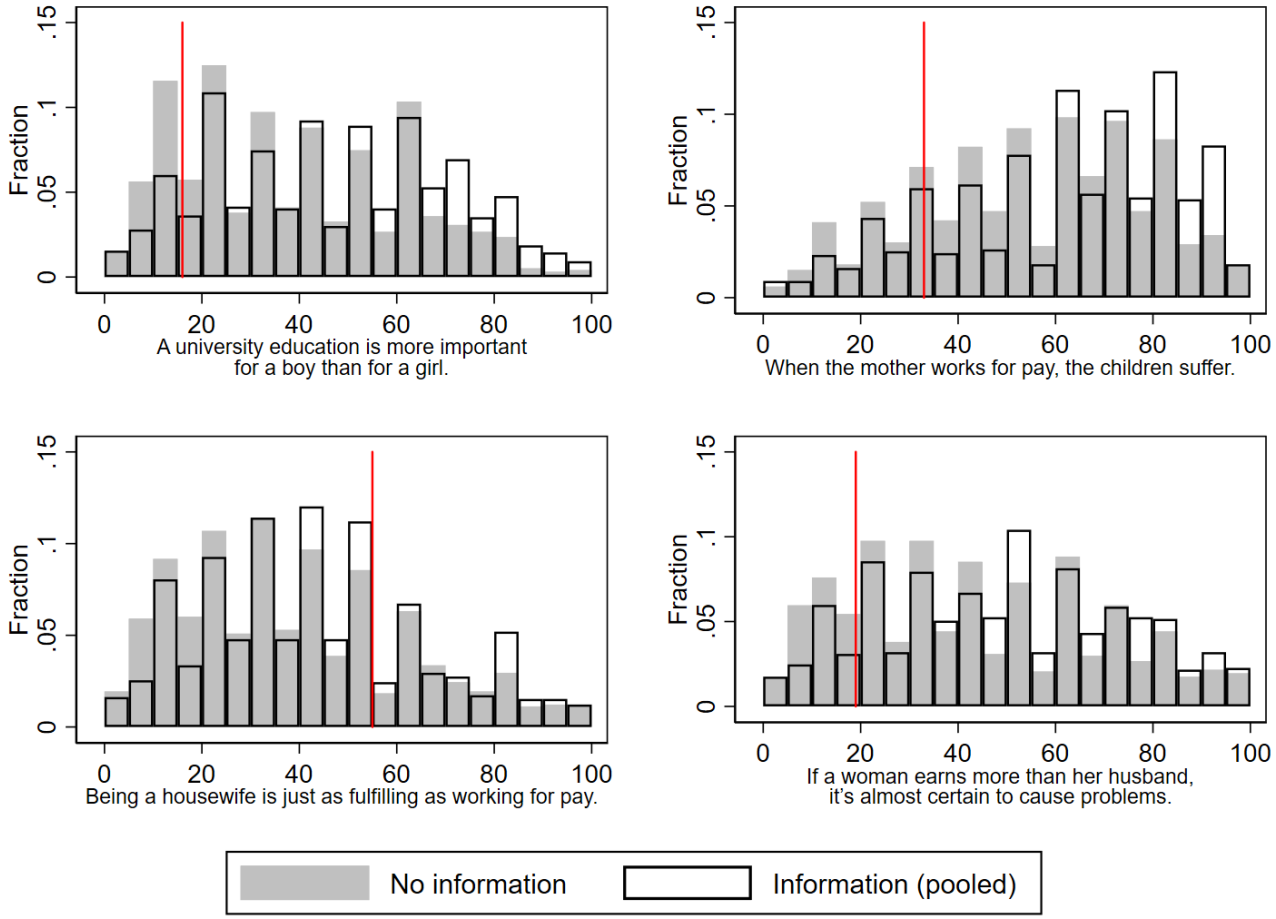


Panel B: Expectations to Work at Least 30 Hours per Week



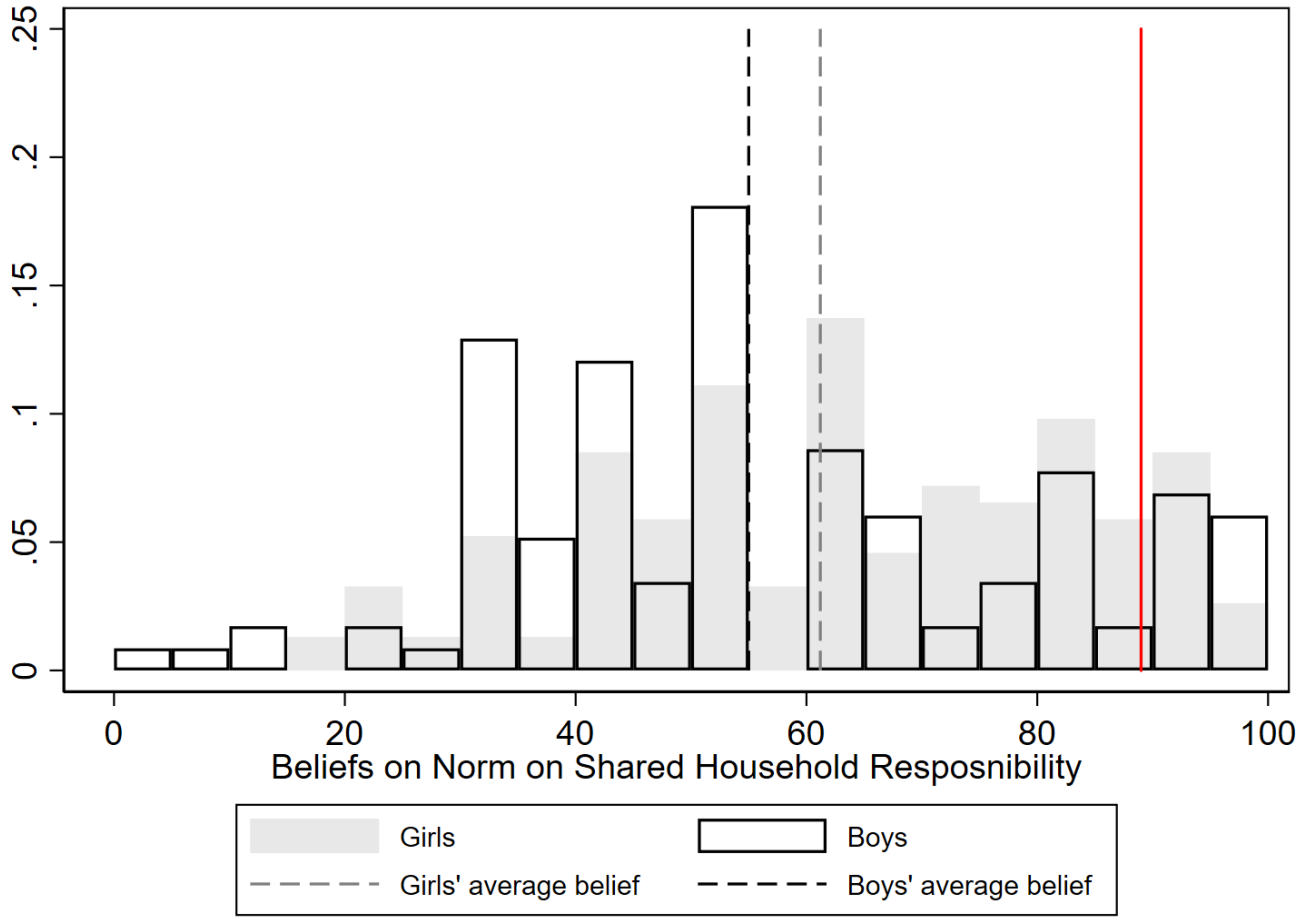
Notes: Responses to the question "Imagine you are 30 years old and you have a child aged between 1 and 6 years with your partner. What do you think, how many hours per week on average would you like to work in order to earn money?" Treatm. (pooled): Respondent in experimental groups *salience*, *information* or *salience and information*. ***/**/* indicate significance at the 1%/5%/10% level from regression according to equation 1.

Figure 5: Distribution of Incentivized Outcomes across Treatments



Notes: Responses to incentivized belief questions on share of Germans agreeing with gender-related statements depicted at x-axis. Correct values indicated by vertical lines. No information: Respondent in experimental groups *control group* or *salience*. Information (pooled): Respondent in experimental groups *information* or *salience and information*.

Figure 6: Distribution of Beliefs about the Norm on Shared Household Responsibility



Notes: Beliefs about the extent of the norm that men should take as much responsibility for the home and the children as women. Correct values indicated by vertical lines. Sample: Respondents in experimental group *salience and egalitarian information*.

Table 1: Treatment Effects on Labor-Supply Expectations

	Self-expected labor supply	Expected within-family gender gap
	(1)	(2)
Panel A: Girls		
Treatments (pooled)	-2.610*** (0.600)	2.240*** (0.766)
p-values: MHT Correction		
Treatments (pooled)	0.000	0.006
Control mean	23.84	7.65
Observations	1164	1164
R-squared	0.099	0.075
Panel B: Boys		
Treatments (pooled)	-1.814*** (0.650)	-1.415 (0.980)
p-values: MHT Correction		
Treatments (pooled)	0.004	0.162
Control mean	31.13	10.88
Observations	836	836
R-squared	0.039	0.050
Panel C: All		
Treatments (pooled)	-2.187*** (0.445)	0.708 (0.608)
Control mean	27.06	9.07
Observations	2000	2000
R-squared	0.191	0.044
Covariates	Yes	Yes

Notes: OLS regressions. Treatments (pooled): Respondent in experimental groups *salience*, *information* or *salience and information*. Dependent variables: (1) Hours expected to work per week when having a child 1-6 (0=0 hours, i.e. not at all; 10= 10 hours; 20= 20 hours, i.e. part-time; 30=30 hours, 40=40 hours, i.e. full-time); (2) Responses to labor supply for both parents (self and partner) with higher values indicating higher labor market supply of men relative to women. Control mean: Mean of the outcome variable in the control group. Covariates include: age, gender, born in Germany, living with parents, currently in school, current school track leading to university entrance degree, mother working full-time, having siblings, West Germany, living in large city, parents with university education, risk, patience, and imputation dummys. MHT Correction refers to the multiple hypothesis testing procedure presented in List et al. (2019) and corrects for two subgroups (girls and boys). Robust standard errors in parentheses. ***/**/* indicate significance at the 1%/5%/10% level.

Table 2: Separate Treatment Effects on Labor-Supply Expectations

	Self-expected labor supply	Expected within-family gender gap
	(1)	(2)
Panel A: Girls		
Salience	-1.965*** (0.731)	2.207** (0.924)
Information	-2.582*** (0.774)	1.969** (0.970)
Salience & information	-3.250*** (0.718)	2.552*** (0.941)
p-values: MHT Correction		
Salience	0.022	0.092
Information	0.004	0.164
Salience & information	0.000	0.045
Control mean	23.84	7.65
Observations	1164	1164
R-squared	0.101	0.076
Panel B: Boys		
Salience	-1.328 (0.823)	-1.233 (1.283)
Information	-1.578* (0.872)	-2.145* (1.295)
Salience & information	-2.418*** (0.786)	-0.951 (1.200)
p-value: MHT Correction		
Salience	0.095	0.538
Information	0.126	0.229
Salience & information	0.014	0.407
Control mean	31.13	10.88
Observations	836	836
R-squared	0.041	0.051
Panel C: All		
Salience	-1.534*** (0.555)	0.779 (0.756)
Information	-2.132*** (0.581)	0.372 (0.785)
Salience & information	-2.827*** (0.534)	0.963 (0.748)
p-value: MHT Correction		
Salience	0.007	0.498
Information	0.000	0.639
Salience & information	0.000	0.444
Control mean	27.06	9.07
Observations	2000	2000
R-squared	0.193	0.045
Covariates	Yes	Yes

Notes: OLS regressions. *Salience*, *Information* and *Salience & information* indicate membership of respective treatment groups. Dependent variables: (1) Hours expected to work per week when having a child 1-6 (0=0 hours, i.e. not at all; 10= 10 hours; 20= 20 hours, i.e. part-time; 30=30 hours, 40=40 hours, i.e. full-time); (2) Responses to labor supply for both parents (self and partner) with higher values indicating higher labor market supply of men relative to women. Control mean: Mean of the outcome variable in the control group. See Table 1 for included covariates. MHT Correction refers to the multiple hypothesis testing procedure presented in List et al. (2019) and corrects for two subgroups (girls and boys) as well as multiple treatments in Panel B. Results from Wald tests, testing for equal coefficients reject *Salience*=*Salience & information* in column (1) of Panel A ($p<0.1$) as well as column (1) of Panel C ($p<0.05$). Equal coefficients within all remaining treatment/outcome/subgroup combinations cannot be rejected. Robust standard errors in parentheses. ***/**/* indicate significance at the 1%/5%/10% level.

Table 3: Persistence of Information Treatment Effects on Self-Expected Labor Supply (Follow-up Sample)

	Self-expected labor supply		
	(1) All	(2) Girls	(3) Boys
Information provision (pooled)	-1.673***	-1.673**	-1.537**
	(0.493)	(0.676)	(0.702)
Information provision (pooled) x follow-up	0.553	0.641	0.438
	(0.518)	(0.691)	(0.788)
Follow-up	0.447	0.503	0.378
	(0.359)	(0.465)	(0.564)
Info provision in follow-up	-1.120**	-1.032	-1.098
Control mean	26.15	22.63	30.48
Covariates	Yes	Yes	Yes
Observations (respondents)	1319	734	585
R-squared	0.199	0.109	0.028

Notes: OLS regressions. Info provision: Respondent in experimental groups *information* or *salience and information*. Dependent variable: Hours expected to work per week when having a child 1-6 (0=0 hours, i.e. not at all; 10= 10 hours; 20= 20 hours, i.e. part-time; 30=30 hours, 40=40 hours, i.e. full-time). *Info provision in follow-up* is the linear combination of the coefficients on *Info provision* plus *Info provision x follow-up*. Control mean: Mean of the outcome variable in the omitted group (i.e. experimental groups *control group* or *salience*) reported in the main survey. See Table 1 for included covariates. Sample: Respondents who participated in the follow-up survey. Robust standard errors, adjusted for clustering at the respondent level, in parentheses. ***/**/* indicate significance at the 1%/5%/10% level.

Table 4: Heterogeneous Treatment Effects on Self-expected Labor Supply by Prior Beliefs (Belief Elicitation Sample)

	Self-expected labor supply		
	(1) All	(2) Girls	(3) Boys
Salience & information	-1.334*	-1.017	-1.297
	(0.687)	(1.241)	(0.845)
Misperception x salience & information	-0.046	-0.145	-0.020
	(0.266)	(0.403)	(0.400)
Misperception	0.292	0.366	0.215
	(0.219)	(0.302)	(0.303)
Covariates	Yes	Yes	Yes
Observations	965	565	400
R-squared	0.197	0.116	0.045

Notes: OLS regressions. Salience & information: Respondents in respective experimental group. Dependent variable: Hours expected to work per week when having a child 1-6 (0=0 hours, i.e. not at all; 10= 10 hours; 20= 20 hours, i.e. part-time; 30=30 hours, 40=40 hours, i.e. full-time). Misperception: Actual share minus guessed share of German adult population holding the opinion that parents (of respondent's gender) should reduce their labor market supply as long as the children are young, divided by 10. See Table 1 for included covariates. Sample: Respondents in experimental groups *salience* and *salience and egalitarian information*. Robust standard errors in parentheses. ***/**/* indicate significance at the 1%/5%/10% level.

Table 5: Information Treatment Effects on Incentivized Outcome

	University education more important for boy.	Children suffer if mother works for pay.	Being a housewife as fulfilling as working for pay.	Causes problems if a woman earns more than her husband.
	(1)	(2)	(3)	(4)
Panel A: Girls				
Information provision (pooled)	0.457***	0.173***	0.095***	0.225***
	(0.085)	(0.042)	(0.024)	(0.075)
p-values: MHT Correction				
Information (pooled)	0.000	0.000	0.000	0.003
Control mean	2.262	1.612	0.668	2.150
Covariates	Yes	Yes	Yes	Yes
Observations	1130	1156	1143	1137
R-squared	0.041	0.053	0.046	0.024
Panel B: Boys				
Information provision (pooled)	0.607***	0.215***	0.069**	0.337***
	(0.102)	(0.049)	(0.029)	(0.087)
p-values: MHT Correction				
Information (pooled)	0.000	0.000	0.014	0.000
Control mean	2.299	1.583	0.678	1.978
Covariates	Yes	Yes	Yes	Yes
Observations	811	829	820	807
R-squared	0.069	0.059	0.057	0.040

Notes: OLS regressions. Info provision: Respondent in experimental groups *information* or *salience and information*. Dependent variables: Beliefs about share of Germany agreeing with the statements that (1) a university education is more important for a boy than for a girl relative to correct value (=16). (2) the children suffer if the mothers works for pay relative to correct value (=33). (3) being a housewife is just as fulfilling as working for pay relative to correct value (=55). (4) it is almost certain to cause problems if a woman earns more than her husband relative to correct value (=20). Results (not shown) from full interaction model between gender and treatment indicators do not reveal any heterogeneous treatment effects by gender. Control mean: Mean of the outcome variable in the control group. See Table 1 for included covariates. MHT Correction refers to the multiple hypothesis testing procedure presented in List et al. (2019) and corrects for multiple subgroups (girls and boys) and multiple outcomes (all 4 outcomes listed). Robust standard errors in parentheses. ***/**/* indicate significance at the 1%/5%/10% level.

Table 6: Effects of the More Egalitarian Norm on the Expected Within-Family Gender Gap (Follow-Up Sample)

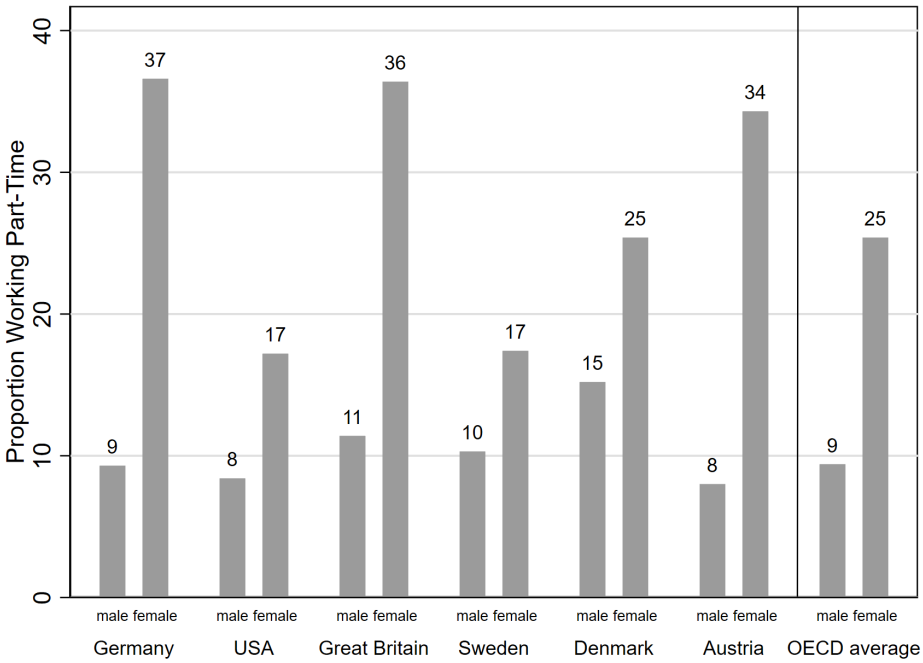
	Expected within-family gender gap		
	(1) All	(2) Girls	(3) Boys
Egalitarian treatments (pooled)	-1.321***	-1.261***	-1.398**
	(0.351)	(0.437)	(0.571)
p-values: MHT Correction			
Treatments (pooled)		0.007	1.000
Individual FEs	Yes	Yes	Yes
Observations (respondents)	1319	734	585
R-squared	0.010	0.011	0.009

Notes: OLS regressions. Egalitarian treatments (pooled): Respondent in experimental groups *egalitarian information* or *salience and egalitarian information*. Dependent variables: Responses to labor supply for both parents (self and partner) with higher values indicating higher labor market supply of men relative to women. Results (not shown) from interaction model between gender and treatment indicator do not reveal any heterogeneous treatment effects by gender. Sample: Follow-up survey participants. Robust standard errors, adjusted for clustering at the respondent level, in parentheses. ***/**/* indicate significance at the 1%/5%/10% level.

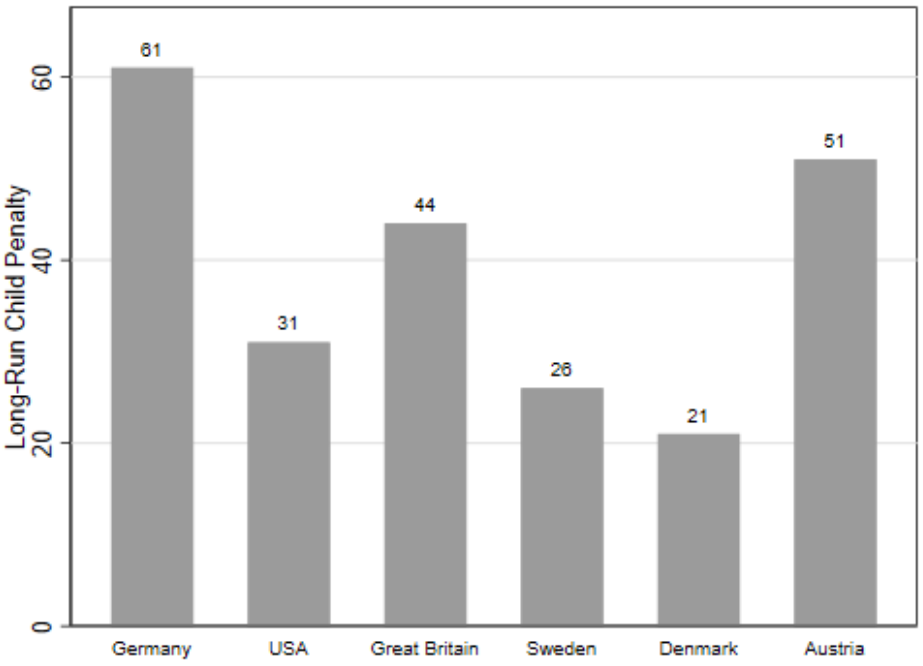
Appendix A: Appendix Figures and Tables

Figure A1: Gender Inequalities in Labor Supply across Countries

Panel A: Gender Gaps in Part-Time Employment across Countries

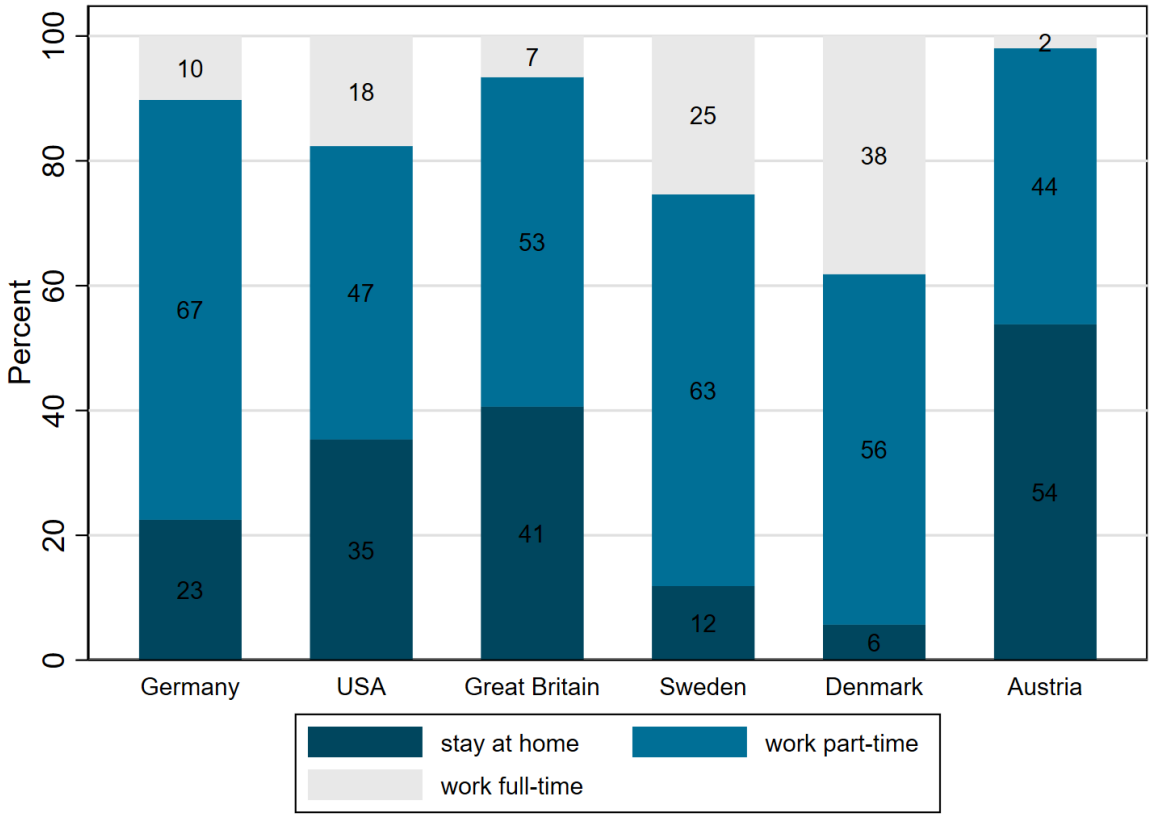


Panel B: Long-run Child Penalties across Countries



Notes: Panel A: Part-time employment rate as proportion of persons employed part-time among all employed persons, by gender. Part-time employment is defined as people in employment (whether employees or self-employed) who usually work less than 30 hours per week in their main job. Data source: OECD, 2018; Panel B: Long-run child penalties. Data source: Kleven et al. (2019).

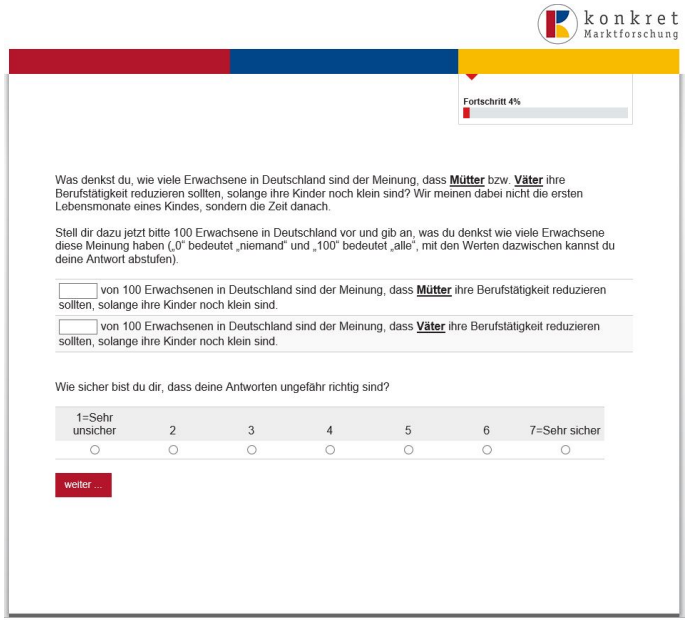
Figure A2: Norm about Maternal Labor Supply across Countries



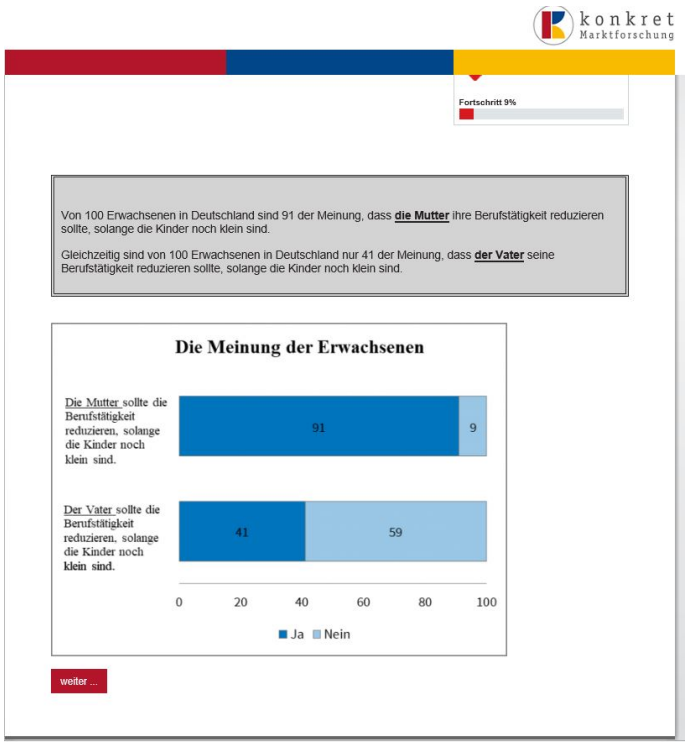
Notes: Response to the question "Do you think women should work outside the home full-time, part-time or not at all under the following circumstances? ...When there is a child under school age." Weighted means. Data source: International Social Survey Program (ISSP) in 2012.

Figure A3: Screenshots on Norm Treatments

Panel A: Belief Elicitation about Norm Prescribing Parental Labor Supply



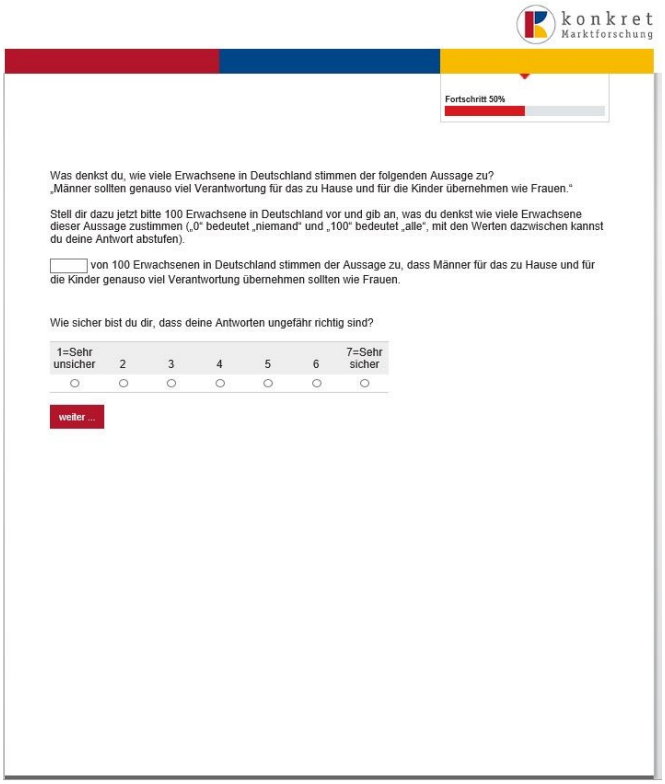
Panel B: Information about Norm Prescribing Parental Labor Supply



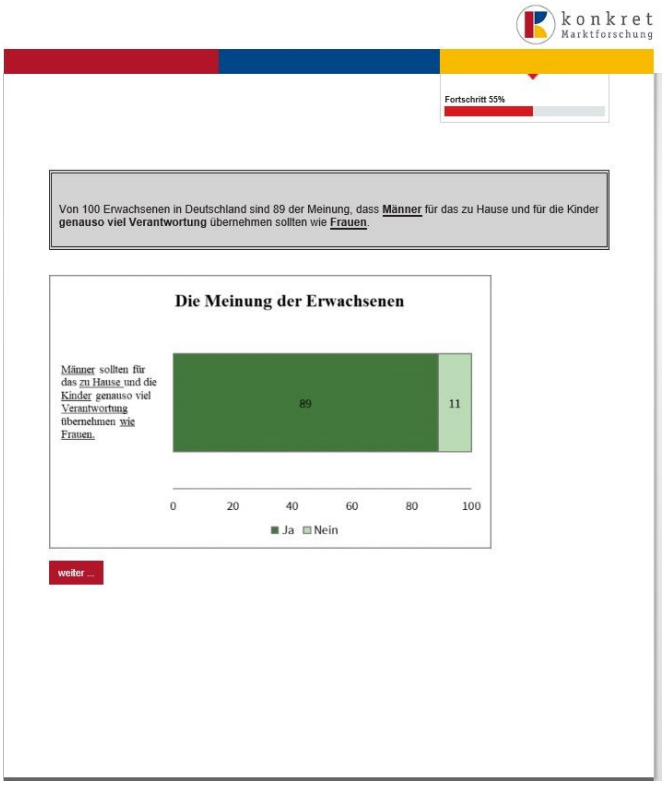
Notes: Panel A: Screenshot on the belief elicitation questions in treatment *salience*. Panel B: Screenshot on the information provision in treatment *information*. Respondents in treatment *salience & information* first receive the belief elicitation question (panel A) and afterwards accurate information (panel B).

Figure A4: Screenshots on More Egalitarian Norm Treatments in Follow-Up Survey

Panel A: Belief Elicitation about Norm on Shared Household Responsibility



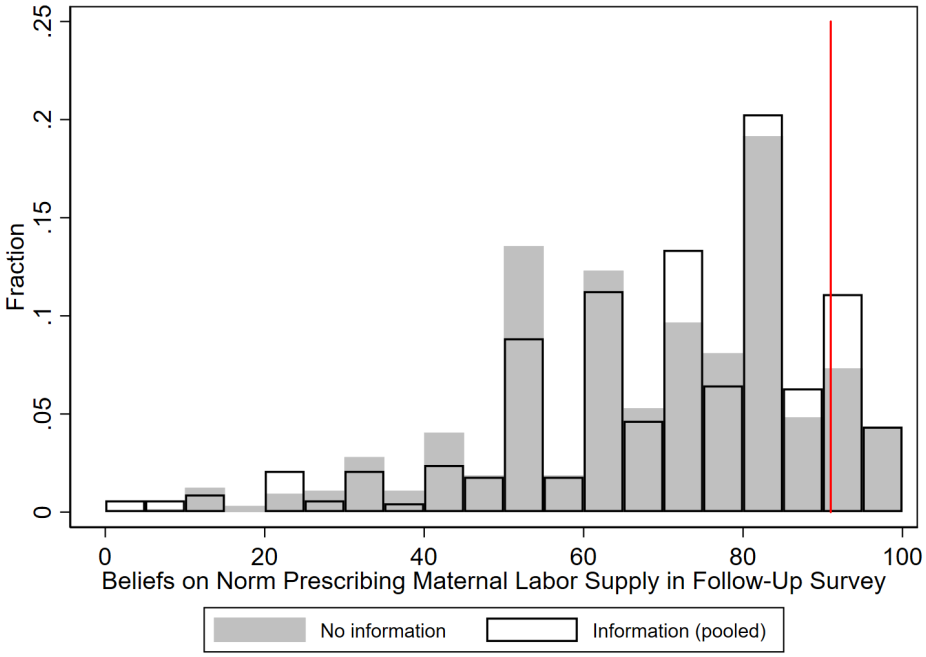
Panel B: Information about Norm on Shared Household Responsibility



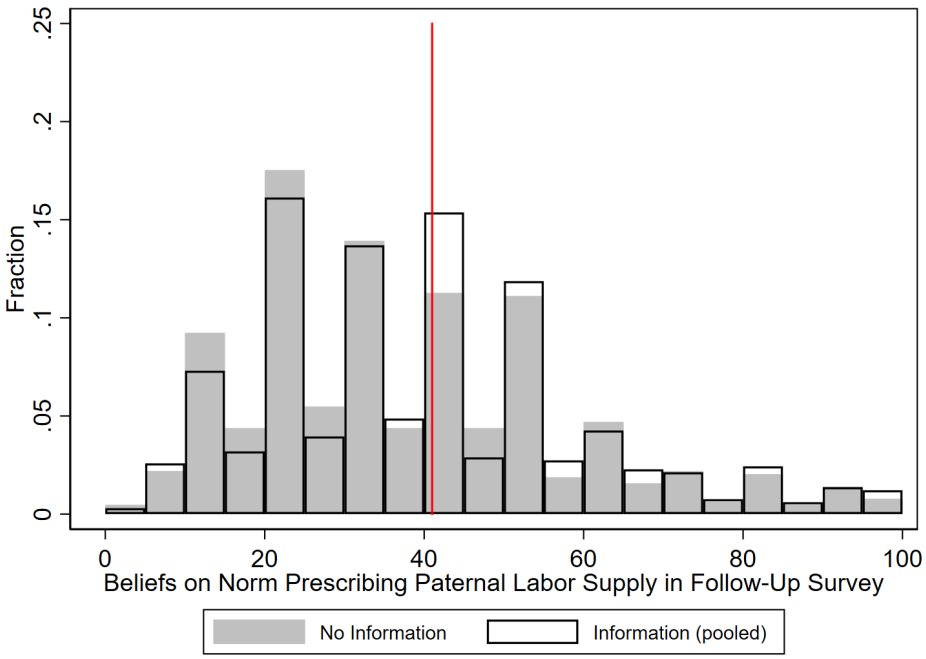
Notes: Respondents in treatment *salience and egalitarian information* first receive the belief elicitation question (panel A) and afterwards accurate information (panel B). Respondents in treatment *egalitarian information* are provided with accurate information (panel B).

Figure A5: Distribution of Beliefs about Norm Prescribing Parental Labor Supply in Follow-Up Survey

Panel A: Norm that Mothers Should Reduce their Labor Supply when Children Are Young



Panel B: Norm that Fathers Should Reduce their Labor Supply when Children Are Young



Notes: Beliefs about the extent of the norm that parents should reduce their labor market supply when children are young elicited in the follow-up survey. Correct values indicated by vertical lines. Information (pooled): Respondents in experimental groups *information* or *Salience and information*). Sample: Follow-up survey participants.

Table A1: Comparison of Survey Sample Characteristics to Microcensus Data

Characteristic	Microcensus 2015 (1)	Sample mean (2)
Female	0.488	0.582
		(0.011)
Age	15.508	15.748
		(0.026)
Living in West Germany (excl. Berlin)	0.847	0.794
		(0.009)
Attending Hauptschule/Realschule (low/middle track)	0.288	0.172
		(0.008)
Attending school with several tracks	0.156	0.138
		(0.008)
Attending Gymnasium (high track)	0.393	0.509
		(0.011)
Living with both parents	0.761	0.723
		(0.010)
At least one parent with uni degree [if living with both]	0.449	0.420
		(0.013)
Mother does not work [if living with both]	0.233	0.173
		(0.010)
Mother works full-time [if living with both]	0.207	0.405
		(0.013)
Father works full-time [if living with both]	0.875	0.914
		(0.008)
Observations	18501	2000

Notes: Column 1: Means based on Microcensus data from 2015. Column 2: Sample means and standard errors (in parentheses) of our survey data. Data source: German population Microcensus 2015 and own survey data.

Table A2: Sociodemographic Characteristics across Treatments

	Control mean (1)	Salience (2)	Diff. (3)	Information (4)	Diff. (5)	Salience & information (6)	Diff. (7)
Female	0.559	0.603	0.043	0.607	0.048	0.563	0.004
Age	15.752	15.758	0.006	15.800	0.048	15.685	-0.067
Living in west Germany	0.796	0.778	-0.019	0.802	0.005	0.799	0.003
City size $\geq$ 100,000	0.681	0.698	0.016	0.697	0.016	0.687	0.006
Born in Germany	0.963	0.957	-0.006	0.973	0.010	0.976	0.013
Currently at school	0.893	0.922	0.030	0.922	0.030*	0.911	0.019
Obtained degree/Current track leads to							
No degree	-0.000	-0.000	0.000	0.002	0.002	0.000	0.000
Basic degree (Hauptschulabschluss)	0.049	0.049	-0.001	0.045	-0.005	0.040	-0.010
Middle school degree (Realschulabschluss)	0.284	0.239	-0.045	0.247	-0.037	0.305	0.021
University entrance degree (Fachabitur)	0.061	0.058	-0.003	0.096	0.035**	0.067	0.006
University entrance degree (Abitur)	0.606	0.654	0.049	0.611	0.005	0.589	-0.017
Living status							
Living with both parents	0.715	0.732	0.017	0.671	-0.044	0.705	-0.010
Living with one parent	0.267	0.251	-0.016	0.292	0.026	0.272	0.005
Living without parents	0.019	0.017	-0.001	0.037	0.018*	0.024	0.005
Having siblings	0.841	0.827	-0.014	0.849	0.008	0.797	-0.043*
At least one parent with university degree	0.369	0.413	0.044	0.389	0.020	0.386	0.017
Maternal employment status							
Mother works full-time	0.419	0.426	0.008	0.423	0.005	0.437	0.018
Mother works part-time	0.367	0.357	-0.010	0.366	-0.001	0.362	-0.004
Mother housewife	0.031	0.026	-0.006	0.033	0.001	0.031	0.000
Paternal employment status							
Father works full-time	0.809	0.814	0.005	0.838	0.029	0.852	0.043*
Father works part-time	0.041	0.043	0.002	0.033	-0.008	0.041	0.001
Risk tolerance (11-point scale)	5.674	5.616	-0.059	5.538	-0.136	5.500	-0.174
Patience (11-point scale)	6.461	6.341	-0.120	6.430	-0.031	6.421	-0.040
Observations	540	463		463		508	

Notes: Group means. "Diff." displays the difference in means between the control group and respective treatment groups. Significance levels of "Diff." from linear regressions of the background variables on the respective treatment indicators. Robust standard errors in parentheses. ***/**/* indicate significance at the 1%/5%/10% level.

Table A3: Participation in the Follow-Up Survey

	Respondent participated in follow-up survey	
	(1)	
Treatments		
Salience	0.022	(0.031)
Information	0.017	(0.031)
Salience & information	0.056*	(0.030)
Covariates		
Female	-0.058***	(0.022)
Age	-0.041***	(0.010)
Living in west Germany	-0.022	(0.027)
City size $\geq$ 100,000	0.060**	(0.024)
Born in Germany	-0.068	(0.059)
Currently at school	0.086*	(0.047)
No degree	0.435***	(0.096)
Middle school degree (Realschulabschluss)	0.029	(0.055)
University entrance degree (Fachabitur)	0.073	(0.067)
University entrance degree (Abitur)	0.026	(0.056)
Living with one parent	-0.034	(0.026)
Living without parents	-0.090	(0.075)
Having siblings	-0.044	(0.028)
At least one parent with uni degree	-0.005	(0.023)
Mother works full-time	0.017	(0.023)
Father works full-time	0.058*	(0.031)
Risk tolerance (11-point scale)	-0.006	(0.005)
Patience (11-point scale)	-0.004	(0.005)
Observations	1901	
R-squared	0.037	

Notes: Dependent variable: Dummy variable coded one if respondent participated in the follow-up survey. Salience/Information/Salience & information indicate membership of respective treatment group. Robust standard errors in parentheses. ***/**/* indicate significance at the 1%/5%/10% level.

Table A4: Sociodemographic Characteristics across the Treatments (Follow-up Sample)

	Control mean (1)	Salience (2)	Diff. (3)	Information (4)	Diff. (5)	Salience & information (6)	Diff. (7)
Female	0.535	0.570	0.035	0.581	0.046	0.544	0.009
Age	15.623	15.694	0.071	15.714	0.091	15.558	-0.065
Living in west Germany	0.789	0.769	-0.021	0.797	0.007	0.803	0.013
City size $\geq$ 100,000	0.690	0.717	0.027	0.733	0.043	0.704	0.014
Born in Germany	0.956	0.958	0.002	0.968	0.012	0.972	0.016
Currently at school	0.918	0.922	0.004	0.949	0.031	0.935	0.017
Obtained degree/Current track leads to							
No degree	0.000	-0.000	-0.000	0.003	0.003	0.000	0.000
Basic degree (Hauptschulabschluss)	0.047	0.054	0.007	0.039	-0.007	0.030	-0.017
Middle school degree (Realschulabschluss)	0.298	0.246	-0.052	0.241	-0.057	0.290	-0.008
University entrance degree (Fachabitur)	0.068	0.074	0.006	0.085	0.016	0.069	0.001
University entrance degree (Abitur)	0.587	0.626	0.039	0.632	0.045	0.611	0.024
Living status							
Living with both parents	0.713	0.746	0.032	0.698	-0.015	0.713	-0.001
Living with one parent	0.263	0.238	-0.025	0.279	0.016	0.268	0.004
Living without parents	0.023	0.016	-0.007	0.022	-0.001	0.020	-0.004
Having siblings	0.822	0.810	-0.011	0.857	0.036	0.772	-0.050
At least one parent with university degree	0.383	0.404	0.021	0.410	0.026	0.386	0.003
Maternal employment status							
Mother works full-time	0.421	0.453	0.032	0.410	-0.012	0.451	0.030
Mother works part-time	0.380	0.332	-0.048	0.397	0.017	0.366	-0.014
Mother housewife	0.035	0.026	-0.009	0.025	-0.010	0.039	0.004
Paternal employment status							
Father works full-time	0.815	0.840	0.025	0.834	0.019	0.868	0.052*
Father works part-time	0.038	0.036	-0.002	0.038	0.000	0.037	-0.002
Risk tolerance (11-point scale)	5.611	5.518	-0.093	5.479	-0.132	5.456	-0.155
Patience (11-point scale)	6.322	6.267	-0.055	6.404	0.083	6.439	0.118
Observations	342	307		315		355	

Notes: Group means. "Diff." displays the difference in means between the control group and respective treatment groups. Significance levels of "Diff." from linear regressions of the background variables on the respective treatment indicators. Robust standard errors in parentheses. Sample: Follow-up survey participants. ***/**/* indicate significance at the 1%/5%/10% level.

Table A5: Sociodemographic Characteristics across the More Egalitarian Treatments (Follow-up Sample)

	Control mean (1)	Egalitarian information (2)	Diff. (3)	Salience & egalitarian information (4)	Diff. (5)
Female	0.552	0.557	0.005	0.567	0.015
Age	15.624	15.675	0.051	15.652	0.028
Living in west Germany	0.794	0.787	-0.007	0.785	-0.008
City size $\geq$ 100,000	0.716	0.689	-0.027	0.726	0.010
Born in Germany	0.963	0.970	0.007	0.956	-0.008
Currently at school	0.922	0.932	0.009	0.952	0.029*
Obtained degree/Current track leads to					
No degree	-0.000	-0.000	0.000	0.004	0.004
Basic degree (Hauptschulabschluss)	0.041	0.037	-0.004	0.050	0.009
Middle school degree (Realschulabschluss)	0.278	0.269	-0.009	0.251	-0.027
University entrance degree (Fachabitur)	0.071	0.077	0.006	0.077	0.007
University entrance degree (Abitur)	0.610	0.617	0.007	0.618	0.008
Living status					
Living with both parents	0.719	0.727	0.008	0.700	-0.019
Living with one parent	0.261	0.249	-0.012	0.285	0.025
Living without parents	0.020	0.025	0.004	0.015	-0.006
Having siblings	0.806	0.814	0.008	0.833	0.027
At least one parent with university degree	0.394	0.393	-0.000	0.400	0.006
Maternal employment status					
Mother works full-time	0.452	0.418	-0.034	0.407	-0.045
Mother works part-time	0.357	0.410	0.053*	0.344	-0.013
Mother housewife	0.026	0.041	0.015	0.033	0.007
Paternal employment status					
Father works full-time	0.840	0.839	-0.001	0.840	-0.000
Father works part-time	0.040	0.038	-0.001	0.030	-0.010
Risk tolerance (11-point scale)	5.577	5.311	-0.265*	5.641	0.064
Patience (11-point scale)	6.374	6.470	0.096	6.178	-0.196
Observations	683	366		270	

Notes: Group means. "Diff." displays the difference in means between the control group and respective treatment groups. Significance levels of "Diff." from linear regressions of the background variables on the respective treatment indicators. Robust standard errors in parentheses. Sample: Follow-up survey participants. Robust standard errors in parentheses. ***/**/* indicate significance at the 1%/5%/10% level.

Table A6: Labor-Supply Expectations across Sociodemographic Characteristics (Respondents in the Control Group)

	Self-expected labor supply (1)		Expected within-family gender gap (2)	
Female	-7.293***	(0.748)	-3.233***	(1.053)
Living in east Germany	3.698***	(0.910)	-4.089***	(1.009)
City size $\geq$ 100,000	0.372	(0.894)	-0.420	(1.175)
Aged 16 or 17	0.111	(0.830)	-0.831	(1.072)
University entrance degree (Abitur)	-0.030	(0.929)	-4.615***	(1.189)
Mother works full-time	0.346	(0.807)	-2.669***	(1.023)
Father works full-time	-0.069	(1.033)	0.225	(1.250)
Having siblings	-0.736	(1.139)	1.112	(1.235)
At least one parent with university degree	-0.510	(0.821)	-2.910***	(1.046)
Patience median or above	0.112	(0.819)	1.158	(1.051)
Risk median or above	0.557	(0.820)	0.187	(1.044)
Mean of the outcome	27.056		9.074	
Observations	540		540	

Notes: Bivariate OLS regressions, each column shows results from a different regression of the respective outcome on the respective sociodemographic characteristic. Dependent variables: (1) Hours expected to work per week with child 1-6 (0=0 hours, i.e. not at all; 10= 10 hours; 20= 20 hours, i.e. part-time; 30=30 hours, 40=40 hours, i.e. full-time). (2) Responses to labor supply for both parents (self and partner) with higher values indicating higher labor market supply of men relative to women. Sample: Respondents in the control group. Robust standard errors in parentheses. ***/**/* indicate significance at the 1%/5%/10% level.

Table A7: Heterogeneous Treatment Effects on Labor-Supply Expectations by Gender

	Self-expected labor supply	Expected within-family gender gap
	(1)	(2)
Panel A: Combined treatment effects		
Treatments (pooled)	-1.805***	-1.453
	(0.658)	(0.975)
Treatments (pooled) x female	-0.675	3.811***
	(0.889)	(1.243)
Female	-7.333***	-3.358***
	(0.749)	(1.045)
Panel B: Seperate treatment effects		
Salience	-1.286	-1.191
	(0.845)	(1.278)
Information	-1.648*	-1.995
	(0.871)	(1.293)
Salience & information	-2.366***	-1.210
	(0.798)	(1.201)
Salience x female	-0.449	3.476**
	(1.120)	(1.577)
Information x female	-0.836	4.113**
	(1.165)	(1.620)
Salience & information x female	-0.824	3.883**
	(1.073)	(1.533)
Female	-7.332***	-3.363***
	(0.750)	(1.046)
Control mean	31.13	10.88
Covariates	Yes	Yes
Observations	2000	2000
R-squared	0.193	0.049

Notes: OLS regressions. Panel A: Treatments (pooled): Respondent in experimental groups *salience*, *information* or *salience and information*. Panel B: *Salience*, *Information* and *Salience & information* indicate membership of respective treatment groups. Dependent variables: (1) Hours expected to work per week when having a child 1-6 (0=0 hours, i.e. not at all; 10= 10 hours; 20= 20 hours, i.e. part-time; 30=30 hours, 40=40 hours, i.e. full-time); (2) Responses to labor supply for both parents (self and partner) with higher values indicating higher labor market supply of men relative to women. Control mean: Mean of the outcome variable in the control group. See Table 1 for included covariates. Robust standard errors in parentheses. ***/**/* indicate significance at the 1%/5%/10% level.

Table A8: Persistence of Separate Treatment Effects on Self-Expected Labor Supply (Follow-up Sample)

	Self-expected labor supply		
	(1) All	(2) Girls	(3) Boys
Salience	-2.034***	-2.613***	-1.845*
	(0.698)	(0.954)	(0.971)
Information	-2.042***	-2.548***	-1.630*
	(0.690)	(0.964)	(0.963)
Salience & information	-3.157***	-3.323***	-2.981***
	(0.648)	(0.896)	(0.920)
Salience x follow-up	2.490***	2.258**	2.774**
	(0.714)	(0.922)	(1.137)
Information x follow-up	1.461**	1.694*	1.108
	(0.694)	(0.958)	(1.008)
Salience & information x follow-up	1.970***	1.793*	2.177**
	(0.731)	(1.007)	(1.072)
Follow-up	-0.731	-0.601	-0.881
	(0.482)	(0.667)	(0.704)
Salience in follow-up	0.456	-0.355	0.929
Information in follow-up	-0.581	-0.854	-0.522
Salience & information in follow-up	-1.187*	-1.531	-0.804
Control mean	27.19	23.66	31.26
Covariates	Yes	Yes	Yes
Observations (respondents)	1319	734	585
R-squared	0.202	0.114	0.033

Notes: OLS regressions. *Salience*, *Information* and *Salience & information* indicate membership of respective treatment groups. Dependent variable: Hours expected to work per week when having a child 1-6 (0=0 hours, i.e. not at all; 10= 10 hours; 20= 20 hours, i.e. part-time; 30=30 hours, 40=40 hours, i.e. full-time). *Salience in follow-up*/*Information in follow-up*/*Salience & information in follow-up* are the linear combinations of the coefficients on the *respective treatment indicators* plus *respective treatment indicator x follow-up*. Control mean: Mean of the outcome variable in the omitted group (i.e. control group and treatment *salience*) reported in the main survey. See Table 1 for included covariates. Sample: Follow-up survey respondents. Robust standard errors, adjusted for clustering at the respondent level, in parentheses. ***/**/* indicate significance at the 1%/5%/10% level.

Table A9: Persistence of Information Treatment Effects on Beliefs about Norms Prescribing Parental Labor Supply (Follow-up Sample)

	Relative belief			
	(1) All	(2) Girls	(3) Boys	(4) All
Panel A: Social norm towards mothers elicited in follow-up				
Information provision (pooled)	0.025**	0.034**	0.013	0.014
	(0.012)	(0.016)	(0.018)	(0.018)
Information provision (pooled) x female				0.019
				(0.024)
Female				0.001
				(0.017)
Control mean	0.73	0.73	0.73	0.73
Covariates	Yes	Yes	Yes	Yes
Observations	1308	731	577	1308
R-squared	0.017	0.035	0.029	0.017
Panel B: Social norm towards fathers elicited in follow-up				
Information provision (pooled)	0.058**	0.038	0.086**	0.085**
	(0.027)	(0.037)	(0.041)	(0.040)
Information provision (pooled) x female				-0.049
				(0.055)
Female				0.100**
				(0.039)
Control mean	0.87	0.91	0.82	0.87
Covariates	Yes	Yes	Yes	Yes
Observations	1296	728	568	1296
R-squared	0.040	0.045	0.048	0.040

Notes: OLS regressions. Info provision: Respondent in the experimental groups *information* or *salience and information*. Dependent variables: Panel A: Belief about social norm towards mothers relative to correct value (=91). Panel B: Belief about social norm towards fathers relative to correct value (=41). Control mean: Mean of the outcome variable in the omitted group (i.e. control group *and* belief elicitation only). See Table 1 for included covariates. Sample: Follow-up survey respondents. Robust standard errors in parentheses. ***/**/* indicate significance at the 1%/5%/10% level.

Table A10: Seperate Treatment Effects on Incentivized Outcomes

	University education more important for boy.	Children suffer if mother works for pay.	Being a housewife as fulfilling as working for pay.	Causes problems if a woman earns more than her husband.
	(1)	(2)	(3)	(4)
Panel A: Girls				
Salience	0.016	-0.014	-0.011	0.013
	(0.117)	(0.058)	(0.034)	(0.107)
Information	0.437***	0.104*	0.086**	0.176*
	(0.117)	(0.058)	(0.034)	(0.105)
Salience & information	0.494***	0.231***	0.095***	0.288***
	(0.118)	(0.058)	(0.034)	(0.107)
p-values: MHT Correction				
Salience	0.989	0.999	1.000	0.900
Information	0.000	0.634	0.134	0.664
Salience & information	0.000	0.000	0.072	0.126
Control mean	2.266	1.629	0.679	2.154
Covariates	Yes	Yes	Yes	Yes
Observations	1130	1156	1143	1137
R-squared	0.041	0.057	0.046	0.025
Panel B: Boys				
Salience	-0.097	-0.107	-0.007	-0.140
	(0.138)	(0.068)	(0.042)	(0.124)
Information	0.412***	0.091	0.043	0.161
	(0.146)	(0.070)	(0.042)	(0.125)
Salience & information	0.691***	0.233***	0.086**	0.369***
	(0.138)	(0.066)	(0.039)	(0.118)
p-values: MHT Correction				
Salience	0.973	0.709	0.999	0.886
Information	0.069	0.843	0.911	0.856
Salience & information	0.000	0.013	0.299	0.014
Control mean	2.329	1.634	0.683	2.040
Covariates	Yes	Yes	Yes	Yes
Observations	811	829	820	807
R-squared	0.073	0.066	0.059	0.045

Notes: OLS regressions. *Salience*, *Information* and *Salience & information* indicate membership of respective treatment groups. Dependent variables: Beliefs about share of Germans agreeing with the statements that (1) a university education is more important for a boy than for a girl relative to correct value (=16). (2) the children suffer if the mothers works for pay relative to correct value (=33). (3) being a housewife is just as fulfilling as working for pay relative to correct value (=55). (4) it is almost certain to cause problems if a woman earns more than her husband relative to correct value (=20). Results (not shown) from full interaction model between gender and treatment indicator reveals significant heterogeneous treatment effects by gender. Control mean: Mean of the outcome variable in the control group. See Table 1 for included covariates. MHT Correction refers to the multiple hypothesis testing procedure presented in List et al. (2019) and corrects for multiple subgroups (girls and boys), multiple treatments as well as multiple outcomes (all 4 outcomes listed). Robust standard errors in parentheses. ***/**/* indicate significance at the 1%/5%/10% level.

Table A11: Treatment Effects on Labor-Supply Expectations Without Child

	Self-expected labor supply	Expected within-family gender gap
	(1)	(2)
Panel A: Girls		
Treatments (pooled)	-0.468	0.267
	(0.531)	(0.360)
p-values: MHT Correction		
Treatments (pooled)	0.375	0.714
Control mean	34.37	1.09
Observations	1164	1164
R-squared	0.023	0.014
Panel B: Boys		
Treatments (pooled)	0.538	-0.391
	(0.581)	(0.631)
p-values: MHT Correction		
Treatments (pooled)	0.590	0.539
Control mean	34.45	3.49
Observations	836	836
R-squared	0.026	0.023
Panel C: All		
Treatments (pooled)	0.002	0.050
	(0.391)	(0.340)
Control mean	34.41	2.15
Observations	2000	2000
R-squared	0.021	0.031
Covariates	Yes	Yes

Notes: OLS regressions. Treatments (pooled): Respondent in experimental groups *salience*, *information* or *salience and information*. Dependent variables: (1) Hours expected to work per week without child (0=0 hours, i.e. not at all; 10= 10 hours; 20= 20 hours, i.e. part-time; 30=30 hours, 40=40 hours, i.e. full-time); (2) Responses to labor supply without child for both spouses (self and partner) with higher values indicating higher labor market supply of men relative to women. Control mean: Mean of the outcome variable in the control group. See Table 1 for included covariates. MHT Correction refers to the multiple hypothesis testing procedure presented in List et al. (2019) and corrects for two subgroups (girls and boys). Robust standard errors in parentheses. ***/**/* indicate significance at the 1%/5%/10% level.

Table A12: Separate Treatment Effects of the More Egalitarian Norm on the Expected Within-Family Gender Gap (Follow-Up Sample)

	Expected within-family gender gap		
	(1) All	(2) Girls	(3) Boys
Egalitarian information	-1.066**	-1.078*	-1.049*
	(0.426)	(0.584)	(0.622)
Salience & egalitarian information	-1.667***	-1.503**	-1.880*
	(0.590)	(0.660)	(1.055)
p-values: MHT Correction			
Egalitarian information	0.013	0.066	1.000
Salience & egalitarian information	0.009	0.063	1.000
Individual FEs	Yes	Yes	Yes
Observations (respondents)	1319	734	585
R-squared	0.011	0.011	0.010

Notes: OLS regressions. *Egalitarian information* and *Salience & egalitarian information* indicate membership of respective treatment groups. Dependent variable: Responses to labor supply for both parents (self and partner) with higher values indicating higher labor market supply of men relative to women. Results (not shown) from interaction model between gender and treatment indicator does not reveals significant heterogeneous treatment effects by gender. Robust standard errors, adjusted for clustering at the respondent level, in parentheses. ***/**/* indicate significance at the 1%/5%/10% level.

Online Appendix O1: Additional Results

O1.1. Results from the ifo Education Survey 2018

The experiment presented in the main part of the paper is an extended version of a previous experiment that was conducted with 1,085 adolescents within the scope of the ifo Education Survey 2018.[38] The main purpose of the ifo Education Survey was to inform the German policy debate about the opinion of adolescents on education-policy topics that are unrelated to social norms or labor-supply expectations. At the end of the questionnaire (which encompassed a total of 30 questions on education policy), we ran a similar but substantially shortened version of the experiment on the social norm prescribing parental labor supply. Particularly, the experiment randomized respondents into one of two experimental groups (the control group and treatment *salience and information*), and then elicited labor-supply expectations. Online Appendix Table O4 presents the results. Treatment effects in this alternative sample are remarkably similar to—and statistically indistinguishable from—those in our main sample (see Table 2 for comparison). Given the importance of replication for the credibility of scientific findings (e.g., Maniadis et al., 2014), it is also reassuring that our treatment effects replicate in two independent samples.

O1.2. Perceived Peer Pressure as Potential Mechanism

A potential complementary mechanism through which the gender norm alters labor-market expectations is perceived peer pressure: Individuals may adhere to the social norm to meet their peer groups expectations, and thereby avoid peer punishment of non-conform behavior (e.g., Benabou and Tirole, 2011).[39]

To study the empirical relevance of this channel, we elicited respondents' beliefs about whether the following groups think that mothers and fathers, respectively, should reduce their labor supply when the children are young: (i) family, (ii) friends, and (iii) the (future) partner. For the analysis we combine these items into a z-standardized index measuring peer-group expectations (Kling et al., 2007). To gauge the relevance of these different peer groups for individual respondents, we additionally asked them how important it is for them to meet

[38] The ifo Education Survey is an annual opinion survey on education policy among representative samples of adults in Germany (see https://www.ifo.de/en/survey/ifo-education-survey). In the 2018 wave, the general-population sample was complemented by a sample of adolescents that was surveyed about education topics (see Woessmann et al., 2018). The sampling and polling was done by the polling firm *Kantar Public*, which drew respondents from a different subject pool than the one used for our main study.

[39] There are at least two further potential reason for why individuals follow norms, namely (i) because they derive direct utility from actions that maintain their identity-based self-concepts (e.g., Akerlof and Kranton, 2000) and (ii) because of "best-practice considerations" in the sense that respondents may think that norms entail meaningful information on how to achieve certain outcomes most effectively (e.g., Cialdini and Goldstein, 2004; Cialdini and Trost, 1998). Focusing on child development and family income as two such outcomes, we find no evidence that the latter channel is relevant in our setting (results available upon request).

the expectations of each of these four groups (Online Appendix O3 presents the question wording).[40]

Treated girls, but not boys are more likely to believe that their peers expect mothers to reduce their labor supply. In columns 1 and 2 of Online Appendix Table O5, we regress the indices of respondents' beliefs about their peer groups' norms towards mothers and fathers, respectively, on the pooled treatment indicator. The treatments increase girls' beliefs that their peer groups demand that mothers reduce their labor supply by 15 percent of a standard deviation, but do not affect their corresponding beliefs regarding fathers. For boys, we find no significant pooled treatment effects.

Next, we study the connection between the importance that adolescents assign to their peer groups' opinions and labor-supply expectations. Online Appendix Table O6 regresses girls' and boys' self-expected labor supply on the pooled treatment indicator and its interaction with the importance respondents assign to their peer groups' expectations.[41] The coefficient on the interaction term reveals that treatment effects for girls (but not for boys) are stronger the more important they consider peer expectations.

Put together, these two sets of results suggest that treatment effects for girls may in fact be driven by their desire to adhere to their peer groups' norms. The peer-pressure channel seems to be less relevant for boys.

O1.3. Importance of Different Job Attributes

To investigate whether the traditional norm affects further labor-market-related preferences that are only indirectly linked to labor supply as young parents, we also study treatment effects on adolescents' preferences for future job attributes. For that purpose, respondents rated the following eight different job attributes on a five-point scale from "very important" to "very unimportant": "The job ..." (i) "... can be reconciled with children.", (ii) "... enables part-time employment.", (iii) "... offers a high salary.", (iv) "... offers good career opportunities.", (v) "... offers job security.", (vi) "... is challenging.", (vii) "... gives me enough leisure time", (viii) "... is enjoyable." (see Online Appendix O3 for exact wording). In the control group, female respondents consider reconciliation with children and the possibility to work part-time more important than males (see Online Appendix Table O7), which is in line with Wiswall and Zafar (2018)'s finding on higher willingness to pay for work flexibility among females. Regressing the importance of job attributes on the pooled treatments reveals that the social norm decreases the importance that females assign to the reconciliation of job and children.[42] While we would expect the treatments to increase and not decrease the

[40] Beliefs about the peer groups' views as well as the groups' importance were elicited after treatment administration. We do not find any treatment effects on stated importance of the different peer groups (results available upon request).

[41] In the regressions, we again computed an index of the importance assigned to the different peer groups using the procedure by Kling et al. (2007).

[42] For the regressions, we z-standardized the five-point scale outcomes.

importance of this factor if it makes girls more likely to expect working part-time, it can be rationalized by the fact that females downward-adjust their fertility expectations in response to the treatment, which in turn makes reconciliation of work and children less important (results available upon request). The regression results in Online Appendix Table O7 suggest that social-norm considerations have limited overall effects on labor-market preferences.

Online Appendix O2: Figures and Tables

Table O1: Treatment Effects on Labor-Supply Expectations among Girls: All Answer Categories

	(1) 0 hours	(2) 10 hours	(3) 20 hours	(4) 30 hours	(5) 40 hours
Panel A: Self-expected labor supply					
Treatments (pooled)	0.034***	0.042***	0.035***	-0.068***	-0.042***
	(0.008)	(0.010)	(0.009)	(0.015)	(0.010)
Salience	0.026***	0.032***	0.027***	-0.052***	-0.032***
	(0.010)	(0.012)	(0.010)	(0.019)	(0.012)
Information	0.032***	0.040***	0.033***	-0.065***	-0.041***
	(0.010)	(0.012)	(0.010)	(0.019)	(0.012)
Salience & information	0.042***	0.052***	0.043***	-0.085***	-0.053***
	(0.010)	(0.012)	(0.011)	(0.019)	(0.013)
Control mean	0.04	0.07	0.49	0.29	0.12
Covariates	Yes	Yes	Yes	Yes	Yes
Observations	1164	1164	1164	1164	1164
Panel B: Expected labor supply for partner					
Treatments (pooled)	0.001	0.003	0.011	0.005	-0.019
	(0.002)	(0.004)	(0.015)	(0.006)	(0.028)
Salience	-0.001	-0.002	-0.005	-0.002	0.010
	(0.002)	(0.005)	(0.019)	(0.008)	(0.034)
Information	0.002	0.004	0.014	0.006	-0.026
	(0.002)	(0.005)	(0.018)	(0.008)	(0.034)
Salience & information	0.003	0.006	0.022	0.009	-0.040
	(0.002)	(0.005)	(0.019)	(0.008)	(0.034)
Control mean	0.00	0.03	0.17	0.40	0.39
Covariates	Yes	Yes	Yes	Yes	Yes
Observations	1164	1164	1164	1164	1164

Notes: Results from an ordered probit model. The table reports the average marginal treatment effects. Treatment (pooled): Respondents in experimental groups *salience*, *information* and *salience and information*. Dependent variable is the answer to the question: *"Imagine you are 30 years old and you have a child aged between 1 and 6 years with your partner. What do you think, how many hours per week on average would you like to work in order to earn money?"* (panel A) or *"And how many hours per week on average would you like your partner to work in order to earn money?"* (panel B). Control mean: Mean of the outcome variable in the control group. See Table 1 for included covariates. Sample: Girls. Robust standard errors in parentheses. ***/**/* indicate significance at the 1%/5%/10% level.

Table O2: Treatment Effects on Labor-Supply Expectations among Boys: All Answer Categories

	(1) 0 hours	(2) 10 hours	(3) 20 hours	(4) 30 hours	(5) 40 hours
Panel A: Self-expected labor supply					
Treatments (pooled)	0.008**	0.013**	0.053***	0.005	-0.081***
	(0.004)	(0.005)	(0.019)	(0.004)	(0.028)
Salience	0.006	0.010	0.040*	0.004	-0.060*
	(0.004)	(0.006)	(0.024)	(0.004)	(0.036)
Information	0.007*	0.011*	0.043*	0.004	-0.066*
	(0.004)	(0.006)	(0.024)	(0.004)	(0.036)
Salience & information	0.012**	0.018***	0.072***	0.007	-0.109***
	(0.005)	(0.007)	(0.023)	(0.005)	(0.034)
Control mean	0.01	0.03	0.16	0.45	0.35
Covariates	Yes	Yes	Yes	Yes	Yes
Observations	836	836	836	836	836
Panel B: Expected labor supply for partner					
Treatments (pooled)	0.007	0.006	0.000	-0.008	-0.006
	(0.014)	(0.013)	(0.001)	(0.016)	(0.012)
Salience	0.002	0.001	0.000	-0.002	-0.001
	(0.018)	(0.016)	(0.001)	(0.020)	(0.015)
Information	-0.010	-0.009	-0.001	0.011	0.008
	(0.018)	(0.016)	(0.001)	(0.020)	(0.015)
Salience & information	0.027	0.023	0.002	-0.030	-0.022
	(0.017)	(0.015)	(0.002)	(0.019)	(0.014)
Control mean	0.11	0.17	0.39	0.23	0.09
Covariates	Yes	Yes	Yes	Yes	Yes
Observations	836	836	836	836	836

Notes: Results from an ordered probit model. The table reports the average marginal treatment effects. Treatment (pooled): Respondents in experimental groups *salience*, *information* and *salience and information*. Dependent variable is the answer to the question: *"Imagine you are 30 years old and you have a child aged between 1 and 6 years with your partner. What do you think, how many hours per week on average would you like to work in order to earn money?"* (panel A) or *"And how many hours per week on average would you like your partner to work in order to earn money?"* (panel B). Control mean: Mean of the outcome variable in the control group. See Table 1 for included covariates. Sample: Boys. Robust standard errors in parentheses. ***/**/* indicate significance at the 1%/5%/10% level.

Table O3: Heterogeneous Treatment Effects on Self-Expected Labor Supply by Sociodemographic Characteristics

	Self-expected labor supply		
	(1) All	(2) Girls	(3) Boys
Region			
Treatments (pooled)	-2.108***	-2.425***	-1.769**
	(0.504)	(0.683)	(0.735)
Treatments (pooled) x living in east Germany	-0.392	-0.952	-0.214
	(1.057)	(1.384)	(1.552)
City size			
Treatments (pooled)	-1.301	-2.167**	0.080
	(0.805)	(1.067)	(1.121)
Treatments (pooled) x city size $\geq$ 100,000	-1.295	-0.644	-2.803**
	(0.964)	(1.292)	(1.376)
Age			
Treatments (pooled)	-2.421***	-3.078***	-1.660
	(0.708)	(0.944)	(1.019)
Treatments (pooled) x aged 16 or 17	0.402	0.801	-0.269
	(0.909)	(1.228)	(1.319)
Educational attainment			
Treatments (pooled)	-2.771***	-2.830**	-2.538**
	(0.855)	(1.232)	(1.144)
Treatments (pooled) x school to uni degree	0.684	0.068	1.115
	(1.008)	(1.419)	(1.416)
Mothers' employment			
Treatments (pooled)	-2.510***	-2.775***	-2.202**
	(0.595)	(0.820)	(0.858)
Treatments (pooled) x mother w. full-time	0.765	0.381	0.946
	(0.889)	(1.199)	(1.303)
Parental education			
Treatments (pooled)	-2.774***	-2.924***	-2.445***
	(0.561)	(0.745)	(0.829)
Treatments (pooled) x parents w. uni degree	1.562*	0.899	1.546
	(0.915)	(1.267)	(1.320)
Recruitment			
Treatments (pooled)	-1.518***	-1.583**	-1.777**
	(0.540)	(0.680)	(0.883)
Treatments (pooled) x recruited via parents	-1.654*	-2.796**	-0.194
	(0.939)	(1.357)	(1.308)

Notes: OLS regressions. Treatments (pooled): Respondent in experimental groups *salience*, *information* or *salience and information*. Dependent variable: Hours expected to work per week when having a child 1-6 (0=0 hours, i.e. not at all; 10= 10 hours; 20= 20 hours, i.e. part-time; 30=30 hours, 40=40 hours, i.e. full-time). Living in east Germany/ city size / aged 16 or 17 / school to uni degree / mother w. full-time / parents w. uni degree / recruited via parents: Respondent belongs to respective subgroup. See Table 1 for included covariates. Robust standard errors in parentheses. ***/**/* indicate significance at the 1%/5%/10% level.

Table O4: Treatment Effects on Self-Expected Labor Supply in the ifo Education Survey

	Self-expected labor supply
	(1)
Panel A: Girls	
Salience & information	-3.114***
	(0.814)
p-values: MHT Correction	
Treatments (pooled)	0.000
Control mean	23.75
Observations	553
R-squared	0.074
Panel B: Boys	
Salience & information	-2.179***
	(0.764)
p-values: MHT Correction	
Treatments (pooled)	0.003
Control mean	31.52
Observations	532
R-squared	0.062
Panel C: All	
Salience & information	-2.566***
	(0.558)
Control mean	27.39
Observations	1085
R-squared	0.200
Covariates	Yes

Notes: OLS regressions. *Salience & information*: Respondent in respective treatment group. Dependent variable: Hours expected to work per week when having a child 1-6 (0=0 hours, i.e. not at all; 10= 10 hours; 20= 20 hours, i.e. part-time; 30=30 hours, 40=40 hours, i.e. full-time). Control mean: Mean of the outcome variable in the control group. See Table 1 for included covariates. MHT Correction refers to the multiple hypothesis testing procedure presented in List et al. (2019) and corrects for multiple subgroups (girls and boys) in Panel A and B. Sample: 2018 survey participants. Robust standard errors in parentheses. ***/**/* indicate significance at the 1%/5%/10% level.

Table O5: Treatment Effects on Perceived Peer Pressure

	Index: Peers' opinion about labor supply of mothers	Index: Peers' opinion about labor supply of fathers
	(1)	(2)
Panel A: Girls		
Treatments (pooled)	0.133**	0.042
	(0.065)	(0.063)
p-values: MHT Correction		
Treatments (pooled)	0.163	0.873
Covariates	Yes	Yes
Observations	1163	1163
R-squared	0.052	0.033
Panel B: Boys		
Treatments (pooled)	0.032	-0.022
	(0.073)	(0.077)
p-values: MHT Correction		
Treatments (pooled)	0.874	0.772
Covariates	Yes	Yes
Observations	836	835
R-squared	0.051	0.031

Notes: OLS regressions. Treatments (pooled): Respondent in experimental groups *salience*, *information* or *salience and information*. Dependent variables: (1) Index summarizing respondents' belief about opinion of parents, friends and (future) partner on appropriate labor market supply for mothers following the methodology in Kling et al. (2007). (2) Index summarizing respondents' belief about opinion of parents, friends and (future) partner on appropriate labor market supply for fathers following the methodology in Kling et al. (2007). Results (not shown) from full interaction model between gender and treatment indicators do not reveal any heterogeneous treatment effects by gender. See Table 1 for included covariates. MHT Correction refers to the multiple hypothesis testing procedure presented in List et al. (2019) and corrects for multiple subgroups (girls and boys) and multiple outcomes (all 2 outcomes listed). Robust standard errors in parentheses. ***/**/* indicate significance at the 1%/5%/10% level.

Table O6: Heterogeneous Treatment Effects on Self-Expected Labor Supply by Respondents' Importance to Conform to Peers' Expectations

	Self-expected labor supply	
	(1) Girls	(2) Boys
Treatments (pooled)	-2.702***	-1.839***
	(0.599)	(0.653)
Treatments (pooled) x conformity index	-1.170*	0.555
	(0.608)	(0.687)
Conformity index	0.967*	-0.378
	(0.527)	(0.562)
Covariates	Yes	Yes
Observations	1164	836
R-squared	0.102	0.039

Notes: OLS regressions. Treatments (pooled): Respondent in experimental groups *salience*, *information* or *salience and information*. Dependent Variable: Hours expected to work per week when having a child 1-6 (0=0 hours, i.e. not at all; 10= 10 hours; 20= 20 hours, i.e. part-time; 30=30 hours, 40=40 hours, i.e. full-time). Conformity index: Index summarizing respondents' stated importance to conform to expectations of parents, friends and (future) partner following the methodology in Kling et al. (2007). See Table 1 for included covariates. Robust standard errors in parentheses. ***/**/* indicate significance at the 1%/5%/10% level.

Table O7: Treatment Effects on Preferences for Job Attributes

	Reconcil. with children	Enables part-time work	High salary	Good career opport.	Job security	Leisure time	Enjoyable	Challenge
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Girls								
Treatments (pooled)	-0.152**	-0.027	0.022	-0.011	0.003	-0.040	0.066	0.026
	(0.066)	(0.062)	(0.063)	(0.063)	(0.058)	(0.064)	(0.059)	(0.066)
p-values: MHT Correction								
Treatments (pooled)	0.287	1.000	1.000	1.000	0.999	1.000	0.983	1.000
Control importance	0.83	0.72	0.91	0.83	0.96	0.91	0.96	0.74
Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1164	1164	1163	1164	1164	1164	1164	1164
R-squared	0.040	0.033	0.022	0.068	0.030	0.023	0.021	0.038
Panel B: Boys								
Treatments (pooled)	-0.037	-0.031	0.037	0.055	-0.008	-0.012	0.013	-0.001
	(0.078)	(0.079)	(0.074)	(0.073)	(0.074)	(0.079)	(0.073)	(0.075)
p-values: MHT Correction								
Treatments (pooled)	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.993
Control importance	0.76	0.53	0.91	0.83	0.95	0.88	0.95	0.78
Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	836	835	836	834	835	836	836	835
R-squared	0.031	0.031	0.020	0.043	0.031	0.029	0.024	0.042

Notes: OLS regressions. Treatments (pooled): Respondent in experimental groups *salience*, *information* or *salience and information*. Dependent variables: (1)-(8) Respondents' stated importance of respective job attribute on a 5 point-scale, standardized (the higher the value, the more important the respective job preference). Results (not shown) from full interaction model between gender and treatment indicators do not reveal any heterogeneous treatment effects by gender. Control importance: Share of respondents in the control group reporting respective job preference to be (very) important. See Table 1 for included covariates. MHT Correction refers to the multiple hypothesis testing procedure presented in List et al. (2019) and corrects for multiple subgroups (girls and boys) and multiple outcomes (all 8 outcomes listed). Robust standard errors in parentheses. ***/**/* indicate significance at the 1%/5%/10% level.

Online Appendix O3: Survey Items

Treatments

Salience

What do you think, how many adults in Germany hold the opinion that mothers and fathers, respectively, should reduce their labor-market supply while the children are young? We do not think of the first months after child birth, but the time thereafter. Now, imagine 100 adults in Germany and indicate, what you think, how many adults agree with this statement. ('0' means 'nobody' and '100' means 'everybody', with the numbers in between you can scale your answer).
Answers: (i) XX adults hold the opinion that mothers should reduce their labor-market supply while the children are young; (ii) XX adults hold the opinion that fathers should reduce their labor-market supply while the children are young.

Information

Out of 100 adults in Germany, 91 hold the opinion that the mother should reduce her labor-market supply while the children are young. At the same time, out of 100 adults in Germany, 41 hold the opinion that the father should reduce his labor-market supply while the children are young.

Salience and Information

What do you think, how many adults in Germany hold the opinion that mothers and fathers, respectively, should reduce their labor-market supply while the children are young? We do not think of the first months after child birth, but the time thereafter. Now, imagine 100 adults in Germany and indicate, what you think, how many adults agree to this statement. ('0' means 'nobody' and '100' means 'everybody', with the numbers in between you can scale your answer).
Answers: XX adults hold the opinion that mothers should reduce their labor-market supply while the children are young. XX adults hold the opinion that fathers should reduce their labor-market supply while the children are young.
Out of 100 adults in Germany, 91 hold the opinion that the mother should reduce her labor-market supply while the children are young. At the same time, out of 100 adults in Germany, 41 hold the opinion that the father should reduce his labor-market supply while the children are young.

Egalitarian Information

Out of 100 adults in Germany, 89 hold the opinion that men should take as much responsibility for the home and the children as women.

Salience and Egalitarian Information

What do you think, how many adults in Germany agree with the following statement? Men should take as much responsibility for the home and the children as women. Now, imagine 100 adults in Germany and indicate, what you think, how many adults agree to this statement. ('0' means 'nobody' and '100' means 'everybody', with the numbers in between you can scale your answer).
Answers: XX adults agree with following statement that men should take as much responsibility for the home and the children as women.

Outcome Questions

Labor-Supply Expectations (with Child)

Imagine you are 30 years old and you have a child aged between 1 and 6 years. What do you think, how many hours per week on average would you like to work in order to earn money?
Answers: 0 hours, i.e. not at all; about 10 hours, about 20 hours; about 30 hours; about 40 hours, i.e. full-time
And how many hours per week on average would you like your partner to work in order to earn money?
Answers: 0 hours, i.e. not at all; about 10 hours, about 20 hours; about 30 hours; about 40 hours, i.e. full-time

Labor-Supply Expectations without Child

Now, imagine you are 30 years old, you live together with your partner and you have no children. What do you think, how many hours per week on average would you like to work in order to earn money?
Answers: 0 hours, i.e. not at all; about 10 hours, about 20 hours; about 30 hours; about 40 hours, i.e. full-time
And how many hours per week on average would you like your partner to work in order to earn money?
Answers: 0 hours, i.e. not at all; about 10 hours, about 20 hours; about 30 hours; about 40 hours, i.e. full-time

Incentivized Outcomes

Screen 1: On the next page, we will pose four guess questions. If your guesses to these questions are accurate, you will receive an additional remuneration of 2 Euros. If you receive an additional remuneration due to a correct guess, we will refund the amount in the course of the next 4 weeks after the data have been analyzed. For further information click here. [If click: After completion of the survey we will randomly select one of the four questions. If your guess to this randomly selected question is accurate, you will receive the additional remuneration. Your chances for the remuneration are highest if you give your best guess to all of the four questions.

Screen 2: If your guesses to these questions are accurate, you will receive an additional remuneration of 2 Euros. What do you think, how many adults in Germany agree with the following statements? Image 100 adults in Germany and indicate, what you think, how many adults hold this opinion (0 means nobody and 100 means all; with the numbers in between you can scale your answers)

Answers: (i) XX of 100 adults hold the opinion that a university education is more important for a boy than for a girl; (ii) XX of 100 adults hold the opinion that when the mother works for pay, the children suffer; (iii) XX of 100 adults hold the opinion that being a housewife is just as fulfilling as working for pay; (iv) XX of 100 adults hold the opinion that if a woman earns more than her husband, it is almost certain to cause problems.

Perceived Peer Pressure

What do you think, which opinion do the following groups hold regarding the labor supply of mothers and fathers, respectively? My parents hold the opinion that (i)... mothers whose children are aged between 1 and 6 years hold should reduce their labor supply, (ii)... fathers whose children are aged between 1 and 6 years old should reduce their labor supply; My friends hold the opinion that (i)... mothers whose children are aged between 1 and 6 years old should reduce their labor supply, (ii)... fathers whose children are aged between 1 and 6 years old should reduce their labor supply; My (future) partner would hold the opinion that (i)... mothers whose children are aged between 1 and 6 years old should reduce their labor supply, (ii)... fathers whose children are aged between 1 and 6 years old should reduce their labor supply.

Answers: holds/hold in any case the opinion; holds/hold rather the opinion; holds/hold rather not the opinion; holds/hold in no case the opinion; neither nor

Importance of Peer Group Opinion

How important is it to you to meet the expectations of the following groups? (i)... my family (e.g., parents, siblings); (ii)... my friends: (iii)... my (future) partner; (iv)... the society in general

Answers: very important; rather important; rather unimportant; very unimportant; neither nor

Preferences for Job Attributes

Now, we turn to the job that you would like to work for later on. For the choice of job, different things can play a role. How important are the following things for you? [Order of items randomized] (i) The job can be reconciled with children; (ii) The job enables part-time employment; (iii) The job offers a high salary; (iv) The job offers good career opportunities; (v) The job offers job security (no unemployment); (vi) The job is challenging; (vii) The job gives me enough leisure time; (viii) The job is enjoyable.

Answers: very important; rather important, rather unimportant; very unimportant; neither nor