IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17370

Dividing Housework between Partners: Individual Preferences and Social Norms

Danilo Cavapozzi
Marco Francesconi
Cheti Nicoletti

OCTOBER 2024

IZA – Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17370

Dividing Housework between Partners: Individual Preferences and Social Norms

Danilo Cavapozzi
University of Venice

Marco Francesconi
University of Essex, CEPR, IZA and CESifo

Cheti Nicoletti
University of York, ISER, University of Essex and IZA

OCTOBER 2024

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Dividing Housework between Partners: Individual Preferences and Social Norms*

Using UK longitudinal data on dual-earner couples, this paper estimates a model of intrahousehold housework decisions, which combines a randomized experimental framework eliciting counterfactual choices with gender norms differences across ethnicities and cohorts to identify the impacts of individual preferences and gender identity norms. Equal sharing of tasks yields greater utility for both men and women, with women disliking domestic chores as much as men. Although couples would want to use housework arrangements to compensate for differentials in labor market involvement, women end up performing a substantially larger share of housework. This is not due to specialization, rather social norms play a key role. Exposure to more egalitarian gender attitudes significantly increases the probability of choosing an equal share of housework. Were attitudes evened up to the most progressive levels observed in the sample, women doing more housework than their partners would stop to be the norm already among present-day households, except for households with children.

JEL Classification: C25, C26, D13, J16, J22

Keywords: intrahousehold allocation of chores, labor supply, vignettes, gender identity norms, gender gaps

Corresponding author:
Marco Francesconi
Department of Economics
University of Essex
Wivenhoe Park
Colchester CO4 3SQ
United Kingdom
E-mail: mfranc@essex.ac.uk

* We are grateful to Sonia Bhalotra, Alberto Bisin, Alison Booth, Thomas Cornelissen, Adeline Delavande, Libertad Gonzales, Paul Hufe, Andrea Ichino, Gizem Koşar, Shelly Lundberg, Julie Moschion, Giacomo Pasini, Bob Pollak, Jean-Marc Robin, Kjell Salvanes, Emma Tominey, Wilbert van der Klaauw, Christine Valente, Alessandra Voena, Stephanie von Hinke, and Simon Weber for their constructive comments. Seminar and workshop participants at several universities provided useful suggestions.

1 Introduction

Despite the enormous increase in female labor force participation observed in many advanced economies since the end of World War II (e.g., Goldin, 2014; Olivetti and Petrongolo, 2016; Greenwood et al., 2024), the sexual division of labor within households remains distinctly gendered, with women shouldering most of the housework, even when their shares of market work and earnings are equal or higher than their partners' (Gershuny, 2000; Bittman et al., 2003; Couprie, 2007; Bertrand et al., 2015; Goussé et al., 2017). The aim of this paper is to provide new evidence on the roles played by gender identity norms and individual preferences to explain intrahousehold differences in the amount of time spent in housework.

Dual-earner couples in today's Britain devote an average of 19 hours to domestic chores every week, about half the time spent in a standard paid full-time job, with more than 70% of women pulling off the largest share of this housework.[1] At the same time, in more than half of the same households, women are involved in market work at least as much as their partners. These figures imply substantial gender inequalities in work-family balance, which may have wide-reaching consequences. Much research emphasizes that, if women bear the greatest burden of household chores and other unpaid work and care, they are likely to experience poorer career opportunities which in turn hinder gender pay equality (e.g., Sullivan, 2019; Cortés and Pan, 2019; Ciasullo and Uccioli, 2024). Growing evidence also documents that the recent COVID-19 crisis has led to a widening of the gender gap in housework, exacerbating women's stress and anxiety (e.g., Seedat and Rondon, 2021). Besides issues of fairness, there is also an efficiency argument at play, according to which a greater gender equality in the labor market (and, by extension, in the division of domestic chores) can improve economic growth (Hsieh et al., 2019; Bertrand, 2020; Boerma and Karabarbounis, 2021).

Becker's classic theory of time allocation explains women's specialization in housework as a result of gender differences in the returns to market work or in the productivity of unpaid domestic activities (Becker, 1965, 1981, 1985). More recent studies, which use collective rather than unitary household models, also lead to a specialization result under the assumption that men's and women's time inputs in housework are perfect substitutes (see, among others, Chiappori, 1997; Pollak, 2013; Chiappori and Lewbel, 2015). Such theories, however, cannot be invoked to justify a traditional allocation of domestic work with women doing most of it and when, at the same time, they earn more or work longer hours in the

[1] Similar figures emerge in other countries, such as the United States, France, Germany, Italy, and Spain (Kimmel and Connelly, 2007; Cardoso et al., 2010; Álvarez and Miles-Touya, 2019; Del Boca et al., 2020). Housework is defined as domestic chores, or core nonmarket work, which include cooking, cleaning, doing the laundry and grocery, shopping, gardening, house related DIY jobs, and ensuring that household bills are paid. It excludes childcare.

market, unless they enjoy an intrinsically greater utility flow from house-working. In our empirical analysis, we find no support for this preference-driven explanation.[2]

An alternative explanation focuses on gender role attitudes. A vast economic literature shows that culture — broadly defined to account for beliefs, values, attitudes and social norms — affects a wide range of economic behaviors, including women's labor market outcomes (Fernández et al., 2004; Fortin, 2005; Guiso et al., 2006; Fernández and Fogli, 2009; Maurin and Moschion, 2009; Bisin and Verdier, 2011; Fernández, 2011; Alesina et al., 2013; Alesina and Giuliano, 2014; Nicoletti et al., 2018; Olivetti et al., 2020; Bertrand et al., 2021; Bursztyn et al., 2024; Doepke et al., 2024; Bau and Fernández, 2024; Albrecht et al., 2024). The link between gender identity norms and couples' division of domestic activities, however, has received much less attention.[3]

Four important exceptions are worth stressing. Bertrand et al. (2015) find that women who are more successful than their partners in the labor market pay for this success at home in order to offset the reversal of the traditional gender roles. Our contribution is to endogenize gender norms and assess their role in conjunction with that played by individual preferences over housework arrangements. Using the estimates obtained from a search model of marriage, Goussé et al. (2017) show that family values have strong effects on home production time inputs which differ across households, with traditional attitudes inducing women to spend more time in home production, and the opposite for men, and liberal (or more egalitarian) attitudes implying an erosion of women's comparative advantage in home production. We use a different model in which, unlike Goussé et al. (2017), we allow family values by gender to be endogenous to housework decisions. We also employ an alternative strategy to identify individual preferences based on a randomized stated preference experiment.

Siminski and Yetsenga (2022) use Australian time use data to test the role played by comparative advantage in explaining intrahousehold allocations of chores. They find that, in couples, women do more domestic work than their male counterparts at every point of the relative wage distribution, and they show that comparative advantage plays little or no role. However, albeit invoked as a main driver of the sexual division of labor within households, gender norms are not explicitly modeled in their framework, but just used to interpret the results. This is instead key in our paper.

[2] Lise and Yamada (2019) do not impose perfect substitutability in home production and find evidence that the woman's utility weight is strongly related to her relative labor market productivity. Their application, which does not account for social norms, relies on data from Japan, a clear case of the male breadwinner model, whereby women's market hours and wages are disproportionately lower than their partners'.

[3] The idea that individuals modify their behavior in such a way as to satisfy gender roles has long been discussed by sociologists. See, for example, Mott and Moore (1979), Fenstermaker Berk (1985), Hochschild and Machung (1989), Blair and Johnson (1992), Greenstein (1995), Gershuny (2000), and Bittman et al. (2003).

The final study by Ichino et al. (2024) proposes a test of gender identity norms based on the response of husbands' and wives' home production time, proxied by temporary parental leave, to changes in their post-tax wages, which change the cost of complying with social norms regarding the division of household tasks. They find that immigrant, married and male breadwinner couples respond more strongly to tax changes that induce a traditional allocation of spouses' time, while the corresponding counterpart households react more strongly to tax changes that imply a more egalitarian division of housework. Unlike this study, our analysis relies on the direct measurement of gender values, rather than inferring them from the observed allocations of time devoted to the market and to domestic activities, and uses a clearer definition of home production, rather than approximating it with the take-up of leave to care for a sick child during working hours.

Our aim is to understand why women tend to do more than their fair share of housework than men. A gendered division of chores does not raise equity concerns if it could be explained by women having preferences for doing more housework. However, if a greater burden of housework on women's shoulders were driven by gender norms, independently of own preferences, concerns about fairness and efficiency would be legitimate. This simple insight motivates our analysis, in which we attempt to assess the separate roles played by gender norms and by preferences. Of course, preferences and social norms are constantly shaping each other and identification of a clear-cut distinction between these two pathways is challenging.

In this paper, we formulate and estimate an empirical model of the sexual division of labor within households in which couples decide over three different divisions of domestic activities, one in which the woman does less than her partner, one in which she does more, and another in which the two share equally the amount of housework. As mentioned, the two key ingredients of the model are preferences and gender role norms. With regards to norms, the model implicitly takes account of the disutility associated with a loss of gender identity or of the utility gain associated with conformity (Akerlof, 1997; Akerlof and Kranton, 2000, 2010). Social norms about the role of women in society and inside the family shape a specific division of housework between partners. Conforming to such norms boosts utility, while deviating from them imposes psychic costs to family members. For each couple in our analysis, social norms are measured by the average gender role attitudes shared by the peers of each of the two partners, with peers defined over sex, birth cohort, and ethnicity (Guiso et al., 2006; Fernández, 2007; Fernández and Fogli, 2009; Desmet et al., 2017).

Acknowledging their endogeneity, we instrument peers' gender role attitudes leveraging variation in the labor force participation of peers' mothers by gender, cohort, and ethnicity. Our identification strategy is closely related to the *epidemiological* approach introduced by

Fernández et al. (2004) and further developed in several other studies (e.g. Fernández and Fogli, 2006; Fernández, 2007; Giuliano, 2007; Fernández and Fogli, 2009; Fernández, 2011; Alesina and Giuliano, 2010; Nollenberger et al., 2016; Holmlund et al., 2023; Boelmann et al., 2024). Furthermore, because individuals from specific ethnicities may sort into areas with characteristics that could affect both labor market and housework decisions, we estimate our first stage equation by using a household fixed effect estimation, which exploits only within couple variation netting out the impact of factors that are invariant between partners, such as area characteristics. We implement this estimation using a control function approach, i.e. including in the model the estimated error terms from the first stage equation. Finally, because our instruments may be correlated with aggregate trends in female labor supply, we control for cohort fixed effects and identify the impact of gender norms through the large variation in peers' labor force participation trends across ethnicity.[4]

Preferences, the other key ingredient of the model, are hard to identify using survey data on satisfaction with housework shared between partners.[5] Our strategy, therefore, is to identify preferences with a randomized vignette-based stated preference experiment conducted within the Innovation Panel of the UK Household Longitudinal Study (UKHLS). In the experiment, people are asked to imagine themselves in different hypothetical scenarios ("vignettes"), each one outlining the sharing of housework and market work between partners, relative gaps in the wage rate, whether there are children present and whether paid help for domestic work is used. The experiment is designed to ensure that scenarios and individual characteristics are randomly matched, and thus uncorrelated to each other. Respondents are then asked to report their satisfaction with the housework arrangement described in each of the hypothetical scenarios. The experiment hinges on generating a large variety of scenarios and randomly allocating them to people, so that there is no correlation between the variables describing each scenario and the respondent's characteristics in real life.

Such design is close to the stated choice experiments with elicited probabilities that were introduced by Blass et al. (2010), and more recently have been used by Wiswall and Zafar (2018), Delavande and Zafar (2019) and Koşar et al. (2022). We use our experimental data

[4]A different perspective is that gender norms are part of individuals' own preferences and become internalized or subsumed into them (see the engaging discussion by Camerer and Fehr, 2004). Our analysis instead makes a clear distinction between the two and identifies them separately. To give an example, a man could have a penchant for cooking and shopping and may end up doing a larger share of housework than his female partner, even though social norms impart that cooking and shopping should be done by women.

[5]There are at least three simple reasons for this. First, satisfaction with real-life housework share can be the result of an ex-post rationalization process. Second, in most of the existing surveys even in advanced economies, it is rare to observe women doing less housework, earning more, and doing more paid work than their male partners. It is thus difficult to reliably assess preferences for such allocations, as it is for all other uncommon situations. Third, controlling for all unobserved characteristics that affect both housework decisions and couples' preferences is implausible.

to back out individual utilities for couples in real life, and include them directly into the model of intrahousehold chore allocation decisions.

Using data on dual-earner couples drawn from the main survey of the UKHLS, we find that both preferences and gender identity norms play a substantial role in explaining the observed division of home production activities between male and female partners. On the preferences side, an equal distribution of tasks yields higher levels of utility for both men and women. Couples tend to use housework arrangements to compensate for imbalances in labor market involvement, with partners — regardless of gender — attaining higher utility levels if they share domestic chores equally when they are equally involved in paid employment and, conversely, if one works more at home when the other works more in the market. Women dislike domestic chores as much as (or even more than) men do, suggesting that the pronounced intrahousehold imbalance in the observed allocation of chores cannot be driven by an innate female preference for housework.

On the norms front, we find that peers' gender role attitudes of the woman matter more than the ones of the male partner. Put differently, it is female gender norms that shape households' decisions on chores allocations. Focusing on her peers, then, we find that exposure to more egalitarian gender attitudes significantly elevates the probability of choosing an equal share of housework. This impact is particularly strong for highly educated couples and households with children. Should gender identity norms continue to grow towards equality in the future as we observe in the data, a fair allocation of housework that mirrors the partners' distribution of labor market involvement could materialize in about 50 years, keeping everything else unchanged. If instead people's gender roles moved now to the most progressive values recorded in our sample, a higher share of housework for women would stop being the norm among present-day households, except for women with children.

Not only do our results contribute to each of the strands of the economic literature mentioned above, they are policy relevant. As many interventions are meant to affect labor market participation differently by gender, for instance through taxation, benefit provision, and childcare support, we are likely to witness a reproduction of gender imbalances unless such policies are accompanied by changes in social norms and individual preferences over gender identity roles. For instance, there is empirical evidence that being exposed to counter-stereotypical parental behaviors in childhood leads to more egalitarian attitudes and behaviors once adult (e.g., Fernández et al., 2004; Bertrand, 2019; Farré et al., 2023). There is more work to do, however, to understand how gender attitudes are shaped (e.g., Becker and Mulligan, 1997; Bowles, 1998; La Ferrara et al., 2012; Carlana, 2019; Porter and Serra, 2019; Dahl et al., 2021; Dhar et al., 2022). This is an important area for future research.

2 An Empirical Model of Housework Decisions

We consider a dual-earner household, i, in which partners decide about housework arrangements. They choose among three discrete mutually exclusive alternatives, j: (a) the man does more housework than the woman ($j = 1$); (b) the woman and the man do the same amount of housework ($j = 2$); or (c) the woman does more housework than the man ($j = 3$). For each household i and alternative $j = \{1, 2, 3\}$, the utility is given by

$$U_{i,j} = V_{i,j} + \epsilon_{i,j}, \tag{1}$$

where $V_{i,j}$ is a deterministic component associated with choice j and $\epsilon_{i,j}$ is a random component which is known to the couple but otherwise unobserved. As standard in an additive random utility framework, the couple chooses the alternative which maximizes its utility, i.e., $j = \text{argmax}_j U_{i,j}$.[6]

Letting y_i denote a variable taking value 1 if the man does more housework, 2 if both partners do (approximately) the same amount of housework, and 3 if the woman does more housework, we can compute the probability that household i chooses alternative j as

$$Pr(y_i = j) = Pr(U_{i,j} \geq U_{i,s}, \text{ for } \ s \neq j) = Pr(\epsilon_{i,s} - \epsilon_{i,j} \leq V_{i,j} - V_{i,s}, \text{ for } \ s \neq j). \tag{2}$$

Assuming that the error component $\epsilon_{i,j}$ follows an i.i.d. Gumbel (type 1 extreme value) distribution with density $f(\epsilon_{i,j}) = exp(-\epsilon_{i,j})exp[-exp(-\epsilon_{i,j})]$, we can derive a multinomial logit model for the discrete choice variable y_i by using maximum likelihood estimation once a functional form for the deterministic component of the utility, $V_{i,j}$, is specified. To this purpose, we assume that $V_{i,j}$ be additively separable in two components. Specifically,

$$V_{i,j} = \rho^E V_{i,j}^E(\mathbf{X}_i) + \rho^N V_{i,j}^N(N_{i.f}, N_{i,m}), \tag{3}$$

[6] In our setup, intrahousehold housework allocations could be seen as a *residual* decision within marriage, after labor supply and other domestic decisions, including childcare, are taken. Childcare and housework specialization, however, may not be separable. For this reason, in subsection 6.1, we perform our analysis separately for couples with children and childless couples. We will show that the impact of preferences on housework decisions is similar across these two groups of households, and that gender norms play a relevant role in both types of households. We believe this framework is meaningful and effectual, as it allows us to gain new insights on the role played by social norms, while we can also address endogeneity issues related to individual preferences, which we can identify through the randomized vignette-based stated preference experiment described in detail below. An alternative strategy would be to formulate a fully structural dynamic model in which partners jointly choose marriage, individual labor supply, fertility, and housework allocations (including domestic help and childcare). This would be an ambitious project, which has never been tried before and is left for future research. The closest to this alternative is the paper by Goussé et al. (2017), which estimates a search model of marriage, labor supply, and home production, while keeping fertility decisions exogenous. In their environment, however, family values by gender are taken as given, whereas in ours they are endogenous to the chores decisions.

where $V_{i,j}^E(\cdot)$ and $V_{i,j}^N(\cdot)$ represent the household's economic utility and the gender role norms utility corresponding to housework arrangement j, respectively. The first term varies with the choice of housework arrangement j and with a set of characteristics $\mathbf{X}_i$. In the empirical analysis, this vector is constructed in such a way to reflect the design of our randomized stated preference experiment, and it is comprised of the two partners' labor supply, their relative hourly earnings, and indicator variables for paid domestic help, and presence and age of children. The second term also varies with j and depends on the gender role attitudes that the couple is exposed to. In the analysis below, this is given by the gender role norms expressed by the woman's and the man's peers, denoted by $N_{i,f}$ and $N_{i,m}$, respectively. The parameters ρ^E and ρ^N measure the weights given to the economic and gender identity utilities, respectively.

Specifying $V_{i,j}$ as the sum of these two utility subcomponents allows us to evaluate the separate contributions played by the two main forces behind the sexual division of housework identified in previous studies. The first is the specialization driven by men's potential comparative advantage in the labor market (Becker, 1985; Chiappori and Lewbel, 2015; Siminski and Yetsenga, 2022). The second is attributed to conformity to gender identity norms or social pressure (Akerlof and Kranton, 2000; Bertrand et al., 2015; Goussé et al., 2017).[7]

Economic Utility — This component of preferences reflects the household's economic fundamentals and it is given by

$$V_{i,j}^E(\mathbf{X}_i) = \alpha^m u_{i,j}^m(\mathbf{X}_i) + \alpha^f u_{i,j}^f(\mathbf{X}_i), \tag{4}$$

where the superscripts m and f refer to the male and female partner, respectively; $u_{i,j}^m(\mathbf{X}_i)$ and $u_{i,j}^f(\mathbf{X}_i)$ are the individual specific economic utilities associated with housework arrangement j for the male and female partners, and α^m and α^f are the corresponding weights which can reflect the bargaining power of the man and the woman, both taking positive values between 0 and 1 and so that $\alpha^f = 1 - \alpha^m$. Individual utilities depends on the opportunity cost of time spent in housework, time spent in the labor market, and time for childcare needs, which are captured by the vector $\mathbf{X}_i$. We infer the mapping of $\mathbf{X}_i$ into each individual utility using the randomized vignette-based stated preference experiment that will be described in Section 3.

Gender Identity Utility — The specification of $V_{i,j}^N(N_{i,m}, N_{i,f})$ follows the insights of the conformist model formulated by Akerlof (1997) and refined by Akerlof and Kranton (2000).

[7] This specification abstracts away from possible interactions between such two forces. In the empirical application, however, we consider a sensitivity analysis where peer's gender role norms are allowed to interact with household utility. See subsection 6.2.

In that setting, each household minimizes the social distance between itself and others in terms of its intrahousehold sexual division of labor or, equivalently, partners lose utility should their household arrangements fail to conform to those suggested by social norms.

An individual's gender identity is measured by the average gender role attitudes observed among his/her peers. Peers are defined to be the group of individuals with the same sex, ethnicity, and birth cohort as the reference person. Gender role attitudes are proxied by a composite index derived from a set of questions on beliefs about the role of women in society, which takes higher values for more egalitarian (less traditional) gender role beliefs. Section 3 will provide more details. $N_{i,m}$ and $N_{i,f}$, therefore, are the sum of the average gender role attitudes of the male and female partners' peers. We assume that the gender identity utility is linear in $N_{i,m}$ and $N_{i,f}$ so that

$$V_{i,j}^{N}(N_{i,m}, N_{i,f}) = \gamma_j^m N_{i,m} + \gamma_j^f N_{i,f},$$

where γ_j^m and γ_j^f capture the effect on the gender identity utility of a one-unit increase in the average gender role attitudes shared by his and her peers, respectively.

Total Utility — Piecing together all the preference components into (3) and (1) yields

$$U_{i,j} = \eta^{E,m} u_{i,j}^m(\mathbf{X}_i) + \eta^{E,f} u_{i,j}^f(\mathbf{X}_i) + \eta_j^{MN} N_{i,m} + \eta_j^{FN} N_{i,f} + \epsilon_{i,j}, \tag{5}$$

where $\eta^{E,m} = \rho^E \alpha^m$, $\eta^{E,f} = \rho^E \alpha^f$, $\eta_j^{MN} = \rho^N \gamma_j^m$ and $\eta_j^{FN} = \rho^N \gamma_j^f$. Under the distributional assumption on $\epsilon_{i,j}$ and the probability statements given in (2), we can derive the following likelihood (see Train, 2009):

$$L = \prod_{i=1}^{n} \left[\frac{1}{1 + exp(V_{i,2} - V_{i,1}) + exp(V_{i,3} - V_{i,1})} \right]^{\mathbb{1}(y_i=1)} \left[\frac{exp(V_{i,2} - V_{i,1})}{1 + exp(V_{i,2} - V_{i,1}) + exp(V_{i,3} - V_{i,1})} \right]^{\mathbb{1}(y_i=2)} \left[\frac{exp(V_{i,3} - V_{i,1})}{1 + exp(V_{i,2} - V_{i,1}) + exp(V_{i,3} - V_{i,1})} \right]^{\mathbb{1}(y_i=3)}, \tag{6}$$

where $\mathbb{1}(y_i = j)$ takes value 1 if $y_i = j$ and 0 otherwise, and

$$V_{i,j} - V_{i,1} = \eta^{E,m}[u_{i,j}^m(\mathbf{X}_i) - u_{i,1}^m(\mathbf{X}_i)] + \eta^{E,f}[u_{i,j}^f(\mathbf{X}_i) - u_{i,1}^f(\mathbf{X}_i)] + \widetilde{\eta}_j^{MN} N_{i,m} + \widetilde{\eta}_j^{FN} N_{i,f},$$

with $\widetilde{\eta}_j^{MN} = \eta_j^{MN} - \eta_1^{MN}$ and $\widetilde{\eta}_j^{FN} = \eta_j^{FN} - \eta_1^{FN}$.

Expression (6) is the likelihood of a conditional multinomial logit model that includes both alternative-variant and alternative-invariant explanatory variables (McFadden, 1974; Chamberlain, 1980; McFadden and Train, 2000; Train, 2009). The individual economic utilities, $u^m_{i,j}(\mathbf{X}_i)$ and $u^f_{i,j}(\mathbf{X}_i)$, are alternative variant and allow us to identify the coefficients $\eta^{E,m}$ and $\eta^{E,f}$. Because $\eta^{E,m} = \rho^E\alpha^m$, $\eta^{E,f} = \rho^E\alpha^f$, $\alpha^m + \alpha^f = 1$, and $0 < \alpha^m < 1$, we can identify ρ^E, α^m, and α^f as $\rho^E = \eta^{E,m} + \eta^{E,f}$, $\alpha^m = \frac{\eta^{E,m}}{\rho^E}$, and $\alpha^f = 1 - \alpha^m$.

A result from the empirical analysis is that we cannot reject the equality of the female and male weights, i.e., $\alpha^m = \alpha^f = 0.5$, so that the random utility model boils down to

$$U_{i,j} = \rho^E V^E_{i,j} + \eta^{MN}_j N_{i,m} + \eta^{FN}_j N_{i,f} + \epsilon_{i,j}, \tag{7}$$

where $V^E_{i,j} = 0.5\big(u^m_{i,j}(\mathbf{X}_i) + u^f_{i,j}(\mathbf{X}_i)\big)$. In our empirical work, then, we will directly consider

$$U^E_{i,j} = u^m_{i,j}(\mathbf{X}_i) + u^f_{i,j}(\mathbf{X}_i), \tag{8}$$

which is simply the sum of his and her individual economic utility. Finally, we cannot identify the weight given to the gender identity utility, ρ^N, nor γ^m_j and γ^f_j separately, but we can identify the effect of an increase in the gender role attitudes of peers on the utility associated with alternative j, i.e., $\widetilde{\eta}^{MN}_j$ and $\widetilde{\eta}^{FN}_j$.

3 Data

We use two different samples from the UK Household Longitudinal Study (UKHLS). One is drawn from the main UKHLS survey.[8] This is a nationally representative panel survey of about 40,000 households in the UK collecting yearly information on a wide range of individual and household socioeconomic, health, and demographic characteristics since 2009–10. The other comes from the Innovation Panel (IP) on a nationally representative sample of approximately 1,500 individuals from England, Wales, and Scotland who are interviewed annually on the same set of questions asked in the main survey as well as on a set of additional items used for experimental and statistical purposes.[9] To ensure representativeness, both surveys use a stratified and geographically clustered sample design with computer assisted personal interviewing.[10]

[8] This is publicly available at the University of Essex, Institute for Social and Economic Research, NatCen Social Research and Kantar Public [producers]: Understanding Society: Waves 1-6, 2009-2015 [computer file]. 8th Edition. Colchester, Essex: UK Data Service [distributor], November 2016. SN: 6614, <http://dx.doi.org/10.5255/UKDA-SN-6614-8>.

[9] Northern Ireland, therefore is not covered by the IP sample. See University of Essex, Institute for Social and Economic Research and TNS BMRB, Understanding Society, Innovation Panel, Waves 1-7, 2008-2014 [computer file]. 5th edition. Colchester, Essex: UK Data Archive [distributor], July 2015, SN: 6849, <http://dx.doi.org/10.5255/UKDA-SN-6849-5>.

[10] For a full explanation of the survey methodology and fieldwork results, we refer to the Technical Reports accompanying the public release of the data and online material available at <https://www.understandingsociety.ac.uk/documentation>.

3.1 IP Sample

We construct our IP sample from the fifth and sixth sweeps of data collected in 2012 and 2013. These are the only waves in which the vignette experiment on housework was performed. Our sample is restricted to 25–60 years old respondents who completed the vignette section in at least one of the two waves. After excluding individuals with missing values on the vignette questions, the sample consists of 1,245 individuals (695 women and 550 men). Of these individuals, 482 are interviewed in only one of the two waves and provide up three vignette evaluations, while the other 763 individuals are in both waves and contribute up to six evaluations.

At the start of the experiment, respondents are informed of the key features of the exercise.[11] Regardless of their marital status, respondents are then asked to imagine they are married or in a live-in partnership and to report their level of satisfaction with three hypothetical random scenarios (or vignettes) related to housework arrangements, which are orthogonal to their own actual arrangements. Each scenario deals with five key factors jointly: (a) the share of housework done by the respondent and their partner (with shares equal to zero, one quarter, half, three quarters, or one); (b) each partner's labor market involvement (both work full-time, both work part-time, respondent works full-time and partner works part-time, and respondent works part-time and partner works full-time); (c) relative hourly pay (respondent's hourly pay is half, approximately equal or double that of their partner); (d) the absence or presence of one child and child age (no child, one child aged 6 months, 5 years, or 15 years); and (e) the potential use of paid housework (none or a cleaner comes one morning a week). Each respondent is asked to report the level of satisfaction with the randomly allocated housework arrangements on a 7-point Likert scale, whereby 1=completely dissatisfied, 2=mostly dissatisfied, 3=somewhat dissatisfied, 4=neither satisfied nor dissatisfied, 5=somewhat satisfied, 6=mostly satisfied, and 7=completely satisfied.[12]

[11] Specifically, the preamble states: "We are interested in the way people feel about how couples share housework. You will be asked about three different scenarios, each describing an imaginary situation relating to different work arrangements between couples. In some of the situations the couple may have children. In these cases you should assume that both partners are involved in their day-to-day care, and that both are happy with the amount of time they spend with their children. We are really only interested in knowing how you feel about the way housework is shared. There are no right or wrong answers." To give each participant privacy and enough time to perform the exercise, the experiment is executed with computer-assisted self-interviews.

[12] An example of a vignette question is as follows: "Imagine that you are married or cohabiting, you and your partner both have part time jobs, and your hourly pay is approximately the same as your partner's. You have one child aged 6 months; your partner does three quarters of the housework while you do one quarter of it, and you do not employ anybody to help with the housework. How satisfied would you say you are with the sharing of the housework?".

There are 160 different versions of the questionnaire, each version containing three scenarios for a total of 480 vignettes.[13] We use a D-efficient sampling technique, which minimizes the correlation between the five factors (a)–(e) described above, and maximizes the variance of each of the factors within the questionnaire versions, therefore guaranteeing a "level balance", i.e. ensuring that each category occurs with approximately equal frequency (see Kuhfeld et al., 1994; Auspurg and Hinz, 2014).[14] Vignettes are randomly assigned to respondents so that the five key factors are by design uncorrelated with respondents' observed and unobserved traits and past decisions, including labor supply and wages. Because of this feature, issues of ex-post rationalization are likely to be minimal. Another advantage of this design is that it allows us to assess the utility that men and women derive from any type of housework arrangement, even arrangements that are unlikely to be observed in the data. By eliciting the size of housework contribution and hourly earnings of one partner *relative* to the other partner, concerns about systematic discrepancies between objective and subjective hours and income distributions are also negligible (Cruces et al., 2013).

A few considerations are worth stressing. First, although sharing some features with stated choice experiments (e.g., McFadden and Train 2000), our exercise is different. Stated choice experiments ask people to choose across different alternatives the best one or to rank their preferred choices, but do not allow to assess how much the best option is far from the second best option or the other alternatives. Our experiment, instead, asks people to rate the different alternatives enabling us to distinguish between cases where the preferred option is highly preferred to the other alternatives from cases in which the preferred option is almost equally ranked with the second best alternative. In this sense, our exercise is closer to the stated choice experiments with elicited probabilities that have been recently used, among others, by Wiswall and Zafar (2018), Delavande and Zafar (2019), and Koşar et al. (2022).

Second, our experimental design is intended to capture individual-specific preferences rather than gender-norm-led tastes. Auspurg et al. (2017) use the same exercise but consider only one wave and show that even for gender primed individuals there are no differentials in preferences revealed by men and women in vignette questions, suggesting that the results are *not* driven by people internalizing gender norms into their own preferences. This holds true also in our more general application.

[13] These 480 hypothetical scenarios come from 5 (housework arrangements) × 4 (labor market involvements) × 3 (relative hourly wages) × 4 (child categories) × 2 (domestic help categories).

[14] The D-efficient design allows respondents to face a set of vignette questions that randomly cover a wide set of different potential scenarios. Presenting male and female partners in the same household with identical scenarios ensures that gender differences in evaluations of the scenarios are not caused by differences in experimental stimuli (Auspurg and Hinz, 2014).

Third, another concern could arise if responses are deemed to not reveal actual (true) individual preferences but instead display inclinations that are considered socially acceptable. However, because we ask people to report their level of satisfaction with housework in a set of scenarios with *ever shifting* characteristics that are likely to matter in real life, differences in the satisfaction reported for changing scenarios are less likely to be influenced by social desirability and more likely to reflect actual preferences over the described scenarios (see Auspurg and Hinz, 2014).

Fourth, it is possible that respondents believe that researchers wish to show that couples share housework fairly and for this reason may decide to report preferences which support researchers' beliefs.[15] This implies that responses would not reflect genuine preferences. To induce experimenter demand effects explicitly, both De Quidt et al. (2018) and Mummolo and Peterson (2019) assess this type of potential bias by informing respondents on what researchers expect. They find evidence suggesting that demand effects are modest. Given similar promptings in the preamble to the exercise mentioned above, we rely on this evidence to strengthen confidence in our approach.

3.2 UKHLS Sample and Descriptive Statistics

This sample comes from waves 2 and 4 of the UKHLS (collected in 2010–2012 and 2012–2014, respectively). These are the waves with information on gender role attitudes. The number of respondents in wave 2 and 4 in the UKHLS main survey are 54,597 and 47,157, respectively. We restrict our analysis to married or cohabiting dual-earner couples, in which both partners are 25–60 years old employees and live in England, Wales or Scotland. The final sample consists of 7,621 households-wave observations, 4,068 couples in wave 2 and 3,553 in wave 4. The sample size reduction is almost entirely driven by sample selection (on age, marital status, and employment) and only marginally by nonresponse.

To reflect the characterization of the model in the previous section, our outcome variable distinguishes three mutually-exclusive housework arrangements: (a) women do less housework; (b) women and men equally share housework; and (c) women do most of the housework. The UKHLS asks respondents to report the number of hours per week devoted to chores or core non-market work, which include cooking, cleaning, doing the laundry, grocery shopping, gardening, house-related DIY jobs, and ensuring that household bills are paid (for similar definitions, see Aguiar and Hurst, 2007; Bertrand et al., 2015). A woman is defined to do less housework if the time she spends in chores is less than 45% of the overall time devoted to housework by both partners. She does instead most of the housework if her time

[15] De Quidt et al. (2018), De Quidt et al. (2019), and Mummolo and Peterson (2019) discuss experimenter demand effects as an additional source of bias in experimental work.

spent in chores constitutes more than 55% of the total housework time, while intermediate arrangements are classified as equal share of housework between partners.[16]

Table 1 shows that in 72% of the household-wave observations women do most of the housework. The remaining 28% of the sample is divided between households in which the man does more of the chores and those in which both partners share equally (14% in each arrangement). These allocations are close to the distributions found in other studies using UK data, such as Gershuny (2000), Couprie (2007), Goussé et al. (2017), and Gimenez-Nadal et al. (2018).[17] Considering the actual weekly housework hours distribution in each category, we see that, at the mean, the gender difference is relatively small if we look at either men or women who work less than their partner (4.2 weekly hours for men as opposed to 5.7 hours for women), and those who work approximately the same amount as their partner (9.21 and 9.29 hours for men and women, respectively). We notice a larger gender gap among those who do more housework than their partner, with an average weekly hours of 11.7 and 15.0 for men and women, respectively.

Table 2 breaks down the housework distribution by partners' labor market involvement. Women's specialization in domestic chores is particularly pronounced in households where the man works full time and the woman works part time, which make up more than 46% of the sample. This is the case also in the remaining household types, even when the man works part time and the woman full time, although such couples represent only 3% of the sample.

From the model in Section 2, our key explanatory variables are $N_{i,m}$ and $N_{i,f}$, the mean gender role attitudes of the male and female partners' peers in each household, and the two individual economic utilities for the male and female partners in household i associated with housework arrangement j, $u^m_{i,j}$ and $u^f_{i,j}$, respectively.

To compute $N_{i,m}$ and $N_{i,f}$, we first use a battery of questions designed to elicit well-established measures of gender identity norms (Fernández et al., 2004; Fortin, 2005). Respondents are asked if they agree with the following five claims: (i) "Pre-school child suffers if mother works"; (ii) "Family suffers if mother works full-time"; (iii) "Husband and wife should contribute to household income"; (iv) "Husband should earn, wife should stay at home"; (v) "Employers should help mothers combine jobs and childcare". Respondents' agreement with each of these statements is rated according to a 5-point scale, where 1=strongly agree,

[16] We have performed a series of sensitivity checks with alternative cut-offs for the definition of housework arrangements. We shall discuss them in subsection 6.2, but we anticipate that our main results remain unchanged.

[17] They are also similar in other advanced economies, including the United States, France, Germany, Italy, and Spain (e.g., Kimmel and Connelly, 2007; Aguiar and Hurst, 2007; Zick et al., 2008; Cardoso et al., 2010; Sevilla-Sanz et al., 2010; Álvarez and Miles-Touya, 2019; Del Boca et al., 2020).

2=agree, 3=neither agree nor disagree, 4=disagree, and 5=strongly disagree. We define each individual's gender role index by summing their separate responses across the five questions after inverting the scale for questions (iii) and (v).[18]

The index therefore varies between 5 and 25, with higher values indicating more egalitarian norms between the sexes and lower values capturing more traditional gender role attitudes. We then compute the average of this score across his and her peers separately. Peers are defined as individuals with the same gender, in the same birth cohort, and of the same ethnicity as the respondent.[19] We end up with 136 groups (68 for women and 68 for men) with a minimum of 4 peers per group and an average size of 358 peers for men and 447 for women.

For each household i, $N_{i,m}$ and $N_{i,f}$ are the sum of his and her peers' average gender role attitudes and they are standardized to have a standard deviation (SD) of 1. $N_{i,m}$ and $N_{i,f}$, therefore, capture gender norms shared with peers of the two partners in household i and leverage a great deal of sample variation in culture, broadly defined over gender, birth cohort, and ethnicity. Our homophily measure echoes what earlier research has done to characterize cultural norms shared with peers defined over ethnicity and country of ancestry (e.g., Guiso et al., 2006; Fernández, 2007; Fernández and Fogli, 2009). Cultural beliefs on how housework should be shared between partners may differ by gender and are likely to vary across countries of origin and birth cohorts. As individuals may conform to, or imitate, the behavior of peers who are culturally more similar to them, we consider peers who are homogeneous in terms of sex, ethnicity, and birth cohort.

Table 1 reports a mean value for $N_{i,m}$ of about 26 points and for $N_{i,f}$ of about 32 points. Albeit not shown for the sake of brevity, we detect pronounced differences in $N_{i,m}$ and $N_{i,f}$ between groups in the population. For instance, women with a degree have gender norms that are 0.11SD (standard deviation) greater than their low-education counterparts, while the gap in gender norms between men with and without a higher degree is only 0.06SD. The differences are particularly notable across cohorts and by ethnicity. Men and women born after 1980 have attitudes that are on average about 2SD more progressive than those

[18] In a sensitivity exercise, we summarize the responses with the first component obtained from a principal component analysis (see subsection 6.2). The results are identical to those shown below. We also redefine $N_{i,m}$ and $N_{i,f}$ using either the third and fourth claims (iii)–(iv) or claims (ii)–(iv) in an attempt to focus more on beliefs about housework arrangements between partners than on beliefs about children and childcare or employers' responsibilities towards workers. Our main results do not change.

[19] We have four birth cohorts (1960 or earlier, 1961–1970, 1971–1980, and 1981 or later). We distinguish 17 ethnic groups: three groups of white individuals (British/English/Scottish/Welsh/Northern Irish, Irish, and any other white background); five groups of Asian/Asian British individuals (Indian, Pakistani, Bangladeshi, Chinese, and any other Asian background); three groups of black individuals (African, black Caribbean, and any other black background); four groups of mixed ethnicity (white and black Caribbean, white and black African, white and Asian, any other mixed background); Arab; and any other ethnic group.

held by corresponding men and women born before 1961. Similarly, gender norms for both white men and women are about 2SD higher than the norms shared among Asian men and women. Finally, and perhaps unsurprisingly, in households that divide domestic tasks equally, both men and women have average gender norms that are 0.2SD significantly higher (p-value<0.01) than those found for households in which the woman does more.

To estimate $u^m_{i,j}(\mathbf{X}_i)$ and $u^f_{i,j}(\mathbf{X}_i)$, we combine the vignette data in the IP sample with the predictors $\mathbf{X}_i$, which include each partner's labor supply, their relative earnings power, indicator variables for the presence of children in different age groups, and a dummy variable for the use of paid housework. In order to perform this combination, we harmonize some of the variables in $\mathbf{X}_i$ to conform to those used in the vignette experiment. In particular, each partner's labor supply is summarized by two dummy variables for full-time employment (=1 if working 35 or more hours a week) and part-time employment (=1 if working fewer than 35 hours a week). Relative earnings power is proxied by a set of indicator variables that specify whether one partner's hourly wages are the same as, less, or more than the other's. Respondents are defined to earn less (more) if their hourly earnings are less than 45% (more than 55%) than their partners'.

In the estimation of (5), we also control for the oldest partner birth cohort and age, the couple's country of residence (i.e., England, Wales or Scotland), and household nonlabor income. Table 1 shows a sharp inequality in the intrahousehold distribution of activities, whereby women do most of the housework in nearly 72% of households while 56% of them earn hourly wages that are equal to, or higher than, their partners'.

4 Identification

Our aim is to estimate the random utility model for household chores decisions summarized in (6). This estimation procedure is in three steps. First, the randomized vignette experiment is used to estimate the relationship between individual utilities for men and women, $u^m_{i,j}(\mathbf{X}_i)$ and $u^f_{i,j}(\mathbf{X}_i)$, and the $\mathbf{X}_i$ characteristics that refer to hypothetical scenarios. Second, real-life individual utilities are backed out from the parameters estimated in the first step after replacing $\mathbf{X}_i$ with the characteristics depicting *actual* conditions for each couple. Third, the random utility model (5) is estimated by using the predicted utilities from the second step and by instrumenting peers' gender norms, $N_{i,m}$ and $N_{i,f}$, with peers' mothers' labor force participation rates. When these steps are completed, we have all the ingredients needed for the estimation of the likelihood function (6). The first and the last steps are presented in greater detail in the next two subsections.[20]

4.1 Endogeneity of Individual Utilities

As discussed in Section 2, individual economic utilities, $u^m_{i,j}(\mathbf{X}_i)$ and $u^f_{i,j}(\mathbf{X}_i)$, depend on the observable predictors $\mathbf{X}_i$. In particular, $\mathbf{X}_i = [X_i^W, X_i^E, X_i^K, X_i^H]$, where X_i^W measures both partners' labor market involvement, X_i^E proxies their relative earnings power, X_i^K refers to the presence and age of children in household i, and X_i^H records the couple's reliance on help for housework.[21] The vector $\mathbf{X}_i$, therefore, is meant to capture variation across households in the opportunity cost of time spent in housework, the time spent in the labor market, and the time for childcare needs.

Besides $\mathbf{X}_i$, individual utilities depend also on housework arrangement y_i. Considering all the combinations of the variables in $\mathbf{X}_i$ and the three housework arrangements which characterize y_i, we end up with 288 different hypothetical domestic scenarios, 3 (housework arrangements) × 4 (labor market involvement) × 3 (relative hourly wages) × 4 (children) × 2 (domestic help). The individual utility identification relies on the vignette-based stated preference experiment, in which each respondent reports his/her level of utility with up to six scenarios randomly selected out of the 288 available scenarios. Letting s denote a specific hypothetical scenario, we estimate the following specification of individual economic utility

$$\begin{aligned} u_{ist} = \lambda_0 + \psi d_i + \sum_{j=1,3} \beta_j \mathbb{1}(y_{ist} = j) + \sum_{j=1,3} \delta_j d_i \times \mathbb{1}(y_{ist} = j) + \mathbf{X}_{ist}\boldsymbol{\varphi} \\ + \sum_{j=1,3} \mathbf{X}_{ist} \times \mathbb{1}(y_{ist} = j)\boldsymbol{\theta}_j + \lambda_t + \mu_i + \nu_{ist}, \end{aligned} \tag{9}$$

where d_i takes value 1 if the respondent in household i is a woman, and 0 otherwise; $\mathbb{1}(y_{ist} = j)$ is an indicator variable taking value 1 if the housework arrangement in hypothetical scenario s in year t is j (=1,3), and 0 otherwise (with $j = 2$ being the reference category, where both

[20]We should stress that, in the analysis, we cannot treat the share of housework as a continuous variable, since individual utilities, $u^m_{i,j}(\mathbf{X}_i)$ and $u^f_{i,j}(\mathbf{X}_i)$, cannot be predicted at each of the potential values of housework share. Such continuous predictions would require to perform a different vignette experiment with an extremely large number of scenarios to span the full set of values of housework share.

[21]Specifically, X_i^W is a vector of four dummy variables indicating whether both partners work part time, whether the woman works full time and the man part time, whether the woman works part time and the man full time, or whether both partners work full time; X_i^E, is a vector of three dummy variables indicating whether the woman's hourly earnings are larger than the man's, whether both partners earn approximately the same hourly wages, or whether woman's hourly earnings are lower than the man's; X_i^K is a vector of four dummy variables indicating whether the couple has no children, whether they have children and the youngest of them is aged between 0 and 2, whether they have children and the youngest is aged between 3 and 11, or whether they have children and the youngest is between 12 and 15 years (in the experiment, the presence of children aged 0–2, 3–11, or 12–15 years are represented by scenarios in which the child is 6 months old, 5 years old, 15 years old, respectively); and X_i^H is an indicator variable taking value 1 if the couple report that they pay for help with their domestic activities.

partners do approximately the same amount of housework); λ_t refers to time fixed effects; μ_i is an unobserved individual fixed effect; and ν_{ist} is an idiosyncratic error term.[22] Using the estimated preference parameters from equation (9) and replacing $\mathbf{X}_{ist}$ with the actual $\mathbf{X}_{it}$ observed in the UKHLS sample (see Table 1), individual economic utilities are predicted and then standardized to have a variance of one over the pooled sample of men and women.

A key identification insight is that, through the vignettes, we randomize specific attributes of each intrahousehold arrangement, while keeping all other conditions identical across arrangements. Sufficient variation in the attributes across choice scenarios permits us to recover the preference parameters and to predict the utilities in equation (5). To ensure this, the D-efficient sampling technique is used to generate and allocate different scenarios to respondents. Such sampling method maximizes the variance of each of the variables characterizing a scenario within the set of vignettes received by the same individuals and minimizes the correlation between the five variables characterizing each scenario.[23] We can consistently estimate such parameters so long as the preference shocks, $\mu_i + \nu_{ist}$, are independent of the housework arrangements. In our context, this condition holds because of the randomized experimental design.

In the next section, we will also show empirically that we cannot reject the equality between random effects and fixed effects versions of specification (9). We shall then use a random effects specification. The lack of correlation between preference shocks and the primitives underlying the individual economic utilities suggests that our estimation is unlikely to be affected by omitted variable bias or ex-post rationalization. Note that, even if the preferences parameters in model (9) were heterogeneous across individuals, the vignette-based experiment allows us to interpret the estimates as expected preference parameters. This implies we do not need to impose any strong assumption on the distribution of the preference parameters echoing the approach used in stated preference experiments with elicited probabilities (see Blass et al., 2010; Koşar et al., 2022).

Unlike that approach, however, we observe the level of utility on a 7-point scale associated with each hypothetical housework scenario. Satisfaction with household arrangements is

[22] In the results presented in Section 5, λ_t is an indicator variable that equals 1 for respondents in 2013, and 0 for respondents in 2012. Furthermore, of the interactions between the indicator for housework arrangement j and $\mathbf{X}_i$, those involving X_i^E, X_i^K, and X_i^H turn out to be always statistically insignificant at conventional levels and are thus not used in the estimation. Including them does not alter our results.

[23] To check whether or not we have enough variation within groups, we consider the IP sample and find there are at least 247 hypothetical scenarios used within each of the following demographic groups: individuals with a university degree, individuals with less than a university degree, individuals with children, individuals with no children, individuals born in 1951–1959, in 1960–1967, 1968–1976 and in 1977–1988. Moreover, the level of utility reported by individuals within each of the above subgroups covers the full distribution of utility values in the sample. This evidence strengthens our priors that there is ample dispersion in both scenario attributes and utility levels reported by IP respondents across different socio-demographic groups.

assumed to increase linearly over the scale, which yields a cardinal measure of utility. This is similar to the elicitation of subjective expectations using up to 10 beans to express the likelihood of an event happening (e.g., Delavande et al., 2011; Delavande and Kohler, 2016).[24] In a sensitivity exercise, we also treat our satisfaction measure as a non-cardinal measure and estimate ordered probit models to predict the probability of well-being above a given threshold. The results from this exercise are reported in Table A1 and are fully in line with those reported in the analysis below. As found by Ferrer-i-Carbonell and Frijters (2004), assuming ordinality or cardinality of self-reported satisfaction (utility) does not make much of a difference, whereas controlling for fixed-effects is crucial. Our utility model contains individual fixed effects and allows us to achieve better comparability across individuals even if their responses might be affected by unobserved traits (e.g., optimism, empathy and envy), as long as these do not change across hypothetical scenarios.

4.2 Endogeneity of Peers' Gender Role Attitudes

Endogeneity of partners' peers' gender role attitudes, $N_{i,m}$ and $N_{i,f}$, can arise as a result of unobservables that affect household decisions over both housework and peers' attitudes. For example, if peers from countries with a more egalitarian culture agglomerate into areas with specific amenities (e.g., areas where formal childcare provision and quality are higher, or access to public services is easier), then the effect of peers may capture unobserved environmental characteristics rather than gender norms. More generally, endogeneity of $N_{i,m}$ and $N_{i,f}$ could be driven by unobserved individual and environmental characteristics shared between partners and their peers, by peers' unobserved characteristics (i.e., group composition), which might directly affect the couple's decisions, by measurement error in the attitude variables, and by standard reflection problem issues, although in our application this latter source of endogeneity is expected to be minor, given the way our analysis is designed.

We address such endogeneity concerns using a control function approach (see, among others, Petrin and Train, 2010; Wooldridge, 2015), which directly solves potential issues of variables omission and classical measurement error in $N_{i,m}$ and $N_{i,f}$. The procedure consists of two stages. In the first stage, we use a couple (household i's) fixed effects estimation to predict his and her peers' gender role attitudes, $N_{i,g}$, $g = \{m, f\}$, i.e.:

$$N_{i,g} = \pi \text{Female}_{i,g} + \mathbf{X}_i^C \boldsymbol{\zeta} + \phi Z_{i,g} + \tau_i + \omega_{i,g}, \tag{10}$$

[24] Manski (2004), Delavande et al. (2011) document that Likert scales and subjective probabilities responses may produce different probability distributions of the events under analysis. In their comparison, however, Delavande and colleagues use coarse Likert measures with only a 4-point scale and, crucially, they are interested in probability distributions rather than cardinal well-being rankings. See also Schröder and Yitzhaki (2017), Bond and Lang (2019), and Jenkins (2020).

where $\text{Female}_{i,g}$ is an indicator variable taking value 1 for women $g = f$ and 0 for men if $g = m$, $\mathbf{X}_i^C$ is a vector of control variables included in the random utility model, $Z_{i,g}$ is the instrumental variable which differs for the man $g = m$ and the woman $g = f$ in the same household i, τ_i denotes household fixed effects, and $\omega_{i,g}$ is an idiosyncratic error term.

$\mathbf{X}_i^C$ includes dummies for the country of residence within the UK, age of the oldest partner, household nonlabor income, and the older partner's birth cohort fixed effects (see Table 1 for descriptive statistics of these controls). As mentioned, $N_{i,m}$ and $N_{i,f}$ denote his and her peers' average gender role index, which is the sum of the average gender role indexes over the man's and the woman's peers, respectively. Peers are individuals of the same sex, ethnicity, and birth cohort as the reference person in each couple. The instruments $Z_{i,m}$ and $Z_{i,f}$ refer, respectively, to the proportion of the man's peers' and the woman's peers' mothers who worked when his and her peers were 14 years old. Both instruments are plausible, given the evidence documenting that gender role attitudes are transmitted from mothers to children (e.g., Blau et al., 2013; Farré and Vella, 2013; Johnston et al., 2014), while direct systematic interactions between an individual and his/her peers' mothers are arguably much less common. Summary statistics for the peers' gender role attitudes and corresponding instruments are provided in Table 1.

Our identification strategy follows the epidemiological approach, which uses cultural homophily with peers defined over ethnicity or country of ancestry (e.g. Fernández and Fogli, 2006; Giuliano, 2007; Alesina and Giuliano, 2010; Holmlund et al., 2023). This approach may be sensitive to the possibility that individuals from specific ethnicities sort into areas with characteristics that could affect labor market and housework decisions. To address (or attenuate) this potential endogeneity issue, model (10) includes household fixed effects, τ_i, which account for unobserved area characteristics and amenities as well as for any unobservables that are shared between partners, e.g., the share of paid work hours done by women over the total of hours of paid work done by the couple, the couple's average education, income, and age. Furthermore, because our instruments may be correlated with aggregate trends in female labor supply, we control for cohort fixed effects,[25] and identify the impact of gender norms through the large variation in peers' labor force participation trends across ethnicity.

We first perform the fixed effect estimation of equation (10) and then fit the alternative-specific conditional multinomial logit model (5), including as additional controls $\mathbf{X}_i^C$ and the predicted gender-specific composite residuals $R_{i,m}(= \widehat{\tau}_i + \widehat{\omega}_{i,m})$ and $R_{i,f}(= \widehat{\tau}_i + \widehat{\omega}_{i,f})$. Keeping

[25] We use the oldest partner's year of birth and consider 4 indicator variables for births occurring in 1960 or earlier (reference category), 1961—1970, 1971—1980, and 1981 or later.

the same notation as in expression (5) for simplicity, this yields

$$U_{i,j} = \eta^{E,m} u^m_{i,j}(\mathbf{X}_i) + \eta^{E,f} u^f_{i,j} + \mathbf{X}^C_i \boldsymbol{\eta}^C + \eta^{MN}_j N_{i,m} + \eta^{FN}_j N_{i,f} + \rho^m_j R_{i,m} + \rho^f_j R_{i,f} + \varepsilon_{i,j}, \quad (11)$$

where ρ^m_j and ρ^f_j, capture the correlations between the error terms in the first stage regressions and the error term in the random utility model. These should be zero if there is no endogeneity of $N_{i,m}$ and $N_{i,f}$.

5 Results

5.1 Experimental Estimates of Individual Economic Utilities

Table 3 presents the results from the random effects estimation of the individual utility model (9) using our experimental design. We re-emphasize that the vignette experiment addresses endogeneity issues, since respondents are asked to report satisfaction with hypothetical scenarios which are unrelated to their real-life situation. The Hausman-test statistic at the bottom of the table does not reject equality between random effects and fixed effects specifications of the model, revealing there is no correlation between individual-specific unobservables and the hypothetical scenarios' characteristics. To provide further evidence that responses given in the experiment are unrelated to real-life variables, we also ran a set of sensitivity analyses where we allowed the coefficients of model (9) to vary across people with different characteristics one at a time. We do not find statistically significant differential effects by birth cohort, education, martial status, presence of children, being or not at work, and work hours. This corroborates the randomization of the experimental design, indicating that the responses do not depend on personal characteristics and there is no correlation between individual-specific unobservables and the variables that describe our hypothetical scenarios, which in turn suggests that omitted variable bias or ex-post rationalization issues are unlikely to affect our estimation.

Looking at the estimates of the individual utility model (9) in Table 3, we find that, for both men and women, equal housework shares lead to significantly higher levels of satisfaction than a counterfactual world in which housework is divided unequally, even when they do less than their partners ($\beta_1, \beta_3 < 0$). This taste for an equal share of housework is more pronounced for women ($\delta_1 < 0$ and $\delta_3 < 0$). Furthermore, men have an aversion for doing more housework, which is only slightly larger than that for doing less housework than their partner. The aversion for doing more household is greater among women, suggesting that women tend to dislike doing more housework even more than men.

Women report higher levels of utility ($\psi > 0$), in line with earlier research on job satisfaction (Clark, 2018), although recent research finds evidence of falling female happiness relative to men's (Stevenson and Wolfers, 2009), something we cannot detect with our short panel of data. Utility goes up if individuals are asked to imagine to rely on external help with housework ($\varphi^H > 0$). This, ceteris paribus, might reflect greater leisure and joint household consumption (e.g., Hamermesh, 2002; Aguiar and Hurst, 2007; Browning et al., 2021; Georges-Kot et al., 2024). Higher hypothetical relative hourly wages also contribute to boost utility ($\varphi_3^E > 0$), whereas satisfaction with housework is affected neither by child presence nor by child age structure ($\varphi_1^K = \varphi_2^K = \varphi_3^K = 0$). This could be due to the fact that interviewees are prompted to assume that both partners share childcare responsibilities equally. Finally, happiness with equality in the intrahousehold allocation of chores is mirrored by happiness with equality in paid work ($\varphi_1^W = 0$ and $\varphi_2^W, \varphi_3^W < 0$).

It is possible that individual utilities be affected by the interaction between the hypothetical housework scenario and the various subcomponents of $\mathbf{X}_{ist}$. We cannot find any significant interaction with X_{ist}^E, X_{ist}^K, or X_{ist}^H. It is interesting to link this to the previous finding according to which individuals with a greater earnings power enjoy a utility boost ($\varphi_3^E > 0$). The zero-interaction result implies that the earlier estimate cannot be associated with an intrahousehold allocation of chores that compensates for earnings differentials between partners. Put differently, a specialization argument based on earnings power cannot explain this result.

There is evidence, instead, that satisfaction with domestic chores arrangements varies considerably with partners' labor market involvement.[26] To ease the interpretation of these results, Figure 1 displays the predicted utilities for men and women with different combinations of hypothetical housework and paid work schedules. Regardless of gender, an equal allocation of domestic work is unambiguously preferred when both partners work either full time or part time. This utility premium is at least 13% for male part-timers and up to 25% for female full-timers with respect to the second-best intrahousehold work allocation. Individuals who work full time whose partners work part time, instead, enjoy a greater utility if their housework share is lower than the partners' and experience a significant disutility of up to 30% if they bear most of the chores. Symmetrically, in scenarios where interviewees are asked to visualize themselves working part time and their partners full time, utility goes down should respondents do less housework than their partners.

These results emphasize that couples are likely to use housework arrangements to compensate for imbalances in labor market involvement, but not earnings differentials. That is, when partners are equally involved in paid employment, they attain higher levels of utility

[26] This is statistically confirmed by a large partial R^2 associated with the interaction terms between domestic chore arrangements and partners' labor market involvement in the individual utility regressions (9). Dropping such interactions would reduce the proportion of variation in y explained by the model by 31%.

if they also share domestic chores equally. Conversely, if one partner works relatively more in the market, the other partner compensates with more work at home. The former set of arrangements provides indirect evidence for intrahousehold consumption complementarities (Stevenson and Wolfers, 2007; Mansour and McKinnish, 2014); while the latter offers support to the notions of equity in time allocation between partners (Hamermesh, 2020) and value of joint leisure (Browning et al., 2021; Georges-Kot et al., 2024).

Except for the result that women have a greater aversion than men towards unequal housework arrangements, we cannot identify other differential responses by gender. This suggests that the pronounced intrahousehold differences observed in the actual allocation of chores in Table 1 are hard to be accounted for by gender differentials in preferences over housework arrangements. Intrahousehold gaps in domestic work may be driven instead by specialization due to gendered comparative advantages or by gender role attitudes.[27] We now turn to consider the role played by these two channels more closely, although the evidence on the former is weak so far.

5.2 Specialization and Gender Norms

Table 4 reports the marginal effects obtained from model (7), which is a variant of the alternative-specific conditional logit model (5). In this variant, we assume that couples give the same weight to his and her individual utility, $\alpha^m = \alpha^f$ or, equivalently, $\eta^{E,m} = \eta^{E,f}$. This is justified by the results found with an unrestricted version of the model with male and female bargaining weights allowed to differ, according to which we cannot reject equality of the weights at standard levels of statistical significance, with a p-value for the equality test of 0.136. The equality of weights given to his and her utility suggests that men and women have a similar bargaining power and the advantage of men with respect to women in the labor market does not play a role in housework decisions.[28]

The results in Table 4 are obtained using a maximum likelihood estimation of the conditional logit model without instrumental variables. The justification for proceeding without instrumental variables is in Table 5, where we report the test for exogeneity of his and her

[27] We should point out that the model does not consider women's potential relative advantage in domestic work. Women may have to specialize in childcare, especially in the first months after childbirth (e.g., breastfeeding). Excluding childcare, as we do in our application, would however lead to a chore production function that does not require gender-specific skills. To emphasize this point, male and female time inputs in housework are assumed to be perfect substitutes, in line with several previous papers (e.g., Chiappori, 1997; Pollak, 2013; Chiappori and Lewbel, 2015). Furthermore, Siminski and Yetsenga (2022) provide convincing evidence that comparative advantage has a small role in explaining intrahousehold labor allocations.

[28] Interestingly, using a sample of British households, Goussé et al. (2017) find evidence of women's slightly greater bargaining power as well as women's slightly lower share of net total private expenditures. Both such parameters are close to 0.5. For comparable estimates for the United States, see Gayle and Shephard (2019).

peers' gender norms, $N_{i,m}$ and $N_{i,f}$, and the F-test for the relevance of the instruments, $Z_{i,m}$ and $Z_{i,f}$. The first row of Table 5 reveals no evidence of endogeneity. As documented in the same table, this is the case also for the all the other models we discuss below.

Our variables of interest in model (7) are $N_{i,m}$ and $N_{i,f}$, his and her peers' average gender role indexes, and $U^E_{i,j} = u^m_{i,j} + u^f_{i,j}$, the couple's economic utility defined in expression (8). $N_{i,m}$ and $N_{i,f}$ are expressed in standard deviations (SDs), with higher values representing more egalitarian norms. Similarly, both $u^m_{i,j}$ and $u^f_{i,j}$ are measured in SDs, with the SD computed on the pooled sample of partners.

In the top panel of Table 4, we show the marginal effects of a one-standard deviation increase in $U^E_{i,j}$, $N_{i,m}$, and $N_{i,f}$ on the probability that the female partner does less housework, shares it equally with, or does more of it than her male partner (column (a), (b) and (c), respectively). For each marginal effect, in square brackets we report the p-values of the Wald test for the null hypothesis of a zero marginal effect.

A one-unit increase in $U^E_{i,j}$ associated with a given housework arrangement j leads to a significant increase in the probability that the household chooses such arrangement. In absolute terms, the largest impact of 3.9 percentage points of a standard deviation emerges for the arrangement in which the woman does more housework, which represents a 5.5% increase in the probability of choosing that alternative over the baseline, i.e., over the observed unconditional probability of a household choosing that specific arrangement. A utility boost associated with each of the other two options (either the woman does less or an equal split of housework) leads to about 17% higher probability. The "off-diagonal" counterfactual utility effects, which we identify through the vignette-based stated preference experiment, are negative (albeit smaller in absolute value), indicating that the probability of choosing a given housework arrangement is inversely related to the utility associated with the other alternative arrangements. Couples therefore choose the option that yields the largest utility. This implies that the probability of choosing arrangement j is increasing in $U^E_{i,j}$ and decreasing in $U^E_{i,s}$ for $s \neq j$, which leads to the negative sign for the off-diagonal effects in Table 4.

These results are in line with the experimental evidence of the previous subsection. As documented there, the economic utility $U^E_{i,j}$ increases when the division of labor between partners is fairer, i.e., when the housework share compensates for intrahousehold imbalances in labor market involvement. The results in Table 4 show that household utility increases in $U^E_{i,j}$, therefore suggesting that household decisions are aligned with individual preferences. Couples in which both partners have comparable labor market involvement prefer, and choose, equal housework arrangements. Couples in which one partner is more involved in market work prefer, and choose, an arrangement in which the other partner does more housework.

Intrahousehold compensation of housework for market work is therefore consistent with standard time allocation theory. But this cannot be the only, perhaps not even the main, explanation for the gendered distribution of housework. Against the backdrop of Table 2, according to which 53.6% of couples either have a similar labor market involvement or the woman works more than her partner, the evidence in Table 4 provides little support for an explanation of the intrahousehold allocation of chores based exclusively on a specialization argument. This argument would hinge on male comparative advantage in the labor market, which is hard to invoke given that women shoulder a substantially greater share of housework in 71.7% of the households in the sample, even if their labor market share is lower than their partner in only 46.0% of cases.

Looking at peers' gender norms, we find that a one-standard deviation increase in her peers' attitudes towards more egalitarian gender norms, $N_{i,f}$, leads to a substantial rise in the probability of the woman doing less housework with respect to the mean by 12.6% and a decrease in the probability of the woman doing more housework by 3.6%, which are both statistically significant at the 5% level. These results emphasize the importance of gender role norms in explaining intrahousehold housework arrangements, aligning well with the evidence presented by Bertrand et al. (2015), Goussé et al. (2017), and Ichino et al. (2024).

In the case of men, there does not seem to be a statistically significant effect of his peers' attitudes and we cannot reject the equality of the effects of peers' attitudes across gender. For this reason, in our preferred specification of model (7), we exclude his peer's attitudes, $N_{i,m}$. The results of this preferred specification are reported in Table 6. They show that the estimated effects remain unchanged and confirm the findings we just discussed.[29]

5.3 What Happens When Gender Norms Become More Egalitarian

Considering the role played by peers' gender norms, $N_{i,f}$, in the previous analysis, we now provide a more precise quantitative assessment of its impact on the division of housework between partners. We perform two simulations.

Before illustrating the simulation exercises, we set the stage by assessing the relationship between $N_{i,f}$ and birth cohort, which gives us a useful benchmark to interpret the simulation results. We find that one additional year of birth leads to a growth in gender identity attitudes by about 0.063 of a standard deviation of female peers attitudes. This means that, ceteris paribus, it would take about 16 years to see an increase in the peers' gender role attitudes by one standard deviation.

[29] We also tested if couples give the same weight to his and her individual utility, $\alpha^m = \alpha^f$, in a model where we include only her peers' gender role attitudes. As before, we cannot reject equality (p-value=0.135).

Our first exercise considers the effect of increasing peers' gender role attitudes, $N_{i,f}$, by one SD, which is equivalent to the average increase in peers' gender role when moving forward in time by 16 years, *ceteris paribus*. Table 7 summarizes the key findings from this simulation. Row (a) reports the observed share of women who do less, equal, or more paid work than their partners as shown in Table 2, whereas row (b) replicates the corresponding observed fractions for housework. In row (c), a "gender fair" housework distribution mirrors the observed paid work distribution in row (a). Since in the sample 3.7% (46.4%) of the women work more (less) than their partners in the market, a fair allocation of chores implies the same 3.7% (46.4%) of women do less (more) of the housework. Subtracting the fair from the actual housework distribution will approximate the excess housework shouldered by women or by men. This, in row (d), reveals that 11.1% of men and 25.3% of women do more housework than their fair share. There is a clear asymmetry in this allocation which penalizes women more than men.

Now, suppose that $N_{i,f}$ increases by one SD. Using the estimated marginal effects of $N_{i,f}$ reported in the fourth row of Table 6, we expect that the probability of a woman doing less housework than her partner would increase by 1.8 percentage points, i.e., 12.6% with respect of the mean of 0.143 (see row (b) in Table 7), the probability of doing a similar amount of housework would increase by 1.0 percentage point (or 7.1%), while the probability of doing more housework would decrease by 2.8 percentage points (3.9%). The corresponding housework distribution after such changes is reported in row (e) of Table 7 and the new counterfactual excess housework distribution is in row (f). The expansion towards more egalitarian gender role attitudes increases the fraction of men doing more than their fair share only slightly from 11.1% to 12.9%, but reduces the proportion of women doing more housework than their fair share from 25.3% to 22.5%. Over a span of 16 years, the increase in egalitarian views alone could then abate the unfair extra housework done by women by about 11% (from 25.3% to 22.5%) and, all else equal, would equalize the probability that men and women do more than their fair share of housework in about 50 years. Assuming a future trend in gender norms similar to the one observed in the sample, a span of 50 years would imply an increase in gender egalitarian views by about 3SD, which would lead to a similar excess housework for men of 0.165 (=0.143+0.018×3–0.032) and for women of 0.164 (=0.717–0.028×3–0.464).

In our second simulation, we use the conditional logit estimates underpinning Table 6 to compute predicted probabilities over the entire distribution of female peers' gender role norms. For this exercise, we predict the housework share for a benchmark household in which both partners work full time, earn the same amount of labor income, do not pay for help with housework, have no children, live in England, have median nonlabor income, with the oldest partner in the couple born in 1981 or later and with an age equal to the median age in this birth-cohort group. The other controls are set to their median values in the overall sample. Figure 2 reports the results, focusing on the probability that women do most of the housework across the $N_{i,f}$ distribution. The left-hand side panel of the graph shows the simulated shares by education i.e., without and with a university degree.[30] The proportion of women who do more housework than their male partner starts at nearly 75% at the bottom of the $N_{i,f}$ distribution and decreases to less than 50% at the very top of the distribution for both groups of women, albeit at a faster pace among highly educated couples.[31]

The right-hand side panel of Figure 2 displays simulated housework shares by motherhood status. The proportion of mothers who do more housework is much larger than the corresponding fraction among childless women. In fact, it is close to two-thirds even when peers' gender role attitudes are most progressive, in sharp contrast with their childless counterparts, for whom the probability to do more housework is just over 40% at the top of the $N_{i,f}$ distribution.

Our simulations are computed on households in which both partners work full time and roughly earn the same amount of labor income. Despite this, they suggest that 70% of women do more housework than their partner when their peers' gender role attitudes are highly traditional. Being exposed to median peers' attitudes does reduce the fraction of women who do more housework below 60% among highly educated women and childless women. However, it affects only marginally the housework share among less educated women and mothers. We know culture is likely to change slowly over time (e.g., Guiso et al., 2006; Bisin and Verdier, 2011; Alesina and Giuliano, 2014). In fact, if progressive gender identity norms continue to evolve as we observe in the data without further interventions, it would take another 50 years before a fair allocation of housework that reflects the partners' distribution of market work could materialize. The cumulative impact of social attitudes on how housework is divided between partners may be substantial nonetheless. If, for instance, gender norms leveled up to the most progressive attitudes observed in the sample, our simulations suggest that women doing more housework than their partners would stop to be the norm already among present-day households, except for those with children.

[30] A household is defined to have 'high' education if at least one partner has a university (or higher) degree.
[31] The curves remain flat in some parts of the distribution because there are ties in the corresponding percentiles.

6 Additional Evidence

6.1 Impact Heterogeneity

We explore two dimensions of potential heterogeneity, one along the education margin and the other by child presence. Couples with different levels of education could have both different preferences over the division of housework and exposure to different peers and, thus, different social norms. We define highly educated households as those in which at least one partner has a university degree and re-estimate our benchmark model separately for high- and low-education couples.

Table 8 reports the marginal effects from this analysis, following the same structure as in Table 6. There are no large differences by education in the marginal effects of an increase in household utility associated with each of the three household arrangements on the probability of choosing those arrangements, except for a slight increase for highly educated households. These couples tend to prefer the intrahousehold division they actually choose more than their low-education counterparts. Peers' social norms seem to matter differently for the two groups of households. An increase by one SD in the gender role index of her peers towards more egalitarian attitudes significantly reduces the likelihood of her doing more housework than her partner by 4% for both low- and high-education couples. This reduction is accompanied in turn by significant increases in the probability of sharing chores equally by 19.5% among low-education couples and in the probability of her doing less housework by 14.7% for highly educated couples.

As motherhood has been identified as one of the key factors slowing down female progress in the labor market (e.g., Francesconi, 2002; Bertrand et al., 2010; Adda et al., 2017; Kleven et al., 2019), this is also likely to affect the intrahousehold allocation of domestic duties. Put differently, childcare and housework specialization may not be separable, as we have assumed so far. The results in Table 9 reveal that the impact of preferences on housework decisions is relatively similar between households with children and households without. Having socially more progressive peers, instead, has different impacts. A one-standard-deviation increase in her peers' gender role index would lead both couples with and without children to heighten the probability of her doing less housework and to decrease the probability of her doing more housework; but the increase in the likelihood of her doing less housework is substantially larger for childless households in both absolute terms (2.6% point versus 0.9% points) and percentage increase relative to the mean (16.3% rather than 7.0%).

Taking stock of this analysis, we reiterate that both preferences over housework arrangements and gender identity norms play a distinctive role in shaping the intrahousehold distribution of chores. On the one hand, preferences seem to have a slightly greater impact among highly educated couples, while they play similar roles among childless couples and households with children. On the other hand, the exposure to gender egalitarian values leads to an overall reduction in the housework shouldered by women, regardless of education or motherhood status. For childless and highly educated couples, however, we observe a more pronounced increase in the likelihood of women doing less housework than their partner.

6.2 Sensitivity Checks

We perform seven sensitivity exercises, which are meant to address a wide set of concerns. First, as discussed in Section 4, our benchmark model assumes individual utilities to be increasing linearly over a 7-point scale associated with each hypothetical housework scenario of the vignette experiment. An alternative is to consider this satisfaction assessment as a non-cardinal measure (e.g., Schröder and Yitzhaki, 2017; Bond and Lang, 2019), and re-run the previous analysis, starting with the estimation of the individual economic utilities using random effects ordered probit regressions. Once we have such estimates, we can then follow the same procedure as the one described earlier to obtain a new set of alternative-specific conditional logit marginal effects. These new estimates are reported in Table A1 in the Online Appendix and are essentially identical to those reported in Table 6.

Second, we use alternative ways of measuring gender role attitudes. Rather than averaging over the five responses to the questions designed to elicit gender identity norms described in subsection 3.2, we construct $N_{i,f}$ with the first component of a principal component analysis and then repeat our standard procedure to estimate the marginal effects. The new results shown in Table A2 are similar to the benchmark estimates of Table 6. We also experiment with two additional definitions of gender role norms. In one definition, we consider only two of the five gender role items listed in subsection 3.2, that is: (iii) "Husband and wife should contribute to household income", and (iv) "Husband should earn, wife should stay at home", which are meant to provide an index less driven by the presence of children and childcare and more driven by work arrangements between partners. The results, shown in panel A of Table A3, lend further credibility to the previous evidence on the separability between childcare and housework. In the other definition, besides items (iii) and (iv), we also include item (ii) (i.e.,"All in all, family life suffers when the woman has a full-time job"). The new results in panel B of the same table corroborate our main findings.

Third, we modify the definition of peers using individuals with the same gender, same birth cohort, and same country of birth (rather than same ethnicity as we do in the benchmark analysis). The marginal effects in panel A of Table 10 uphold the findings displayed in Table 6. We confirm that peers' gender roles attitudes matter, especially when considering female partner's peers. A one-standard deviation increase in women's gender role norms continues to significantly raise the probability of women doing less housework and significantly abate the probability that women do more housework.

Fourth, to moderate the possible confounding influence of small peer groups, we change the minimum size of the group of peers, from 4 to 16 (see subsection 3.2). This reduces the sample size marginally, but the results in panel B of Table 10 remain close to the corresponding estimates shown in Table 6.

Fifth, our benchmark model assumes that the weight given to the economic utility be the same across couples with different gender role norms. We relax this assumption by allowing the model's coefficients to change between households with his peers' gender role attitudes above/below the median, and then similarly for her peers. The results are reported in Online Appendix Tables A4 and A5 and confirm that economic utility matters for both progressive and traditional households. The weight on economic utility tends to be larger for men and women with more gender egalitarian views, indicating that the allocation of housework between partners becomes more gender neutral in such couples. The estimates for the male and female peers' gender role norms, however, become less precisely estimated, making it difficult to draw definitive conclusions, although it is clear that women who have more gender egalitarian peers are also less likely to be influenced by their peers' gender attitudes.

Sixth, we modify the cut-offs used to define the three categories for the chores division variable. Specifically, a woman is re-defined as doing less housework than her partner if the time she spends in domestic activities is less than 40% (rather than 45%) of the total housework time by both partners, and as doing more housework when her time is above 60% (rather than 55%) of the total housework time. The estimates in panel A of Table A6 confirm our main findings. Finally, in our seventh sensitivity analysis, we relax the assumption of linearity in peers gender roles attitudes by considering polynomials of order two in $N_{i,f}$. Panel B of Table A6 shows marginal effects that are similar to our benchmark results.[32]

7 Conclusion

In more than 70% of British households, women shoulder most of the domestic chores. The canonical argument to explain women's specialization in housework hinges on greater male returns to market work. Although this explanation may still be valid for some couples, it

[32]Furthermore, to address potential issues of omitted variables bias in the estimation of equation (7), we repeated the analysis after adding hours worked by each of the two partners. The estimates, not shown, confirm our main results.

loses its bite when, as we find in our sample, women earn more or work longer hours in the labor market than their male partners. This paper leverages a rich empirical model of intrahousehold housework decisions, which combines a randomized experimental framework with an epidemiological approach to identify the separate impacts of two other channels, that is, individual preferences and gender identity norms.

Both channels play an important role. On preferences, we emphasize three results. First, for both men and women, an equal distribution of domestic chores within households leads to significantly higher levels of utility. Second, couples tend to use housework arrangements to compensate for imbalances in labor market involvement, with partners attaining higher levels of utility if they share domestic chores equally when they are equally involved in paid employment and, conversely, if one works more at home when the other works more in the market. Third, women dislike domestic chores as much as (or even more than) men do, suggesting that the large intrahousehold imbalance in the actual allocation of chores cannot be solely accounted for by gender differentials in preferences over housework arrangements.

The influence of social norms on the division of home production activities between male and female partners is substantial. We stress three findings. First, exposure to more egalitarian gender norms significantly decreases the probability that women do more housework. It also raises the probability of women doing less housework, with this effect being larger for highly educated couples and childless households. For less educated couples, we also find a significant increase in their likelihood of choosing an equal share of housework when they are exposed to more progressive gender role attitudes. Second, should gender identity norms continue to evolve towards equality in the future as we observe in the data, a fair allocation of housework that mirrors the partners' distribution of market work would materialize in 50 years, without further interventions. Third, even when considering households in which both partners work full time and earn the same amount of labor income, women are predicted to do more housework than their male partner in at least 70% of cases if their peers' gender role norms were staunchly traditional. Were attitudes moved to the median, the percentage of women doing more housework would fall below 60% for women with a university degree and for childless women; but it would remain above 65% for less educated women and for mothers. An equal division of housework would be reached only when peers' social norms are highly progressive, but it would still be unachievable for mothers.

Several areas for future research seem to be promising. One is to broaden our semi-structural framework to model labor market and housework decisions jointly. This will require a richer experimental design and a fully structural model, substantially expanding the framework by Goussé et al. (2017). Another is to see the extent to which an unfair allocation of work at home and in the market affects the chances of having (more) children and the odds of divorce, possibly amplifying gender inequality over time. This line of work would combine the insights highlighted by Doepke et al. (2024) with some of the key ingredients of a Gayle and Shephard (2019) type of model. Another is to account for assortativeness in partnership formation, which will account for the possibility that men and women who live together today are a more selected population than their counterparts in the past, with different preferences and different gender identity norms.

Finally, it would be valuable to understand how housework-relevant gender role identity can be shaped among men and women who are in a partnership already. Existing evidence shows that paternity leave extensions can lead to more egalitarian gender role attitudes (Kotsadam and Finseraas, 2011; Farré et al., 2023) and to a greater share of childcare done by fathers (e.g., Albrecht et al., 2024). However, there is also compelling evidence that documents a regression toward traditional gender role norms during the recent COVID-19 lockdowns and in periods of recession that affect women more negatively than men (e.g., Del Boca et al., 2020; Farré et al., 2022; Andrew et al., 2022; Boring and Moroni, 2023).

References

Adda, J., C. Dustmann, and K. Stevens (2017). The Career Cost of Children. *Journal of Political Economy 125*(2), 293–337.

Aguiar, M. and E. Hurst (2007). Measuring Trends in Leisure: The Allocation of Time over Five Decades. *Quarterly Journal of Economics 122*(3), 969–1006.

Akerlof, G. A. (1997). Social Distance and Social Decisions. *Econometrica 65*(5), 1005–1027.

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

Akerlof, G. A. and R. E. Kranton (2010). *Identity Economics: How Our Identities Shape Our Work, Wages, and Well-Being.* Princeton, NJ: Princeton University Press.

Albrecht, J., P.-A. Edid, R. Fernández, J. Lee, P. Thoursie, and S. Vroman (2024). Parental Leave: Economic Incentives and Cultural Change. IZA Discussion Paper No. 17210.

Alesina, A. and P. Giuliano (2010). The Power of the Family. *Journal of Economic Growth 15*, 93–125.

Alesina, A. and P. Giuliano (2014). Family Ties. In P. Aghion and S. N. Durlauf (Eds.), *Handbook of Economic Growth*, Volume 2A, Chapter 4, pp. 177–215. Amsterdam: Elsevier.

Alesina, A., P. Giuliano, and N. Nunn (2013). On the Origins of Gender Roles: Women and the Plough. *Quarterly Journal of Economics 128*(2), 469–530.

Álvarez, B. n. and D. Miles-Touya (2019). Gender Imbalance in Housework Allocation: A Question of Time. *Review of Economics of the Household 17*(4), 1257–1287.

Andrew, A., S. Cattan, M. Costa Dias, C. Farquharson, L. Kraftman, S. Krutikova, A. Phimister, and A. Sevilla (2022). The Gendered Division of Paid and Domestic Work under Lockdown. *Fiscal Studies 43*(4), 325–340.

Auspurg, K. and T. Hinz (2014). *Factorial Survey Experiments*, Volume 175. Sage Publications.

Auspurg, K., M. Iacovou, and C. Nicoletti (2017). Housework Share between Partners: Experimental Evidence on Gender-Specific Preferences. *Social Science Research 66*, 118–139.

Bau, N. and R. Fernández (2024). Culture and the Family. In S. Lundberg and A. Voena (Eds.), *Handbook of the Economics of the Family*, Volume 1, Chapter 1, pp. 1–47. Amsterdam: Elsevier.

Becker, G. S. (1965). A Theory of the Allocation of Time. *Economic Journal 75*(299), 493–517.

Becker, G. S. (1981). *A Treatise on the Family*. Cambridge, MA: Harvard University Press (1991 enlarged edition).

Becker, G. S. (1985). Human Capital, Effort, and the Sexual Division of Labor. *Journal of Labor Economics 3*(1, Part 2), S33–S58.

Becker, G. S. and C. B. Mulligan (1997). The Endogenous Determination of Time Preference. *Quarterly Journal of Economics 112*(3), 729–758.

Bertrand, M. (2019). The Gender Socialization of Children Growing Up in Nontraditional Families. *AEA Papers and Proceedings 109*, 115–121.

Bertrand, M. (2020). Gender in the Twenty-First Century. *American Economic Review Papers and Proceedings 110*(May), 1–24.

Bertrand, M., P. Cortes, C. Olivetti, and J. Pan (2021). Social Norms, Labor Market Opportunities, and the Marriage Gap for Skilled Women. *Review of Economic Studies 88*(4), 1936–1978.

Bertrand, M., C. Goldin, and L. 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, M., E. Kamenica, and J. Pan (2015). Gender Identity and Relative Income within Households. *Quarterly Journal of Economics 130*(2), 571–614.

Bisin, A. and T. Verdier (2011). The Economics of Cultural Transmission and Socialization. In J. Benhabib, A. Bisin, and M. Jackson (Eds.), *Handbook of Social Economics*, Volume 1, Chapter 9, pp. 339–416. Amsterdam: Elsevier.

Bittman, M., P. England, L. Sayer, N. Folbre, and G. Matheson (2003). When Does Gender Trump Money? Bargaining and Time in Household Work. *American Journal of Sociology 109*(1), 186–214.

Blair, S. L. and M. P. Johnson (1992). Wives' Perceptions of Fairness and the Division of Household Labor: The Intersection of Housework and Ideology. *Journal of Marriage and the Family 54*(3), 570–581.

Blass, A. A., S. Lach, and C. F. Manski (2010). Using Elicited Choice Probabilities to Estimate Random Utility Models: Preferences for Electricity Reliability. *International Economic Review 51*(2), 421–440.

Blau, F. D., L. M. Kahn, A. Y.-H. Liu, and K. L. Papps (2013). The Transmission of Women's Fertility, Human Capital, and Work Orientation Across Immigrant Generations. *Journal of Population Economics 26*(2), 405–435.

Boelmann, B., A. Raute, and U. Schönberg (2024). Wind of Change? Cultural Determinants of Maternal Labor Supply. *American Economic Journal: Applied Economics, forthcoming*.

Boerma, J. and L. Karabarbounis (2021). Inferring Inequality With Home Production. *Econometrica 89*(5), 2517–2556.

Bond, T. N. and K. Lang (2019). The Sad Truth about Happiness Scales. *Journal of Political Economy 127*(4), 1629–1640.

Boring, A. and G. Moroni (2023). Turning back the clock: Beliefs about gender roles during lockdown. *Labour Economics 84*, 102363.

Bowles, S. (1998). Endogenous Preferences: The Cultural Consequences of Markets and Other Economic Institutions. *Journal of Economic Literature 36*(1), 75–111.

Browning, M., O. Donni, and M. Gørtz (2021). Do You Have Time To Take a Walk Together? Private and Joint Time Within the Household. *Economic Journal 131*(635), 1051–1080.

Bursztyn, L., A. W. Cappelen, B. Tungodden, A. Voena, and D. Yanagizawa-Drott (2024). How Are Gender Norms Perceived? Unpublished paper.

Camerer, C. F. and E. Fehr (2004). Measuring Social Norms and Preferences Using Experimental Games: A Guide for Social Scientists. In J. Henrich, R. Boyd, S. Bowles, C. Camerer, E. Fehr, and H. Gintis (Eds.), *Foundations of Human Sociality*, Chapter 3, pp. 55–95. Oxford: Clarendon Press.

Cardoso, A. R., E. Fontainha, and C. Monfardini (2010). Children's and Parents' Time Use: Empirical Evidence on Investment in Human Capital in France, Germany and Italy. *Review of Economics of the Household 8*(4), 479–504.

Carlana, M. (2019). Implicit Stereotypes: Evidence from Teachers' Gender Bias. *Quarterly Journal of Economics 134*(3), 1163–1224.

Chamberlain, G. (1980). Analysis of Covariance with Qualitative Data. *Review of Economic Studies 47*(1), 225–238.

Chiappori, P.-A. (1997). Introducing Household Production in Collective Models of Labor Supply. *Journal of Political Economy 105*(1), 191–209.

Chiappori, P.-A. and A. Lewbel (2015). Gary Becker's "A Theory of the Allocation of Time". *Economic Journal 125*(583), 410–442.

Ciasullo, L. and M. Uccioli (2024). What Works for Working Couples? Work Arrangements, Maternal Labor Supply, and the Division of Home Production. IZA Discussion Paper No. 16991.

Clark, A. E. (2018). Four Decades of the Economics of Happiness: Where Next? *Review of Income and Wealth 64*(2), 245–269.

Cortés, P. and J. Pan (2019). When Time Binds: Substitutes for Household Production, Returns to Working Long Hours, and the Skilled Gender Wage Gap. *Journal of Labor Economics 37*(2), 351–398.

Couprie, H. (2007). Time Allocation Within the Family: Welfare Implications of Life in a Couple. *Economic Journal 117*(516), 297–305.

Cruces, G., R. Perez-Truglia, and M. Tetaz (2013). Biased Perceptions of Income Distribution and Preferences for Redistribution: Evidence from a Survey Experiment. *Journal of Public Economics 98*(C), 100–112.

Dahl, G., A. Kotsadam, and D.-O. Rooth (2021). Does Integration Change Gender Attitudes? The Effect of Randomly Assigning Women to Traditionally Male Teams. *Quarterly Journal of Economics 2*(136), 987–1030.

De Quidt, J., J. Haushofer, and C. Roth (2018). Measuring and Bounding Experimenter Demand. *American Economic Review 108*(11), 3266–3302.

De Quidt, J., L. Vesterlund, and A. J. Wilson (2019). Experimenter Demand Effects. In A. Schram and U. Aljaž (Eds.), *Handbook of Research Methods and Applications in Experimental Economics*, Chapter 20, pp. 384–400. Edward Elgar Publishing.

Del Boca, D., N. Oggero, P. Profeta, and M. Rossi (2020). Women's and Men's Work, Housework and Childcare, Before and During COVID-19. *Review of Economics of the Household 18*(4), 1001–1017.

Delavande, A., X. Giné, and D. McKenzie (2011). Measuring Subjective Expectations in Developing Cuntries: A Critical Review and New Evidence. *Journal of Development Economics 94*(2), 151–163.

Delavande, A. and H.-P. Kohler (2016). HIV/AIDS-related Expectations and Risky Sexual Behaviour in Malawi. *Review of Economic Studies 83*(1), 118–164.

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

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

Dhar, D., T. Jain, and S. Jayachandran (2022). Reshaping Adolescents' Gender Attitudes: Evidence from a School-Based Experiment in India. *American Economic Review 112*(3), 899–927.

Doepke, M., A. Hannusch, F. Kindermann, and M. Tertilt (2024). The Economics of Fertility: A New Era. In S. Lundberg and A. Voena (Eds.), *Handbook of the Economics of the Family*, Volume 1, Chapter 4, pp. 151–254. Amsterdam: Elsevier.

Farré, L., Y. Fawaz, L. González, and J. Graves (2022). Gender Inequality in Paid and Unpaid Work During Covid-19 Times. *Review of Income and Wealth 68*(2), 323–347.

Farré, L., C. Felfe, L. González, and P. Schneider (2023). Changing gender norms across generations: Evidence from a paternity leave reform. IZA Discussion Paper No. 16341, IZA Institute of Labour Economics.

Farré, L. and F. Vella (2013). The Intergenerational Transmission of Gender Role Attitudes and Its Implications for Female Labour Force Participation. *Economica 80*(318), 219–247.

Fenstermaker Berk, S. (1985). *The Gender Factory: The Apportionment of Work in American Households*. New York: Plenum Press.

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

Fernández, R. (2011). Does Culture Matter? In J. Benhabib, M. O. Jackson, and A. Bison (Eds.), *Handbook of Social Economics*, Volume 1A, Chapter 11, pp. 481–510. Amsterdam: Elsevier.

Fernández, R. and A. Fogli (2006). Fertility: The Role of Culture and Family Experience. *Journal of the European Economic Association 4*(2-3), 552–561.

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

Fernández, R., A. Fogli, and C. Olivetti (2004). Mothers and Sons: Preference Formation and Female Labor Force Dynamics. *Quarterly Journal of Economics 119*(4), 1249–1299.

Ferrer-i-Carbonell, A. and P. Frijters (2004, 06). How Important is Methodology for the Estimates of the Determinants of Happiness? *The Economic Journal 114*(497), 641–659.

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

Francesconi, M. (2002). A Joint Dynamic Model of Fertility and Work of Married Women. *Journal of Labor Economics 20*(2), 336–380.

Gayle, G.-L. and A. Shephard (2019). Optimal Taxation, Marriage, Home Production, and Family Labor Supply. *Econometrica 87*(1), 291–326.

Georges-Kot, S., D. Goux, and E. Maurin (2024). The Value of Leisure Synchronization. *American Economic Journal: Applied Economics 16*(1), 351–376.

Gershuny, J. I. (2000). *Changing Times, Work and Leisure in Post Industrial Society*. Oxford: Oxford University Press.

Gimenez-Nadal, J. I., J. A. Molina, and Y. Zhu (2018). Intergenerational Mobility of Housework Time in the United Kingdom. *Review of Economics of the Household 16*(4), 911—-937.

Giuliano, P. (2007). Living Arrangements in Western Europe: Does Cultural Origin Matter? *Journal of the European Economic Association 5*(5), 927–952.

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

Goussé, M., N. Jacquemet, and J.-M. Robin (2017). Marriage, Labor Supply, and Home Production. *Econometrica 85*(6), 1873–1919.

Greenstein, T. N. (1995). Gender Ideology, Marital Disruption, and the Employment of Married Women. *Journal of Marriage and the Family 57*(1), 31–42.

Greenwood, J., N. Guner, and R. Marto (2024). The Great Transition: Kuznets Facts for Family-Economists. In S. Lundberg and A. Voena (Eds.), *Handbook of the Economics of the Family*, Volume 1, Chapter 7, pp. 389–441. Amsterdam: Elsevier.

Guiso, L., P. Sapienza, and L. Zingales (2006). Does Culture Affect Economic Outcomes? *Journal of Economic Perspectives 20*(2), 23–48.

Hamermesh, D. S. (2002). Timing, Togetherness and Time Windfalls. *Journal of Population Economics 15*(4), 601–623.

Hamermesh, D. S. (2020). Life Satisfaction, Loneliness and Togetherness, with an Application to Covid-19 Lock-downs. *Review of Economics of the Household 18*(4), 983–1000.

Hochschild, A. and A. Machung (1989). *The Second Shift: Working Parents and the Revolution at Home*. New York, NY: Viking.

Holmlund, H., H. Rainer, and P. Reich (2023). All Geared towards Success? Cultural Origins of Gender Gaps in Student Achievement. *Journal of Economic Behavior & Organization 209*, 222–242.

Hsieh, C.-T., E. Hurst, C. I. Jones, and P. J. Klenow (2019). The Allocation of Talent and U.S. Economic Growth. *Econometrica 87*(5), 1439–1474.

Ichino, A., M. Olsson, B. Petrongolo, and P. Skogman Thoursie (2024). Taxes, Childcare and Gender Identity Norms. *Journal of Labor Economics, forthcoming*.

Jenkins, S. P. (2020). Better Off? Distributional Comparisons for Ordinal Data about Personal Well-being. *New Zealand Economic Papers 54*(3), 211–238.

Johnston, D. W., S. Schurer, and M. A. Shields (2014). Maternal Gender Role Attitudes, Human Capital Investment, and Labour Supply of Sons and Daughters. *Oxford Economic Papers 66*(3), 631–659.

Kimmel, J. and R. Connelly (2007). Determinants of Mothers' Time Choices in the United States: Caregiving, Leisure, Home Production, and Paid Work. *Journal of Human Resources 42*(3), 643–681.

Kleven, H., C. Landais, and J. E. Søgaard (2019). Children and Gender Inequality: Evidence from Denmark. *American Economic Journal: Applied Economics 11*(4), 181–209.

Koşar, G., T. Ransom, and W. Van der Klaauw (2022). Understanding Migration Aversion Using Elicited Counterfactual Choice Probabilities. *Journal of Econometrics 231*(1), 123–147.

Kotsadam, A. and H. Finseraas (2011). The state intervenes in the battle of the sexes: Causal effects of paternity leave. *Social science research 40*(6), 1611–1622.

Kuhfeld, W. F., R. D. Tobias, and M. Garratt (1994). Efficient Experimental Design with Marketing Research Applications. *Journal of Marketing Research 31*(4), 545–557.

La Ferrara, E., A. Chong, and S. Duryea (2012). Soap Operas and Fertility: Evidence from Brazil. *American Economic Journal: Applied Economics 4*(4), 1–31.

Lise, J. and K. Yamada (2019). Household Sharing and Commitment: Evidence from Panel Data on Individual Expenditures and Time Use. *Review of Economic Studies 86*(5), 2184–2219.

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

Mansour, H. and T. McKinnish (2014). Couples' Time Together: Complementarities in Production versus Complementarities in Consumption. *Journal of Population Economics 27*(4), 1127–1144.

Maurin, E. and J. Moschion (2009). The Social Multiplier and Labour Market Participation of Mothers. *American Economic Journal: Applied Economics 1*(1), 251–272.

McFadden, D. (1974). Conditional Logit Analysis of Qualitative Choice Behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics*, pp. 105–142. New York: Academic Press.

McFadden, D. and K. Train (2000). Mixed MNL Models for Discrete Response. *Journal of Applied Econometrics 15*(5), 447–470.

Mott, F. L. and S. F. Moore (1979). The Causes of Marital Disruption Among Young American Women: An Interdisciplinary Perspective. *Journal of Marriage and the Family 41*(2), 355–365.

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

Nicoletti, C., K. G. Salvanes, and E. Tominey (2018). The Family Peer Effect on Mothers' Labor Supply. *American Economic Journal: Applied Economics 10*(3), 206–234.

Nollenberger, N., N. Rodríguez-Planas, and A. Sevilla (2016). The Math Gender Gap: The Role of Culture. *American Economic Review 106*(5), 257–261.

Olivetti, C., E. Patacchini, and Y. Zenou (2020). Mothers, Peers, and Gender-Role Identity. *Journal of the European Economic Association 18*(1), 266–301.

Olivetti, C. and B. Petrongolo (2016). The Evolution of the Gender Gap in Industrialized Countries. *Annual Review of Economics 8*(September), 405–434.

Petrin, A. and K. Train (2010). A Control Function Approach to Endogeneity in Consumer Choice Models. *Journal of Marketing Research 47*(1), 3–13.

Pollak, R. A. (2013). Allocating Household Time: When Does Efficiency Imply Specialization? NBER Working Paper No. 19178, National Bureau of Economic Research.

Porter, C. and D. Serra (2019). Gender Differences in the Choice of Major: The Importance of Female Role Models. *American Economic Journal: Applied Economics 12*(3), 226–254.

Schröder, C. and S. Yitzhaki (2017). Revisiting the Evidence for Cardinal Treatment of Ordinal Variables. *European Economic Review 92*, 337–358.

Seedat, S. and M. Rondon (2021). Women's Wellbeing and the Burden of Unpaid Work. *BMJ 374*(1972).

Sevilla-Sanz, A., J. I. Gimenez-Nadal, and C. Fernández (2010). Gender Roles and the Division of Unpaid Work in Spanish Households. *Feminist Economics 16*(4), 137–184.

Siminski, P. and R. Yetsenga (2022). Specialization, Comparative Advantage, and the Sexual Division of Labor. *Journal of Labor Economics 40*(4), 851–887.

Stevenson, B. and J. Wolfers (2007). Marriage and Divorce: Changes and Their Driving Forces. *Journal of Economic Perspectives 21*(2), 27–52.

Stevenson, B. and J. Wolfers (2009). The Paradox of Declining Female Happiness. *American Economic Journal: Economic Policy 1*(2), 190–225.

Sullivan, O. (2019). Gender Inequality in Work-Family Balance. *Nature Human Behaviour 3*, 201–203.

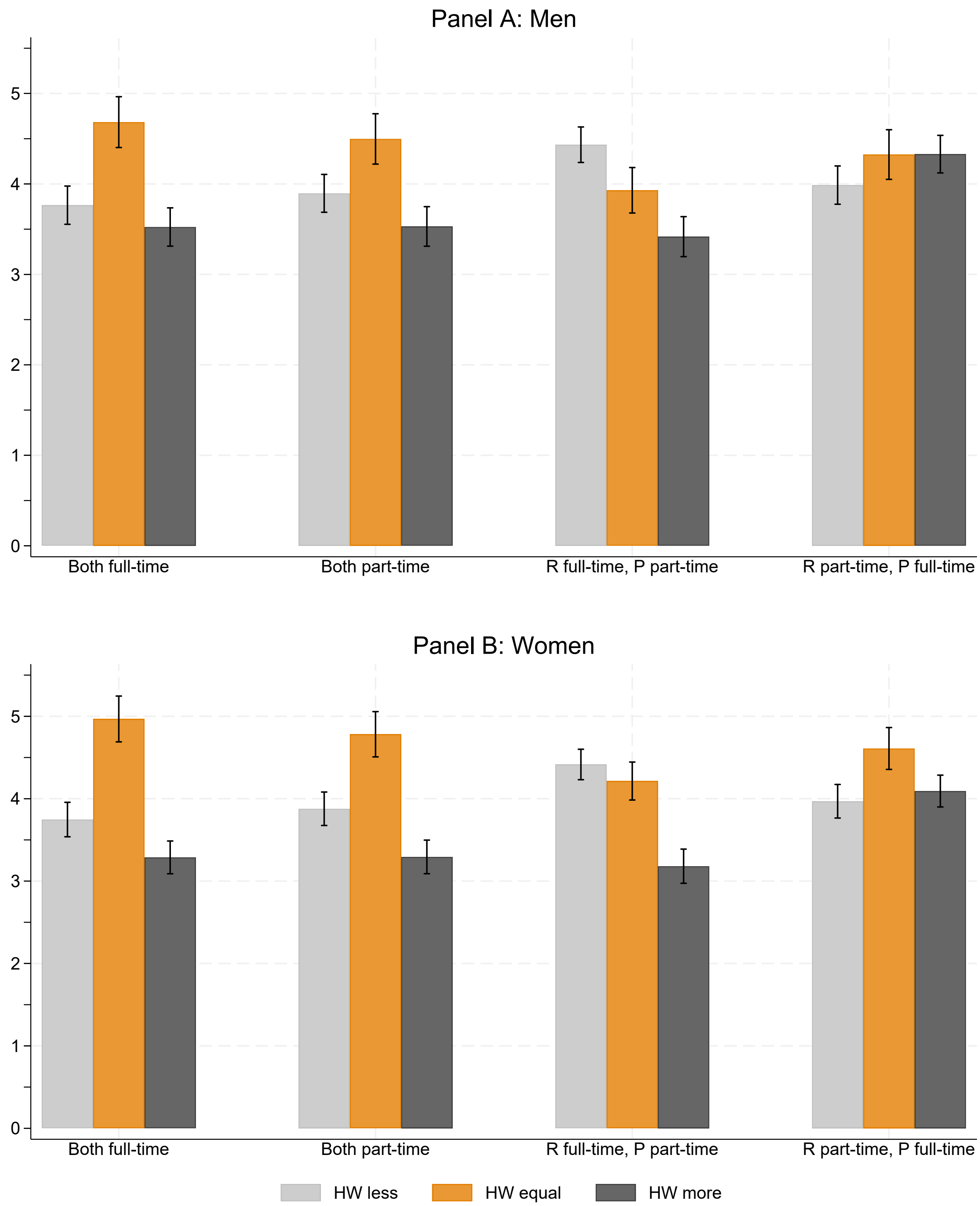
Train, K. E. (2009). *Discrete Choice Methods with Simulation.* Cambridge University Press.

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

Wooldridge, J. M. (2015). Control Function Methods in Applied Econometrics. *Journal of Human Resources 50*(2), 420–445.

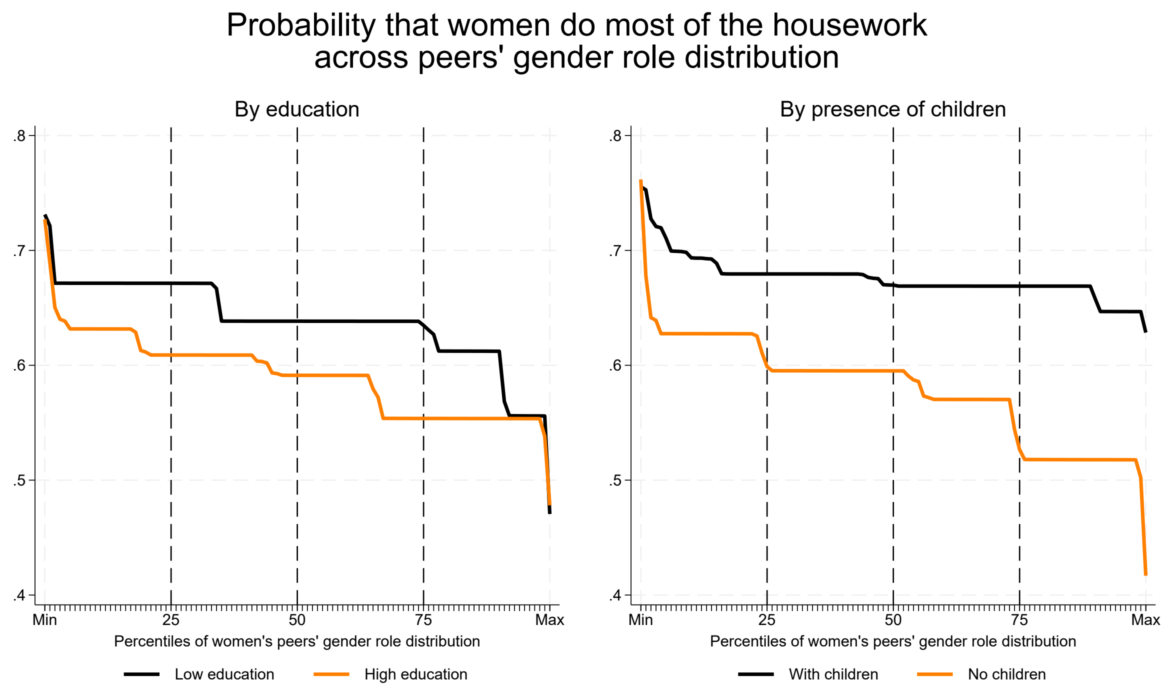
Zick, C. D., W. K. Bryant, and S. Srisukhumbowornchai (2008). Does Housework Matter Anymore? The Shifting Impact of Housework on Economic Inequality. *Review of Economics of the Household 6*(1), 1–28.

Figure 1: Predicted Individual Utility over Domestic Chores Arrangements by Hypothetical Housework, Labor Market Involvement, and Gender



Notes: Predictions are obtained for couples with no children, equal hourly wages, and no reliance on paid housework. The whiskers in each bar display the 95% confidence interval around each prediction. "HW" refers to housework, "R" to respondent, and "P" to partner.

Figure 2: Simulated Housework Shares across the Distribution of Peers' Gender Role Norms, by Education and Motherhood Status



Notes: A household is defined to have 'high' education if at least one partner has a university (or higher) degree.

Table 1: Summary Statistics

Variable	Mean	SD
Housework arrangements ($y_{i,j}$):		
Woman does less (j = 1)	0.143	
Equal division (j = 2)	0.141	
Woman does more (j = 3)	0.717	
Key determinants:		
Mans' peers' gender role norms ($N_{i,m}$)	26.472	1.000
Womans' peers' gender role norms ($N_{i,f}$)	32.295	1.000
Man's economic utility, standardized ($u^m_{i,j}$)	8.663	0.846
Woman's economic utility, standardized ($u^f_{i,j}$)	8.230	1.091
Household utility ($U^E_{i,j} = (u^m_{i,j} + u^f_{i,j})$)	16.893	1.793
Instrumental variables:		
Man's peers' mothers working (proportion)	0.638	
Woman's peers' mothers working (proportion)	0.663	
Individual utility predictors ($\mathbf{X}_i$):		
Paid work indicators (base category: Both partners work full-time):		
Both partners work part-time	0.045	
Man works full-time, woman works part-time	0.464	
Man works part-time, woman works full-time	0.032	
Hourly wages indicators (base category: About equal):		
Man's hourly wages lower than his partner	0.329	
Woman's hourly wages lower than her partner	0.445	
Indicators for child presence and age (base category: No child):		
Youngest child, 0–2 years of age	0.155	
Youngest child, 3–11 years of age	0.274	
Youngest child, 12–15 years of age	0.101	
Household pays for help with housework (=1, 0 otherwise)	0.097	
Control variables:		
Household nonlabor income	0.213	0.289
Country of residence (base category: England):		
Wales	0.073	
Scotland	0.106	
Oldest partner's year of birth (base category: ≤1960):		
1961–1970	0.363	
1971–1980	0.289	
1981+	0.073	
Age of oldest partner	43.879	8.873
Number of (household×wave) observations	7,621	
Number of couples	5,150	

Sources: Innovation Panel, 2012 and 2013; UKHLS, waves 2 (2010–12) and 4 (2012–14).
Notes: u^m_i and u^f_i are predicted using the estimates from (9) (see also Table 3). Standardized utilities are computed dividing each of them by the overall standard deviation in the main sample (not separately by gender). The peers' gender role norms ($N_{i,m}$ and $N_{i,f}$) are standardized so as to have SD equal to 1. Nonlabor income is adjusted using the OECD equivalence scale and transformed using the inverse hyperbolic sine transformation.

Table 2: Distribution of Housework by Partners' Labor Market Involvement

	Woman less	Equal	Woman more	Observations (col. pct'age)
Both FT	0.187	0.179	0.635	3,502 (0.460)
Both PT	0.116	0.157	0.727	344 (0.045)
Man FT, Woman PT	0.086	0.099	0.815	3,533 (0.464)
Man PT, Woman FT	0.368	0.182	0.450	242 (0.032)
Observations [row pct'age]	1,087 [0.143]	1,073 [0.141]	5,461 [0.717]	7,621 [100]

Sources: UKHLS, waves 2 (2010–12) and 4 (2012–14).
Notes: Each cell shows a row probability. Figures in parentheses in the right-most column are column percentages and correspond to the market work distribution shown at the top of the fourth panel of Table 1. Figures in square brackets in the bottom row are row percentages and correspond to the unconditional distribution of housework arrangements reported in the top panel of Table 1. FT and PT stand for full-time and part-time employment, respectively.

Table 3: Individual Economic Utility: Experimental Random Effects Estimates

Variable	Coefficient	(S.E.)
λ_0	4.684***	(0.147)
λ_t	-0.001	(0.058)
Female (ψ)	0.285**	(0.120)
Housework arrangements [$\mathbb{1}(y_{ist} = j)$]:		
Lower share ($j = 1$; β_1)	-0.918***	(0.154)
Higher share ($j = 3$; β_3)	-1.159***	(0.151)
Labor market involvement ($\mathbf{X}^W$)		
Both partners work PT (φ_1^W)	-0.185	(0.161)
Respondent works FT, partner works PT (φ_2^W)	-0.753***	(0.147)
Respondent works PT, partner works FT (φ_3^W)	-0.358**	(0.162)
Relative hourly wages ($\mathbf{X}^E$):		
Respondent earns half (φ_1^E)	0.075	(0.052)
Respondent earns twice (φ_3^E)	0.122**	(0.057)
Paid housework ($\mathbf{X}^H$) (φ^H)	0.136***	(0.044)
Presence and age of children ($\mathbf{X}^K$):		
6-month (φ_1^K)	-0.000	(0.062)
5-year (φ_2^K)	0.084	(0.060)
15-year (φ_3^K)	-0.018	(0.058)
Housework arrangements by gender [$d_i \times \mathbb{1}(y_{ist} = j)$]:		
Lower housework share × female (δ_1)	-0.303**	(0.134)
Higher housework share × female (δ_3)	-0.521***	(0.134)
Housework arrangements by work involvement [$\mathbf{X}_i^W \times \mathbb{1}(y_{ist} = j)$]:		
Lower housework share, both PT (θ_{11}^W)	0.317*	(0.191)
Lower housework share, FT-PT (θ_{12}^W)	1.422***	(0.177)
Lower housework share, PT-FT (θ_{13}^W)	0.580***	(0.191)
Higher housework share, both PT (θ_{31}^W)	0.191	(0.190)
Higher housework share, FT-PT (θ_{32}^W)	0.646***	(0.180)
Higher housework share, PT-FT (θ_{33}^W)	1.163***	(0.192)
Hausman test [p-value]	7.518	[0.995]
F-test [p-value]	363.213	[0.000]
Number of (household×wave) observations	5,955	
Number of individuals	1,245	

Source: Innovation Panel from UKHLS.
Notes: Dependent variable is the respondent's reported level of satisfaction with household arrangements. Robust standard errors clustered at the individual level are in parentheses. FT and PT stand for full-time and part-time employment, respectively. In square brackets are p-values of tests. Hausman test is for the null hypothesis of equality of fixed effects and random effects specifications. F-test is for the null hypothesis of zero effect for all predictors. To ease the reading of the results, variables and parameters used in equation (9) are reported close to the variable names. See the notes to Table 1 and the text for the rest of the notation.
* p<0.10, ** p<0.05, *** p<0.01

Table 4: Marginal Effects from the Alternative-Specific Conditional Logit Model

	Average marginal effects on $P(y=j)$		
	(a) Woman does less $\frac{\partial P(y=1)}{\partial x}$	(b) Equal share $\frac{\partial P(y=2)}{\partial x}$	(c) Woman does more $\frac{\partial P(y=3)}{\partial x}$
$U_{i,j}^{E}$ (household utility) associated with:			
Woman does less ($y=1$)	0.024	-0.004	-0.020
	[0.000]	[0.000]	[0.000]
Equal share ($y=2$)	-0.004	0.024	-0.019
	[0.000]	[0.000]	[0.000]
Woman does more ($y=3$)	-0.020	-0.019	0.039
	[0.000]	[0.000]	[0.000]
$N_{i,m}$ (man's peers' gender norms)	0.000	0.005	-0.006
	[0.966]	[0.422]	[0.542]
$N_{i,f}$ (woman's peers' gender norms)	0.018	0.008	-0.025
	[0.002]	[0.285]	[0.003]
Baseline unconditional probability	$P(y=1) = 0.143$	$P(y=2) = 0.141$	$P(y=3) = 0.717$
Cross-gender differences on peers' gender norms coefficients			
Woman does less ($y=1$)		0.78 [0.378]	
Woman does more ($y=3$)		0.19 [0.661]	
Significance of peers' gender norms coefficients			
$N_{i,m}$ (man's peers' gender norms)		0.64 [0.725]	
$N_{i,f}$ (woman's peers' gender norms		12.07 [0.002]	
Number of (household×wave) observations		7,621	

Notes: The figures in each cell of the top panel are average marginal effects from the estimation of the alternative-specific conditional logit model (7). In square brackets are their corresponding p-values. Inference is heteroscedasticity-robust as we allow clustered errors within cohort of birth and ethnicity. The analysis includes all the control variables listed in Table 1. See the notes to Table 1 and the text for the rest of the notation.

Table 5: Tests for instrumental variable estimation

	First stage F-test	Exogeneity test
Relaxing the exogeneity assumption of peer's gender norms in the specification in		
Table 4	91.37	2.38
	[0.000]	[0.666]
Table 6	91.37	1.77
	[0.000]	[0.412]
Table 8, Panel A	60.29	2.01
	[0.000]	[0.366]
Table 8, Panel B	95.24	2.18
	[0.000]	[0.336]
Table 9, Panel A	96.90	3.95
	[0.000]	[0.139]
Table 9, Panel B	63.12	0.24
	[0.000]	[0.885]

Notes: The second column reports the F-statistic and the corresponding p-value testing the joint insignificance of the instruments in the first-stage equations. The third column reports the Chi-squared statistic and the corresponding p-value testing the joint insignificance of the control functions in the second stage equations. Inference is heteroscedasticity-robust as we allow clustered errors within cohort of birth and ethnicity. The analysis includes all the control variables listed in Tables 4, 6, 8, 9.

Table 6: Marginal Effects from the Alternative-Specific Conditional Logit Model

	Average marginal effects on $P(y=j)$		
	(a) Woman does less $\frac{\partial P(y=1)}{\partial x}$	(b) Equal share $\frac{\partial P(y=2)}{\partial x}$	(c) Woman does more $\frac{\partial P(y=3)}{\partial x}$
$U^E_{i,j}$ (household utility) associated with:			
Woman does less ($y=1$)	0.024	-0.004	-0.020
	[0.000]	[0.000]	[0.000]
Equal share ($y=2$)	-0.004	0.024	-0.019
	[0.000]	[0.000]	[0.000]
Woman does more ($y=3$)	-0.020	-0.019	0.039
	[0.000]	[0.000]	[0.000]
$N_{i,f}$ (woman's peers' gender norms)	0.018	0.010	-0.028
	[0.001]	[0.101]	[0.000]
Baseline unconditional probability	$P(y=1)=0.143$	$P(y=2)=0.141$	$P(y=3)=0.717$
Number of (household×wave) observations		7,621	

Notes: The figures in each cell of the top panel are average marginal effects from the estimation of the alternative-specific conditional logit model (7). In square brackets are their corresponding p-values. Inference is heteroscedasticity-robust as we allow clustered errors within cohort of birth and ethnicity. The analysis includes all the control variables listed in Table 1. See the notes to Table 1 and the text for the rest of the notation.

Table 7: Quantifying the Impact of an Increase in Peers' Gender Role Attitudes

	Woman less	Equal share	Woman more
(a) Observed paid work distribution	0.464	0.505	0.032
(b) Observed housework distribution	0.143	0.141	0.717
(c) "Gender fair" housework allocation	0.032	0.505	0.464
(d) Excess housework [(b)-(c)]	for men 0.111	/	for women 0.253
Simulation: $\uparrow N_{i,f}$ by 1SD			
(e) Counterfactual housework distribution	0.161	0.151	0.689
(f) Counterfactual excess housework [(e)-(c)]	for men 0.129	/	for women 0.225

Notes: For rows (a) and (b), refer to the summary statistics shown in Table 1. The "gender fair" allocation in row (c) mirrors the observed paid work distribution in row (a). "Excess housework" in row (d) is the difference between the figures in row (b) and the corresponding figures in row (c). In the bottom panel, row (e) reports the counterfactual housework distribution found after increasing her peers' gender role norms, $N_{i,f}$, by 1SD. Row (f) shows the counterfactual excess housework distribution, found by subtracting (c) from (e). See the text for other details.

Table 8: Marginal Effects from the Alternative-Specific Conditional Logit Model by Education

	Average marginal effects on $P(y=j)$		
	(a) Woman does less $\frac{\partial P(y=1)}{\partial x}$	(b) Equal share $\frac{\partial P(y=2)}{\partial x}$	(c) Woman does more $\frac{\partial P(y=3)}{\partial x}$
Panel A: Low-education households			
$U^E_{i,j}$ (household utility) associated with:			
Woman does less ($y=1$)	0.020	-0.003	-0.017
	[0.000]	[0.000]	[0.000]
Equal share ($y=2$)	-0.003	0.023	-0.020
	[0.000]	[0.000]	[0.000]
Woman does more ($y=3$)	-0.017	-0.020	0.037
	[0.000]	[0.000]	[0.000]
$N_{i,f}$ (peers' gender norms)	0.006	0.025	-0.031
	[0.470]	[0.018]	[0.037]
Baseline unconditional probability	$P(y=1)=0.105$	$P(y=2)=0.128$	$P(y=3)=0.767$
Number of (household×wave) observations		2,643	
Panel B: High-education households			
$U^E_{i,j}$ (household utility) associated with:			
Woman does less ($y=1$)	0.026	-0.005	-0.021
	[0.000]	[0.000]	[0.000]
Equal share ($y=2$)	-0.005	0.023	-0.019
	[0.000]	[0.000]	[0.000]
Woman does more ($y=3$)	-0.021	-0.019	0.039
	[0.000]	[0.000]	[0.000]
$N_{i,f}$ (man's peers' gender norms)	0.024	0.005	-0.029
	[0.005]	[0.460]	[0.001]
Baseline unconditional probability	$P(y=1)=0.163$	$P(y=2)=0.147$	$P(y=3)=0.690$
Number of (household×wave) observations		4,978	

Notes: High-education households are those in which at least one partner has a university degree. For other details, see the notes to Table 4.

Table 9: Marginal Effects from the Alternative-Specific Conditional Logit Model by Child Presence

	Average marginal effects on $P(y=j)$		
	(a) Woman does less $\frac{\partial P(y=1)}{\partial x}$	(b) Equal share $\frac{\partial P(y=2)}{\partial x}$	(c) Woman does more $\frac{\partial P(y=3)}{\partial x}$
Panel A: Households with children			
$U^E_{i,j}$ (household utility) associated with:			
Woman does less ($y=1$)	0.020	-0.003	-0.017
	[0.000]	[0.000]	[0.000]
Equal share ($y=2$)	-0.003	0.019	-0.016
	[0.000]	[0.000]	[0.000]
Woman does more ($y=3$)	-0.017	-0.016	0.033
	[0.000]	[0.000]	[0.000]
$N_{i,f}$ (peers' gender norms)	0.009	0.006	-0.016
	[0.111]	[0.293]	[0.065]
Baseline unconditional probability	$Pr(y=1)=0.128$	$Pr(y=2)=0.121$	$Pr(y=3)=0.751$
Number of (household×wave) observations		4,040	
Panel B: Childless households			
$U^E_{i,j}$ (household utility) associated with:			
Woman does less ($y=1$)	0.023	-0.005	-0.018
	[0.000]	[0.000]	[0.000]
Equal share ($y=2$)	-0.005	0.023	-0.018
	[0.000]	[0.000]	[0.000]
Woman does more ($y=3$)	-0.018	-0.018	0.036
	[0.000]	[0.000]	[0.000]
$N_{i,f}$ (woman's peers' gender norms)	0.026	0.012	-0.038
	[0.010]	[0.206]	[0.003]
Baseline unconditional probability	$Pr(y=1)=0.159$	$Pr(y=2)=0.163$	$Pr(y=3)=0.677$
Number of (household×wave) observations		3,581	

Notes: For details, see the notes to Table 4.

Table 10: Sensitivity Checks: Changing Definition of Peers

	Average marginal effects on $P(y=j)$		
	(a) Woman does less $\frac{\partial P(y=1)}{\partial x}$	(b) Equal share $\frac{\partial P(y=2)}{\partial x}$	(c) Woman does more $\frac{\partial P(y=3)}{\partial x}$
Panel A: Peers' group defined by gender, cohort of birth and country of birth			
$U^E_{i,j}$ (household utility) associated with:			
Woman does less ($y=1$)	0.024	-0.004	-0.019
	[0.000]	[0.000]	[0.000]
Equal share ($y=2$)	-0.004	0.023	-0.019
	[0.000]	[0.000]	[0.000]
Woman does more ($y=3$)	-0.019	-0.019	0.039
	[0.000]	[0.000]	[0.000]
$N_{i,f}$ (woman's peers' gender norms)	0.017	0.010	-0.027
	[0.001]	[0.071]	[0.000]
Baseline unconditional probability	$Pr(y=1)=0.140$	$Pr(y=2)=0.141$	$Pr(y=3)=0.719$
Number of (household×wave) observations		7,005	
Panel B: Peers' groups with a minimum size of 16			
$U^E_{i,j}$ (household utility) associated with:			
Woman does less ($y=1$)	0.024	-0.004	-0.020
	[0.000]	[0.000]	[0.000]
Equal share ($y=2$)	-0.004	0.023	-0.019
	[0.000]	[0.000]	[0.000]
Woman does more ($y=3$)	-0.020	-0.019	0.039
	[0.000]	[0.000]	[0.000]
$N_{i,f}$ (woman's peers' gender norms)	0.017	0.010	-0.027
	[0.002]	[0.107]	[0.000]
Baseline unconditional probability	$Pr(y=1)=0.142$	$Pr(y=2)=0.140$	$Pr(y=3)=0.717$
Number of (household×wave) observations		7,582	

Notes: Besides England, Wales, Scotland, and Northern Ireland (which make up 87% of the sample), 23 other countries contribute to in the definition of peers. For other details see the text and the notes to Table 4.

Online Appendix

Table A1: Marginal Effects from the Alternative-Specific Conditional Logit Model. Couple Members' Utilities Estimated by Ordered Probit Regressions

	Average marginal effects on $P(y=j)$		
	(a) Woman does less $\frac{\partial P(y=1)}{\partial x}$	(b) Equal share $\frac{\partial P(y=2)}{\partial x}$	(c) Woman does more $\frac{\partial P(y=3)}{\partial x}$
$U^E_{i,j}$ (household utility) associated with:			
Woman does less ($y=1$)	0.024	-0.004	-0.020
	[0.000]	[0.000]	[0.000]
Equal share ($y=2$)	-0.004	0.024	-0.019
	[0.000]	[0.000]	[0.000]
Woman does more ($y=3$)	-0.020	-0.019	0.039
	[0.000]	[0.000]	[0.000]
$N_{i,f}$ (woman's peers' gender norms)	0.018	0.010	-0.028
	[0.001]	[0.101]	[0.000]
Baseline unconditional probability	$P(y=1) = 0.143$	$P(y=2) = 0.141$	$P(y=3) = 0.717$
Number of (household×wave) observations		7,621	

Notes: For details, see the notes to Table 4.

Table A2: Sensitivity Check: Aggregating Gender Norms Dimensions According to the First Component of the Principal Component Analysis

	Average marginal effects on $P(y=j)$		
	(a) Woman does less $\frac{\partial P(y=1)}{\partial x}$	(b) Equal share $\frac{\partial P(y=2)}{\partial x}$	(c) Woman does more $\frac{\partial P(y=3)}{\partial x}$
$U^E_{i,j}$ (household utility) associated with:			
Woman does less ($y=1$)	0.025	-0.004	-0.020
	[0.000]	[0.000]	[0.000]
Equal share ($y=2$)	-0.004	0.024	-0.019
	[0.000]	[0.000]	[0.000]
Woman does more ($y=3$)	-0.020	-0.019	0.040
	[0.000]	[0.000]	[0.000]
$N_{i,f}$ (woman's peers' gender norms)	0.014	0.010	-0.023
	[0.006]	[0.055]	[0.000]
Baseline unconditional probability	$P(y=1)=0.143$	$P(y=2)=0.141$	$P(y=3)=0.717$
Number of (household×wave) observations		7,621	

Notes: For details, see the notes to Table 4.

Table A3: Sensitivity Check: Defining Gender Role Attitudes Focusing on Specific Sub-Items.

	Average marginal effects on $P(y=j)$		
	(a) Woman does less $\frac{\partial P(y=1)}{\partial x}$	(b) Equal share $\frac{\partial P(y=2)}{\partial x}$	(c) Woman does more $\frac{\partial P(y=3)}{\partial x}$
Panel A: Peers' gender role indicator considering only items about work arrangements between partners			
$V_{i,j}^{E}$ (household utility) associated with:			
Woman does less ($y=1$)	0.024	-0.004	-0.020
	[0.000]	[0.000]	[0.000]
Equal share ($y=2$)	-0.004	0.023	-0.019
	[0.000]	[0.000]	[0.000]
Woman does more ($y=3$)	-0.020	-0.019	0.039
	[0.000]	[0.000]	[0.000]
$N_{i,f}$ (woman's peers' gender norms)	0.014	0.004	-0.017
	[0.003]	[0.510]	[0.008]
Baseline unconditional probability	$P(y=1) = 0.143$	$P(y=2) = 0.141$	$P(y=3) = 0.717$
Number of (household×wave) observations		7,621	
Panel B: Peers' gender role indicator considering items on work arrangements between partners and on women's work hours			
$V_{i,j}^{E}$ (household utility) associated with:			
Woman does less ($y=1$)	0.024	-0.004	-0.020
	[0.000]	[0.000]	[0.000]
Equal share ($y=2$)	-0.004	0.024	-0.019
	[0.000]	[0.000]	[0.000]
Woman does more ($y=3$)	-0.020	-0.019	0.039
	[0.000]	[0.000]	[0.000]
$N_{i,f}$ (woman's peers' gender norms)	0.017	0.007	-0.024
	[0.001]	[0.199]	[0.001]
Baseline unconditional probability	$P(y=1) = 0.143$	$P(y=2) = 0.141$	$P(y=3) = 0.717$
Number of (household×wave) observations		7,621	

Notes: Panel A: Peers' gender role indicator based on the items "Both the husband and wife should contribute to the household income" and "A husband's job is to earn money, a wife's job is to look after the home and family". Panel B: Peers' gender role indicator based on the items of Panel A plus "All in all, family life suffers when the woman has a full-time job".

Table A4: Marginal Effects from the Alternative-Specific Conditional Logit Model by Men's Peers' Gender Role

	Average marginal effects on $P(y = j)$		
	(a) Woman does less $\frac{\partial P(y=1)}{\partial x}$	(b) Equal share $\frac{\partial P(y=2)}{\partial x}$	(c) Woman does more $\frac{\partial P(y=3)}{\partial x}$
Panel A: Households with men's peers' having egalitarian gender role			
$U^E_{i,j}$ (household utility) associated with:			
Woman does less ($y = 1$)	0.027	-0.005	-0.022
	[0.000]	[0.000]	[0.000]
Equal share ($y = 2$)	-0.005	0.025	-0.020
	[0.000]	[0.000]	[0.000]
Woman does more ($y = 3$)	-0.022	-0.020	0.042
	[0.000]	[0.000]	[0.000]
$N_{i,f}$ (woman's peers' gender norms)	0.012	0.003	-0.015
	[0.217]	[0.714]	[0.144]
Baseline unconditional probability	$P(y=1) = 0.155$	$P(y=2) = 0.149$	$P(y=3) = 0.696$
Number of (household×wave) observations		3,550	
Panel B: Households with men's peers' having non-egalitarian gender role			
$U^E_{i,j}$ (household utility) associated with:			
Woman does less ($y = 1$)	0.022	-0.004	-0.018
	[0.000]	[0.000]	[0.000]
Equal share ($y = 2$)	-0.004	0.022	-0.018
	[0.000]	[0.000]	[0.000]
Woman does more ($y = 3$)	-0.018	-0.018	0.037
	[0.000]	[0.000]	[0.000]
$N_{i,f}$ (woman's peers' gender norms)	0.019	0.014	-0.033
	[0.010]	[0.104]	[0.002]
Baseline unconditional probability	$P(y=1) = 0.132$	$P(y=2) = 0.134$	$P(y=3) = 0.734$
Number of (household×wave) observations		4,071	

Notes: Men's peers with egalitarian gender roles are those with gender role higher than the median. For other details, see the notes to Table 4.

Table A5: Marginal Effects from the Alternative-Specific Conditional Logit Model by Women's Peers' Gender Role

	Average marginal effects on $P(y=j)$		
	(a) Woman does less $\frac{\partial P(y=1)}{\partial x}$	(b) Equal share $\frac{\partial P(y=2)}{\partial x}$	(c) Woman does more $\frac{\partial P(y=3)}{\partial x}$
Panel A: Households with women's peers' having egalitarian gender role			
$U^E_{i,j}$ (household utility) associated with:			
Woman does less ($y=1$)	0.029	-0.006	-0.023
	[0.000]	[0.000]	[0.000]
Equal share ($y=2$)	-0.006	0.028	-0.022
	[0.000]	[0.000]	[0.000]
Woman does more ($y=3$)	-0.023	-0.022	0.045
	[0.000]	[0.000]	[0.000]
$N_{i,f}$ (woman's peers' gender norms)	0.011	0.015	-0.025
	[0.305]	[0.380]	[0.137]
Baseline unconditional probability	$P(y=1)=0.157$	$P(y=2)=0.154$	$P(y=3)=0.689$
Number of (household×wave) observations		3,757	
Panel B: Households with women's peers' having non-egalitarian gender role			
$U^E_{i,j}$ (household utility) associated with:			
Woman does less ($y=1$)	0.021	-0.003	-0.017
	[0.000]	[0.000]	[0.000]
Equal share ($y=2$)	-0.003	0.020	-0.017
	[0.000]	[0.000]	[0.000]
Woman does more ($y=3$)	-0.017	-0.017	0.034
	[0.000]	[0.000]	[0.000]
$N_{i,f}$ (woman's peers' gender norms)	0.023	0.014	-0.037
	[0.021]	[0.141]	[0.009]
Baseline unconditional probability	$P(y=1)=0.129$	$P(y=2)=0.128$	$P(y=3)=0.744$
Number of (household×wave) observations		3,864	

Notes: Women's peers with egalitarian gender roles are those with gender role indicator higher than the median. For other details, see the notes to Table 4.

Table A6: Sensitivity Check: Using Alternative Cut-Offs for Housework Division and Relaxing Linearity

	Average marginal effects on $P(y=j)$		
	(a) Woman does less $\frac{\partial P(y=1)}{\partial x}$	(b) Equal share $\frac{\partial P(y=2)}{\partial x}$	(c) Woman does more $\frac{\partial P(y=3)}{\partial x}$
Panel A: Alternative cut-offs for housework division (Equal share if the share lies between 40% and 60%)			
$U_{i,j}^E$ (household utility) associated with:			
Woman does less ($y=1$)	0.016	-0.005	-0.011
	[0.000]	[0.000]	[0.000]
Equal share ($y=2$)	-0.005	0.037	-0.031
	[0.000]	[0.000]	[0.000]
Woman does more ($y=3$)	-0.011	-0.031	0.042
	[0.000]	[0.000]	[0.000]
$N_{i,f}$ (woman's peers' gender norms)	0.015	0.013	-0.028
	[0.000]	[0.090]	[0.001]
Baseline unconditional probability	$P(y=1) = 0.093$	$P(y=2) = 0.277$	$P(y=3) = 0.630$
Number of (household×wave) observations		7,621	
Panel B: Including second order polynomial of woman's peers' gender role			
$U_{i,j}^E$ (household utility) associated with:			
Woman does less ($y=1$)	0.024	-0.004	-0.020
	[0.000]	[0.000]	[0.000]
Equal share ($y=2$)	-0.004	0.024	-0.019
	[0.000]	[0.000]	[0.000]
Woman does more ($y=3$)	-0.020	-0.019	0.039
	[0.000]	[0.000]	[0.000]
$N_{i,f}$ (woman's peers' gender norms)	0.016	0.011	-0.027
	[0.003]	[0.188]	[0.002]
Baseline unconditional probability	$P(y=1) = 0.143$	$P(y=2) = 0.141$	$P(y=3) = 0.717$
Number of (household×wave) observations		7,621	

Notes: For other details, see the notes to Table 4.