Journal of Economic Behavior and Organization 186 (2021) 164–188

Contents lists available at ScienceDirect

Journal of Economic Behavior and Organization

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

Boys don't cry (or do the dishes): Family size and the housework gender gap☆

Giorgia Menta*, Anthony Lepinteur

University of Luxembourg, Luxembourg

A R T I C L E I N F O

Article history:
Received 8 December 2020
Revised 12 March 2021
Accepted 24 March 2021

JEL classification:
J13
J16
D13

Keywords:
Housework gender gap
Family size
Instrumental variables

A B S T R A C T

We here use data from the British Cohort Study (BCS) to link family size to age-16 children's contribution to household chores and the adult housework gender gap. Assuming that home production is an increasing function of family size and using an instrument to account for the endogeneity of fertility, we show that larger families have a different effect on boys and girls at age 16: girls in large families are significantly more likely to contribute to housework, with no effect for boys. We then show that childhood family size affects the housework gender gap between the cohort members and their partners at age 34. Women who grew up in larger families are more likely to carry out a greater share of household tasks in adulthood, as compared to women from smaller families. In addition, growing up in a large family makes cohort members more likely to sort into households with a wider housework gender gap as adults. We show that the persistent effect of family size is due to the adoption of behaviours in line with traditional gender roles: a lower likelihood of employment and shorter commutes for women, along with a higher employment probability for their partners.

1. Introduction

Recent decades have seen a shift in the distribution of housework within couples. The time women devote to household chores has fallen, while men's participation in housework has risen. However, the housework gender gap is yet to be closed in most countries: the cross-national trends in Altintas and Sullivan (2016) show that convergence has stalled since the 1980s, especially in those countries where the gap was initially smaller. The burden of housework and childcare continues to disproportionately weigh on women, with consequences in terms of labour-market outcomes and well-being. Using data from the Multinational Time Use survey, Sayer (2010) shows that women in the early 2000s carried out 1.5 to 2 times as

☆ We would like to thank Michel Beine, Sonia Bhalotra, Liyousew Borga, Alessandra Casarico, Pierre-André Chiappori, Conchita D'Ambrosio, Arnaud Dupuy, Daniela Sonedda, Alexandros Theloudis and Philippe Van Kerm for their useful comments. We are particularly grateful to Andrew Clark for the constructive remarks and for helping us make the paper a less painful read. We are also grateful to the participants of the 11th Workshop on Labour Economics (IAAEU Trier), the 33rd Annual ESPE Conference, the 8th ECINEQ Meeting, the 31st EALE Conference, the University of Luxembourg PhD Jamboree, the MINLAB PhD Workshop in Labour Economics, the Workshop on Gender and Economics (CREA), and the ISER Internal Seminar Series. We gratefully acknowledge financial support from the Fonds National de la Recherche Luxembourg FNR/P11/05, FNR/P11/05BIS, and funding 10949242. The authors would like to assure the reader that equal housework participation was ensured throughout the drafting of this paper.

* Corresponding author.
E-mail address: giorgia.menta@uni.lu (G. Menta).

https://doi.org/10.1016/j.jebo.2021.03.035
0167-2681/

much housework as men in developed countries. Extending the analysis to the more recent waves of the American Time Use Survey, Bianchi et al. (2012) confirm that a decade later American women were still responsible for about 1.6 times more housework than men. Along the same lines, McMunn et al. (2020) show that, in 2010, women in 93% of British couples spent more time on housework than their partners.

Standard models of household decision-making suggest that differences in bargaining power, from labour-market earnings and non-market work productivity, help determine intra-household time allocation (Chiappori, 1992; Van Klaveren et al., 2008). Female labour-force participation and educational outcomes are at a historical high in most OECD economies (International Labour Office, 2018; World Economic Forum, 2018), with considerable consequences for women's bargaining power and the quality of their outside options (Antman, 2014; Bittman et al., 2003). A large body of empirical work shows that the time spent in home production falls as either absolute or relative earnings rise (Bittman et al., 2003; Gupta, 2007; Gupta and Ash, 2008; Bertrand et al., 2015) and that, conditional on being employed, educated women participate less in housework (Baxter et al., 2008). Despite this progress, why do women then still devote disproportionately more time to housework than men?

Many researchers have turned their attention towards explanations based on gender identity formation (Akerlof and Kranton, 2000; 2010). In societies where the definition of masculinity is based on the principle that men should be the family breadwinners and should not engage in 'feminine' forms of housework, individuals will find it costly to adopt behaviours that deviate from this prescription, as it would be at odds with their identity and would translate into a utility loss. This is consistent with empirical work showing that women who are more educated or earn more than their partners, and who so deviate from gender-role prescriptions, compensate via a more traditional division of housework (Bittman et al., 2003; Lyonette and Crompton, 2015; Bertrand et al., 2015). Considering the housework gender gap as a by-product of utility-maximising behaviour is however at odds with the literature on subjective well-being: Flèche et al. (2018, 2020) show that the housework gender gap can be perceived as unfair and, as such, it produces lower levels of happiness and marital stability.

Comparatively little is known about the role of childhood characteristics in shaping adult differences in housework participation. Based on the intergenerational cultural socialisation framework of Bisin and Verdier (2001), some authors have shown that children's perception of gender roles are directly linked to their parents' attitudes, contributing to the persistence of unequal gender norms (Farré and Vella, 2013). Children's socialisation into traditional gender roles can also happen indirectly, as a result of the household's demographic structure. While we know that decisions involving marital status and fertility affect adults' labour-force participation (Angrist and Evans, 1998; Cruces and Galiani, 2007; Bloom et al., 2009; Baxter et al., 2008), the concomitant effects on intra-household time allocation may well involve not only parents but also children. We here consider the role of family size: assuming that the amount of housework rises with family size (Blundell et al., 2005; Cherchye et al., 2012), then the time that parents move out of the labour market may not suffice to satisfy the greater demand for home production, so that children may be asked to step in and contribute more to housework (Brody and Steelman, 1985). If the effect of family size on children's housework contribution depends on their gender, a larger family size might then feed through to the adult housework gender gap, through factors such as educational achievement, future labour-market outcomes, fertility and gender attitudes.

To the best of our knowledge, the causal impact of family size on the allocation of childhood household tasks, and the persistence of this effect in adulthood, has not been explored. We address the endogeneity of family size via an instrumental-variables approach, as in Angrist and Evans (1998). In the latter the impact of fertility on women's labour supply in the US is considered using an instrument reflecting parental preferences for child sex composition: parents whose first two children are of the same sex are more likely to have a third. Similarly, we restrict the analysis to families with two or more children and exploit parents' preferences for variety in the sex mix of the offspring to predict the number of children in the household. When presenting our results, we extensively discuss the validity of the instrument in our context, following Conley et al. (2012) and carrying out a number of additional tests.

In our sample of the 1970 British Cohort Study (BCS), a larger family size during childhood increases the share of housework performed by girls at age 16, but not that of boys. This conclusion is robust to different measures of housework. Girls also consistently spend less time on other activities, namely homework and leisure. The effect of family size is mostly found in low-SES and conservative households. We then show that family size at age 16 also affects the division of household tasks in adulthood: at age 34, women in the BCS who grew up in large families are more likely to perform a larger share of housework as compared to women from smaller families, and they additionally sort into households where the housework gender gap is significantly larger. We again find that the effect of childhood family size is significantly higher for cohort members who grew up in low-SES and conservative households. The results at age 42 are similar. We then argue that this persistence is in large part due to the adoption of behaviours conforming to traditional gender roles: women who grew up in large families are more likely to be not-employed and to have an employed husband.

Our paper contributes to the existing literature in a number of ways. To the best of our knowledge, we are the first to use both cohort data and an instrumental-variable strategy to estimate the causal effect of family size on the contribution of children to household tasks. The richness of our data allows us to explore the effect of family size on the time spent in other activities, such as leisure or homework. Second, we use the same instrumental variable strategy to estimate whether the effect of family size at age 16 is persistent and affects the housework gender gap of the cohort members once partnered at age 34. Last, we consider some of the channels through which childhood family size affects the adult division of housework, namely education, labour-market outcomes and fertility.

The remainder of the paper is organized as follows. Section 2 reviews two strands of the literature: the first on the link between family size and children's contribution to housework, and the second on the influence of family size on a set of determinants of the adult housework gender gap. Section 3 then describes the data and identification strategy, and the empirical results at age 16 appear in Section 4. The results at age 34 are then discussed in Section 5. Last, Section 6 concludes.

2. Literature review

2.1. The determinants of the division of housework among children

Theoretical models of household time-allocation usually consider that only adults carry out household tasks, while children, if anything, create the need for more housework (Blundell et al., 2005; Cherchye et al., 2012). However, time-use surveys reveal that children actually spend a significant amount of time performing household tasks (Peters and Haldeman, 1987; Bianchi and Robinson, 1997). The contribution of children to housework can first be explained by parental time constraints: employed parents may not have sufficient time to handle the housework load and may ask their children to help them. We expect children to be imperfect substitutes for their parents, as they are likely to be less productive than adults and can only contribute to a limited set of tasks. It can also be argued that parents ask their children to help with household tasks as they wish to transmit a set of skills to them and foster their human capital (Blair, 1992a).

The empirical literature on children's contribution to household tasks is small and mostly non-causal. Using US data, Gager et al. (1999) show that girls aged between 3 and 11 spend more time on housework than boys do. Girls also carry out more household tasks when their mother is employed full-time (Peters and Haldeman, 1987; Blair, 1992a), while the evidence is inconclusive for boys (Blair, 1992b). Antill et al. (1996) find that parental involvement in household tasks positively predicts children's housework participation.

We here focus on the effect of family size on the allocation of housework among children: according to Brody and Steelman (1985), this is ambiguous. An additional household member increases the housework load and may lead to parents asking their children to participate to a greater extent. At the same time, an additional child also increases the number of potentially helping hands in the household. The net effect of family size on the housework load per child will then be positive (negative) if the new household member's contribution is higher (lower) than the marginal increase in housework her presence entails.

Using US samples of children aged from 3 to 11 and 12 to 16 respectively, both Bianchi and Robinson (1997) and Gager et al. (1999) find a positive relationship between family size and children's time spent on housework. These papers are the most-closely related to our first empirical question here, but neither addresses endogenenity. Family size is considered as a simple control variable. However, fertility decisions are not random and depend on confounding factors that may also be directly related to the allocation of housework among children.

2.2. The effects of family size in childhood on intra-household time allocation in adulthood

The housework gender gap can be defined as the difference between women and men in the time spent on housework. A body of theoretical and empirical work has aimed to understand why women still devote more time to household tasks than men do. On the theoretical side, both unitary and collective models of household decision-making suggest that the partner with the lowest earnings should spend relatively more time on housework (Stratton, 2015). Bittman et al. (2003); Gupta (2007) and Gupta and Ash (2008) confirm this prediction empirically: women contribute less to housework as their earnings rise. As earnings are positively correlated with human capital, we expect the housework gender gap to be smaller in households where the wife is highly-educated. Baxter et al. (2008) use Australian data to show that women with a Bachelor's degree spend less time on average on household tasks than do women without a Bachelor's degree, conditional on being employed.

While the education gap between men and women has almost closed (World Economic Forum, 2018), the housework gender gap remains. The stream of literature burgeoning from the seminal work on gender identity by Akerlof and Kranton (2000, 2010) attributes part of this persistence to gender norms. In Bittman et al. (2003), couples that deviate from the norm that "*a husband should make more money than his wife*" compensate by a more traditional division of housework in the US and Australia. This finding is corroborated in Bertrand et al. (2015), who also show that, controlling for the absolute level of income, women with a higher probability of out-earning their husbands are less likely to participate in the labour force. An extensive Sociological literature has confirmed that, holding earnings constant, egalitarian attitudes about the gender division of labour are associated with a smaller housework gender gap (see Carlson and Lynch, 2013, for a detailed review).

Baxter (2005) and Baxter et al. (2008) emphasize the role of life-course transitions in the housework gender gap: while men's contribution to household tasks is relatively insensitive to marital status and the number of children, marriage and motherhood significantly increase that of women. Using respectively British and German data, Schober (2011) and Grunow et al. (2012) confirm the asymmetric effect of parenthood on parental contributions to housework. Here again, Schober (2011) shows that parents with more egalitarian gender attitudes share housework more equally.

The three groups of determinants of the housework gender gap described above (i.e. education and earnings, gender norms and demographics) have one thing in common: they are all likely to be influenced by childhood family structure and, as such, are good candidates for mediating an effect of childhood family size on the adulthood division of household tasks.

The paragraphs below review some of the literature describing the relationship between childhood family size and adult outcomes.

Björklund and Salvanes (2011) note that a number of contributions have found large and robust negative associations between family size and different measures of child quality, such as educational achievement and adult labour-market outcomes. This is in line with the theoretical literature on the trade-off between child quality and quantity (Becker, 1960; Becker and Lewis, 1973). However, the use of instrumental variables to address the endogeneity of fertility decisions produces more nuanced results. In Angrist et al. (2010) and Åslund and Grönqvist (2010) there is no causal effect of family size on adult educational achievement or labour-market outcomes, while other authors find negative and significant effects on private-school attendance (Conley and Glauber, 2006; Cáceres-Delpiano, 2006) and IQ (Black et al., 2010).

Anderton et al. (1987), Booth and Kee (2009), Kolk (2014), and Fasang and Raab (2014) find evidence supporting of the intergenerational transmission of fertility decisions. Instrumenting for family size in Norwegian data, Cools and Hart (2017) find a differential effect of childhood family size on adult fertility by gender: an additional sibling increases male fertility but reduces female fertility. The authors argue that this difference comes from mothers reducing their labour supply relatively less when they have a daughter than when they have a son. Cools and Hart (2017) also provides descriptive evidence of a substitution effect, as girls are more likely than boys to help with housework as family size rises. Girls then become more aware of the associated strain of large families and limit their own number of children in adulthood.

One may also expect adulthood gender attitudes to be influenced by childhood family size. We know that family structure and parental background play a role in the intergenerational transmission of gender attitudes. Vella (1994) uncovers a relationship between young womens attitudes towards female employment and her parents' educational backgrounds and labor-market behaviour. Using differences in the male draft across US states as an exogenous source of variation in mothers' labour-force participation, Fernández et al. (2004) argue that men who grew up in families with working mothers develop less stereotypical gender attitudes and are less likely to be the household breadwinner. The intergenerational transmission of gender attitudes can be tested directly by correlating parental gender attitudes with those of their children. Using the NLSY1979, Farré and Vella (2013) find that the mother's views of the role of women, both in the family and in the labour market, affect the views of her children. They also show that mothers with less-traditional views about the role of women are more likely to have working daughters and working daughters-in-law (consistent with Fernández et al., 2004). Using the British Cohort Study, Johnston et al. (2013) test for the external validity of Farré and Vella (2013) and find similar results. Last, Giménez-Nadal et al. (2019) use Russian panel data to infer the gender norms of the parents from the share of housework carried out by the mother, and find that conservative parents have sons and sons-in-law who perform less housework in adulthood. As it affects the allocation of household tasks among boys and girls, family size is then likely to affect children's gender norms.

3. Data and empirical strategy

3.1. The British Cohort Study (BCS)

Our empirical analysis is based on the British Cohort Study (BCS). The 1970 BCS follows the lives of more than 17,000 people born in England, Scotland and Wales in a single week of 1970. Over the course of the lives of cohort members, the 1970 BCS has collected information on, amongst others things, physical, educational and social development, health, economic circumstances and gender attitudes. Since the birth wave of the survey in 1970, there have been nine other waves ('sweeps') at ages 5, 10, 16, 26, 30, 34, 38, 42, and 46. At each sweep, different sources and methods were used to gather information on the cohort members. In the birth survey, the main questionnaire was completed by the midwife present at birth and supplementary information was obtained from clinical records. As the cohort members aged, questionnaires were administered to parents, teachers and, eventually, cohort members themselves. Medical examinations were also carried out and cohort members participated in thorough assessments.

The first outcome variable of interest during childhood is cohort member's contribution to household tasks. This is derived from the question "What kind of things do you help with at home?", asked when the cohort member is about 16 years old. The question is followed by a set of twelve items, each depicting a particular area of contribution to housework. The items are listed in the questionnaire as follows: "Shopping", "Washing up", "Cleaning the house", "Making the bed", "Cooking", "Looking after elderly relatives", "Looking after pets", "Washing and/or ironing clothes", "Gardening", "Cleaning car if any", "Painting or decorating" and "Looking after younger children if any". The possible answers are "Regularly", "Sometimes" or "Rarely or never".

A second housework outcome variable refers to the cohort member and their partner at age 34. Cohort members married to or cohabiting with a partner were asked to report who does most of the following household tasks: "Shopping", "Washing up", "Cleaning the house", "Cooking", "Paying the bills", "Looking after children when they are ill", "Washing and/or ironing clothes" and "Looking after the children in general'. The possible answers were "I do most of it", "My partner does most of it", "We share more or less equally" or "Someone else does it".[1]

[1] Survey-derived housework measures are usually biased as compared to the more accurate time-use diaries (Geist, 2010). Kan (2008) documents a systematic inflation in housework measures derived from stylised questionnaires rather than time-use data, the gap being larger for men than for women.

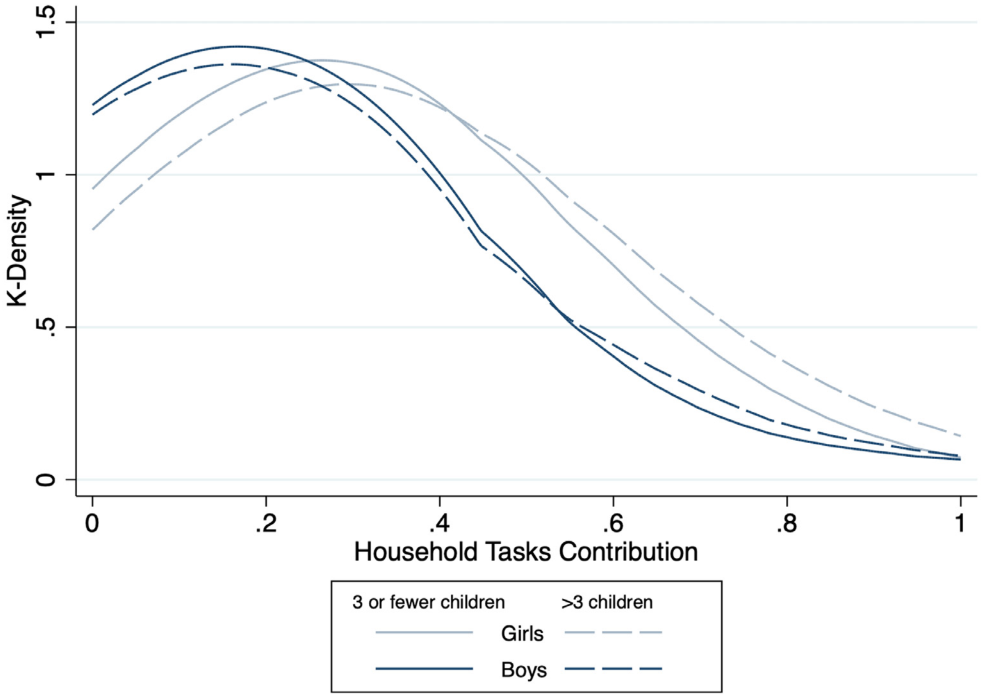

Fig. 1. Distribution of household tasks, by sex and family size at age 16.
Note: These figures refer to our estimation sample. We use the epanechnikov kernel function and bandwidth of 0.2.

We further combine information on household composition at birth and age 16 to create a measure of family size. We do so by adding the number of younger siblings of the cohort member at age 16 to the number of her older siblings as reported in the birth sweep.[2] We also know the gender and birth date of all of the cohort member's siblings, which we will use to create our instrumental variable.

3.2. The endogeneity of family size

Our first goal is to estimate the impact of family size on teenagers' contributions to household tasks. To do so, we first show the Kernel density of the contribution to household chores of BCS cohort members at age 16 (calculated as the share of household tasks the cohort member helps with "Regularly", as opposed to "Sometimes" or "Rarely or never"), by gender and family size (Fig. 1). Consistent with the extant literature, we find that, for any family size, girls contribute more to household tasks than boys do. The descriptive results in Fig. 1 further suggest that while family size does not much affect boys' contribution to housework, girls in larger families spend more time on household tasks than girls in smaller families do.

The evidence in Fig. 1 is suggestive of a role of family size for girls, but does not address endogeneity. The distribution of fertility across households cannot be assumed to be random, as it depends on a set of both observable and unobservable household characteristics that may well be correlated with household tasks both during childhood and adulthood. For example, BCS family size at age 16 is larger when the mother is not employed and has conservative opinions about maternal employment. Being on average less educated and less likely to be employed, mothers in large families will mechanically have more time to spend on housework, which in turn has a crowding-out effect on childrens' own contribution. Naive specifications that do not account for negative selection into parenthood and the time-use of mothers may thus underestimate the true effect of family size on childrens' contribution to housework.

In a similar spirit, Achen and Stafford (2005) show that the amount of time married men report spending in housework is larger than the time their wives report for them. However, if reporting biases are orthogonal to the instrumental variable we use in our empirical analysis, there is no reason to believe that our main estimate will be affected. Second, in contrast with the literature, we here do not measure hours spent in housework, but the relative contribution of individuals as compared to their partners – a measure which we believe to be less susceptible to measurement error.

[2] As our identification strategy relies on the gender composition of the two first-born children in the household, we measure family size as the total number of siblings. We are aware that this measure may include siblings who had already left the household when the cohort member is age 16. We address this potential concern by using the number of children living in the household at the fourth survey sweep as an alternative measure of family size. The use of this alternative measure produces even stronger results (available upon request), although, as expected, the instrument appears to be slightly weaker.

The endogeneity of family size is commonly addressed via instrumental variables. A first popular strategy is to instrument the size of families with at least two children by the sex composition of the two first-born children. To the best of our knowledge, Angrist and Evans (1998) is the first influential work to use this strategy, and estimates the causal impact of family size on women's labour supply. The rationale here is that parents have a preference for variety: a couple with the first two children of the same sex is more likely to try and have a third, relative to a couple whose first two children are a boy and a girl. As the sex mix of children can be seen as random, the instrument provides the exogenous variation necessary for plausible identification. This approach has been widely-used in the literature to assess the impact of family size on a variety of child outcomes, such as education, fertility and labour market outcomes (Angrist et al., 2010; Black et al., 2010; Cools and Hart, 2017). Section 4.2 provides a discussion of the validity of the same-sex instrument applied to our context.

Multiple births can also be seen as a source of exogenous change in family size. A number of articles have used twin births as an instrument to estimate the causal impact of family size on outcomes such as women's labour supply (Rosenzweig and Wolpin, 2000; Angrist et al., 2010) and children's education (Black et al., 2005; Cáceres-Delpiano, 2006; Åslund and Grönqvist, 2010). However, using individual data on 17 million births over 72 countries, Bhalotra and Clarke (2019) underline that twin births are systematically positively correlated with maternal health. This finding is robust to a battery of tests and casts doubt on the validity of multiple births as an instrument for family size.

3.3. Empirical strategy

We account for the endogeneity of family size by following the instrumental-variable approach in Angrist and Evans (1998). Our instrument is a dummy for the first two children of a couple being of the same sex. We do not make use of multiple births in our main analysis for a number of reasons. First, as noted above, Bhalotra and Clarke (2019) suggests that multiple births may not be random and can reflect positive selection into motherhood that could bias our estimates. Second, our estimation sample is of limited size and the lack of statistical power may be prejudicial to our analysis; a similar concern is raised by Black et al. (2005) when considering the results from estimation samples of sizes comparable to ours.[3]

We first estimate the following model by Two-Stages Least Squares (2SLS):

$$FamSize_i^{16} = \alpha_1 SameSex_i + \delta_1 X_i + \epsilon_i$$

$$HhTasks_i^{16} = \alpha_2 \widehat{FamSize}_i^{16} + \delta_2 X_i + \mu_i \tag{1}$$

where $FamSize_i^{16}$ is family size at age 16, calculated as the total number of siblings of the cohort member plus one, and $HhTasks_i^{16}$ is the contribution to household tasks of individual i at age 16, calculated as the share of household tasks the cohort member helps with "Regularly" (as opposed to "Sometimes" or "Rarely or never"). We use $SameSex_i$, a dummy for the first- and second-born children being of the same sex, as an instrument for $FamSize_i^{16}$. X_i is a vector of standard controls, including dummies for sex and the child being of European descent. We measure parental education by the age at which the parents left school, and include a dummy for the child's parents still living together in 1986. We additionally exploit information on the study child's father's socio-economic status (SES), observed at child age 0, as proxied by his occupational status (following the Registrar General's 1966 Classification of Occupation). In particular, we control for a dummy equal one for being in either a professional, managerial, or non-manual skilled occupation ('high SES'), and zero for being either a skilled manual, semi-skilled, or unskilled worker, or unemployed ('low SES'). We also include an index constructed by the data providers to measure the mother's attitudes towards maternal employment, when the study child was 5 years old. Controlling for this variable attenuates the possibility of the $SameSex_i$ instrument to be spuriously capturing differences in fertility responses correlated with heterogeneous adherence to conservative gender norms. We control for a potential independent effect of the gender mix of all the siblings by adding a dummy named *Balanced* for there being at least two siblings of different sex in the family. We finally control for birth order dummies.

Our first estimation sample covers individuals from families with at least two children and with valid information on both the household tasks performed at age 16 and the controls. This produces 3389 observations. One reported task out of four is performed "regularly" and the average family size in our estimation sample is 2.8. The full descriptive statistics on this estimation sample can be found in Table A1 in Appendix A. Only 6349 out of roughly 13,000 solicited families completed and returned the questionnaire measuring children's contribution to housework (the "Document G: Home and All That"). We ask in Table A2 whether children from our estimation sample differ significantly from those with similar characteristics but who did not complete "Document G: Home and All That" (i.e. children with at least one sibling and with valid information on the controls, but no information on the household tasks performed at age 16). Children in our estimation sample have on average a better family background (higher-SES households, more educated parents and a more stable parental relationship) and are mostly girls. This is not surprising as male survey respondents, as well as respondents

[3] Despite its limitations, we do ultimately use twin births as an instrument for family size as a supplementary robustness check, as suggested in Angrist et al. (2010). While the estimates are in line with our main results, none are significantly different from zero at conventional levels. As suggested by Black et al. (2005) and Cools and Hart (2017), this lack of precision may come from the limited sample size. The results from this additional identification approach are available upon request.

whose parents have low levels of education, typically have higher attrition rates and non-response rates with respect to females (Mostafa and Wiggins, 2015).[4] We find a similar pattern of selection when comparing our estimation sample to the overall BCS population with non-missing information on the controls.

We then ask whether family size at age 16 continues to influence the time devoted to housework at age 34, via a second 2SLS model:

$$FamSize_i^{16} = \beta_1 SameSex_i + \delta_3 X_i + \epsilon_i$$

$$Y_i^{34} = \beta_2 \widehat{FamSize}_i^{16} + \delta_4 X_i + \mu_i \tag{2}$$

Here Y_i^{34} corresponds to one of the three following dependent variables measured at age 34: the share of household tasks carried out by the wife (or female partner), the share of household tasks carried out by the husband (or male partner) and the housework gender gap (the difference between these two shares). The vector X_i includes the same control variables as in model 1. We do not control for the socio-demographic characteristics of the cohort members at age 34 (e.g. labour-force status, number of children) as we suspect that these may mediate the effect of family size in childhood and, as such, are 'bad controls' (Angrist and Pischke, 2008). We explore this potential mediation in Section 5.2 using the decomposition approach in Gelbach (2016).

Our second estimation sample covers individuals who are in a partnership at age 34, with at least one sibling at age 16, and with valid information on the household tasks performed at age 34 and on the controls. This produces a sample of 3200 observations.[5] The cohort members in our estimation sample sort on average into couples where about half of housework is only carried out by women. Only 15 percent of the tasks are only carried out by men. Additional descriptive statistics for this sample are shown in Table A3.

4. Family size and the contribution to household tasks at age 16

4.1. Main results

Table 1 shows both the OLS and 2SLS estimates of model 1. The first two columns refer to the whole sample of households with at least two children, while the sub-samples by child sex appear in columns (3) through (6). The main variable of interest is family size. In the OLS estimates in column (1), an additional household member has a positive and significant impact on the contribution of the study child to household tasks. When we instrument *Family size* by *Same sex*, the 2SLS results in column (2) also reveal a positive and significant coefficient on *Family size*.[6] This result is in line with Bianchi and Robinson (1997) and Gager et al. (1999): larger family size increases the contribution of children to household tasks. Looking at the estimates in column (2), one additional sibling increases the share of tasks performed "regularly" by 5.4 percentage points. This effect is equal to 25 percent of a standard deviation of the dependent variable. In terms of magnitude, the effect of family size on the share of housework is equivalent to about 60% of the effect of being a girl.[7]

Fig. 1 suggested that the effect of family size was mainly found for girls. We formally check whether there is a difference between boys and girls in columns (3) to (6) of Table 1. Both OLS and 2SLS estimates confirm that an increase in family size translates into a significantly higher contribution of girls to household tasks at age 16, while it does not affect the contribution of boys. The positive family size coefficient in columns (1) and (2) is thus mostly driven by girls.[8] Note that when we do control for variables that are arguably endogenous (e.g. children's cognitive skills and mothers' labour force participation), estimates for family size are not statistically different from results in Table 1.

While both the OLS and 2SLS estimates are positive and significantly different from zero at 5% level, both in the overall sample and in the subsample of girls, the OLS estimate is smaller than that from IV. This is in line with our hypothesis that the negative selection into parenthood reduces the true effect of family size in OLS. The difference between the OLS and 2SLS estimates might also be due to the fact that the former is expected to capture the influence of family size over the whole estimation sample while the latter is a Local Average Treatment Effect (LATE) for the compliers, that is children whose parents are more likely to have a third child if the first two are of the same sex. One could argue that the preferences for variety in the offspring sex-mix are not randomly distributed across the population: individuals with more conservative

[4] Additionally, based on observable characteristics, we do not find any evidence that girls and boys select into non-response in systematically different ways (except for father's age: having an older father significantly increases the probability of returning the questionnaire for boys, while it plays no role for girls). These results are available upon request.

[5] The estimation samples in childhood and adulthood are not the same size. This is because only a sub-sample of the cohort members reported their contribution to housework tasks at age 16, while the age-34 housework questionnaire was administered to all partnered cohort members. Keeping only the cohort members who appear in both estimation samples reduces the number of observations drastically (under 1500 observations). While such selection does not affect the size of our estimates, the smaller sample size does reduce their precision. The results are available upon request.

[6] We have also re-estimated all our regressions using a dummy "Having at least three children" rather than all discrete values of family size, with the results remaining qualitatively unchanged.

[7] One may argue that we should not compare the effect of a discrete variable to that of dummies. When we dichotomise family size as a dummy for having at least three children, we find that the estimated coefficient has the same magnitude of that of the female dummy.

[8] We also follow Wooldridge (2010) and use same-sex and its interaction with the cohort member's gender to instrument family size and its interaction with a female dummy. The difference between boys and girls continues to be significantly different from zero at the 5% level.

Table 1
Family size and share of household tasks done 'Regularly' at age 16: OLS and 2SLS results.

	All		Girls		Boys	
	OLS (1)	2SLS (2)	OLS (3)	2SLS (4)	OLS (5)	2SLS (6)
Family Size	0.014***	0.054**	0.018**	0.084***	0.008	0.014
	(0.005)	(0.023)	(0.007)	(0.029)	(0.008)	(0.036)
Female	0.088***	0.090***				
	(0.007)	(0.007)				
Observations	3389	3389	1935	1935	1454	1454
F-stat (first stage)		137		80		58

Notes: Robust standard errors in parentheses. 'Family size' indicates the number of siblings of the cohort member at age 16, plus one. All regressions control for the ethnicity of the cohort member, birth order dummies, a dummy indicating whether the cohort member's parents are still living in the same household, a dummy for the father having a high SES, years of education of the cohort member's parents, age of the parents at birth of the cohort member, an index measuring the attitude of the mother regarding maternal employment, a dummy indicating whether the gender composition of the siblings is balanced and regional dummies. Family size is instrumented by a dummy equal one if the first two children in the household are of the same sex. Statistical significance is coded following the standard notation: *** if the *p*-value is lower than 0.01, ** if the *p*-value is lower than 0.05, * if the *p*-value is lower than 0.01.

views may care more about a mixed-sex offspring pool, as compared to progressive individuals. If that was the case, parents whose first two children are of the same sex and who decide not to have a third one may hold systematically more progressive views than parents in the same position who instead decide to have a third child.

What does this mean for the interpretation of our LATE? The effect sizes found in columns 2 and 4 might be partly driven by the fact that larger weights are attached to the second-stage treatment effect of those families in which fertility decisions are more sensitive to the same-sex instrument – that is, arguably more conservative families. As conservativism is likely correlated with a more gender-stereotypical assignment of housework to children, our 2SLS estimates might be thus inflated. We address this potential source of concern in two ways: first, all our regression control for the mother's attitudes towards maternal employment, a proxy of the family's conservativism. Second, we check whether we find any evidence of conservative parents holding stronger preferences for a mixed-sex offspring. As families with the same-sex instrument equal to zero already achieved a composite gender mix, we here focus only on families in which the first two children are of the same sex. Using a range of indices capturing the mother's gender attitudes in 1975, Table A4 shows that, conditional on being assigned to the treatment (i.e. same-sex equal one), women who have a third child do not hold systematically more conservative views than those who stop at two children. This suggests that preferences for mixed-sex offspring are not systematically driven by conservative attitudes in our estimation samples.

One may suspect a smaller effect of family size in families that outsource their home-production in the market. The outsourcing of housework is not accurately measured in the BCS, so we use the father's SES as a proxy for the probability of hiring help. We expect the effect of family size to be smaller in high-SES families, as they are more likely to hire in help for home production, thus decreasing the need for children to contribute to housework. Net of concerns on families'" reactivity to the same-sex instrument, we may also expect gender attitudes to be a source of heterogeneity for the effect of family size – the effect being arguably larger for children from conservative households. Traditional parents might believe, for instance, that their daughters (but not their sons) will face a marriage-market premium when endowed with a set of domestic skills (this is consistent with the matching model developed by Chiappori et al., 2009, under the assumption of traditional household roles). We empirically test the presence of these two sets of channels in Table 2, where we split the estimation sample by gender, father's SES, and mother's adherence to conservative gender norms (above or below the third quartile of the distribution of the measure of mothers' attitudes towards maternal employment discussed above).[9] Panel A of the table shows that family size significantly increases the contribution of girls to housework in low-SES families but not in high-SES families. In Panel B, the effect of family size on boys' contribution to household tasks remains insignificant in both cases. We also see that girls with conservative mothers are more likely to contribute to household tasks as family size rises; there is no significant effect for girls with non-conservative mothers or for boys.[10]

In Appendix B we show that our results are robust to alternative measures of children's contribution to housework, as well as to the aggregation of subsets of tasks according to their characteristics (e.g. 'feminine' vs 'masculine' tasks). Additionally, we explore the time children spend in other activities (homework and leisure) and find consistent results suggesting that girls from larger families dedicate less time to homework and leisure activities.

4.2. Assessment of the identification assumptions

For a valid causal interpretation of our 2SLS estimates, the following five assumptions need to hold (Angrist et al., 1996):

[9] The results using the median of the mothers' gender attitudes distribution are qualitatively similar.

[10] We have also looked at other sources of heterogeneity, and find that the effect of family size is stronger for girls who are first- or second-born and whose mother is young.

Table 2
Family size and share of household tasks done 'Regularly' at age 16: heterogeneity - 2SLS results.

	Low SES (1)	High SES (2)	Conserv. Mothers (3)	Non-Conserv. Mothers (4)
Panel A. Girls				
Family Size	0.116***	−0.017	0.116**	0.055
	(0.035)	(0.054)	(0.047)	(0.035)
Observations	1303	632	490	1445
F-stat (first stage)	61	17	38	48
Panel B. Boys				
Family Size	−0.004	0.043	0.105	−0.030
	(0.048)	(0.043)	(0.077)	(0.040)
Observations	925	529	402	1052
F-stat (first stage)	36	33	13	48

Notes: Robust standard errors in parentheses. 'Family size' indicates the number of siblings of the cohort member at age 16, plus one. All regressions control for the ethnicity of the cohort member, birth order dummies, a dummy indicating whether the cohort member's parents are still living in the same household, a dummy for the father having a high SES, years of education of the cohort member's parents, age of the parents at birth of the cohort member, an index measuring the attitude of the mother regarding maternal employment, a dummy indicating whether the gender composition of the siblings is balanced and regional dummies. Family size is instrumented by a dummy equal one if the first two children in the household are of the same sex. Statistical significance is coded following the standard notation: *** if the *p*-value is lower than 0.01, ** if the *p*-value is lower than 0.05, * if the *p*-value is lower than 0.01.

- Assumption 1, SUTVA: in our context this is equivalent to saying that the time spent in household tasks is not affected by the sex composition of the first two children born in other families. Additionally, SUTVA requires the treatment (an increase in family size after the first two children) to be the same across all individuals.
- Assumption 2, Independence: the sex composition of the two first-born children is uncorrelated with any confounder of the association between family size and household tasks.
- Assumption 3, Relevance: having two first-born children of the same sex predicts a significant increase in family size.
- Assumption 4, Monotonicity: there are no parents whose preferences are such that they have more children only if their first two are of different sex, and stop having children if their first two are of the same sex.
- Assumption 5, Exclusion Restriction: once family size is taken into account, the time spent in household tasks is not affected by the sex composition of the two first-born children.

4.2.1. From assumption 1 to 4

We here discuss how plausible the assumptions above are in our context. Assumption 1 is uncontroversial: our setting does not exploit within-family variations and we do not expect major general equilibrium effects from the sex-composition instrument. Except in the rare cases of multiple birth, we can also safely assume that the treatment (an increase in family size after the first two children) is the same across individuals. Sex at birth being random, Assumption 2 is also indisputable: conditional on having at least two children, the sex-composition of the first two is random. The relevance assumption can be shown to hold by looking at the first-stage regressions of model (1), as summarised by the top section of Table A5. The instrument's coefficient is positive and highly-significant in all subsamples and the Cragg-Donald Wald F-statistics (also reported at the bottom of Table 1) confirm the instrument's strength. Last, Assumption 4 holds if they are no defiers. Defiers in our context would be parents with a strict preference for having at least two children of the same sex, therefore choosing to have a third child or more only if the first two are of different sex. While the existence of this category is unlikely in our case (a 98%-white British sample), it would not be surprising in different cultural context, where son preference is for instance a wide-spread social norm (Lee, 2008; Rosenzweig and Wolpin, 2000). In order to rule out any systematic preference for one particular gender, we split the same-sex instrument into two dummies: one for the first two children being boys and one for them being girls. Both instruments are equally predictive of subsequent fertility behaviour, suggesting that indeed British parents in our estimation sample do not display preferences for a particular gender of their offspring.

4.2.2. The exclusion restriction: a discussion

Our identification strategy would prove problematic if the same-sex instrument were to be correlated with the dependent variable through a channel other than family size. While Angrist and Evans (1998) used mothers' labour-force participation as their main outcome, we here consider a child-level outcome that may well be correlated with the sex composition of the two first-born children. One way this could happen is via a crowding-out effect: if children of one particular gender systematically contributes more to housework, then having a sibling of such gender may reduce the residual amount of housework to be carried out by the child. If we assume ex-ante one of our key findings, that is girls contribute more to housework than boys (Gager et al., 1999), and keeping everything else constant, having a sister will always decrease the residual pool of housework a child can contribute to. While we partially take sibling composition into account by including a dummy for there being at least a girl and a boy in the household, we cannot fully rule out empirically the presence of such channel.

Before discussing this channel any further, we first investigate whether the data suggest in any way the presence of a direct effect of the instrument on the outcome variable. We do so by carrying out a placebo test and then looking into systematic correlations between the instrument and a number of observed covariates. The first test relies on the following intuition: if the instrument were to be correlated with the dependent variable through a channel other than family size, we would expect to find a significant 2SLS family-size estimate even for household tasks that should not be affected by family size. We here appeal to the estimates shown in the fourth and fifth rows of Panel B of Table B1. These show, respectively, the estimated family-size coefficients for household tasks that are likely increasing in family size and those that are not. None of the 2SLS family-size estimates are significant for the latter (in the fifth row) while most of the estimates in the fourth row are significant, suggesting that the instrument is unlikely to significantly affect the contribution to household tasks other than via its impact on family size.[11] We then look for the presence of any systematic correlation between the instrument and observable characteristics of the cohort members at age 16 and their parents. We follow Angrist et al. (2010) and Falck et al. (2014) and derive reduced-form estimates from the regression of cohort members' and their parents' characteristics at age 16 (normally used as control variables in our baseline regressions) on the same-sex instrument and all other controls. If observable characteristics were to be correlated with our instrument, we may expect the same for unobservable characteristics too. Table A6 shows that the only covariate that is significantly explained by the same-sex instrument is mother's age at the birth of the cohort member. While we may worry about this correlation, the estimates are arguably small in economic terms: for example, in column (1) mothers whose two first-born children are of the same sex are on average five months younger than mothers whose first two children are of different sex. The absence of selection into consecutive same-sex pregnancies based on observable characteristics mitigates our concerns about systematic correlations with unobservable characteristics.

While reassuring, the absence of evidence suggesting the instrument has a direct effect on housework contribution alone is not sufficient to completely rule out any threat to the exclusion restriction. We then draw from Conley et al. (2012) and relax the exclusion restriction, by allowing the instrument to be only 'plausibly' exogenous. The intuition behind the method in Conley et al. (2012) is to allow the instrument to have a direct effect on the dependent variable in the second-stage of the 2SLS estimation. We here follow the application in Nybom (2017) and express model 1 as follows:

$$FamSize_i^{16} = \alpha_1 SameSex_i + \delta_1 X_i + \epsilon_i$$

$$HhTasks_i^{16} = \alpha_2 \widehat{FamSize}_i^{16} + \lambda\gamma SameSex_i + \delta_2 X_i + \mu_i. \tag{3}$$

where γ is the instrument's reduced-form effect and λ is a parameter varying between 0 and 1. As explained in Nybom (2017), the adjusted effect of family size mechanically converges towards zero as λ converges to one. Fig. A1 shows the estimates of family size instrumented by the sex-composition of the first two children in the household, first for the whole sample and then separately for girls and boys. If λ is greater than 0.38 for the whole sample and greater than 0.45 for girls, the 2SLS estimates of instrumented family size are no longer significantly different from zero. This is equivalent to saying that, as long as the direct effect of the instrument is respectively smaller than 38% and 45% of the instrument's reduced form effect, the effect of family size on housework will remain significantly different from zero for the whole sample and for girls.

Conceptually, and in line with the plausible exogeneity exercise, the bias induced by a violation of the exclusion restriction would be problematic (i.e. result in an over-estimation of the effect of family size) only in case the same-sex instrument caused an increase in housework participation, through a channel other than family size. A plausible channel, as mentioned above, could be via the crowding-out effects coming from the sisters' housework participation. Table A7 presents a topology of all possible sibling compositions, given the value of the same-sex instrument. Consider first the case in which the cohort member is the first- or second-born. If the cohort member is a boy, then the instrument Z will take value 0 if he has a sister (case *a*) and 1 if he has a brother (case *b*).[12] If we assume that girls contribute more than boys to household tasks, then the cohort member will be more likely to perform a larger share of housework in case *b* than in case *a*. Hence, with 2SLS we would tend to overestimate the family-size coefficient due to the positive correlation between the instrument and the dependent variable, given all other covariates. This would be problematic if we found a positive effect of family size on the share of housework performed by boys, as we would not be able to distinguish whether the effect comes from the real association between the variables or the violation of the exclusion restriction. However, since we find no statistically-significant effect of family size on the share of housework for boys, the real effect should be either zero or negative - which in either case corroborates the finding that the effect of family size is larger for girls than for boys.

Now consider the case where the cohort member is a first- or second-born girl. The instrument takes value 0 when she has a brother (case *c*) and 1 when she has a sister (case *d*). With the same assumption as above, the cohort member will be more likely to contribute more to housework in case *c* than in case *d*. Here, conditional on the controls, the instrument would be negatively correlated with the dependent variable. We again are not particularly worried about this potential violation of the exclusion restriction, as it would bias the coefficient of family size towards zero.

We finally consider the case where the cohort member is neither the first- nor the second-born. The instrument now takes the value of one in two occurrences: either the two first-born children are both boys (cases b_2 and d_2) or both girls

[11] This is assuming that all tasks are potentially equally-affected by the instrument, a condition under which our restriction of the analysis to a set of tasks that is arguably insensitive to family size comes without loss of generality.

[12] Here, when talking about brothers and sisters in relation to first- or second-born cohort members, we refer to the brother or sister that, together with the cohort member, makes up the pool of the two first-born children.

(cases b_1 and d_1). Irrespective of his or her gender, a cohort member with two older sisters would tend to perform relatively less housework compared to the case where the instrument is zero (cases a_1 and c_1). As argued above, the 2SLS estimate of family size would then be biased toward zero. Instead, when the cohort member has two older brothers (cases b_2 and d_2) there will be comparatively more housework to do and he or she might be asked to contribute relatively more than in the case where the two eldest siblings are of opposite sexes. We may here expect the 2SLS estimate to overestimate the effect of family size. This is the most worrying case, as the effect size we estimate is potentially inflated. To check whether the sample of cohort members with this particular sibling mix is behind our results, we replicate our main results from Table 1 excluding the 236 cohort members who have two older brothers. Consistent with the results in Table 1, the 2SLS point estimate of family size is 0.061 for the whole sample and 0.088 for girls (both significant at the 1% level).

What do we empirically know about the effect of the sibship sex-mix on children's contribution to housework? While there are plenty of papers relying on the same-sex instrument, the number of studies linking it to housework contribution in childhood is scarce. One exception is the descriptive work of Schulz (2021), which assesses the influence of a variety of factors (among which, the siblings sex composition) on the time children spend performing household chores. Using the German Time Use Study, Schulz (2021) shows that, while girls spend on average more time in housework than boys, each child's contribution to housework is independent of their siblings' gender (similar results are also found in Cordero-Coma and Esping-Andersen, 2018). Assuming these conclusions would hold in our context as well (as somewhat already suggested by Tables A6 and B1), the direct effect of the same-sex instrument on the dependent variable would neither be positive (making the positive upper bounds derived from Fig. A1 implausibly high) nor negative (in contrast with what we would expect theoretically from most of the cases described in our topology above), but rather nil.[13]

5. Family size and contribution to household tasks at age 34

5.1. Main results

We now ask whether the effect of family size at age 16 persists into adulthood and affects the division of housework in households formed by BCS respondents and their partners at age 34. To do so, we replicate our 2SLS analysis using as the dependent variables the share of household tasks performed by the female partner, by the male partner, and the housework gender gap (i.e. the difference between the two shares). Before presenting our estimates, it is important to verify that growing up in a large family in childhood (instrumented by the gender composition of the first two siblings) does not influence the probability of being in a partnership at age 34 and, hence, being in our estimation sample. We rule out this concern of endogenous selection in Table A8 by showing that the instrumented family size at age 16 does not affect significantly the probability of being in a partnership at age 34.

Having found no evidence of selection into cohabitation based on the same-sex instrument, we now turn to the main results in our adult sample described in Table 3. The 2SLS estimates for the whole sample (Panel A) confirm a persistent effect of family size at age 16 on the division of household tasks at age 34. Larger families at age 16 predict a greater share of household tasks done by women, while the male share remains unchanged. As expected, column (3) then shows that the larger the family at age 16, the greater the housework gender gap at age 34. As such, cohort members who grew up in larger families sort into couples that conform more to stereotypical gender roles and in which the housework gender gap is even larger.

In Panels B and C of Table 3 we then ask whether this result is stronger for men or women. It appears that only women sort into households with a significantly larger housework gender gap as family size at age 16 rises. As revealed in columns (1) and (2) of Panel B, this is mostly explained by a significantly higher share of household tasks predominantly carried out by the wife.[14] In the bottom Panel of Table 3, there is some evidence that male cohort members who grew up in larger families have a larger housework gender gap in their adult household, although the estimated coefficient here is not statistically significant.

The previous section established that there is no significant effect of family size on the contribution to household tasks for respondents who grew-up in high-SES families or with relatively progressive mothers. We again look at this kind of heterogeneity, with the results for females appearing in the first row of Table 4. The pattern here is similar: larger families have no impact on the contribution to household tasks and the housework gender gap at age 34 for women from relatively well-off families and non-conservative families. Consistent with the childhood results, women raised in large low-SES families contribute significantly (at the 5% level) more to household tasks and sort into couples with a higher housework

[13] It is important to stress that the exclusion restriction might also be violated for infra-marginals (always-takers and never-takers). If this is the case, our 2SLS estimates would be biased (see Jones, 2015, for technical details). Based on the adjustment via calibration formula proposed by Jones (2015) and assuming no direct effect of the same-sex instrument on the outcome for never-takers, we can express the bias as the product of the effect of the instrument on the outcome for always-takers and the share of always-takers over the share of compliers, i.e. $\eta_{AT} \frac{\pi_{AT}}{\pi_C}$. Taking estimates for π_{AT} and π_C from Kowalski (2019), we can derive the minimum η_{AT} such that the true effect of family size on girls' housework is not statistically different from zero, that is $\eta_{AT} = 0.05$. In other words, as long as $\eta_{AT} < 0.05$ the true effect of family size remains positive and significant. As we argued above, most of the cases we derived from the topology in Table A7 suggest that the direct effect of the same-sex instrument on the outcome should be theoretically negative. Following the evidence in Schulz (2021) and Cordero-Coma and Esping-Andersen (2018), η_{AT} has been descriptively shown to be zero in contexts similar to ours.

[14] We also find that the share of tasks to which men and women contribute equally and the share of tasks performed by someone else are unaffected by family size in childhood. These results are available upon request.

Table 3
Family size at age 16 and household tasks at age 34 - 2SLS results.

	Wife HH Tasks (1)	Husband HH Tasks (2)	Housework Gender Gap (3)
Panel A. Whole sample			
Family Size	0.052**	0.003	0.049*
	(0.025)	(0.012)	(0.029)
Female	0.125***	−0.045***	0.170***
	(0.009)	(0.004)	(0.011)
Observations	3200	3200	3200
F-stat (first stage)	140	140	140
Panel B. Women			
Family Size	0.070**	−0.001	0.071*
	(0.034)	(0.015)	(0.040)
Observations	1731	1731	1731
F-stat (first stage)	73	73	73
Panel C. Men			
Family Size	0.033	0.009	0.023
	(0.037)	(0.019)	(0.044)
Observations	1469	1469	1469
F-stat (first stage)	64	64	64

Notes: Robust standard errors in parentheses. 'Family size' indicates the number of siblings of the cohort member at age 16, plus one. All regressions control for the ethnicity of the cohort member, birth order dummies, a dummy indicating whether the cohort member's parents are still living in the same household, a dummy for the father having a high SES, years of education of the cohort member's parents, age of the parents at birth of the cohort member, an index measuring the attitude of the mother regarding maternal employment, a dummy indicating whether the gender composition of the siblings is balanced and regional dummies. Family size is instrumented by a dummy equal one if the first two children in the household are of the same sex. Statistical significance is coded following the standard notation: *** if the *p*-value is lower than 0.01, ** if the *p*-value is lower than 0.05, * if the *p*-value is lower than 0.01.

gender gap. We find no significant results for men (see Table A9 in the Appendix). We can replicate this analysis for cohort members at age 42, to ensure that our estimates are not a statistical artifact driven by the choice of a particular survey year: the results in Table A10 are qualitatively similar.

The assumptions for a valid causal interpretation of all the estimates in this section are the same as the ones discussed in Section 4.2. While the same considerations made about the first four assumptions (SUTVA, independence, relevance and monotonicity) hold regardless of the different outcomes we consider here, it may be relevant to discuss how the exclusion restriction translates in the adulthood context. As cohort members are now living with a partner and not anymore with their childhood family members, we can rule out the presence of any sibling crowding-out effects on the pool of housework to be performed. Sibling sex-composition could however still affect individuals' contribution to housework in adulthood via peer effects: interacting with a sister (brother) who formed a household where she (he) does most (none) of the chores might set an example for our cohort members. However there is only limited evidence that this mechanism might be in place (Nicoletti et al., 2018) and, if in place, we believe its effects to be of second order.

5.2. Channels

Why does family size at age 16 continue to explain the individual's contribution to housework 18 years or more later? In Table 4 we explore the role of family size on adult characteristics which are likely to help shaping the housework gender gap of female cohort members (we replicate the exercise in Table A9 for male cohort members).

Section 2.2 suggested that we might expect children who grew up in larger families to have lower education and thus worse labour-market outcomes. We investigate this in Table 4 by estimating the causal effect of family size at age 16 on the school-leaving age and the probability of having at least an A-level at age 34. Most of the estimated coefficients on family size are not significantly different from zero. This finding continues to hold with other measures of educational attainment and is consistent with the results in Cáceres-Delpiano (2006), Black et al. (2005) and Angrist et al. (2010). Note that women who grew up in large families with non-conservative mothers are the only exception, as they are more likely to have an A-level and left full-time education at an older age.

We then look at the effect of age-16 family size on a set of labour-market outcomes, namely employment, the monthly wage (in logs), weekly working hours and commuting time. Commuting in BCS is measured in time bands, and we here create a dummy for commuting time of over 30 minutes. Only women from low-SES families have a significantly lower probability of being employed and, when employed, spend less time commuting. This is in line with the burgeoning literature on gendered preferences over workplace amenities (Mas and Pallais, 2017) and local labour markets (Manning and Petrongolo, 2017). Our results provide indirect evidence that the definition of local labour market might differ by gender,

Table 4
Family size at age 16 and adult women's outcomes: 2SLS results.

	Effect of Family Size for Women				
	All Women (1)	Low SES (2)	High SES (3)	Conserv. Mothers (4)	Non-Conserv. Mothers (5)
Housework gender gap	0.071*	0.086*	0.021	0.158	0.042
	(0.040)	(0.046)	(0.095)	(0.098)	(0.044)
Educational Attainment (age 34)					
Age left FT education	0.688	0.333	1.999	−0.713	1.096**
	(0.488)	(0.514)	(1.405)	(1.097)	(0.546)
At least A-level	0.070	0.085	0.028	−0.147	0.125*
	(0.062)	(0.070)	(0.148)	(0.144)	(0.069)
Labour Market Outcomes (age 34)					
Not employed	0.096	0.118*	−0.035	0.094	0.091
	(0.060)	(0.068)	(0.144)	(0.149)	(0.065)
Monthly wage (log)†	0.183	0.150	0.838	0.176	0.213
	(0.459)	(0.498)	(1.226)	(1.180)	(0.490)
Weekly working hours†	−3.145	−3.357	−1.007	0.443	−3.855
	(2.294)	(2.529)	(5.831)	(5.815)	(2.480)
Commuting time†	−0.053	−0.094*	0.130	0.088	−0.096*
	(0.048)	(0.048)	(0.150)	(0.131)	(0.052)
Employed partner	0.078***	0.082**	0.098**	0.047	0.079**
	(0.030)	(0.037)	(0.048)	(0.068)	(0.033)
Demographic characteristics (age 34)					
Married	0.110*	0.066	0.268*	−0.031	0.126*
	(0.061)	(0.067)	(0.152)	(0.148)	(0.067)
Having a least one child	0.017	0.047	−0.104	0.131	−0.012
	(0.055)	(0.058)	(0.163)	(0.140)	(0.060)
Number of children	0.023	0.086	−0.176	−0.123	0.054
	(0.152)	(0.156)	(0.479)	(0.336)	(0.172)

Notes: Robust standard errors in parentheses. The Table reports 2SLS estimates of the coefficient for 'Family size' (the number of siblings of the cohort member at age 16, plus one) for different dependent variables. All regressions control for the ethnicity of the cohort member, birth order dummies, a dummy indicating whether the cohort member's parents are still living in the same household, a dummy for the father having a high SES, years of education of the cohort member's parents, age of the parents at birth of the cohort member, an index measuring the attitude of the mother regarding maternal employment, a dummy indicating whether the gender composition of the siblings is balanced and regional dummies. Family size is instrumented by a dummy equal one if the first two children in the household are of the same sex. Statistical significance is coded following the standard notation: *** if the p-value is lower than 0.01, ** if the p-value is lower than 0.05, * if the p-value is lower than 0.01.
†The regressions based on these outcomes are based on a subsample of employed cohort members. Results for these outcomes are similar when including also individuals who are not employed and conditioning on employment.

due to the different costs associated with distance from the workplace: women might face social constraints that confine them to 'even more local' labour markets.

Only limited information is available on the partners of BCS respondents. However, we can estimate the causal effect of family size on the probability of having an employed partner. We find here positive and significant estimates in almost all our samples: women who grew up in large families tend to sort into couples where their partner is more likely to be employed.

We now turn to life-course transitions. According to Baxter et al. (2008) the housework gender gap does not change with marriage but does increases as individuals enter parenthood, and it has been shown that fertility is transmitted across generations (Anderton et al., 1987; Booth and Kee, 2009; Kolk, 2014; Fasang and Raab, 2014). The effect of family size at age 16 on the housework gender gap might therefore transit via the cohort members' own number of children. Table 4 asks whether family size affects the probability of being married at age 34, as well as the probability of being a parent and the number of children. There is some evidence that family size increases the probability of marriage. Our fertility results are somewhat in line with Cools and Hart (2017): only men's fertility decisions are positively influenced by their own family size in childhood, but not at conventional significance levels (see Table A9). On the contrary, women's fertility decisions are not affected by their number of siblings and hence do not lie behind the effect of family size at age 16 on the housework gender gap at age 34.[15] As for the exclusion restriction, a discussion similar to that in Section 5.1 could be applied to this context as well. See Angrist et al. (2010) for an additional extensive discussion of the same-sex instrument's validity when

[15] One may worry that the significance of our estimates in Table 4 results from multiple hypothesis testing. Following Falck et al. (2014), we assume each regression to be an independent draw (where the significance at conventional levels of the family size coefficient is seen as 'success') and exploit the properties of a binomial distribution to derive how likely it is that the coefficients in Table 4 are statistically significant only by chance. In the Table, 14 out of 55 estimated coefficients are statistically significant at least at a 10% level. The probability that 14 or more out of 55 coefficients are significant at the 10% level by chance is only 0.08%. In the subsample of regressions based on women who grew up in low-SES families – for whom our results appear to be stronger – the probability that 4 or more coefficients are significantly different from zero at the 10% level by chance is 1.85%.

Table 5
Family size at age 16 and the channels: 2SLS results - Women.

	Whole sample			Low SES			Conservative Mothers		
	Base (1)	Full (2)	Expl. (3)	Base (4)	Full (5)	Expl. (6)	Base (7)	Full (8)	Expl. (9)
Family Size	0.071*	0.033	0.038***	0.086*	0.044	0.042***	0.158	0.148	0.010
	(0.040)	(0.038)	(0.012)	(0.046)	(0.043)	(0.014)	(0.098)	(0.092)	(0.023)
Contributions									
Not employed			0.011			0.014*			0.008
			(0.007)			(0.008)			(0.014)
Commuting time			0.005			0.010*			−0.003
			(0.005)			(0.005)			(0.005)
Employed partner			0.015**			0.016*			0.006
			(0.007)			(0.008)			(0.017)
Married			0.007			0.003			−0.002
			(0.004)			(0.004)			(0.009)
Observations	1731	1731	1731	1179	1179	1179	430	430	430
F-stat (first stage)	73	74		55	55		12	13	

Notes: Robust standard errors in parentheses. The Table reports 2SLS estimates of the coefficient for 'Family size' (the number of siblings of the cohort member at age 16, plus one) under different sample specifications. All regressions control for the ethnicity of the cohort member, birth order dummies, a dummy indicating whether the cohort member's parents are still living in the same household, a dummy for the father having a high SES, years of education of the cohort member's parents, age of the parents at birth of the cohort member, an index measuring the attitude of the mother regarding maternal employment, a dummy indicating whether the gender composition of the siblings is balanced and regional dummies. Family size is instrumented by a dummy equal one if the first two children in the household are of the same sex. Columns 3, 6, and 9 perform a decomposition analysis of the difference between the effect of family size in the baseline and full model specifications, following the approach of Gelbach (2016). Statistical significance is coded following the standard notation: *** if the p-value is lower than 0.01, ** if the p-value is lower than 0.05, * if the p-value is lower than 0.01.

using labour market outcomes, educational outcomes, and marriage and fertility decisions (see also Cools and Hart, 2017, for a discussion on the latter).

The results above suggest that women who grew up in large families, and particularly low-SES families, sort into partnerships with more traditional gender roles, i.e. with a lower probability of employment and shorter commuting time for the woman, a higher probability of employment for the husband, and a greater probability of being married rather than cohabiting. But does the adoption of these gender roles fully explain the persistence of the housework gender gap? In order to answer this question we here follow the decomposition approach developed by Gelbach (2016). The decomposition relies on the omitted-variables bias formula and can be used to attribute a portion of the treatment effect to potential mediators. Results from this exercise are presented in Table 5, first for all women in our sample and then only for women from low-SES families and families with conservative mothers. For each of the three subsamples, the Table reports three columns: the first, 'Base', reports the coefficient for family size from the baseline regression (model 2 in Section 3.3); the second, 'Full', further augments the model specifications by adding the four potential mediators (not in employment, commuting time, having an employed partner, being married); last, in column 'Expl.' the difference between the family size coefficient in the baseline specification and the full specification is decomposed into portions that can be explained by each of the potential mediators – conditional on controlling for all of them simultaneously.

The first three columns of Table 5 show that about half of the estimated effect of family size on the housework gender gap for women can be explained by the channels we included in the full model specification. While all channels display a positive contribution, only having an employed partner appears to be a significant mediator. Moving on to women from low-SES backgrounds, columns 4 to 6 show a similar narrative – with labour market outcomes (labour force participation and commuting time) significantly contributing to explaining the persistent effect of family size on the adult housework gender gap. No change is observed instead across the two specifications in the subsample of women who grew up with conservative mothers. Although only descriptive, these results confirm that the long-lasting influence of family size in childhood on the housework gender gap can be largely attributed to the adoption of behaviours that conform to traditional gender roles.[16] Results for the subsample of women from high-SES families and women with non-conservative mothers are displayed in Table A11. We also replicate this analysis for men in Tables A12 and A13.

6. Conclusion

In this paper we have assessed the impact of childhood family size on the allocation of household tasks of British Cohort Study cohort members at age 16 and then at age 34. We account for the endogeneity of fertility by exploiting parents' preferences for variety in the sex mix of their offspring, and use the sex composition of the first two children as an instrumental-variable predictor of family size.

[16] Additionally including the other adult outcomes shown in Table 4 in our mediation analysis does not appear to further attract part of the estimated effect of family size.

We find that family size significantly increases the probability that adolescents contribute to housework at the age of 16. However, we show that our estimates substantially differ by gender: only girls do more housework as the family size increases. This finding is not sensitive to the measurement of housework, and girls also spend relatively less time on leisure and homework in larger families. There is also heterogeneity by father's SES, as only girls whose father has low SES do more housework as family size rises. This is consistent with high-SES parents being more likely to outsource housework and less likely to ask their children to help with the chores. We also find that the effect of family size on housework at age 16 is larger for girls whose mothers hold conservative attitudes.

The effect of family size in childhood is persistent: at age 34, female cohort members who grew up in large families are more likely to sort into couples in which the housework gender gap is significantly larger with respect to women from smaller families. We again find that women from low-SES families and with conservative mothers are behind this finding. We show that the long-term effect of childhood family size is explained by the adoption of behaviours that are in line with more conservative gender roles. First, women who grew up in large families are less likely to be employed, and when they are employed their commuting time is significantly shorter. They are also more likely to be married to employed partners who, in return, have less time to spend on household chores.

Contextualising our results in an identity formation framework à la Akerlof and Kranton (2000, 2010), it can be argued that women who grew up in large families maximise their utility by respecting the behavioural prescriptions of the traditional gender attitudes into which they were socialised. If this were the case, women would find it fair (or, at least, not sub-optimal) to do more housework than their male partners and there would be no direct cost in terms of welfare (Flèche et al., 2018; 2020). However, identity can be seen as a narrative, and as such can be interpreted as a flexible concept (Sveningsson and Alvesson, 2003; Ashforth, 2000). More specifically, it can adapt to act as a buffer against adverse life events (as in Ibarra, 2003, where changes in working identity are seen as a coping mechanism for unexpected changes in employment status).

Following the same line of thought, one may suggest that girls who grew up in larger families are more likely to adopt a conservative gender identity, in order to rationalise the fact that they are asked to contribute more to chores as the housework load increases. We have shown that these girls perform significantly more housework than their partners when they turn 34 and have worse labour market outcomes: conservative identities of women who grew up in large families, which can partly form as a childhood coping mechanism, have then the potential to develop into a set of constraining norms as in Collier (2016). This is in line with the literature showing that women in charge of most the housework load have limited opportunities for career and skills enhancements (Hirsch, 2005; Manning and Petrongolo, 2008; Russo and Hassink, 2008; Evertsson, 2013). Additionally, as argued by Mandel and Semyonov (2005) and Pettit and Hook (2005), conservative norms have the power to institutionalise economic inequality between women and men.

Declaration of Competing Interest

The authors declare no conflict of interest.

Appendix A. Additional figures and tables

Table A1
Descriptive statistics (childhood estimation sample).

	Mean	SD	Min	Max
Age 16:				
HH tasks	0.25	0.22	0	1
Homework	0.34	0.47	0	1
Leisure	0.41	0.23	0	1
Family Size	2.79	1.11	2	10
Same Sex	0.50		0	1
Balanced	0.65		0	1
Two Natural Parents in Household	0.72	0.45	0	1
Age 5:				
Mother's gender attitude	−0.08	1.02	−3	2
Age 0:				
Female	0.57		0	1
White	0.98		0	1
Birth Order	2.15	1.14	1	8
Age mother left school	15.80	2.02	0	27
Age father left school	16.13	2.59	0	33
Mother's age at cohort member's birth	26.55	5.04	18	46
Father's age at cohort member's birth	29.32	5.90	16	67
High SES	0.34		0	1
Observations	3389			

Table A2
Sample composition at age 16: differences in means.

	(1)	(2)	(1)-(2)	(1)	(3)	(1)-(3)
Family Size	2.785	2.964	−0.179***	2.785	2.555	0.231***
	[1.106]	[1.215]	(0.030)	[1.106]	[1.266]	(0.025)
Same Sex	0.497	0.518	−0.022	0.497	0.506	−0.010
	[0.500]	[0.500]	(0.013)	[0.500]	[0.500]	(0.011)
Balanced	0.654	0.663	−0.009	0.654	0.658	−0.004
	[0.476]	[0.473]	(0.012)	[0.476]	[0.474]	(0.010)
Two natural parents in household	0.835	0.777	0.058***	0.835	0.798	0.038***
	[0.371]	[0.416]	(0.011)	[0.371]	[0.402]	(0.009)
Mother's gender attitudes	−0.081	−0.016	−0.065	−0.081	−0.036	−0.044
	[1.019]	[1.017]	(0.026)	[1.019]	[1.026]	(0.021)
Female	0.571	0.404	0.167***	0.571	0.498	0.073***
	[0.495]	[0.491]	(0.013)	[0.495]	[0.500]	(0.010)
White	0.981	0.983	−0.001	0.981	0.983	−0.002
	[0.135]	[0.131]	(0.003)	[0.135]	[0.128]	(0.003)
Birth order	2.078	2.262	−0.185***	2.078	1.967	0.111***
	[0.95]	[0.981]	(0.025)	[0.950]	[0.984]	(0.020)
Age mother left education	15.802	15.475	0.328***	15.802	15.709	0.094
	[2.018]	[1.582]	(0.047)	[2.018]	[1.847]	(0.040)
Age father left education	16.127	15.669	0.458***	16.127	15.958	0.169***
	[2.586]	[1.971]	(0.060)	[2.586]	[2.348]	(0.050)
Mother's age at cohort member's birth	26.551	26.253	0.298	26.551	26.128	0.423***
	[5.045]	[5.230]	(0.132)	[5.045]	[5.133]	(0.106)
Father's age at cohort member's birth	29.331	28.915	0.416***	29.331	28.912	0.419***
	[6.029]	[6.200]	(0.161)	[6.029]	[6.229]	(0.131)
High SES	0.343	0.245	0.098***	0.343	0.240	0.103***
	[0.475]	[0.430]	(0.012)	[0.475]	[0.421]	(0.010)

Notes: The columns labeled '(1)' refer to the estimation sample at age 16. The column labeled '(2)' refers to the sample of cohort members living in households with at least two children but with missing information on household tasks. The column with label '(3)' instead refers to the overall BCS population with non-missing information on the covariates shown in the table. Columns '(1)-(2)' and '(1)-(3)' refer respectively to the differences in means between column (1) and column (2) and between column (1) and column (3). Standard deviations are in square brackets, while standard errors are reported in parentheses. Statistical significance is coded following the standard notation: *** if the *p*-value is lower than 0.01, ** if the *p*-value is lower than 0.05, * if the *p*-value is lower than 0.01.

Table A3
Descriptive statistics (adulthood estimation sample).

	Mean	SD	Min	Max
Age 34:				
Wife HH Tasks	0.47	0.26	0	1
Husband HH Tasks	0.15	0.12	0	1
Housework Gender Gap	0.32	0.32	−1	1
Married	0.73		0	1
Employed	0.84		0	1
A-level	0.44		0	1
Number of children	1.31	1.09	0	8
Age 16:				
Family Size	2.82	1.10	2	11
Same Sex	0.49		0	1
Balanced	0.66		0	1
Two natural parents in household	0.73	0.44	0	1
Age 5:				
Mother's gender attitude	−0.09	1.02	−3	2
Age 0:				
Female	0.54		0	1
White	0.99		0	1
Birth Order	2.18	1.16	1	8
Age mother left school	15.75	1.86	0	27
Age father left school	16.01	2.41	0	32
Mother's age at cohort member's birth	26.43	5.10	18	46
Father's age at cohort member's birth	29.08	5.87	16	63
High SES	0.32		0	1
Observations	3200			

Table A4
Maternal attitudes at age 5 (only households where the same-sex instrument is equal one).

	Two children (1)	Three plus (2)	Difference (3)
Panel A. Whole sample			
Mother's attitudes toward:			
Maternal employment	−0.033	−0.131	0.098
	[1.010]	[1.003]	(0.063)
Sex equality	0.146	0.082	0.065
	[0.935]	[1.078]	(0.061)
Better life for women	−0.029	−0.080	0.050
	[0.950]	[1.087]	(0.062)
Anti-authoritarian child rearing	0.175	0.097	0.078
	[0.938]	[1.003]	(0.060)
Panel B. Girls			
Mother's attitudes toward:			
Maternal employment	−0.020	−0.133	0.113
	[0.976]	[0.960]	(0.082)
Sex equality	0.118	0.077	0.041
	[0.954]	[1.103]	(0.084)
Better life for women	−0.072	−0.035	−0.037
	[1.025]	[1.048]	(0.087)
Anti-authoritarian child rearing	0.142	0.098	0.044
	[0.942]	[1.043]	(0.082)
Panel C. Boys			
Mother's attitudes toward:			
Maternal employment	−0.050	−0.128	0.078
	[1.054]	[1.053]	(0.098)
Sex equality	0.184	0.087	0.097
	[0.908]	[1.051]	(0.089)
Better life for women	0.025	−0.131	0.157*
	[0.841]	[1.131]	(0.087)
Anti-authoritarian child rearing	0.218	0.096	0.122
	[0.933]	[0.958]	(0.087)

Notes: The table reports differences in means across family types. While we cannot distinguish compliers from infra-marginal individuals, we can expect compliers to populate column 2 (where the same-sex instrument is equal one and family size is above two). All variables are PCA-derived z-scores capturing maternal attitudes toward different topics (maternal employment; sex equality; needs for better life for women; anti-authoritarian child-rearing) at child age 5, computed by the data providers. High scores are to be interpreted as egalitarian/liberal views. Standard deviations are in square brackets, while standard errors are reported in parentheses. Statistical significance is coded following the standard notation: *** if the *p*-value is lower than 0.01, ** if the *p*-value is lower than 0.05, * if the *p*-value is lower than 0.01.

Table A5
First-stage regressions.

	All (1)	Females (2)	Males (3)
Child samples (age 16)			
Same sex	0.054***	0.054***	0.054***
	(0.005)	(0.006)	(0.007)
F-test	138	80	58
Observations	3389	1935	1454
Adult samples (age 34)			
Same sex	0.054***	0.054***	0.053***
	(0.005)	(0.006)	(0.007)
F-test	140	73	64
Observations	3200	1731	1469

Notes: The Table reports first-stage coefficients and Cragg-Donald Wald F-statistics for the same-sex instrument in the childhood estimation samples (age 16) and the adulthood estimation samples (age 34). Standard errors in parentheses. Statistical significance is coded following the standard notation: *** if the *p*-value is lower than 0.01, ** if the *p*-value is lower than 0.05, * if the *p*-value is lower than 0.01.

Table A6
Instrument and controls in childhood: OLS results.

	All (1)	Girls (2)	Boys (3)
A. Child characteristics			
Female	−0.025	-	-
	(0.026)		
White	−0.005	−0.018	−0.010
	(0.012)	(0.013)	(0.009)
B. Parent's characteristics			
Two natural parents in household	0.005	−0.011	0.029
	(0.018)	(0.023)	(0.027)
Mother's gender attitude	−0.080	−0.069	−0.101
	(0.053)	(0.070)	(0.079)
Age mother left school	0.137	0.194*	0.077
	(0.091)	(0.116)	(0.141)
Age father left school	0.036	−0.013	0.105
	(0.110)	(0.138)	(0.174)
Mother's age at cohort member's birth	−0.532***	−0.389*	−0.721***
	(0.160)	(0.219)	(0.242)
Father's age at cohort member's birth	−0.214	−0.094	−0.280
	(0.201)	(0.267)	(0.314)
High SES	−0.002	−0.025	0.028
	(0.014)	(0.019)	(0.022)

Notes: Robust standard errors in parentheses. Each cell shows the reduced-form effect of the instrument 'Same-sex' for each of our controls. All regressions control for the ethnicity of the cohort member, birth order dummies, the cognitive and non-cognitive skills of the cohort member at age 16, a dummy indicating whether the cohort member's parents are still living in the same household, a dummy for the father having a high SES, years of education of the cohort member's parents, age of the parents at birth of the cohort member, an index measuring the attitude of the mother regarding maternal employment, a dummy indicating whether the gender composition of the siblings is balanced and regional dummies. Statistical significance is coded following the standard notation: *** if the p-value is lower than 0.01, ** if the p-value is lower than 0.05, * if the p-value is lower than 0.01. The lowest robust F-statistics is 24.4 and belongs to the 2SLS baseline regression with at least 10 reported tasks, for the boys subsample.

Table A7
A topology of sibling compositions.

Birth order:	First or second		Third +	
Sex of cohort member:	Boy	Girl	Boy	Girl
Instrument:				
$Z = 0$	a	c	a_1	c_1
$Z = 1$ (Two girls)	-	d	b_1	d_1
$Z = 1$ (Two boys)	b	-	b_2	d_2

Table A8
Family size and probability to be partnered at age 34 - 2SLS results.

	Probability to be partnered
Family Size	0.027
	(0.031)
Female	−0.024**
	(0.012)
Observations	4227
F-stat (first stage)	183

Notes: Robust standard errors in parentheses. 'Family size' indicates the number of siblings of the cohort member at age 16, plus one. All regressions control for the ethnicity of the cohort member, birth order dummies, a dummy indicating whether the cohort member's parents are still living in the same household, a dummy for the father having a high SES, years of education of the cohort member's parents, age of the parents at birth of the cohort member, an index measuring the attitude of the mother regarding maternal employment, a dummy indicating whether the gender composition of the siblings is balanced and regional dummies. Family size is instrumented by a dummy equal one if the first two children in the household are of the same sex. Statistical significance is coded following the standard notation: *** if the p-value is lower than 0.01, ** if the p-value is lower than 0.05, * if the p-value is lower than 0.01.

Table A9
Family size at age 16 and adult men's outcomes: 2SLS results.

	Effect of Family Size for Men				
	Whole Sample (1)	Low SES (2)	High SES (3)	Conserv. Mothers (4)	Non-Conserv. Mothers (5)
Housework gender gap	0.023	−0.010	0.065	−0.046	0.059
	(0.044)	(0.060)	(0.059)	(0.079)	(0.054)
Educational Attainment (age 34)					
Age left FT education	0.158	0.028	0.745	−0.822	0.534
	(0.513)	(0.676)	(0.784)	(0.870)	(0.617)
At least A-level	−0.026	−0.053	0.058	−0.096	−0.009
	(0.063)	(0.084)	(0.096)	(0.105)	(0.078)
Labour Market Outcomes (age 34)					
Not employed	0.021	0.021	0.038	−0.047	0.052
	(0.039)	(0.054)	(0.049)	(0.050)	(0.053)
Monthly wage (log)†	−0.031	−0.029	0.033	0.099	−0.046
	(0.105)	(0.114)	(0.231)	(0.263)	(0.140)
Weekly working hours†	1.068	−0.227	1.843	4.799	−0.553
	(2.355)	(3.274)	(2.816)	(3.094)	(3.168)
Commuting time†	−0.083	−0.064	−0.147	−0.266**	0.011
	(0.071)	(0.094)	(0.104)	(0.117)	(0.089)
Employed partner	−0.091	−0.090	−0.066	−0.006	−0.162*
	(0.068)	(0.093)	(0.088)	(0.099)	(0.090)
Demographic characteristics (age 34)					
Married	0.118*	0.176*	0.007	0.145	0.122
	(0.068)	(0.097)	(0.091)	(0.104)	(0.088)
Having a least one child	0.040	0.015	0.061	0.002	0.082
	(0.070)	(0.095)	(0.098)	(0.106)	(0.089)
Number of children	0.219	0.324	0.035	0.192	0.285
	(0.154)	(0.217)	(0.196)	(0.239)	(0.194)

Notes: Robust standard errors in parentheses. The Table reports 2SLS estimates of the coefficient for 'Family size' (the number of siblings of the cohort member at age 16) for different dependent variables. All regressions control for the ethnicity of the cohort member, birth order dummies, a dummy indicating whether the cohort member's parents are still living in the same household, a dummy for the father having a high SES, years of education of the cohort member's parents, age of the parents at birth of the cohort member, an index measuring the attitude of the mother regarding maternal employment, a dummy indicating whether the gender composition of the siblings is balanced and regional dummies. Family size is instrumented by a dummy equal one if the first two children in the household are of the same sex. Statistical significance is coded following the standard notation: *** if the *p*-value is lower than 0.01, ** if the *p*-value is lower than 0.05, * if the *p*-value is lower than 0.01.
† The regressions based on these outcomes are based on a subsample of employed cohort members. Results for these outcomes are similar when including also individuals who are not employed and conditioning on employment.

Table A10
Family size at age 16 and household gender gaps at age 42.

	Low SES (1)	High SES (2)	Conserv. Mothers (3)	Non-Conserv. Mothers (4)
Panel A. Women				
Family Size	0.090**	0.082	0.223*	0.049
	(0.046)	(0.106)	(0.129)	(0.043)
Observations	1179	552	430	1301
F-stat (first stage)	55	17	12	61
Panel B. Men				
Family Size	−0.056	0.043	−0.014	−0.016
	(0.049)	(0.052)	(0.061)	(0.046)
Observations	996	473	391	1078
F-stat (first stage)	34	39	24	44

Notes: Robust standard errors in parentheses. 'Family size' indicates the number of siblings of the cohort member at age 16, plus one. All regressions control for the ethnicity of the cohort member, birth order dummies, the cognitive and non-cognitive skills of the cohort member at age 16, a dummy indicating whether the cohort member's parents are still living in the same household, a dummy for the father having a high SES, years of education of the cohort member's parents, age of the parents at birth of the cohort member, an index measuring the attitude of the mother regarding maternal employment, a dummy indicating whether the gender composition of the siblings is balanced and regional dummies. Family size is instrumented by a dummy equal one if the first two children in the household are of the same sex. Statistical significance is coded following the standard notation: *** if the *p*-value is lower than 0.01, ** if the *p*-value is lower than 0.05, * if the *p*-value is lower than 0.01.

Table A11
Family size at age 16 and the channels: 2SLS results - Women.

	High SES			Non-Conservative Mothers		
	Base (1)	Full (2)	Expl. (3)	Base (4)	Full (5)	Expl. (6)
Family Size	0.021	−0.008	0.029	0.042	−0.004	0.046***
	(0.095)	(0.094)	(0.035)	(0.044)	(0.042)	(0.015)
Contributions						
Not employed			−0.004			0.010
			(0.015)			(0.008)
Commuting time			−0.013			0.012*
			(0.015)			(0.007)
Employed partner			0.017			0.015**
			(0.014)			(0.007)
Married			0.029			0.008
			(0.019)			(0.005)
Observations	552	552	552	1301	1301	1301
F-stat (first stage)	17	17		61	60	

Notes: Robust standard errors in parentheses. The Table reports 2SLS estimates of the coefficient for 'Family size' (the number of siblings of the cohort member at age 16) under different sample specifications. All regressions control for the ethnicity of the cohort member, birth order dummies, a dummy indicating whether the cohort member's parents are still living in the same household, a dummy for the father having a high SES, years of education of the cohort member's parents, age of the parents at birth of the cohort member, an index measuring the attitude of the mother regarding maternal employment, a dummy indicating whether the gender composition of the siblings is balanced and regional dummies. Family size is instrumented by a dummy equal one if the first two children in the household are of the same sex. Statistical significance is coded following the standard notation: *** if the p-value is lower than 0.01, ** if the p-value is lower than 0.05, * if the p-value is lower than 0.01.

Table A12
Family size at age 16 and the channels: 2SLS results - Men.

	Whole sample			Low SES			Conservative Mothers		
	Base (1)	Full (2)	Expl. (3)	Base (4)	Full (5)	Expl. (6)	Base (7)	Full (8)	Expl. (9)
Family Size	0.023	0.003	0.020	−0.010	−0.034	0.024	−0.046	−0.069	0.023
	(0.044)	(0.043)	(0.013)	(0.060)	(0.059)	(0.017)	(0.079)	(0.077)	(0.028)
Contributions									
Not employed			−0.002			−0.002			0.004
			(0.004)			(0.006)			(0.005)
Commuting time			−0.004			−0.004			−0.017
			(0.003)			(0.005)			(0.011)
Employed partner			0.017			0.019			0.019
			(0.011)			(0.014)			(0.022)
Married			0.008			0.010			0.018
			(0.005)			(0.007)			(0.013)
Observations	1469	1469	1469	996	996	996	391	391	391
F-stat (first stage)	64	64		34	34		24	22	

Notes: Robust standard errors in parentheses. The Table reports 2SLS estimates of the coefficient for 'Family size' (the number of siblings of the cohort member at age 16) under different sample specifications. All regressions control for the ethnicity of the cohort member, birth order dummies, a dummy indicating whether the cohort member's parents are still living in the same household, a dummy for the father having a high SES, years of education of the cohort member's parents, age of the parents at birth of the cohort member, an index measuring the attitude of the mother regarding maternal employment, a dummy indicating whether the gender composition of the siblings is balanced and regional dummies. Family size is instrumented by a dummy equal one if the first two children in the household are of the same sex. Statistical significance is coded following the standard notation: *** if the p-value is lower than 0.01, ** if the p-value is lower than 0.05, * if the p-value is lower than 0.01.

Table A13
Family size at age 16 and the channels: 2SLS results - Men.

	High SES			Non-Conservative Mothers		
	Base (1)	Full (2)	Expl. (3)	Base (4)	Full (5)	Expl. (6)
Family Size	0.065 (0.059)	0.056 (0.056)	0.010 (0.020)	0.059 (0.054)	0.040 (0.053)	0.019 (0.016)
Contributions						
Not employed			−0.001 (0.003)			−0.005 (0.005)
Commuting time			−0.002 (0.004)			0.000 (0.003)
Employed partner			0.012 (0.015)			0.017 (0.014)
Married			0.001 (0.009)			0.007 (0.006)
Observations	473	473	473	1078	1078	1078
F-stat (first stage)	39	39		44	42	

Notes: Robust standard errors in parentheses. The Table reports 2SLS estimates of the coefficient for ‘Family size’ (the number of siblings of the cohort member at age 16) under different sample specifications. All regressions control for the ethnicity of the cohort member, birth order dummies, a dummy indicating whether the cohort member’s parents are still living in the same household, a dummy for the father having a high SES, years of education of the cohort member’s parents, age of the parents at birth of the cohort member, an index measuring the attitude of the mother regarding maternal employment, a dummy indicating whether the gender composition of the siblings is balanced and regional dummies. Family size is instrumented by a dummy equal one if the first two children in the household are of the same sex. Statistical significance is coded following the standard notation: *** if the p-value is lower than 0.01, ** if the p-value is lower than 0.05, * if the p-value is lower than 0.01.

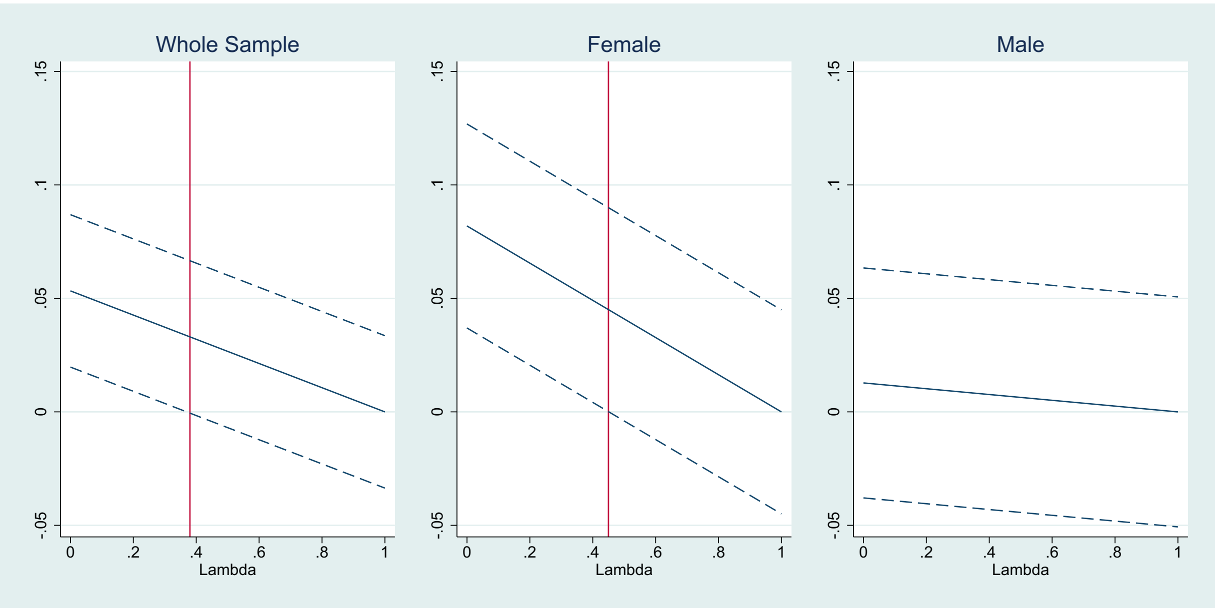

Fig. A1. Relaxing Instrument Exogeneity Note: The solid lines show the estimates of family size instrumented by the sex of the first two children as a function of the value of λ. The dashed lines indicate the 90% confidence intervals. The vertical lines show the minimal value of λ for which the family-size estimates are no longer significantly different from zero.

Appendix B. Robustness checks

B.1. Sensitivity to the definition of the share of housework

We measure the contribution to housework at age 16 using the share of household tasks the cohort member helps with "Regularly". Cohort members report their contribution to twelve different household tasks. Due to missing information and survey filters, the average number of reported tasks in our estimation sample is 9.3 and the median is 10. It can be argued that the number of reported tasks partially drives our estimates. We address this concern in two different ways: we first add the number of reported tasks as an additional control variable, and then re-estimate our main regressions using only cohort members reporting at least ten tasks out of twelve. The results, compared to those from our baseline estimation, appear in the first three rows of Table B1. The first row shows our baseline estimates of family size for the whole sample and then for girls and boys separately. In the second row, controlling for the number of reported tasks makes no difference. We then show the estimated 2SLS coefficients of individuals reporting at least ten tasks in the third row of Table B1. Here, the effect of family size for girls remains unchanged, that for the whole sample is somewhat smaller, and that for boys is negative but not statistically different from zero.

Rather than using tasks that are performed "Regularly" as opposed to "Sometimes" or "Rarely or never", we can also look at the intermediate category "Sometimes". To do so, we assign a score of 1 to tasks performed "Regularly", a score of 0.5 to those performed "Sometimes" and a score of 0 otherwise. As for our original dependent variable, we calculate the share as the average score across the reported tasks. Using this new dependent variable does not affect our conclusions: as revealed by the last line of the Part A of Table B1, an additional family member still increases the whole sample contribution to household tasks and, once again, the result is mostly driven by girls.

B.2. Different definitions of tasks

Our main measure of household contribution uses a set of twelve tasks that have different features. As revealed in Table B2, most tasks are gender-specific. We consider a task to be 'feminine' ('masculine') if the share of girls (boys) reporting doing the task "Regularly" is statistically larger than the share of boys (girls) at the 5% level. Girls spend significantly more time shopping, washing up, cleaning, making the bed, cooking, looking after pets, washing and ironing, and looking after younger siblings, while boys spend more time gardening, cleaning the car, and in DIY activities. The share of girls looking after older people "regularly" is slightly larger than the share of boys, but the difference is not statistically significant at

Table B1
Family size and share of housework at age 16: Sensitivity analysis.

	All (1)	Girls (2)	Boys (3)
A. Sensitivity of the share of housework			
Baseline	0.054**	0.084***	0.014
	(0.023)	(0.029)	(0.036)
Baseline (controlling for no. tasks)	0.061***	0.090***	0.022
	(0.022)	(0.029)	(0.034)
Baseline (at least 10 reported tasks)	0.020	0.082**	−0.043
	(0.022)	(0.032)	(0.029)
Share of housework, counting approach	0.028	0.040*	0.008
	(0.019)	(0.023)	(0.030)
B. Different definitions of tasks			
Feminine	0.064**	0.106***	0.011
	(0.026)	(0.036)	(0.039)
Masculine	0.034	0.051*	0.011
	(0.025)	(0.027)	(0.047)
Unconditional	0.063**	0.104***	0.011
	(0.026)	(0.036)	(0.037)
Increasing in family size	0.063**	0.108***	0.008
	(0.027)	(0.037)	(0.038)
Insensitive to family size	0.027	0.033	0.009
	(0.027)	(0.033)	(0.046)
Excluding care-giving	0.051**	0.091***	−0.001
	(0.024)	(0.032)	(0.035)

Notes: Robust standard errors in parentheses. The Table reports 2SLS estimates of the coefficient for 'Family size' (the number of siblings of the cohort member at age 16) under different definitions of the dependent variable. All regressions control for the ethnicity of the cohort member, birth order dummies, a dummy indicating whether the cohort member's parents are still living in the same household, a dummy for the father having a high SES, years of education of the cohort member's parents, age of the parents at birth of the cohort member, an index measuring the attitude of the mother regarding maternal employment, a dummy indicating whether the gender composition of the siblings is balanced and regional dummies. Family size is instrumented by a dummy equal one if the first two children in the household are of the same sex. Statistical significance is coded following the standard notation: *** if the p-value is lower than 0.01, ** if the p-value is lower than 0.05, * if the p-value is lower than 0.01. The lowest robust F-statistics is 24.4 and belongs to the 2SLS baseline regression with at least 10 reported tasks, for the boys sub-sample.

Table B2
Distribution of household tasks at age 16.

	Girls	Boys	Difference
1. Shopping	0.293	0.168	0.125***
	[0.455]	[0.374]	(0.016)
2. Washing up	0.503	0.332	0.171***
	[0.500]	[0.471]	(0.018)
3. Cleaning	0.302	0.133	0.170***
	[0.459]	[0.339]	(0.015)
4. Making the bed	0.493	0.319	0.174***
	[0.500]	[0.466]	(0.018)
5. Cooking	0.259	0.099	0.160***
	[0.438]	[0.299]	(0.014)
6. Looking after elders	0.084	0.060	0.025*
	[0.278]	[0.237]	(0.013)
7. Looking after pets	0.510	0.438	0.072***
	[0.500]	[0.496]	(0.021)
8. Washing and/or ironing	0.320	0.061	0.259***
	[0.467]	[0.239]	(0.015)
9. Gardening	0.035	0.147	−0.112***
	[0.184]	[0.354]	(0.011)
10. Cleaning car	0.042	0.149	−0.107***
	[0.200]	[0.356]	(0.011)
11. Painting or decorating	0.039	0.107	−0.067***
	[0.195]	[0.309]	(0.010)
12. Looking after youngsters	0.352	0.190	0.162***
	[0.478]	[0.393]	(0.024)
Observations	1935	1454	

Notes: Each household task is reduced to a dummy equal one if the task is performed regularly, zero otherwise. Standard deviations are in square brackets, while standard errors are reported in parentheses. Statistical significance is coded following the standard notation: *** if the *p*-value is lower than 0.01, ** if the *p*-value is lower than 0.05, * if the *p*-value is lower than 0.01.

Table B3
Family size and time use at age 16: 2SLS results.

	Did homework yesterday			Participation to activities		
	All (1)	Girls (2)	Boys (3)	All (4)	Girls (5)	Boys (6)
Family Size	−0.049	−0.131*	0.018	−0.017	−0.058*	0.038
	(0.053)	(0.078)	(0.072)	(0.023)	(0.030)	(0.036)
Female	0.035*			-0.001		
	(0.020)			(0.008)		
Observations	2177	1202	975	3263	1873	1390
F-stat (first stage)	103	57	49	138	81	58

Notes: Robust standard errors in parentheses. 'Family size' indicates the number of siblings of the cohort member at age 16, plus one. All regressions control for the ethnicity of the cohort member, birth order dummies, a dummy indicating whether the cohort member's parents are still living in the same household, a dummy for the father having a high SES, years of education of the cohort member's parents, age of the parents at birth of the cohort member, an index measuring the attitude of the mother regarding maternal employment, a dummy indicating whether the gender composition of the siblings is balanced and regional dummies. Family size is instrumented by a dummy equal one if the first two children in the household are of the same sex. Statistical significance is coded following the standard notation: *** if the *p*-value is lower than 0.01, ** if the *p*-value is lower than 0.05, * if the *p*-value is lower than 0.01.

the 5% level. In the first two rows of Part B of Table B1 we check whether the effect of family size affects the contribution of cohort members to 'feminine' and 'masculine' tasks differently. We find that, as family size rises, girls perform a significantly larger share of both 'feminine' and 'masculine' tasks (although their contribution to the former is larger than to the latter), while boys do not spend more time in any type of tasks. Note that this partition of housework into 'feminine' and 'masculine' almost perfectly overlaps with the intrinsic periodicity of the tasks (e.g. cooking and making the bed are daily activities, while a car needs to be cleaned less frequently), so that our results can also be interpreted in terms of frequency.

In addition, some of the tasks require the presence of particular items or person in the household. This is the case, for instance, for tasks involving care-giving or those such as cleaning the car and tending to the garden. We cannot of course assume that all households in our sample satisfy the pre-conditions for these kind of tasks to be performed. We then exclude these in row three of Part B of Table B1, where we construct the share of household tasks carried out by cohort members based only on 'unconditional' tasks, i.e. tasks that can be carried out in any household. We find that the effect of family size is even stronger when using this measure of housework contribution.

By pooling together the twelve types of household tasks to create one single measure, we also implicitly assume that all tasks increase equally in family size. This is not unrealistic for some of our tasks, such as shopping, washing up, cleaning, making the bed, cooking, washing and ironing, and looking after youngsters (Bawa and Ghosh, 1999). It is however more difficult to believe that looking after the elderly and pets, gardening, cleaning the car, and painting or decorating are tasks that are more likely to be regularly performed in families with more children. We then expect our main estimates to be driven by the first set of tasks, while the second set can be seen more as a placebo test. This is confirmed in the fourth and fifth rows in Part B of Table B1. A girl who grew up in a large family contributes significantly more to those tasks for which demand likely rises in family size. However, there is no significant effect for girls when considering their contribution to the second set of tasks that we expect to be less sensitive to family size. We continue to find no effect for boys regarding either kind of tasks. The last row of Table B1 excludes care-giving activities and only considers the contribution to household tasks that do not involve social interactions. Again, we find results that are in line with our baseline estimates: a larger family increases girls' contribution to housework but not that of boys.

B.3. Other measures of time: homework and leisure

Time is a finite resource. As family size rises and girls contribute more to household tasks, we expect a reduction in the time they spend on homework and leisure. The overall time allocation of boys should instead remain unchanged. We check this in the BCS by looking at time spent on a variety of activities. We measure time spent on homework from the following question: "How much time did you spend doing homework yesterday?". The respondents were asked to use different time categories. Since more than two-thirds of our estimation sample reported doing no homework, we create a dummy for the cohort member having done at least some homework. Cohort members were also asked to report whether they read at least one book during the four weeks before the interview and if they were members of a sports club, a religious organisation, or any other youth organisation over the last 12 months. We construct an index of leisure activities as the share of activities a cohort member engaged in.

We re-estimate our main model using our measures of homework and leisure as the dependent variables. Table B3 shows the results for the whole sample and by gender. Consistent with our main results, girls spend relatively less time doing homework and are less likely to engage into leisure activities in larger families. As expected, there is no effect for boys.[17] As in Table 2, the effect of family size is stronger for girls who grew up in low-SES families and with conservative mothers (these results are available upon request).

References

Achen, A., Stafford, F., 2005. Data quality of housework hours in the panel study of income dynamics: who really does the dishes?Available at http://www.psidonline.isr.umich.edu/Publications/Papers/achenproxyreports04.pdf.
Akerlof, G., Kranton, R., 2000. Economics and identity. Q. J. Econ. 115, 715–753.
Akerlof, G., Kranton, R., 2010. Identity economics. Econ. Voice 7.
Altintas, E., Sullivan, O., 2016. Fifty years of change updated: cross-national gender convergence in housework. Demogr. Res. 35, 455–469.
Anderton, D., Tsuya, N., Bean, L., Mineau, G., 1987. Intergenerational transmission of relative fertility and life course patterns. Demography 24, 467–480.
Angrist, J., Evans, W., 1998. Children and their parents' labor supply: evidence from exogenous variation in family size. Am. Econ. Rev. 88, 450–477.
Angrist, J., Lavy, V., Schlosser, A., 2010. Multiple experiments for the causal link between the quantity and quality of children. J. Labor Econ. 28, 773–824.
Angrist, J., Pischke, J.-S., 2008. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press.
Angrist, J.D., Imbens, G.W., Rubin, D.B., 1996. Identification of causal effects using instrumental variables. J. Am. Stat. Assoc. 91, 444–455.
Antill, J., Goodnow, J., Russell, G., Cotton, S., 1996. The influence of parents and family context on children's involvement in household tasks. Sex Roles 34, 215–236.
Antman, F., 2014. Spousal employment and intra-household bargaining power. Appl. Econ. Lett. 21, 560–563.
Ashforth, B., 2000. Role Transitions in Organizational Life: An Identity-based Perspective. Routledge.
Åslund, O., Grönqvist, H., 2010. Family size and child outcomes: is there really no trade-off? Labour Econ. 17, 130–139.
Bawa, K., Ghosh, A., 1999. A model of household grocery shopping behavior. Mark. Lett. 10, 149–160.
Baxter, J., 2005. To marry or not to marry: marital status and the household division of labor. J. Fam. Issues 26, 300–321.
Baxter, J., Hewitt, B., Haynes, M., 2008. Life course transitions and housework: marriage, parenthood, and time on housework. J. Marr. Fam. 70 (2), 259–272.
Becker, G., 1960. An economic analysis of fertility. In: Demographic and Economic Change in Developed Countries. Columbia University Press, pp. 209–240.
Becker, G., Lewis, G., 1973. On the interaction between the quantity and quality of children. J. Polit. Econ. 81, S279–S288.
Bertrand, M., Kamenica, E., Pan, J., 2015. Gender identity and relative income within households. Q. J. Econ. 130, 571–614.
Bhalotra, S., Clarke, D., 2019. Twin birth and maternal condition. Rev. Econ. Stat. 101, 853–864.
Bianchi, S., Robinson, J., 1997. What did you do today? children's use of time, family composition, and the acquisition of social capital. J. Marr. Fam. 332–344.
Bianchi, S., Sayer, L., Milkie, M., Robinson, J., 2012. Housework: who did, does or will do it, and how much does it matter? Soc. Forces 91, 55–63.
Bisin, A., Verdier, T., 2001. The economics of cultural transmission and the dynamics of preferences. J. Econ. Theory 97, 298–319.
Bittman, M., England, P., Sayer, L., Folbre, N., Matheson, G., 2003. When does gender trump money? bargaining and time in household work. Am. J. Sociol. 109, 186–214.
Björklund, A., Salvanes, K., 2011. Education and Family Background: Mechanisms and Policies. In: Handbook of the Economics of Education, 3. Elsevier, pp. 201–247.
Black, S., Devereux, P., Salvanes, K., 2005. The more the merrier? the effect of family size and birth order on children's education. Q. J. Econ. 120, 669–700.
Black, S., Devereux, P., Salvanes, K., 2010. Small family, smart family? family size and the iq scores of young men. J. Hum. Res. 45, 33–58.
Blair, S.L., 1992. Children'S participation in household labor: child socialization versus the need for household labor. J. Youth Adolesc. 21, 241–258.
Blair, S.L., 1992. The sex-typing of children's household labor: parental influence on daughters' and sons' housework. Youth Soc. 24, 178–203.
Bloom, D., Canning, D., Fink, G., Finlay, J., 2009. Fertility, female labor force participation, and the demographic dividend. J. Econ. Growth 14, 79–101.

[17] The large difference in sample size between the first three columns of Table B3 and Table 1 reflects that 'time spent in homework yesterday' and 'contribution to housework' were measured using different questionnaires. According to the data provider's documentation, the response rate of the former was much lower than that of the latter. Our measures of 'participation in activities' and 'contribution to housework' were measured using the same questionnaire, and the difference of approximately 100 observations here is due to missing information.

Blundell, R., Chiappori, P.-A., Meghir, C., 2005. Collective labor supply with children. J. Polit. Econ. 113, 1277–1306.
Booth, A., Kee, H.J., 2009. Intergenerational transmission of fertility patterns. Oxf. Bull. Econ. Stat. 71, 183–208.
Brody, C., Steelman, L.C., 1985. Sibling structure and parental sex-typing of children's household tasks. J. Marr. Fam. 265–273.
Cáceres-Delpiano, J., 2006. The impacts of family size on investment in child quality. J. Hum. Res. 41, 738–754.
Carlson, D., Lynch, J., 2013. Housework: cause and consequence of gender ideology? Soc. Sci. Res. 42, 1505–1518.
Cherchye, L., De Rock, B., Vermeulen, F., 2012. Married with children: a collective labor supply model with detailed time use and intrahousehold expenditure information. Am. Econ. Rev. 102, 3377–3405.
Chiappori, P.-A., 1992. Collective labor supply and welfare. J. Polit. Econ. 100, 437–467.
Chiappori, P.-A., Iyigun, M., Weiss, Y., 2009. Investment in schooling and the marriage market. Am. Econ. Rev. 99, 1689–1713.
Collier, P., 2016. The cultural foundations of economic failure: a conceptual toolkit. J. Econ. Behav. Org. 126, 5–24.
Conley, D., Glauber, R., 2006. Parental educational investment and childrens academic risk estimates of the impact of sibship size and birth order from exogenous variation in fertility. J. Hum. Res. 41, 722–737.
Conley, T., Hansen, C., Rossi, P., 2012. Plausibly exogenous. Rev. Econ. Stat. 94, 260–272.
Cools, S., Hart, R.K., 2017. The effect of childhood family size on fertility in adulthood: new evidence from iv estimation. Demography 54, 23–44.
Cordero-Coma, J., Esping-Andersen, G., 2018. The intergenerational transmission of gender roles: Children's contribution to housework in germany. J. Marr. Fam. 80, 1005–1019.
Cruces, G., Galiani, S., 2007. Fertility and female labor supply in latin america: new causal evidence. Labour Econ. 14, 565–573.
Evertsson, M., 2013. The importance of work: changing work commitment following the transition to motherhood. Acta Sociol. 56, 139–153.
Falck, O., Gold, R., Heblich, S., 2014. E-Lections: voting behavior and the internet. Am. Econ. Rev. 104, 2238–2265.
Farré, L., Vella, F., 2013. The intergenerational transmission of gender role attitudes and its implications for female labour force participation. Economica 80, 219–247.
Fasang, A.E., Raab, M., 2014. Beyond transmission: intergenerational patterns of family formation among middle-class american families. Demography 51, 1703–1728.
Fernández, R., Fogli, A., Olivetti, C., 2004. Mothers and sons: preference formation and female labor force dynamics. Q. J. Econ. 119, 1249–1299.
Flèche, S., Lepinteur, A., Powdthavee, N., 2018. Gender norms and relative working hours: why do women suffer more than men from working longer hours than their partners? AEA Pap. Proc. 108, 163–168.
Flèche, S., Lepinteur, A., Powdthavee, N., 2020. Gender norms, fairness and relative working hours within households. Labour Econ. 65.
Gager, C., Cooney, T., Call, K.T., 1999. The effects of family characteristics and time use on teenagers' household labor. J. Marr. Fam. 982–994.
Geist, C., 2010. MenS and womenS reports about housework. In: Treas, J., Drobnič, S. (Eds.), Dividing the Domestic: Men, Women, and Household Work in Cross-National Perspective. Stanford University Press: Stanford, CA, pp. 217–240.
Gelbach, J., 2016. When do covariates matter? and which ones, and how much? J. Labor Econ. 34, 509–543.
Giménez-Nadal, J.I., Mangiavacchi, L., Piccoli, L., 2019. Keeping inequality at home: the genesis of gender roles in housework. Labour Econ. 58, 52–68.
Grunow, D., Schulz, F., Blossfeld, H.-P., 2012. What determines change in the division of housework over the course of marriage? Int. Sociol. 27, 289–307.
Gupta, S., 2007. Autonomy, dependence, or display? the relationship between married womens earnings and housework. J. Marr. Fam. 69, 399–417.
Gupta, S., Ash, M., 2008. Whose money, whose time? a nonparametric approach to modeling time spent on housework in the united states. Fem. Econ. 14, 93–120.
Hirsch, B., 2005. Why do part-time workers earn less? the role of worker and job skills. Ind. Labor Relat. Rev. 58, 525–551.
Ibarra, H., 2003. Working Identity: Unconventional Strategies for Reinventing your Career. Harvard Business School Press.
International Labour Office, 2018. World Employment and Social Outlook: Trends for Women 2018 Global Snapshot. ILO, Geneva, Switzerland.
Johnston, D., Schurer, S., Shields, M., 2013. Maternal gender role attitudes, human capital investment, and labour supply of sons and daughters. Oxf Econ Pap 66, 631–659.
Jones, D., 2015. The economics of exclusion restrictions in iv models. NBER Working Paper Series 21391.
Kan, M.Y., 2008. Measuring housework participation: the gap between "stylisedg questionnaire estimates and diary-based estimates. Soc. Indic. Res. 86, 381–400.
Kolk, M., 2014. Understanding transmission of fertility across multiple generations–socialization or socioeconomics? Res. Soc. Stratif. Mobil. 35, 89–103.
Kowalski, A. E., 2019. Counting defiers. NBER Working Paper Series 25671.
Lee, J., 2008. Sibling size and investment in childrens education: an asian instrument. J. Popul. Econ. 21, 855–875.
Lyonette, C., Crompton, R., 2015. Sharing the load? partners relative earnings and the division of domestic labour. Work, Employ. Soc. 29, 23–40.
Mandel, H., Semyonov, M., 2005. Family policies, wage structures, and gender gaps: sources of earnings inequality in 20 countries. Am. Sociol. Rev. 70, 949–967.
Manning, A., Petrongolo, B., 2008. The part-time pay penalty for women in britain. Econ. J. 118, F28–F51.
Manning, A., Petrongolo, B., 2017. How local are labor markets? evidence from a spatial job search model. Am. Econ. Rev. 107, 2877–2907.
Mas, A., Pallais, A., 2017. Valuing alternative work arrangements. Am. Eco. Rev. 107, 3722–3759.
McMunn, A., Bird, L., Webb, E., Sacker, A., 2020. Gender divisions of paid and unpaid work in contemporary uk couples. Work, Employ. Soc. 34, 155–173.
Mostafa, T., Wiggins, R., 2015. The impact of attrition and non-response in birth cohort studies: a need to incorporate missingness strategies. Longit. Life Course Stud. 6, 131–146.
Nicoletti, C., Salvanes, K.G., Tominey, E., 2018. The family peer effect on mothers' labor supply. Am. Econ. J.: Appl. Econ. 10, 206–234.
Nybom, M., 2017. The distribution of lifetime earnings returns to college. J. Labor Econ. 35, 903–952.
Peters, J., Haldeman, V., 1987. Time used for household work: a study of school-age children from single-parent, two-parent, one-earner, and two-earner families. J. Fam. Issues 8, 212–225.
Pettit, B., Hook, J., 2005. The structure of women's employment in comparative perspective. Soc. Forces 84, 779–801.
Rosenzweig, M., Wolpin, K., 2000. Natural "natural experiments" in economics. J. Econ. Lit. 38, 827–874.
Russo, G., Hassink, W., 2008. The part-time wage gap: a career perspective. Economist (Leiden) 156, 145–174.
Sayer, L.C., 2010. Trends in Housework. In: Treas, J., Drobnič, S. (Eds.), Dividing the Domestic: Men, Women, and Household Work in Cross-National Perspective. Stanford University Press: Stanford, CA, pp. 19–38.
Schober, P., 2011. The parenthood effect on gender inequality: explaining the change in paid and domestic work when british couples become parents. Eur. Sociol. Rev. 29, 74–85.
Schulz, F., 2021. Housework time within family households: Mothers', fathers', and siblings' contributions. doi:10.31219/osf.io/4exzs.
Stratton, L., 2015. The determinants of housework time. IZA World of Labor.
Sveningsson, S., Alvesson, M., 2003. Managing managerial identities: organizational fragmentation, discourse and identity struggle. Hum. Relat. 56, 1163–1193.
Van Klaveren, C., Van Praag, B., van den Brink, H.M., 2008. A public good version of the collective household model: an empirical approach with an application to british household data. Rev. Econ. Househ. 6, 169–191.
Vella, F., 1994. Gender roles and human capital investment: the relationship between traditional attitudes and female labour market performance. Economica 191–211.
Wooldridge, J., 2010. Econometric Analysis of Cross Section and Panel Data. MIT press.
World Economic Forum, 2018. The global gender gap report. The Global Gender Gap Report. World Economic Forum, Geneva.