

CREA
Discussion
Paper
2020-03

Economics

Center for Research in Economics and Management
University of Luxembourg

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

available online : http://www.fr.uni.lu/recherche/fdef/crea/publications/discussion_papers

Anthony Lepinteur, FHSE, Université du Luxembourg, L
Giorgia Menta, FHSE, Université du Luxembourg, L

January, 2020

For editorial correspondence, please contact: crea@uni.lu
University of Luxembourg
Faculty of Law, Economics and Finance
6, Rue Richard Coudenhove-Kalergi
L-1359 Luxembourg

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

Anthony Lepinteur
University of Luxembourg

Giorgia Menta
University of Luxembourg

January 6, 2020

Abstract

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, larger families at age 16 make cohort members more likely to sort into households with a larger housework gender gap. We show that the persistent effect of family size is due to the adoption of behaviours and attitudes in line with traditional gender roles: a lower likelihood of employment and shorter commutes for women, along with a higher employment probability for men. Last, the family-size effect here is mostly driven by low-income families, so that the outsourcing of household tasks may help to sustain fertility without aggravating the gender housework gap.

Keywords: Housework Gender Gap, Family Size, Instrumental Variables.
JEL codes: J13, J16, D13.

*We would like to thank Michel Beine, 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, and the Workshop on Gender and Economics (CREA). 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.

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 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.* \(forthcoming\)](#) 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](#)). Gender identity develops according to the behavioural prescriptions that are commonly accepted in a given society. 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).

Even though gender-identity theory suggests that the housework gender gap is a by-product of utility-maximising behaviour, the subjective well-being literature shows that the housework gender gap can be perceived as unfair and, as such, it produces lower levels of happiness. Using British, German and U.S. data, Lepinteur *et al.* (2016) and Flèche *et al.* (2018) demonstrate that women report lower subjective well-being when they work longer hours than their partners. This effect mostly reflects that men do not step up and contribute more to housework as their partner's working hours rise, and women perceive this as unfair. The authors also show that these women are more likely to divorce and leave the labour force (or to find jobs with lower weekly working hours).

However, 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). Indirect evidence comes from the literature on the influence of parental labour-force status on the labour-market outcomes and attitudes of their children (Vella, 1994; Fernández *et al.*, 2004).

There are other ways apart from direct exposure to parental gender norms in which children can be socialised into traditional gender roles within the family. One factor that may well play a role in this sense is 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, then 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 who 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. The Appendix provides an extensive discussion of 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 consistently also spend less time on other activities, namely homework and leisure. The effect of family size is mostly found in low-income and conservative households.

We then show that family size at age 16 also affects the division of household tasks in adulthood: at age 34, BCS cohort members from large families are more likely to be in households in which women spend more time than men in household tasks, and 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-income 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 unemployed and to have an employed husband. Conditional on being employed, they are also less likely to commute. We also find consistent evidence of more conservative opinions among individuals who grew up in large families.

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 also 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, fertility and gender norms.

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 endogeneity. 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 [Cáceres-Delpiano \(2006\)](#); [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 [Conley and Glauber \(2006\)](#) and [Black *et al.* \(2010\)](#) find negative and significant effects on private-school attendance and IQ respectively.

[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 women’s 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, physical, educational and social development, health, economic circumstances and gender attitudes. Since the birth wave of the survey in 1970, there have been eight other waves ("sweeps") at ages 5, 10, 16, 26, 30, 34, 38 and 42. 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 the contribution of the cohort member 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”.

We further combine information on household composition at birth and age 16 to create a measure of family size. We do so by combining the younger siblings of the cohort member when she was 16 years old to her older siblings as reported in the birth sweep.¹ 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 (Figure 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 Figure 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 Figure 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. Naïve specifications that do not account

¹As 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.

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.

The endogeneity of family size is commonly addressed via instrumental variables (IV). 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](#)).

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 \(forthcoming\)](#) 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 \(forthcoming\)](#) 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.²

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

$$\begin{aligned} FamSize_i^{16} &= \alpha_1 SameSex_i + \delta_1 X_i + \epsilon_i \\ HhTasks_i^{16} &= \alpha_2 \widehat{FamSize}_i^{16} + \delta_2 X_i + \mu_i \end{aligned} \tag{1}$$

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, the child being of European descent, and the mother having a regular paying job when the study child is 16 years old. We also control for measures of cognitive skills (a dummy for the respondent having an O-level) and non-cognitive skills (the malaise score). 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 control for household income when the study child was 10 years old (as it is not measured at child age 16). 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. Last, we also control for a potential independent effect of the gender mix of all of the siblings by adding a dummy named *Balanced* for there being at least two siblings of different sex in the family.

Similarly to [Cools and Hart \(2017\)](#), we might think that $SameSex_i$ will not satisfy the exclusion restriction and have an impact on cohort members’ allocation of housework that prescinds from family size, despite the rich set of controls included in the analysis. As the contribution to household tasks differs by gender, it might be the case that having a sister rather than a brother will have a negative impact on the contribution to housework of the cohort member via a crowding out effect. The question

²Despite 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.

here is the extent to which plausible failures of the exclusion restriction might bias our 2SLS estimates. Appendix B provides an extensive discussion of the validity of our instrument, additional robustness checks and an analysis of “plausible exogeneity” à la [Conley et al. \(2012\)](#). We argue there that potential violations of the exclusion restriction are unlikely to drive our results – if anything, they could attenuate them.

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 3,389 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 6,349 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). Our estimation-sample children have on average a better family background (richer 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 whose parents have low levels of education, typically have higher attrition rates and non-response rates with respect to females ([Mostafa and Wiggins, 2015](#)). 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 and 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 will consider this potential mediation in Section 5.2. We have arguably fewer reasons to think that the exclusion restriction is violated in this second specification. It is indeed possible that having a sister rather than a brother would make our cohort members more likely to contribute to household tasks in adulthood – if women contribute more to housework in their childhood family then they may have accumulated a set of skills that can be transmitted to their siblings in the form of knowledge spillovers. However we believe that such mechanisms are of second-order and are hence unlikely to affect our estimate in a sizeable way.

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 3,200 observations.³ 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

³The 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 1,500 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.

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*.⁴ 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 5.6 percentage points. This effect is equal to 25 percent of a standard deviation of the dependent variable. One can also notice that the effect of an increase in family size lies between the effect of having a working mother and being a woman.⁵ While both the OLS and IV estimates are positive and significantly different from zero at 5% level, 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.

Figure 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). 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.⁶

We may suspect a smaller effect of family size in families that outsource their housework in the market. The outsourcing of housework is not accurately measured in the BCS, so we use household income as a proxy for the probability of hiring help. We thus expect the effect of family size to be smaller in richer families, as they are more likely to hire in help for home production. We also expect the effect of family size to be stronger in families with traditional opinions regarding gender attitudes. Parents here might believe that their daughters (but not their sons) will face a marriage-

⁴As the instrument is standard in the literature, we only report the F-statistics to confirm that the instrument is strong and valid. The first-stage regressions show a positive and highly-significant estimate for the instrument, ranging from 0.15 to 0.3 across the different estimation samples and sets of controls (the t-test statistics are always over 7). All of the first-stage regressions are available upon request. 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.

⁵One 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.

⁶We 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.

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). Mothers of BCS members were asked their opinions on a variety of subjects in 1980 and the data providers used a principal component analysis to calculate an index of attitudes towards maternal employment. The attitudes and opinions of the fathers were not collected but, under the assumption of assortative mating, we can assume them to be correlated with those of the mother.

The results appear in [Table 2](#), where we split the estimation sample by gender, family income (above or below the third quartile) and mother’s adherence to conservative gender norms (above or below the third quartile of the distribution of the index of attitudes towards maternal employment).⁷ Family size significantly increases the contribution of girls to housework in low-income families but not in richer families. 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.⁸

4.2 Robustness checks

4.2.1 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 A4](#). 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

⁷The results using the median of the distributions of income and mothers’ gender attitudes are qualitatively similar.

⁸We 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.

coefficients of individuals reporting at least ten tasks in the third row of Table A4. 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 A4, an additional family member still increases the whole sample contribution to household tasks and, once again, the result is mostly driven by girls.

4.2.2 Different definitions of tasks

Our main measure of household contribution uses a set of twelve tasks that have different features. As revealed in Table A5, 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 the 5% level. In the first two rows of Part B of Table A4 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 A4, 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 A4. 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 A4 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.

4.2.3 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 A6 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.⁹ As in Table 2, the effect of family size is stronger for girls who grew up in low-income families and with conservative mothers (these results are available upon request).

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 the housework in households formed by BCS respondents and their partners at age 34. To do so, we replicate our IV 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).

The estimates in the first three columns of Table 3 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

⁹The large difference in sample size between the first three columns of Table A6 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 measure 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.

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.

Columns (4) to (9) then ask whether this result is stronger for female or male cohort members. 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 (4) and (5), this is mostly explained by a significantly higher share of household tasks predominantly carried out by the wife.¹⁰ In the last three columns 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-income families or with relatively less conservative 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-income families contribute significantly (at the 5% level) more to household tasks and sort into couples with a higher housework gender gap. We find no significant results for men (see Table A8 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 A9 are qualitatively similar.

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 A8 for male cohort members).

Section 2.2 suggested that we might expect children who grew up in larger families to have lower

¹⁰We 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.

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-income 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, 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 increase 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 the fertility decisions of men are positively influenced by their own family size in childhood, but not at conventional significance

levels (see Table A8). On the contrary, the fertility decisions of women 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.

We may then conclude that the disproportionate rise in housework contribution between girls and boys resulting from larger families enforces the adherence to conservative gender norms and so explains the results in Table 4. Individuals who grew up in families in which their housework participation depended on their gender might internalise the norm that women should perform the lion's share of housework. As such, they may not support women spending most of their time in the labour market. While gender attitudes are not measured in the BCS at age 34, the questionnaire at age 30 includes some opinion questions. To measure the degree to which a respondent holds conservative gender attitudes we use the extent of agreement to the following question: *“Family life suffers if the mother works full-time”*. The possible responses here are: *“Strongly Agree”*, *“Agree”*, *“Neither Agree of Disagree”*, *“Disagree”* and *“Strongly Disagree”*. In particular, we use a dummy for the respondent saying they *“Strongly Agree”* or *“Agree”* with this statement to see whether a stronger adherence to conservative gender norms is causally affected by family size at age 16. The results appear in the last rows of Table 4 and can be interpreted as a convergent validity test of a shift towards more traditional behaviour and attitudes. A larger family increases the probability that female respondents have conservative gender norms, although the estimate is only significant for women who grew up in low-income families. Women who grew up in large families are also more likely to disagree with the following statements *“Kids benefit if mum has a job outside home”* and *“A mother and her family are happier if she goes out to work”* (results are available upon request).¹¹

The results above suggest that women who grew up in large families, and particularly low-income 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? We show the effects

¹¹One may worry that the significance of our estimates in Table 4 is a result of multiple hypothesis testing. In the Table, 16 out of 60 estimated coefficients are statistically significant at least at a 10% level. The probability that 16 or more out of 60 coefficients are significant at the 10% level by chance is only 0.02%. We are also reassured by the fact that in the subsample of regressions based on women who grew up in low-income families - for whom our results appear to be stronger - the probability that 4 or more coefficients are significantly different from zero at the 5% level by chance is 0.22%.

of childhood family size on the housework gender gap across our different samples in the first column of Table 5. We then add one by one the different adult characteristics that are influenced by family size in childhood, in order to appraise them as potential mediators. As revealed in columns (2) to (4), controlling for the employment probability and commuting time of the cohort members, as well as the employment status of their husbands, is sufficient to explain the persistence of the housework gender gap. Women in Panel 2 (i.e. from low-income families) are an exception: the effect of family size persists even when we hold constant these measures of labour-market outcomes. However, this estimated effect is no longer different from zero once marital status is kept constant. These results confirm that the long-lasting influence of family size in childhood on the housework gender gap is mostly due to the adoption of behaviours that conform to traditional gender roles.¹²

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.

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 differ substantially 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 household income, as girls from low-income families do more housework as family size rises, but not those in richer households. This is consistent with richer 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.

¹²Controlling for the other adult outcomes shown in Table 4 does not change the estimated coefficient on family size. Note also that none of the estimates in Panel 3 and 5 were significantly different from zero in the first column and the inclusion of the different channels unsurprisingly does not change them qualitatively. We replicate this analysis for men in Table A10.

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-income 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.

What is then the role of public policy given our results. In a model of identity formation 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 ([Lepinteur et al., 2016](#); [Flèche et al., 2018](#)). However, identity can be seen as a narrative, and as such can be interpreted as a flexible concept ([Sveningsson and Alvensson, 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.

While identity and attitudes cannot realistically be changed in the short to medium run, we can

arguably foresee a role for policies to intervene in alleviating such constraints for low-income families. As an example, rendering the outsourcing of housework easier might help to sustain fertility without aggravating the housework gender gap. Further evidence is needed for the precise evaluation of the potential welfare losses associated with the persistence of the housework gender gap.

References

- Akerlof, George, and Rachel Kranton. 2000. "Economics and identity." *Quarterly Journal of Economics*, 115: 715–753.
- Akerlof, George, and Rachel Kranton. 2010. "Identity economics." *The Economists' Voice*, 7.
- Altintas, Evrim, and Oriel Sullivan. 2016. "Fifty years of change updated: Cross-national gender convergence in housework." *Demographic Research*, 35: 455–469.
- Anderton, Douglas, Noriko O. Tsuya, Lee Bean, and Geraldine Mineau. 1987. "Intergenerational transmission of relative fertility and life course patterns." *Demography*, 24: 467–480.
- Angrist, Joshua, and Jörn-Steffen Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.
- Angrist, Joshua, and William Evans. 1998. "Children and their parents' labor supply: Evidence from exogenous variation in family size." *American Economic Review*, 88: 450–77.
- Angrist, Joshua, Victor Lavy, and Analia Schlosser. 2010. "Multiple experiments for the causal link between the quantity and quality of children." *Journal of Labor Economics*, 28: 773–824.
- Antill, John, Jacqueline Goodnow, Graeme Russell, and Sandra Cotton. 1996. "The influence of parents and family context on children's involvement in household tasks." *Sex Roles*, 34: 215–236.
- Antman, Francisca. 2014. "Spousal employment and intra-household bargaining power." *Applied Economics Letters*, 21: 560–563.
- Ashforth, Blake. 2000. *Role Transitions in Organizational Life: An Identity-based Perspective*. Routledge.
- Åslund, Olof, and Hans Grönqvist. 2010. "Family size and child outcomes: Is there really no trade-off?" *Labour Economics*, 17: 130–139.
- Bawa, Kapil, and Avijit Ghosh. 1999. "A model of household grocery shopping behavior." *Marketing Letters*, 10: 149–160.
- Baxter, Janeen. 2005. "To marry or not to marry: Marital status and the household division of labor." *Journal of Family Issues*, 26: 300–321.
- Baxter, Janeen, Belinda Hewitt, and Michele Haynes. 2008. "Life course transitions and housework: Marriage, parenthood, and time on housework." *Journal of Marriage and Family*, 70(2): 259–272.
- Becker, Gary. 1960. "An economic analysis of fertility." In *Demographic and Economic Change in Developed Countries*. 209–240. Columbia University Press.
- Becker, Gary, and Gregg Lewis. 1973. "On the Interaction between the Quantity and Quality of Children." *Journal of Political Economy*, 81: S279–S288.
- Bertrand, Marianne, Emir Kamenica, and Jessica Pan. 2015. "Gender identity and relative income within households." *Quarterly Journal of Economics*, 130: 571–614.
- Bhalotra, Sonia, and Damian Clarke. forthcoming. "Twin birth and maternal condition." *Review of Economics and Statistics*.
- Bianchi, Suzanne, and John Robinson. 1997. "What did you do today? Children's use of time, family composition, and the acquisition of social capital." *Journal of Marriage and the Family*, 332–344.

- Bianchi, Suzanne, Liana Sayer, Melissa Milkie, and John Robinson.** 2012. "Housework: Who did, does or will do it, and how much does it matter?" *Social Forces*, 91: 55–63.
- Bisin, Alberto, and Thierry Verdier.** 2001. "The economics of cultural transmission and the dynamics of preferences." *Journal of Economic Theory*, 97: 298–319.
- Bittman, Michael, Paula England, Liana Sayer, Nancy Folbre, and George Matheson.** 2003. "When does gender trump money? Bargaining and time in household work." *American Journal of Sociology*, 109: 186–214.
- Björklund, Anders, and Kjell Salvanes.** 2011. "Education and family background: Mechanisms and policies." In *Handbook of the Economics of Education*. Vol. 3, 201–247. Elsevier.
- Black, Sandra, Paul Devereux, and Kjell Salvanes.** 2005. "The more the merrier? The effect of family size and birth order on children's education." *The Quarterly Journal of Economics*, 120: 669–700.
- Black, Sandra, Paul Devereux, and Kjell Salvanes.** 2010. "Small family, smart family? Family size and the IQ scores of young men." *Journal of Human Resources*, 45(1): 33–58.
- Blair, Sampson Lee.** 1992a. "Children's participation in household labor: Child socialization versus the need for household labor." *Journal of Youth and Adolescence*, 21: 241–258.
- Blair, Sampson Lee.** 1992b. "The sex-typing of children's household labor: Parental influence on daughters' and sons' housework." *Youth and Society*, 24: 178–203.
- Bloom, David, David Canning, Günther Fink, and Jocelyn Finlay.** 2009. "Fertility, female labor force participation, and the demographic dividend." *Journal of Economic Growth*, 14: 79–101.
- Blundell, Richard, Pierre-André Chiappori, and Costas Meghir.** 2005. "Collective labor supply with children." *Journal of Political Economy*, 113: 1277–1306.
- Booth, Alison, and Hiau Joo Kee.** 2009. "Intergenerational transmission of fertility patterns." *Oxford Bulletin of Economics and Statistics*, 71: 183–208.
- Brody, Charles, and Lala Carr Steelman.** 1985. "Sibling structure and parental sex-typing of children's household tasks." *Journal of Marriage and the Family*, 265–273.
- Cáceres-Delpiano, Julio.** 2006. "The impacts of family size on investment in child quality." *Journal of Human Resources*, 41: 738–754.
- Carlson, Daniel, and Jamie Lynch.** 2013. "Housework: Cause and consequence of gender ideology?" *Social Science Research*, 42: 1505–1518.
- Cherchye, Laurens, Bram De Rock, and Frederic Vermeulen.** 2012. "Married with children: A collective labor supply model with detailed time use and intrahousehold expenditure information." *American Economic Review*, 102: 3377–3405.
- Chiappori, Pierre-André.** 1992. "Collective labor supply and welfare." *Journal of Political Economy*, 100: 437–467.
- Chiappori, Pierre-André, Murat Iyigun, and Yoram Weiss.** 2009. "Investment in schooling and the marriage market." *American Economic Review*, 99: 1689–1713.
- Collier, Paul.** 2016. "The cultural foundations of economic failure: A conceptual toolkit." *Journal of Economic Behavior and Organization*, 126.
- Conley, Dalton, and Rebecca Glauber.** 2006. "Parental educational investment and children's academic risk estimates of the impact of sibship size and birth order from exogenous variation in fertility." *Journal of Human Resources*, 41: 722–737.

- Conley, Timothy, Christian Hansen, and Peter Rossi.** 2012. "Plausibly exogenous." *Review of Economics and Statistics*, 94: 260–272.
- Cools, Sara, and Rannveig Kaldager Hart.** 2017. "The effect of childhood family size on fertility in adulthood: New evidence from IV estimation." *Demography*, 54: 23–44.
- Cruces, Guillermo, and Sebastian Galiani.** 2007. "Fertility and female labor supply in Latin America: New causal evidence." *Labour Economics*, 14: 565–573.
- Evertsson, Marie.** 2013. "The importance of work: Changing work commitment following the transition to motherhood." *Acta Sociologica*, 56: 139–153.
- Falck, Oliver, Robert Gold, and Stephan Heblich.** 2014. "E-lections: Voting Behavior and the Internet." *American Economic Review*, 104: 2238–65.
- Farré, Lidia, and Francis Vella.** 2013. "The intergenerational transmission of gender role attitudes and its implications for female labour force participation." *Economica*, 80: 219–247.
- Fasang, Anette Eva, and Marcel Raab.** 2014. "Beyond transmission: Intergenerational patterns of family formation among middle-class American families." *Demography*, 51: 1703–1728.
- Fernández, Raquel, Alessandra Fogli, and Claudia Olivetti.** 2004. "Mothers and sons: Preference formation and female labor force dynamics." *Quarterly Journal of Economics*, 119: 1249–1299.
- Flèche, Sarah, Anthony Lepinteur, and Nattavudh Powdthavee.** 2018. "Gender norms and relative working hours: Why do women suffer more than men from working longer hours than their partners?" *AEA Papers and Proceedings*, 108: 163–68.
- Gager, Constance, Teresa Cooney, and Kathleen Thiede Call.** 1999. "The effects of family characteristics and time use on teenagers' household labor." *Journal of Marriage and the Family*, 982–994.
- Giménez-Nadal, J Ignacio, Lucia Mangiavacchi, and Luca Piccoli.** 2019. "Keeping inequality at home: The genesis of gender roles in housework." *Labour Economics*, 58: 52–68.
- Grunow, Daniela, Florian Schulz, and Hans-Peter Blossfeld.** 2012. "What determines change in the division of housework over the course of marriage?" *International Sociology*, 27: 289–307.
- Gupta, Sanjiv.** 2007. "Autonomy, dependence, or display? The relationship between married women's earnings and housework." *Journal of Marriage and Family*, 69: 399–417.
- Gupta, Sanjiv, and Michael Ash.** 2008. "Whose money, whose time? A nonparametric approach to modeling time spent on housework in the United States." *Feminist Economics*, 14(1): 93–120.
- Hirsch, Barry.** 2005. "Why do part-time workers earn less? The role of worker and job skills." *Industrial and Labor Relations Review*, 58: 525–551.
- Ibarra, Herminia.** 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, David, Stefanie Schurer, and Michael Shields.** 2013. "Maternal gender role attitudes, human capital investment, and labour supply of sons and daughters." *Oxford Economic Papers*, 66: 631–659.
- Kolk, Martin.** 2014. "Understanding transmission of fertility across multiple generations—Socialization or socioeconomics?" *Research in Social Stratification and Mobility*, 35: 89–103.

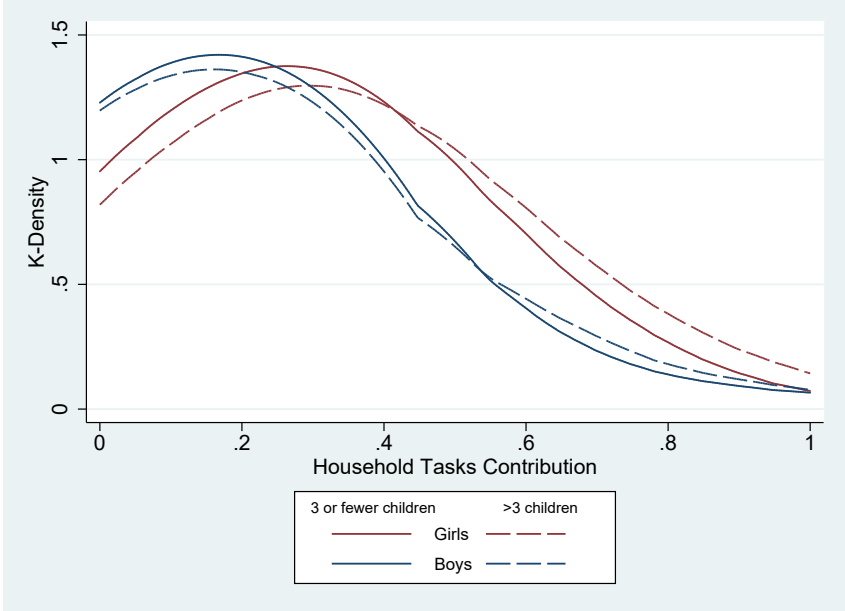
- Lepinteur, Anthony, Sarah Flèche, and Nattavudh Powdthavee.** 2016. “My Baby Takes the Morning Train: Gender Identity, Fairness, and Relative Labor Supply Within Households.” IZA Discussion Papers 10382, Institute for the Study of Labor (IZA).
- Lyonette, Clare, and Rosemary Crompton.** 2015. “Sharing the load? Partners’ relative earnings and the division of domestic labour.” *Work, Employment and Society*, 29: 23–40.
- Mandel, Hadas, and Moshe Semyonov.** 2005. “Family policies, wage structures, and gender gaps: Sources of earnings inequality in 20 countries.” *American Sociological Review*, 70: 949–967.
- Manning, Alan, and Barbara Petrongolo.** 2008. “The part-time pay penalty for women in Britain.” *The economic journal*, 118: F28–F51.
- Manning, Alan, and Barbara Petrongolo.** 2017. “How local are labor markets? Evidence from a spatial job search model.” *American Economic Review*, 107: 2877–2907.
- Mas, Alexandre, and Amanda Pallais.** 2017. “Valuing alternative work arrangements.” *American Economic Review*, 107: 3722–59.
- McMunn, Anne, Lauren Bird, Elizabeth Webb, and Amanda Sacker.** forthcoming. “Gender divisions of paid and unpaid work in contemporary UK couples.” *Work, Employment and Society*.
- Mostafa, Tarek, and Richard Wiggins.** 2015. “The impact of attrition and non-response in birth cohort studies: a need to incorporate missingness strategies.” *Longitudinal and Life Course Studies*, 6: 131–146.
- Nybom, Martin.** 2017. “The distribution of lifetime earnings returns to college.” *Journal of Labor Economics*, 35: 903–952.
- Peters, Jeanne, and Virginia Haldeman.** 1987. “Time used for household work: A study of school-age children from single-parent, two-parent, one-earner, and two-earner families.” *Journal of Family Issues*, 8: 212–225.
- Pettit, Becky, and Jennifer Hook.** 2005. “The structure of women’s employment in comparative perspective.” *Social Forces*, 84: 779–801.
- Rosenzweig, Mark, and Kenneth Wolpin.** 2000. “Natural “natural experiments” in economics.” *Journal of Economic Literature*, 38: 827–874.
- Russo, Giovanni, and Wolter Hassink.** 2008. “The part-time wage gap: A career perspective.” *De Economist*, 156: 145–174.
- Sayer, Liana C.** 2010. “Trends in housework.” In *Dividing the domestic: Men, women, and household work in cross-national perspective.*, ed. Judith Treas and Sonja Drobnič, Chapter 2, 19–38. Stanford University Press: Stanford, CA.
- Schober, Pia.** 2011. “The parenthood effect on gender inequality: Explaining the change in paid and domestic work when British couples become parents.” *European Sociological Review*, 29: 74–85.
- Stratton, Leslie.** 2015. “The determinants of housework time.” *IZA World of Labor*.
- Sveningsson, Stefan, and Mats Alvesson.** 2003. “Managing managerial identities: Organizational fragmentation, discourse and identity struggle.” *Human Relations*, 56: 1163–1193.
- Van Klaveren, Chris, Bernard Van Praag, and Henriette Maassen van den Brink.** 2008. “A public good version of the collective household model: an empirical approach with an application to British household data.” *Review of Economics of the Household*, 6: 169–191.
- Vella, Francis.** 1994. “Gender roles and human capital investment: The relationship between traditional attitudes and female labour market performance.” *Economica*, 191–211.

Wooldridge, Jeffrey. 2010. *Econometric Analysis of Cross Section and Panel Data*. MIT press.

World Economic Forum. 2018. "The Global Gender Gap Report." World Economic Forum, Geneva.

Figures and Tables

Figure 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.

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

	All		Females		Males	
	OLS (1)	2SLS (2)	OLS (3)	2SLS (4)	OLS (5)	2SLS (6)
Family Size	0.014** (0.005)	0.056** (0.023)	0.017** (0.007)	0.087*** (0.030)	0.007 (0.008)	0.014 (0.036)
Working Mother	0.008 (0.008)	0.016* (0.010)	-0.000 (0.011)	0.015 (0.013)	0.023* (0.014)	0.024 (0.015)
Female	0.088*** (0.008)	0.090*** (0.008)				
Observations	3389	3389	1935	1935	1454	1454
F-stat (first stage)						
<i>Family size</i>		135		75		57

Notes: Robust standard errors in parentheses. “Family size” indicates the number of siblings of the cohort member at age 16. All the regressions control for the ethnicity of the cohort member, 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, family income dummies, 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 2: Family size and share of household tasks done “Regularly” at age 16: heterogeneity - 2SLS results

	Females				Males			
	(1) Low Fam. Inc	(2) High Fam. Inc	(3) Conserv. Mothers	(4) Non-Conserv. Mothers	(5) Low Fam. Inc	(6) High Fam. Inc	(7) Conserv. Mothers	(8) Non-Conserv. Mothers
Family Size	0.098*** (0.037)	0.074 (0.051)	0.112** (0.048)	0.056 (0.037)	0.012 (0.041)	0.036 (0.072)	0.115 (0.083)	-0.025 (0.040)
Working Mother	0.025 (0.017)	0.001 (0.020)	-0.000 (0.021)	0.014 (0.016)	0.029 (0.018)	0.009 (0.028)	0.045* (0.027)	0.013 (0.019)
Observations	1356	579	490	1445	984	470	402	1052
F-stat (first stage)	50	29	39	45	39	24	11	48
<i>Family size</i>								

Notes: Robust standard errors in parentheses. “Family size” indicates the number of siblings of the cohort member at age 16. All the regressions control for the ethnicity of the cohort member, 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, family income dummies, 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 0.10.

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

	Whole sample						Females			Males		
	Wife HH Tasks (1)	Husband HH Tasks (2)	Housework Gender Gap (3)	Wife HH Tasks (4)	Husband HH Tasks (5)	Housework Gender Gap (6)	Wife HH Tasks (7)	Husband HH Tasks (8)	Housework Gender Gap (9)			
Family Size	0.052** (0.025)	0.003 (0.012)	0.049* (0.029)	0.071** (0.034)	-0.002 (0.015)	0.074* (0.040)	0.033 (0.036)	0.010 (0.019)	0.023 (0.044)			
Working Mother	0.009 (0.012)	0.004 (0.006)	0.006 (0.015)	0.016 (0.017)	-0.002 (0.007)	0.018 (0.020)	0.002 (0.019)	0.011 (0.010)	-0.009 (0.024)			
Female	0.123*** (0.009)	-0.045*** (0.005)	0.167*** (0.011)									
Observations	3200	3200	3200	1731	1731	1731	1469	1469	1469			
F-stat (first stage)												
<i>Family size</i>	145	145	145	74	74	74	69	69	69			

Notes: Robust standard errors in parentheses. "Family size" indicates the number of siblings of the cohort member at age 16. All the regressions control for the ethnicity of the cohort member, 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, family income dummies, 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 4: Family size at age 16 and adult outcomes: 2SLS results

	Effect of Family Size for Females				
	All	Low Fam.	High Fam.	Conserv.	Non-Conserv.
	Women	Inc	Inc	Mothers	Mothers
	(1)	(2)	(3)	(4)	(5)
Housework gender gap	0.074*	0.110**	-0.013	0.169*	0.041
	(0.040)	(0.045)	(0.086)	(0.101)	(0.043)
<i>Educational Attainment (age 34)</i>					
Age left FT education	0.499	0.605	0.054	-1.154	0.916*
	(0.443)	(0.534)	(0.663)	(0.931)	(0.497)
At least A-level	0.047	0.026	0.109	-0.220	0.108*
	(0.057)	(0.065)	(0.111)	(0.139)	(0.063)
<i>Labour Market Outcomes (age 34)</i>					
Not employed	0.092	0.136*	-0.002	0.074	0.085
	(0.061)	(0.072)	(0.111)	(0.148)	(0.065)
Monthly wage (log) [†]	0.035	-0.296	0.099	0.043	0.211
	(0.125)	(0.146)	(0.262)	(0.263)	(0.137)
Weekly working hours [†]	0.072	1.458	-3.394	5.326	-1.025
	(1.723)	(1.881)	(4.087)	(4.313)	(1.863)
Commuting time [†]	-0.067	-0.131**	0.049	0.109	-0.113*
	(0.047)	(0.063)	(0.137)	(0.137)	(0.064)
Employed partner	0.080***	0.079**	0.084*	0.040	0.080**
	(0.030)	(0.036)	(0.048)	(0.067)	(0.032)
<i>Demographic characteristics (age 34)</i>					
Married	0.109*	0.090	0.178	-0.037	0.117*
	(0.062)	(0.070)	(0.124)	(0.148)	(0.066)
Having a least one child	0.020	0.020	0.017	0.124	-0.010
	(0.055)	(0.062)	(0.115)	(0.136)	(0.060)
Number of children	0.032	0.107	-0.145	-0.144	0.061
	(0.150)	(0.176)	(0.250)	(0.333)	(0.169)
<i>Attitudes (age 30)</i>					
Conservative attitudes	0.114*	0.183**	0.070	-0.057	0.109
	(0.066)	(0.081)	(0.080)	(0.128)	(0.111)

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 the regressions control for the ethnicity of the cohort member, 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, family income dummies, 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 5: Family size at age 16 and the channels: 2SLS results - Females

	Effect of Family Size on the HGG				
	(1)	(2)	(3)	(4)	(5)
Panel 1: Whole Sample	0.074* (0.040)	0.062 (0.040)	0.058 (0.040)	0.041 (0.038)	0.036 (0.038)
Panel 2: Low Family Income	0.110** (0.045)	0.094** (0.045)	0.087** (0.044)	0.072* (0.042)	0.069 (0.042)
Panel 3: High Family Income	-0.013 (0.086)	-0.013 (0.086)	-0.006 (0.086)	-0.050 (0.080)	-0.061 (0.080)
Panel 4: Conservative Mother	0.169* (0.101)	0.161 (0.101)	0.163 (0.101)	0.158 (0.099)	0.159 (0.099)
Panel 5: Non-Conservative Mother	0.041 (0.043)	0.030 (0.043)	0.021 (0.043)	0.002 (0.042)	-0.004 (0.042)
<i>Controlling for:</i>					
<i>Not Employed</i>	.	✓	✓	✓	✓
<i>Commuting time</i>	.	.	✓	✓	✓
<i>Employed partner</i>	.	.	.	✓	✓
<i>Married</i>	✓

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 the regressions control for the ethnicity of the cohort member, 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, family income dummies, 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 0.1.

Appendix A: Additional tables

Table A1: Descriptive statistics (childhood estimation sample)

	Mean	SD	Min	Max
<i>Age 16:</i>				
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	0.50	0	1
Balanced	0.65	0.48	0	1
O-level	0.63	0.40	0	1
Malaise Score	7.68	6.12	0	40
Working Mother	0.48	0.50	0	1
Two Natural Parents in Household	0.72	0.45	0	1
<i>Age 10:</i>				
Banded weekly family income				
<i>Above £250</i>	0.07	0.25	0	1
<i>£200 - £249</i>	0.06	0.24	0	1
<i>£150 - £199</i>	0.18	0.38	0	1
<i>£100 - £149</i>	0.36	0.48	0	1
<i>£50 - £99</i>	0.29	0.45	0	1
<i>£35 - £49</i>	0.03	0.17	0	1
<i>under £35</i>	0.01	0.10	0	1
<i>Age 5:</i>				
Mother's gender attitude	-0.08	1.02	-3	2
<i>Age 0:</i>				
Female	0.57	0.50	0	1
White	0.98	0.14	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
Observations	3389			

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

	(1)	(2)	(1)-(2)	(1)	(3)	(1)-(3)
Family Size	2.785 [1.106]	2.964 [1.215]	-0.179*** (0.030)	2.785 [1.106]	2.555 [1.266]	0.231*** (0.025)
Same Sex	0.497 [0.500]	0.518 [0.500]	-0.022 (0.013)	0.497 [0.500]	0.506 [0.500]	-0.010 (0.011)
Balanced	0.654 [0.476]	0.663 [0.473]	-0.009 (0.012)	0.654 [0.476]	0.658 [0.474]	-0.004 (0.010)
O-level	0.630 [0.396]	0.656 [0.306]	-0.026*** (0.009)	0.630 [0.396]	0.641 [0.362]	-0.011 (0.008)
Malaise Score	9.209 [5.559]	8.656 [5.193]	0.553 (0.495)	9.209 [5.559]	9.118 [5.508]	0.091 (0.139)
Working mother	0.616 [0.487]	0.613 [0.487]	0.002 (0.017)	0.616 [0.487]	0.624 [0.484]	-0.008 (0.012)
Two natural parents in household	0.835 [0.371]	0.777 [0.416]	0.058*** (0.011)	0.835 [0.371]	0.798 [0.402]	0.038*** (0.009)
Banded weekly family income:						
<i>Above £250</i>	0.068 [0.251]	0.045 [0.208]	0.022*** (0.006)	0.068 [0.251]	0.056 [0.230]	0.012 (0.005)
<i>£200 - £249</i>	0.062 [0.242]	0.060 [0.238]	0.002 (0.006)	0.062 [0.242]	0.066 [0.248]	-0.004 (0.005)
<i>£150 - £199</i>	0.180 [0.384]	0.163 [0.370]	0.016 (0.010)	0.180 [0.384]	0.176 [0.381]	0.004 (0.008)
<i>£100- £149</i>	0.364 [0.481]	0.354 [0.478]	0.010 (0.012)	0.364 [0.481]	0.359 [0.480]	0.004 (0.010)
<i>£50- £99</i>	0.287 [0.452]	0.313 [0.464]	-0.026 (0.012)	0.287 [0.452]	0.290 [0.454]	-0.003 (0.009)
<i>£35- £49</i>	0.029 [0.168]	0.049 [0.215]	-0.019*** (0.005)	0.029 [0.168]	0.040 [0.196]	-0.011*** (0.004)
<i>Under £35</i>	0.011 [0.103]	0.016 [0.127]	-0.006 (0.003)	0.011 [0.103]	0.014 [0.116]	-0.003 (0.002)
Mother's gender attitudes	-0.081 [1.019]	-0.016 [1.017]	-0.065 (0.026)	-0.081 [1.019]	-0.036 [1.026]	-0.044 (0.021)
Female	0.571 [0.495]	0.404 [0.491]	0.167*** (0.013)	0.571 [0.495]	0.498 [0.500]	0.073*** (0.010)
White	0.981 [0.135]	0.983 [0.131]	-0.001 (0.003)	0.981 [0.135]	0.983 [0.128]	-0.002 (0.003)
Birth order	2.078 [0.95]	2.262 [0.981]	-0.185*** (0.025)	2.078 [0.950]	1.967 [0.984]	0.111*** (0.020)
Age mother left education	15.802 [2.018]	15.475 [1.582]	0.328*** (0.047)	15.802 [2.018]	15.709 [1.847]	0.094 (0.040)
Age father left education	16.127 [2.586]	15.669 [1.971]	0.458*** (0.060)	16.127 [2.586]	15.958 [2.348]	0.169*** (0.050)
Mother's age at cohort member's birth	26.551 [5.045]	26.253 [5.230]	0.298 (0.132)	26.551 [5.045]	26.128 [5.133]	0.423*** (0.106)
Father's age at cohort member's birth	29.331 [6.029]	28.915 [6.200]	0.416*** (0.161)	29.331 [6.029]	28.912 [6.229]	0.419*** (0.131)

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
<i>Age 34:</i>				
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	0.45	0	1
Employed	0.84	0.36	0	1
A-level	0.44	0.50	0	1
Number of children	1.31	1.09	0	8
<i>Age 16:</i>				
Family Size	2.82	1.10	2	11
Same Sex	0.49	0.50	0	1
Balanced	0.66	0.47	0	1
O-level	0.66	0.39	0	1
Malaise Score	4.68	5.91	0	38
Working Mother	0.42	0.49	0	1
Two natural parents in household	0.73	0.44	0	1
<i>Age 10:</i>				
Banded weekly family income				
<i>Above £250</i>	0.06	0.24	0	1
<i>£200 - £249</i>	0.07	0.25	0	1
<i>£150 - £199</i>	0.18	0.39	0	1
<i>£100 - £149</i>	0.37	0.48	0	1
<i>£50 - £99</i>	0.28	0.45	0	1
<i>£35 - £49</i>	0.03	0.16	0	1
<i>under £35</i>	0.01	0.11	0	1
<i>Age 5:</i>				
Mother's gender attitude	-0.09	1.02	-3	2
<i>Age 0:</i>				
Female	0.54	0.50	0	1
White	0.99	0.10	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
Observations	3200			

Table A4: 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	0.056** (0.023)	0.087*** (0.030)	0.014 (0.036)
Baseline (controlling for no. tasks)	0.063*** (0.023)	0.093*** (0.030)	0.021 (0.035)
Baseline (at least 10 reported tasks)	0.015 (0.013)	0.054*** (0.018)	-0.038* (0.022)
Share of housework, counting approach	0.018* (0.010)	0.028** (0.012)	0.001 (0.017)
B. Different definitions of tasks			
Feminine	0.039*** (0.014)	0.065*** (0.018)	0.000 (0.022)
Masculine	0.016 (0.014)	0.029** (0.014)	-0.007 (0.027)
Unconditional	0.039*** (0.014)	0.066*** (0.018)	-0.003 (0.021)
Increasing in family size	0.039*** (0.014)	0.069*** (0.019)	-0.002 (0.022)
Insensitive to family size	0.011 (0.014)	0.017 (0.017)	-0.006 (0.026)
Excluding care-giving	0.031** (0.013)	0.058*** (0.016)	-0.010 (0.020)

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, 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, family income dummies, 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 0.1. 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 A5: Distribution of household tasks at age 16

	Girls	Boys	Difference
1. Shopping	0.293 [0.455]	0.168 [0.374]	0.125*** (0.016)
2. Washing up	0.503 [0.500]	0.332 [0.471]	0.171*** (0.018)
3. Cleaning	0.302 [0.459]	0.133 [0.339]	0.170*** (0.015)
4. Making the bed	0.493 [0.500]	0.319 [0.466]	0.174*** (0.018)
5. Cooking	0.259 [0.438]	0.099 [0.299]	0.160*** (0.014)
6. Looking after elders	0.084 [0.278]	0.060 [0.237]	0.025* (0.013)
7. Looking after pets	0.510 [0.500]	0.438 [0.496]	0.072*** (0.021)
8. Washing and/or ironing	0.320 [0.467]	0.061 [0.239]	0.259*** (0.015)
9. Gardening	0.035 [0.184]	0.147 [0.354]	-0.112*** (0.011)
10. Cleaning car	0.042 [0.200]	0.149 [0.356]	-0.107*** (0.011)
11. Painting or decorating	0.039 [0.195]	0.107 [0.309]	-0.067*** (0.010)
12. Looking after youngsters	0.352 [0.478]	0.190 [0.393]	0.162*** (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 0.1.

Table A6: Family size and time use at age 16: 2SLS results

	Did homework yesterday			Participation to activities		
	All (1)	Females (2)	Males (3)	All (4)	Females (5)	Males (6)
Family Size	-0.049 (0.054)	-0.133* (0.079)	0.025 (0.072)	-0.016 (0.023)	-0.060** (0.030)	0.042 (0.036)
Working Mother	-0.027 (0.026)	-0.015 (0.036)	-0.048 (0.038)	-0.018* (0.010)	-0.024* (0.014)	-0.015 (0.017)
Female	0.042** (0.020)			-0.003 (0.008)		
Observations	2177	1202	975	3263	1873	1390
F-stat (first stage)						
<i>Family size</i>	101	54	50	136	77	58

Notes: Robust standard errors in parentheses. “Family size” indicates the number of siblings of the cohort member at age 16. All the regressions control for the ethnicity of the cohort member, 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, family income dummies, 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 0.1.

Table A7: Family size and the probability of being partnered at age 34 - 2SLS results

	Probability to be partnered
Family Size	0.038 (0.036)
Working Mother	0.003 (0.018)
Female	0.033** (0.014)
Observations	4227
F-stat (first stage)	
<i>Family size</i>	183

Notes: Robust standard errors in parentheses. “Family size” indicates the number of siblings of the cohort member at age 16. All the regressions control for the ethnicity of the cohort member, 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, family income dummies, 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 0.1.

Table A8: Family size at age 16 and adult outcomes: 2SLS results

	Effect of Family Size for Males				
	Whole Sample (1)	Low Fam. Inc (2)	High Fam. Inc (3)	Conserv. Mothers (4)	Non-Conserv. Mothers (5)
Housework gender gap	0.023 (0.044)	0.012 (0.050)	0.031 (0.092)	-0.045 (0.079)	0.061 (0.055)
<i>Educational Attainment (age 34)</i>					
Age left FT education	0.293 (0.478)	0.739 (0.576)	-0.729 (0.701)	-0.378 (0.771)	0.462 (0.584)
At least A-level	-0.011 (0.057)	0.009 (0.064)	-0.049 (0.125)	-0.026 (0.095)	-0.030 (0.072)
<i>Labour Market Outcomes (age 34)</i>					
Not employed	0.021 (0.038)	0.014 (0.042)	0.055 (0.087)	-0.053 (0.050)	0.058 (0.054)
Monthly wage (log) [†]	-0.031 (0.105)	-0.029 (0.114)	0.033 (0.231)	0.099 (0.263)	-0.046 (0.140)
Weekly working hours [†]	2.020 (1.483)	0.801 (1.601)	6.561* (3.846)	2.662 (1.974)	2.371 (2.060)
Commuting time [†]	-0.076 (0.072)	-0.057 (0.078)	-0.173 (0.162)	-0.238** (0.111)	0.042 (0.094)
Employed partner	-0.092 (0.068)	-0.057 (0.075)	-0.198 (0.158)	0.005 (0.099)	-0.161* (0.091)
<i>Demographic characteristics (age 34)</i>					
Married	0.115* (0.067)	0.153** (0.076)	0.021 (0.147)	0.137 (0.100)	0.106 (0.089)
Having a least one child	0.049 (0.070)	0.029 (0.078)	0.112 (0.153)	-0.006 (0.106)	0.077 (0.091)
Number of children	0.236 (0.153)	0.287 (0.178)	0.073 (0.302)	0.178 (0.234)	0.272 (0.197)
<i>Attitudes (age 30)</i>					
Conservative attitudes	0.095 (0.065)	0.070 (0.080)	0.109 (0.111)	0.207* (0.121)	0.041 (0.074)

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 the regressions control for the ethnicity of the cohort member, 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, family income dummies, 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 A9: Family size at age 16 and household gender gaps at age 42

	Females				Males			
	Low Fam. Inc (1)	High Fam. Inc (2)	Conserv. Mothers (3)	Non-Conserv. Mothers (4)	Low Fam. Inc (5)	High Fam. Inc (6)	Conserv. Mothers (7)	Non-Conserv. Mothers (8)
Family Size	0.087* (0.047)	0.084 (0.083)	0.236* (0.130)	0.042 (0.042)	-0.017 (0.040)	-0.025 (0.078)	-0.003 (0.060)	-0.018 (0.047)
Working Mother	0.005 (0.024)	-0.098*** (0.034)	-0.054 (0.039)	-0.021 (0.023)	-0.002 (0.023)	-0.057 (0.036)	-0.042 (0.035)	-0.007 (0.024)
Observations	1207	524	430	1301	995	474	391	1078
F-stat (first stage)	51	29	12	65	50	19	28	45
<i>Family size</i>								

Notes: Robust standard errors in parentheses. "Family size" indicates the number of siblings of the cohort member at age 16. All the regressions control for the ethnicity of the cohort member, 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, family income dummies, 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 0.1.

Table A10: Family size at age 16 and the channels: 2SLS results - Males

	Effect of Family Size on the HGG				
	(1)	(2)	(3)	(4)	(5)
Panel 1: Whole Sample	0.023 (0.044)	0.025 (0.044)	0.028 (0.044)	0.011 (0.043)	0.004 (0.042)
Panel 2: Low Family Income	0.012 (0.050)	0.014 (0.050)	0.016 (0.050)	0.003 (0.048)	-0.005 (0.048)
Panel 3: High Family Income	0.031 (0.092)	0.032 (0.092)	0.046 (0.092)	0.014 (0.091)	0.014 (0.089)
Panel 4: Conservative Mother	-0.045 (0.079)	-0.052 (0.079)	-0.038 (0.080)	-0.066 (0.077)	-0.071 (0.076)
Panel 5: Non-Conservative Mother	0.061 (0.055)	0.066 (0.056)	0.065 (0.056)	0.050 (0.055)	0.043 (0.054)
<i>Controlling for:</i>					
<i>Not Employed</i>	.	✓	✓	✓	✓
<i>Commuting time</i>	.	.	✓	✓	✓
<i>Employed partner</i>	.	.	.	✓	✓
<i>Married</i>	✓

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 the regressions control for the ethnicity of the cohort member, 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, family income dummies, 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 0.10.

Appendix B: Discussion of the exclusion restriction

We would have problems with our identification strategy 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. If we assume ex-ante one of our key findings, that is girls contribute more to housework than boys ([Gager et al., 1999](#)), the sex composition of the two first-born children in the household may well directly affect all siblings' contribution to housework via a substitution effect: relative to having a brother, a sister will always reduce the residual amount of housework to be carried out. This in turn will affect the cohort member's housework contribution, through a channel other than family size. 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 this threat to identification. We below present an extensive discussion of the conditions under which the exclusion restriction of our instrument might not hold, a set of tests and an analysis of the plausible exogeneity of the instrument following the approach of [Conley et al. \(2012\)](#). We argue that in most cases the potential violation would bias our estimates towards zero and provide some descriptive evidence against one problematic case.

We first carry out a placebo test. 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 related to family size. We here appeal to the estimates shown in the fourth and fifth rows of Panel B of Table [A4](#). These respectively show the estimated family-size coefficients for household tasks that probably do rise in family size and those that likely do 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. This suggests that the instrument is unlikely to affect the contribution to household tasks other than via impact on family size.¹³

We now go one step further and explore whether our instrument is systematically correlated with observable characteristics of the cohort members and their parents at age 16. We follow [Falck et al. \(2014\)](#) and derive reduced-form estimates from the regression of cohort members' and their parents'

¹³This 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.

characteristics at age 16 (normally used as control variables in our baseline regressions) on the same-sex instrument and all other controls. A significant correlation here would indicate a potential violation of the exclusion restriction. More specifically, if observable characteristics are correlated with our instrument, we may expect that unobservable characteristics are too.

Table B1 shows the results for the whole sample and then by the cohort member’s gender. Our instrument does not seem to be systematically correlated with the different characteristics reported. Note that we do not report the correlation between the same-sex instrument and birth order as birth order mechanically rises with family size. Out of the 52 estimates, only eight are significantly different from zero, the majority of which are so only at the 10% level. The only highly significant coefficients are those for 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 were born on average six months later than other mothers with at least two children. The absence of selection into consecutive same-sex pregnancies based on observable characteristics mitigates our concerns about systematic correlations with unobservable characteristics.

We can go one step further and allow for direct effects of the instrument on the dependent variable. As argued by Conley *et al.* (2012), the exclusion restriction is debatable for most instrumental variables, which is why they developed inference methods that are consistent with instruments being 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 regression of 2SLS estimation. We here follow Nybom (2017) and relax the exogeneity of the instrument as follows:

$$\begin{aligned}
 FamSize_i^{16} &= \alpha_1 SameSex_i + \delta_1 X_i + \epsilon_i \\
 HhTasks_i^{16} &= \alpha_2 \widehat{FamSize}_i^{16} + \gamma SameSex_i + \delta_2 X_i + \mu_i
 \end{aligned}
 \tag{3}$$

where γ , the direct effect of the instrument, varies between -0.05 and 0.05. Note that the interval ranges between values of the same magnitude of the effect of family size in column (2) of Table 1. γ can be interpreted as the share of the reduced-form effect of the instrument that is independent of family size. Our objective here is to identify the threshold at which our main 2SLS estimates become

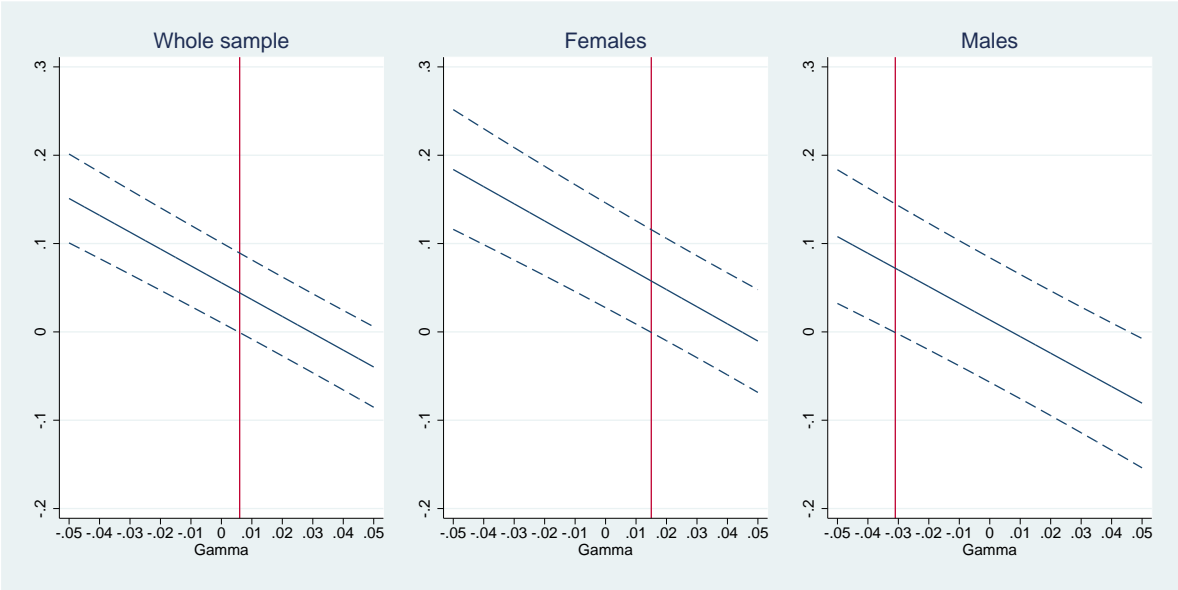
Table B1: Instrument and controls in childhood: OLS results

	All (1)	Girls (2)	Boys (3)
A. Child characteristics			
Female	-0.025 (0.026)	-	-
White	-0.005 (0.012)	-0.018 (0.013)	-0.010 (0.009)
O-Level	0.009 (0.018)	0.006 (0.025)	0.014 (0.028)
Malaise Score	-0.119 (0.253)	-0.160 (0.334)	-0.127 (0.390)
B. Parent's characteristics			
Working mother	-0.039* (0.023)	-0.068** (0.031)	-0.005 (0.033)
Two natural parents in household	0.005 (0.018)	-0.011 (0.023)	0.029 (0.027)
Banded weekly family income			
<i>Above £250</i>	-0.024** (0.011)	-0.023 (0.015)	-0.028 (0.018)
<i>£200 - £249</i>	0.003 (0.012)	0.005 (0.018)	-0.004 (0.016)
<i>£150 - £199</i>	-0.001 (0.019)	-0.018 (0.025)	0.024 (0.030)
<i>£100 - £149</i>	-0.018 (0.025)	-0.004 (0.033)	-0.033 (0.038)
<i>£50 - £99</i>	0.029 (0.024)	0.039 (0.032)	0.016 (0.035)
<i>£35 - £49</i>	0.009 (0.010)	-0.008 (0.012)	0.029* (0.016)
<i>under £35</i>	0.002 (0.006)	0.007 (0.008)	-0.004 (0.008)
Mother's gender attitude	-0.080 (0.053)	-0.069 (0.070)	-0.101 (0.079)
Age mother left school	0.137 (0.091)	0.194* (0.116)	0.077 (0.141)
Age father left school	0.036 (0.110)	-0.013 (0.138)	0.105 (0.174)
Mother's age at cohort member's birth	-0.532*** (0.160)	-0.389* (0.219)	-0.721*** (0.242)
Father's age at cohort member's birth	-0.214 (0.201)	-0.094 (0.267)	-0.280 (0.314)

Notes: Robust standard errors in parentheses. "Family size" indicates the number of siblings of the cohort member at age 16. All the regressions control for the ethnicity of the cohort member, 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, family income dummies, 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 subsample.

insignificant at the 10% level. Figure B1 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. Unsurprisingly, negative values of γ make our 2SLS estimates stronger. However, a positive γ means that the instrument has a direct and positive effect on the contribution to household tasks, that in turn may inflate our main estimates. If γ is greater than 0.01 for the whole sample and greater than 0.02 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 smaller than roughly 20% of our main 2SLS estimate, the effect of family size on housework will remain significantly different from zero.

Figure B1: Relaxing Instrument Exogeneity



Note: The blue lines show the estimates of family size instrumented by the sex of the first two children as a function of the value of γ . The dashed lines indicate the 90% confidence intervals. The red lines show the minimal value of γ for which the family-size estimates are no longer significantly different from zero.

We now argue that if sibling composition has a direct impact on the contribution to housework, it is unlikely to be positive. Table B2 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).¹⁴ 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

¹⁴Here, 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.

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 (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 0.063 for the whole sample and 0.093 for girls (both significant at the 1% level).

Table B2: A topology of sibling compositions

<i>Birth order:</i>	<u>First or second</u>		<u>Third +</u>	
<i>Sex of cohort member:</i>	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