Cheaper by the Dozen: Parental Time Use and Returns to Scale in Raising Children

Todd R. Jones, Ezra Karger, and Valerie Michelman*

Draft first posted on March 20, 2025
This version: March 20, 2025

Abstract

Policies designed to support parents and children vary in whether and to what extent benefits scale with family size. In this paper, we document how the "child penalty"—the effect of having a child on parental outcomes—differs across the intensive and extensive margins of fertility. Using rich individual-level longitudinal data from Australia, we show that parental time spent on childcare and the associated forgone income are highly nonlinear in family size, with the first child incurring significantly larger marginal "costs" than subsequent children. We validate these patterns using pregnancy loss and twin births as instruments for fertility. To illustrate the importance of this return-to-scale in raising children, we provide a conceptual framework in which parents jointly make choices about fertility, work, and childcare. The fixed and marginal costs of each childcare strategy are crucial inputs into the effects of different family policies—such as child tax credits and subsidized childcare—on fertility and parental welfare.

*Jones: Mississippi State University, IZA, and CESifo. toddrjones.com. Karger: Federal Reserve Bank of Chicago. ezrakarger.com. Michelman: University of Delaware. sites.google.com/view/valeriemichelman
Any views expressed do not necessarily reflect those of the Federal Reserve Bank of Chicago or the Federal Reserve System. The authors do not have any relevant and material financial relationships that bear on this research. This paper uses unit record data from Household, Income and Labour Dynamics in Australia Survey [HILDA] conducted by the Australian Government Department of Social Services (DSS). The findings and views reported in this paper, however, are those of the authors and should not be attributed to the Australian Government, DSS, or any of DSS' contractors or partners. DOI: 10.26193/KXNEBO.

1 Introduction

Developed countries use a latticework of family policies to subsidize costs incurred by child-rearing. The goals of these policies vary widely and include increasing fertility, improving child welfare, firming up parental labor force attachments, and decreasing parenthood-induced gender inequality. However, a common feature across policies is that they must, sometimes implicitly but often formulaically, take a stance on how benefit amounts and eligibility scale by household size.[1] Taken at face value, the varied approaches across policies reflect contradictory understandings of the relative costs (benefits) of marginal children to a household (and society).

Despite its crucial role in policy design, there is little systematic evidence about the relative costs to parents of having different numbers of children. Intuitively, the case for some returns-to-scale in child-rearing inputs is clear. A parent with two children can spend a single hour watching both of them. However, the magnitude of returns-to-scale is not obvious. For example, a parent of a one and four-year-old may be able to send the four year old to pre-K for several hours during the day, while still spending those hours caring for the younger child. The burgeoning literature on child-penalties has illuminated differences in magnitudes of "costs" of parenthood, agnostic of family size, between mothers and fathers and across countries (Kleven et al., 2023). Far less is known about the magnitude of costs by family size. What are the magnitudes of returns to scale? Are they comparable across inputs? And what are the resulting implications for designing policy to serve children and families?

To address these questions, we draw on the Household, Income and Labour Dynamics in Australia (HILDA) Survey from Australia to study the returns to scale of raising

[1] For example, in the United States the maximum EITC credit for the 2025 tax year is $649, $4328, $7152, and $8,046 for households with zero, one, two, or at least three qualified children respectively. The federal poverty threshold, which is used in formulae to determine eligibility for many benefits, is $15,060, $20,440, $25,820, $31,200 for families of one, two, three, and four respectively (2024 numbers). The child tax credit is a constant per-child amount for those eligible. For policies such as paid parental leave, the lack of cap or adjustment by number of children is itself a choice to offer equivalent benefit eligibility to parents per child, independent of prior household size.

children. This comprehensive longitudinal household survey has several features that make it a particularly powerful tool for understanding the effects of children on parents. First, as a long-running household panel survey, we can track families over two decades (2001-present), enabling us to observe people from before they have children through their prime childbearing years. This is crucial for analyses in which we want to differentiate effects for small and large households (by completed fertility), which would not be possible in a shorter panel. Second, in addition to standard household survey items of earnings, expenditures, and consumption, HILDA contains a time use module. Time use is rarely gathered in panel surveys, let alone surveys spanning decades. This provides us with unprecedented longitudinal data on non-market labor inputs and childcare responsibilities. Finally, HILDA includes several detailed questions of particular value to our research, such as detailed breakdowns of childcare arrangements and expenses, fertility expectations, and subjective well-being.

We conduct four major empirical exercises. In the first exercise, we use event studies in the style of Kleven et al. (2019) to estimate changes in key outcomes following the time of first birth. Leveraging the long nature of our panel, we conduct these analyses separately by completed fertility (the (predicted) total number of children each parent has had by the end of prime childbearing years). Our empirical approach builds on Kleven et al. (2019) in two key ways. First, we measure the effects of having a child by completed family size to construct baseline estimates of the marginal cost of children. Second, we are not limited to labor market outcomes, but instead paint a fuller picture by also measuring the effects of children on household time use. We document the striking descriptive fact that changes in key parental outcomes in the year following the first birth are similar by completed fertility, before becoming distinguishable in the longer-run. Yet, the magnitude of long-run differences in income and time use effects by completed fertility is dwarfed by the much larger initial "fixed cost" of having a first child. We interpret this descriptive pattern as suggestive of much larger costs of first versus subsequent children.

Our first exercise alone does not rule out that the pattern may be driven by parents of different means and preferences selecting into having different numbers of children. To address that selection question, we conduct two complementary exercises. First, using the timing of reported pregnancies and births, we compare pregnancies that resulted in a birth to those that did not.[2] We separately conduct this pregnancy loss analysis for people who had no prior children and those who have one prior child, allowing us to separately identify the marginal effect of a first child and the marginal effect of a second child on parents. Second, we use the birth of twins as an alternative approach to identify the marginal effect of a second child relative to a first child. Across approaches, we find the implied effects of a first child on parents are substantial and the implied effects of a second child are much smaller.

Our last empirical exercise is a novel specification that leverages the long nature of our panel. We compare the relative short-run response to a first and second birth among people who have the same completed fertility. This allows us to address some of the selection concerns remaining from the twins and pregnancy loss approaches. While those approaches help address selection for the marginal child, they still rely on comparing the effects between different populations, i.e., people with no children who have a pregnancy with no prior children to people with one prior child who have a pregnancy. This approach produces the same pattern of large responses to first birth and much smaller responses to subsequent births.

Our analyses focuses on two key parental inputs: forgone income and time spent on childcare (household production). By looking at income, we capture one of the primary costs of parenthood and our results speak directly to the literature on child penalties, female labor force participation, and fertility. By looking at time spent on childcare, our analysis is not limited to labor market outcomes but instead also includes non-market

[2]There are two important threats to identification here that we discuss at length in Section 2.2. The first is that we cannot distinguish voluntary abortion from miscarriage. The second is that miscarriage may be correlated with potential outcomes and may itself have impacts on potential parents.

work, which is a crucial but rarely measured input to child-rearing.[3] We supplement these core measures with a more holistic picture of costs related to child-rearing, including household consumption and expenditures on market-provided childcare. Our central finding is that the marginal cost of children is not constant: the costs of a first child along the key dimensions of time and income are much larger than those of subsequent children.

Our work contributes to the literature on the impact of children on their parents. Empirically, previous studies of the impact of fertility on parents' labor market outcomes have primarily focused on the impact of second and third children. Much of this focus is due to the common use of twin births and parental preference for having both a boy and a girl as empirical strategies to recover the impact of second and third births.[4] Empirical work on the impact of the first child has only taken off in the past few years, spurred by Lundborg et al. (2017)'s use of IVF as an instrumental variable for female labor-force participation[5] and Kleven et al. (2019)'s event study estimation of the effects of having children on parental income.[6]

Motivated by patterns we document empirically, in the final parts of our paper we explore implications for the following questions: How do the relative costs of raising children vary by childcare strategy? And what are the resulting implications for theory and policy?

The possibility that children cost parents different amounts of time and money at dif-

[3]To the best of our knowledge, Aguilar-Gomez et al. (2022) and Koopmans et al. (2024), are the only other papers to date to document child-penalties on non-market work with time-use data. Neither of these papers investigates patterns by completed fertility, which is our primary focus.

[4]See Clarke (2018) for an excellent summary of empirical work using these and other quasi-experimental strategies.

[5]Lundborg et al. (2017)'s view is that the "current literature on the extensive fertility margin is in its infancy" and "While standard theories of household production, child quality, and economies of scale predict that labor market consequences are stronger at the extensive fertility margin, empirical evidence is rare."

[6]Kleven et al. (2019)'s event study approach isolates the impact of the first child in the short run and estimates the combined effect of all children in the long run. The authors include a version of their main figure split by completed fertility in the Appendix (Figure A.III), and suggest the sustained effect of parenthood on earnings for women 10 years after their first birth is about one-third to two-thirds larger per additional child relative to the effect for mothers of one child.

ferent points in the birth order is touched upon, but not explored in depth, in much of the literature on fertility.[7] Motivated by our empirical results, we bring a variety of marginal costs to the foreground.[8] We consider how choice of childcare strategy may allow parents to trade off fixed and marginal costs. Specifically, raising children requires inputs of typical consumption goods (groceries, clothes, housing, etc.) as well as some combination of parental time, market-based care, and care provided by extended family. Parents choose among childcare strategies that require different mixes of these inputs, and, critically, strategies vary in both average and marginal costs.

To fix ideas, consider a couple consisting of a primary earner and a secondary earner. Parents must obtain or provide children with full-time care, at significant cost, until they are school-aged. Center-based childcare is priced roughly linearly, per-child per-year. Home production of childcare, however, has significant returns to scale. If the secondary earner leaves the labor force to be a primary caregiver, having a second child two years after the first will not result in as many years of missed earnings given that they already left the labor force for the first child. In this case, subsidizing market-provided childcare could decrease fertility if the price reduction lands below the cost of home care for a first child but above the marginal cost of home care for subsequent children. Understanding the structure of these costs is crucial for predicting the impact of different policies on fertility.

[7]Economists have studied questions surrounding fertility for decades. Gary Becker's foundational quantity-quality model posed a solution to the demographic puzzle of his day: what can explain why richer individuals have smaller families, if children are not inferior goods? Becker's solution—as income rises, demand for child "quality" overtakes quantity (Becker and Lewis, 1973)—shaped the following decades of economic work on fertility, birthing hundreds of papers extending and applying the quantity-quality model. However, more recent empirical patterns herald the necessity to shift the direction of inquiry (Doepke et al., 2023). Developed nations have now passed the nadir of what appears to be a U-shaped relationship between income and fertility, and completed fertility is now increasing in income in these nations.

[8]Our work also relates to the strand of fertility research focused on childlessness. Theorists who engage in a serious consideration of childlessness tend to find it valuable to separate fixed and marginal costs. For example, Gobbi (2013), includes a fixed cost of parenthood as a key parameter to which she calibrates her model, and Baudin et al. (2015) analogously estimates a fixed cost of parenthood, separate from the time cost of children, in their structural model.

2 Data

2.1 Australian Household Survey Data

Our analysis uses data from the Household, Income and Labour Dynamics in Australia Survey, or HILDA (Melbourne Institute of Applied Economic and Social Research, 2021).[9] HILDA is an annual household-based longitudinal dataset that has followed individuals for more than 20 years, since its first wave in 2001. HILDA collects data on many features of panelists including time use, employment, income, and expenditures. It also contains detailed information on family structure and other key variables related to children: including fertility expectations, information about pregnancies, and data on who is responsible for childcare (parents, relatives, or daycare).

HILDA is similar in size and structure to the widely-used Panel Study of Income Dynamics (PSID) and National Longitudinal Survey of Youth (NLSY) surveys. As a long-lasting longitudinal survey, households in HILDA attrit, but they do so at rates consistent with other commonly used panels. The initial 2001 wave contained 7,682 households with a total of 19,914 household members, of whom 13,969 were interviewed; 4,787 were children under age 15, and were not interviewed. The survey added new "top-up" households in 2011 to increase the sample size. In 2019,[10] there were 9,664 responding households, with 17,462 people interviewed. In this wave there was a 60.7% response rate among respondents from the initial 2001 wave (who were in scope for the survey), and a 74% response rate among top-up survey respondents.

An advantage to HILDA is that it asks about time use in categories in each survey wave. Respondents report the number of hours they spend *in a typical week* on various categories of activities.[11]. Because the respondent is asked how many hours they spend

[9]We draw on Summerfield et al. (2023) for information in this section.

[10]We end our main sample in 2019 in order to avoid the Covid-19 pandemic's disruption to usual patterns in time use; however, we make use of years after 2019 to measure lifetime fertility.

[11]The categories are: a) Paid employment; b) Traveling to and from a place of paid employment; c) Household errands, such as shopping, banking, paying bills, and keeping financial records (but do not include driving children to school and to other activities); d) Housework, such as preparing meals, washing

in a typical week, their responses are likely more comparable across survey waves than if they had been asked to complete a 24-hour time diary (as in the American Time Use Survey), which can yield more idiosyncratic approaches depending on which day they consider. The categories of time use are also consistent over each survey wave. On the other hand, the categories are not exhaustive. For example, they do not cover leisure or sleep.[12] Respondents are instructed (beginning in wave 2) not to double count time across categories.[13]

2.2 Defining Completed Fertility, Twins, and Pregnancy Loss

In this subsection, we provide an overview of variable construction choices for key variables. More details on these and other variables are available in Appendix B.

Central to several of our analyses is our ability to look at patterns early in adulthood by completed fertility that is realized years later. Ideally, we would be able to track individuals well beyond their childbearing years to ensure the perfect measurement of completed fertility. In practice, we strike a balance between allowing enough years for most people to have finished having children and including more cohorts of survey respondents. We do this by augmenting observed completed fertility for individuals we observe at or beyond the ages when 99% of children have been born (43 years old for women and 51 years old for men) with predicted completed fertility for people whom we observe in the later part of childbearing years (at least 35 years old) but not at the very end. We use the maximum value of completed fertility for individuals we observe at or beyond the ages

dishes, cleaning house, washing clothes, ironing and sewing; e) Outdoor tasks, including home maintenance (repairs, improvements, painting etc.), car maintenance or repairs and gardening; f) Playing with your children, helping them with personal care, teaching, coaching or actively supervising them, or getting them to child care, school and other activities; g) Looking after other people's children (aged under 12 years) on a regular, unpaid basis.; h) Volunteer or charity work (for example, canteen work at the local school; unpaid work for a community club or organization); and i) Caring for a disabled spouse or disabled adult relative, or caring for elderly parents or parents-in-law

[12] Though questions about sleep are asked in some of the survey waves.

[13] However, as with any survey measure, we expect some noise due to potentially misinterpreted instructions.

when 99% of children have been born (43 years old for women and 51 years old for men). For individuals whom we observe as being in the later part of childbearing years (at least 35 years old) but not at the very end (43 years old for women and 51 years old for men), we predict completed fertility as the sum of their last-observed completed fertility value and the number of children they expect to have in the future.

We identify children as twins using two sets of variables. The first method uses the twin identifier variable that is available for respondents, meaning individuals age 15 or older living in a sampled household. This allows us to identify twins among the population of children born in earlier survey years. The second method uses parents' response to a survey question asked in later waves whether they have twins. This question is asked of parents of children who aged 15 and younger. Because of the age restrictions of these two sets of variables are complementary, we combine them to identify twins across the span of survey years.

We define the variable we refer to as pregnancy loss by combining questions about life events in consecutive survey years. Each year, respondents are asked about several potential life events, including whether they or their partner became pregnant and whether they or their partner gave birth or adopted a child. In addition to a yes or no, respondents are asked to indicate how many months ago the event happened, in three-month spans. This allows us to narrow down the timing of when pregnancies started and observe whether or not a birth is reported in months surrounding the projected due date (either in that survey wave or the next), adjusting for the amount of time that elapsed between waves. We use this variable as a proxy for pregnancy loss, overcoming the obstacle that surveyors are specifically instructed not to ask about miscarriage.[14] People we identify as having lost a pregnancy are more likely to report struggles with infertility in the supplementary fertility module, which is asked once every four years.

[14]This instruction is given in several of the waves.

2.3 Descriptive Statistics

Among individuals who appear in at least one wave in which they are at least 30 years old, Table 1 shows that the average age is 36, roughly half are female, most (85%) live in urban areas, roughly three-quarters were born in Australia, one-quarter are single, and the average predicted completed fertility is 2.1.

Fertility patterns and time use in our sample of Australian families are typical of those found in many developed countries, with small family sizes and mothers spending more time on childcare and less time on market work than fathers. Of the 3,875 women who appear in the panel at some point over the age of 40, 14% have no children, 13% have one child, and 37% have two children (see Table 2). Of the 4,596 women who have multiple children (not limited to being at least 40 years old), the median spacing between the first and second child is 2 years, with almost 90% of second children being born within 5 years of the first child (see Appendix Table A.2).

In Table 3, we summarize income, consumption, and time use of our sample by completed fertility. We see that women's completed fertility is inversely related to their income, while men's completed fertility is virtually unrelated to income. These effects are consistent for both parents with partners (Panel a) and single parents (Panel b). Consumption categories related to children are, as expected, larger for adults with children, while categories not specific to specific to children are much more comparable across parents and adults with zero children.[15] Consumption on child-related goods (clothing) increase with number of children, while expenditure on services, specifically education and childcare, are relatively comparable across families of 2, 3, and 4 or more children. Women and men without children display similar patterns of time use. But among women and men with children, we see sharp differences in childcare and housework, with women spending roughly double the number of hours each week on childcare as their male coun-

[15]We adjust variables in Australian dollars for inflation using the OECD consumer price index for Australia (Organization for Economic Co-operation and Development, 2024).

terparts and 2.5 as many hours per week on housework. Hours devoted to childcare and housework increase with the number of children for both men and women, with a steeper gradient of increase for women.

In Panel (a) of Figure 1 we plot, by age, the average weekly income for men and women with different completed fertility. We see that among men, income increases consistently and peaks at age 40 before plateauing and declining through age 60. This pattern is consistent for men with no children and for men who end up with one or more children. But for women, we see that childless women experience similar patterns of income gains over the lifecycle when compared to men, while women with children experience clear declines in income during the years when they are most likely to have children. This pattern is most dramatic for women with more than one child, and the loss of income in the cross-section is large as household size grows. These results are consistent with the large literature on the child penalty, discussed above, including Kleven et al. (2019).

Moving to time use, in Panel (b) and (c) of Figure 1, we see that women with one child[16] spend significantly less time each week working and more time each week on childcare relative to women with no children, with time spent on childcare peaking at an average of 20 hours per week when women are in their late 30s. Comparing women with one child to women with more children, we see that having a second child (and a third, or fourth child) is correlated with a small increase in childcare time use relative to the much larger difference in childcare time spent by women who have one versus zero children. We see the same relative pattern for men in Panel (f), but at much lower levels, with childcare time use peaking at half the level of childcare time use for women. The differences by sex are reflected in average weekly hours spent on work.[17] Panels (c) and

[16]We use phrases such as "women with one child" for expositional convenience. When we refer to populations by the number of children they have, we are always referring to the person-level characteristic of *completed* fertility, not the person-year-level characteristic of the current number of children, unless noted otherwise.

[17]The stark difference we observed in women's childcare hours between the difference between zero versus one child and subsequent numbers of children in Panel (e) is more muted in hours of employment in Panel (c). However, as we can see in Appendix Figure A.1, the pattern emerges when we only compare women who were in the same five-year age range when they had their first child. See Appendix Figure A.3

(d) are consistent with the large literature on the child penalty by gender. We see a large difference for women, while men with children exhibit virtually no difference in work hours over the life cycle relative to men with no children. These results are correlational and represent averages of time use for a changing set of families in the sample over time who have children at different ages, and in different cohorts.

One key difference between people who go on to have different numbers of children is that, on average, they start having children at different ages. To show how employment and time use change when people become parents, Appendix Figure A.2 mirrors Figure 1 but stacks the sample in 'event time' relative to the year of first birth. We assign the median age at first birth to childless individuals in order to continue to include them as a point of comparison. In the years following the birth of a child, women see large declines in income and hours spent working, and spend on average around 45 hours per week on childcare, with men contributing around 18 hours per week to childcare. Focusing on parents who have only one child, these effects decline roughly linearly as the child ages, with parents spending virtually no time on childcare once their child reaches adulthood. While parents who have additional children spend more time on childcare in the years following their first child than parents with fewer children, the peak in the year of first birth is very similar regardless of the number of subsequent children. Furthermore, the number of additional childcare hours reported by parents with two or more children is small relative to the baseline number of hours reported by parents with one child. We see similar (mirrored) patterns for hours of employment (the middle panels), with women (but not men) experiencing large declines in hours spent working after they have their first child, and much smaller marginal differences for the second, third, and fourth children. These patterns are further reflected in other time use categories related to work and childcare such as commuting and household errands (see Appendix Figure A.4).[18]

for the analogous figure for childcare hours.

[18]See Appendix Figure A.5 for additional time use categories.

3 Empirical Strategies and Results

3.1 Event Studies of First Birth by Completed Fertility

Up to this point, we have presented descriptive patterns of income and time use for parents and other adults. Overall, we see that changes in income and childcare time following the birth of a first child are large. Families who go on to have subsequent children exhibit slightly larger and more sustained changes, suggesting that the impact of a first child is much larger than the marginal impact of subsequent children. While the patterns are striking, they rely on an unbalanced panel of households and reflect correlations. The differences in magnitudes are sufficiently strong that they seem unlikely to be driven entirely by selection, but in the previous section we did not attempt to rule out the possibility that selection into different levels of completed fertility drives these effects. We now begin to address this point by estimating event study models, exploring the evolution of key outcomes relative to the year of first birth separately by completed fertility.

We implement a two-stage regression specification following Gardner (2022) to address concerns about standard two-way fixed effect models (as described in Baker et al. (2022)).We use a first-stage regression to estimate the relationship between the outcome and covariates (person, wave, and age fixed effects) among observations in our panel from households that never have children (the 'never treated'), and households that have not yet had children (the 'not-yet-treated'):

$$Y_{it} = \alpha + \delta_t + \eta_i + \epsilon_{it} \tag{1}$$

where Y_{it} is the outcome of interest for individual i at time t; δ_t is a survey wave (year) fixed effect; η_i is an age fixed effect; and ϵ_{it} is the error term. We then run a second regression, taking out the residuals from Equation 1 for households that have a child at some point in our panel, and estimating:

$$Resid(Y)_{it} = \sum_{k=-5}^{15} \beta_k RelYear_{i,t-k} + \mu_{it} \quad (2)$$

where $Resid(Y)_{it}$ is the residual outcome variable from Equation 1; $RelYeari, t-k$ is an indicator for whether individual i had a child at time $t-k$, where k represents event time; β_k is the coefficient of interest, representing the effect of having a child at event time k on outcome Y_{it}; and μ_{it} is the error term. We run these regressions separately for men and women, and we calculate 95% confidence intervals by jointly bootstrapping both stages of that two-stage regression. We perform 1,000 iterations, and use the 26th and 975th observations to construct the confidence intervals, while clustering at the person-level. As in Kleven et al. (2019) , the effect of having a child on the outcome of interest is identified if the timing of having a child is unrelated to non-child-related changes in time use (and other outcomes). We require the person to have been observed at least once at age 35 or older.[19]

We run these models separately for men and women and (in the second stage) by completed fertility. We plot results for income, paid employment hours, and childcare hours in Figure 2. Several patterns emerge consistent with what we saw in the raw data. Women see large declines in income and hours of employment in the years following the birth of their first child, with these declines occurring immediately after the birth of the first child and then recovering slowly over the 15-year time horizon. While the initial effects of having a child on income and employment are similar for women who will have one, two, three, or four or more children, these effects diverge in the 5–10 years following the birth of the first child, which is consistent with smaller (but measurable) effects of the second, third, and fourth child on income and hours of employment. Women also spend significantly more time in the years following the birth of their first child caring for their child(ren), with childcare declining in magnitude, albeit gradually, as children

[19]This leads to pre-period observations coming from women who were present in the first several sample waves.

age. Among men, we see little response in either income or hours of paid employment. Men's hours devoted to childcare also increase after first birth, but by less than half the magnitude of women's. For men, separation by completed fertility over time is minimal.

To test for robustness, we present a version in Appendix Figure A.6 in which we include person fixed effects and require the person to have been observed in at least two waves between relative waves -2 and -5, and at least one time between relative waves 10 and 15. Results are generally similar.

Taken together, these results point to clear patterns of substitution from market to non-market labor (for women) that confirm prior work of Kleven et al. (2023) on the income and employment dimensions, establish new evidence of similar patterns for time use, and establish new facts by completed fertility. Our event studies above, like prior work, are centered around the the year when parents have their first child. However, by estimating these models separately by completed fertility, we highlight two novel patterns. First, we show that short-run responses are comparable across completed fertility groups. Second, we show that only moderate differences (relative to the magnitude of the short-run response to first birth) emerge in later years. These patterns combine to rule out the straightforward selection-based explanation of long-run difference being driven entirely by parents of more children investing less per child instead of differing marginal costs of higher order births. For selection to be the sole mechanism under constant marginal costs per child, we would see smaller immediate responses to first birth by women who go on to have larger completed fertility.

3.2 First versus Second Child Effect Estimates from Pregnancy Loss and Twins

While our event studies relative to first birth provide evidence of returns to scale in child-rearing, that approach does not provide clean distinct estimates for the marginal cost of a first child, second child, etc. We now discuss two additional identification strategies to estimate the causal effect of having a child - at a specific place in the birth order- on employment and time use for parents. Specifically, we use variation in whether a reported pregnancy resulted in a birth or not and variation in whether a pregnancy results in a single child or in twins.

We present these approaches in tandem because they allow us to compare estimates of the marginal effect of a first child and the marginal effect of a second child on parental employment, income, and time use. We use pregnancy loss for women who have no prior children as a strategy to identify the effect of a first child, pregnancy loss for women who already have one child to identify the effect of a second child, and births of twins versus singletons to also identify the effect of a second child.

Our implementation and interpretation of these strategies have important limitations. For pregnancy loss, in our data we cannot distinguish between miscarriage and abortions. As a result, the effects we estimate are averaged across women experiencing both types of pregnancy loss. For future iterations of this paper, we will continue to look to find ways to separate these populations or bound effects. Furthermore, miscarriage may also be correlated to underlying maternal health and have its own direct effects on parents. For an extended discussion of these concerns, see Biro et al. (2023). Twins as a quasi-experiment also has limitations. There has long been an understanding that the chance of twins is increased by IVF treatment. More recent evidence from Bhalotra and Clarke (2019) also documents correlations with maternal health and suggests selective miscarriage as a mechanism. In our case, while neither the pregnancy loss nor twins specifications have detectably different pretrends across groups, balance checks still reveal differences across some characteristics (Tables 4, 5, and 6). We present the results of these two strategies as part of our broader exploration of returns to scale, finding a consistent story across multiple approaches with different flaws.

In Figure 3 we present the "first stage" of our pregnancy loss and twins strategies. Panel (a) displays estimates of children total children ever had for the ten years surrounding a pregnancy for women with no prior children. Estimates from pregnancies that result in a child are plotted in yellow and estimates from pregnancies that meet are definition of loss are plotted in red. Panel (b) presents an analogous figure for women who have one child already at the time of the pregnancy. Panel (c) compares pregnancies that result in a first child to pregnancies that in a first and second child as twins. Each of these approaches result in the 'treatment' groups having an additional child in the first year after pregnancy. That gap persists over the next few years but decreases in magnitude as members of both groups continue to have children.

We next turn to our outcomes of interest. In Figure 4 and Figure 5, panels (a), (b), and (c) present analogous exercises to their counterpart panels in Figure 3, estimating income and hours devoted to childcare respectively.[20] Here we see substantial differences between lost and completed pregnancies conditional on prior children for both outcomes - this is consistent with a large response to a first child. On the other hand, we see small to no detectable differences for childcare hours and income respectively for pregnancy loss conditional on one child and twins — the two specifications estimating the impact of a second child. In panel (d), we plot the implied effect sizes from these three approaches, calculated by scaling the difference between the outcome series by the difference in the corresponding first stages.[21] As a point of comparison, we also reproduce the baseline completed fertility of one series from Figure 2.[22] We see, especially in short run, the two approaches for estimating the response to a first child yield estimates similar to each other and much larger than the estimates from the two approaches estimating the response to a second child.

[20]We present each of these tables with a 15 year (noisier) post-period in Appendix Figures A.7–A.9.

[21]We do not scale for the periods before time 0 as the differences are very small (not always exactly 0 due to measurement error), and scaling by a very small number creates nonsensical estimates.

[22]The baseline completed fertility of one series is conditional on observing the person in a wave when they are at least 35 years old. We do not make this restriction for twins and pregnancy loss.

3.3 First versus Later Child short-run Effects By Completed Fertility

Our fourth approach compares the short-run effects of different order births among individuals who have the same completed fertility. This exercise addresses an aspect of selection that — even if they were randomly assigned — pregnancy loss and twins cannot. When we compare pregnancy loss and twins estimates for people with 0 or 1 prior child, it is still possible that people who become pregnant after already having one child are different than the people who become pregnant without previous children (despite the former being a subset of the latter group). In our fourth approach, we make within completed fertility comparisons by re-estimating our initial event study specifications relative to first birth with a key modification. Event time before first birth is defined relative to first birth as before. However, after the first year after first birth, we assign all subsequent years before second birth to a single dummy to realign post-second birth event time relative to second birth. For individuals with completed fertility of three, we also do this again for third birth. This allows us to compare levels in the year following second birth to the year following first birth despite different numbers of years elapsing between births for different individuals (see Appendix Table A.2). We estimate these series separately for parents with less than three years between first and second births and three or more years, and find similar results. We plot these results in Figure 6.

For individuals who have completed fertility of two, we find levels the year after the second birth to be largely comparable to the year after the first birth. For those with complete fertility of three, the levels following the second and third birth are, if anything, lower than those after the first birth. Thus, in order for the marginal short-run impact of a second birth to be similar magnitude to that of the first, the medium-run impact of the first birth would need to be close to zero – something we consider implausible given that levels following the second (and in this population final) birth do not quickly revert to zero but instead follow the same long gradual decline we document across populations and specifications.

This fourth approach also allows us to look at heterogeneity by birth spacing. When estimating the response to a second child, our twins and pregnancy loss approaches mechanically estimate the effects for differently spaced births. Twins have a spacing of zero years between births whereas our estimates from pregnancy loss are for a second child born at least one year after the first. As expected, given that those two approaches yield similar and very small effects, we see small differences in levels on average between mothers with shorter and longer spacing between first and second birth.

3.4 Patterns in Consumption Data

Costs are multi-dimensional. We focus on labor income and hours spent on childcare as our two primary outcomes to reflect the crucial time allocation decisions parents face when allocating time across labor, leisure, and home production. These variables also have the benefit of being most responsive in the years following birth – allowing an event study approach to be informative. Our focus thus far on earning and time use does not account for the final crucial piece of the puzzle – consumption. While we expect there are also returns to scale in some, but not all, key consumption variables (e.g., being able to buy food at a discount in bulk, passing down clothes, being offered more college financial aid due to siblings), the approaches we take in this paper are less well suited to identify magnitudes. For example, a family with one child living in a two-bedroom house may want an additional bedroom for their second child; however, the timing of that demand shift is not concentrated at the time of birth in the same way that hours of childcare is. The family has much more discretion to temporally substitute housing demand and move several years before or after the birth of their second child. The family can move to a three-bedroom house in anticipation of a second child or delay that move only once that second child is a few years old. With the caveat of these additional challenges of interpretation, we present event studies relative to the year of first birth for key consumption categories to provide a complete descriptive picture.

In Figure 7, we plot the results of event studies relative to first birth by completed fertility for child-related consumption categories. Panel (a) shows that families with more children eventually live in residences with more bedrooms than those with fewer children, and housing space increases with the number of children. However, in Panel (b) we see that rent and mortgage payments do not track use of bedrooms, consistent with larger families moving to housing stock that is less expensive on a per-bedroom basis. But while (monetary) housing costs do not vary significantly by household size, we show in Panel (c) that expenditures on groceries do: households with more children spend (monotonically) more money on groceries, with the marginal cost of a child along the grocery margin decreasing in the number of children. In Panel (d) we see similar effects for spending on children's clothes, with more children leading to additional spending on clothing, but with decreasing marginal costs for households with more children. In other words, reuse and less expensive buying patterns make the first child more costly than subsequent children. Lastly, in Panel (f) we show that education costs do not follow the same pattern: households with fewer children do spend more money on education than do households with 4+ children.

Overall, while some of the consumption series follow the same pattern of differences between completed fertility levels being smaller than the effects for families with a completed fertility of one, the patterns are more varied. We are cautious in interpreting these event studies in light of the greater ease of inter-temporal substitution for some consumption categories, especially goods, and the temporal distance between greatest demand and time of birth for others, especially services.

4 Theoretical Framework

In this section, we use a simple framework to consider the implications of the patterns we document empirically. We draw on the following facts in particular. First, while parents of young children are working, day care centers and grandparents are the two most commonly reported sources of childcare (Appendix Table A.3). Second, hours of employment and hours of childcare provided by parents are much more responsive to first births than they are to subsequent births (Figure 2). Finally, expenditures on total childcare for families who rely primarily on private or community daycare are nearly double among families who have two children versus one child, while the difference by number of children is minimal for families who primarily rely on relatives for childcare (Figure 8).

Following Doepke et al. (2023), we start with a model in which childcare may be provided by a parent at the price of forgone wages or by the market at a market price. Crucially, we allow for not only the initial but also the marginal cost of childcare to vary by provider type. We also introduce a third option: childcare provided by extended family. In our model, while childcare provided by the market is priced linearly (e.g., center-based daycare is priced per child per hour), we allow there to be returns to scale in home production of childcare (e.g., a parent who leaves the workforce to care for a child from birth to school age forgoes five years of wages whereas a parent who leaves the workforce to care for two children born two years apart from birth to school age forgoes seven years of wages) and returns to scale for childcare provided by an extended family member (e.g., the family must move to be near grandparents for them to care for grandchildren, but grandparents do not require monetary compensation per child).

Formally, consider a unitary household consisting of two adults $i \in \{1, 2\}$, who can earn wages $w_1 \geq w_2$. We will refer to adult 1 as the primary earner and adult 2 as the secondary earner. The household decides how much to consume (c), how many children to have (n), and their method of providing childcare (s) if they have children. Each child requires ϕ units of childcare. Childcare for n children can be purchased from the market at linear price p; provided at home by a parent for forgone wages $w_2 h(n)$ where $h'(n) > 0$ and $h''(n) \leq 0$ for $n \geq 1$; or provided by an extended family member for $f(n)$, where $f'(n) > 0$, and $f''(n) < 0$. Utility derived from children and consumption is separable, with δ weight attached to the utility of children and ν baseline utility of being childless. The household maximization problem is as follows:

$$\max_{c,n} \ \log(c) + \delta \log(\nu + n)$$

$$\text{s.t.} \ \ c + pn\phi \leq w_1 + w_2 \text{ if using market care}$$

$$\text{s.t.} \ \ c \leq w_1 + w_2[1 - h(n)\phi] \text{ if using home care}$$

$$\text{s.t.} \ \ c + f(n)\phi \leq w_1 + w_2 \text{ if using extended family}$$

Guided by the patterns we see in the data, we consider the case where home production of childcare has a constant marginal cost after the first child. Specifically, $h(n) = 1 + \alpha(n-1)$ where $\alpha \in (0, 1)$ for $n \geq 1$. Analogously, we allow care provided by extended family to have a different cost for the first and subsequent children, $f(n) = \gamma + \beta(n-1)$. Note, these functional forms require us to consider childlessness ($n = 0$) as a separate case where $h(0) = 0$ and $f(0) = 0$.

This model allows us to compare the effects (on fertility) of different policies designed to either increase fertility or dampen the effects of having a child on the labor market decisions of parents. First, we consider a subsidy that decreases the market price of childcare from p to p'. How should we expect this policy to impact fertility and labor force participation? And how does its impact depend on the returns to scale of home provision of childcare (parameterized by α) and the availability of care from extended family (parameterized by γ and β)?

We start by establishing the baseline case of market and home care as the only options and we assume there are no returns to scale. If $\alpha = 1$, then optimal childcare strategy depends only on whether $p > w_2$. Going from $p > w_2$ to $p' < w_2$ would cause a switch from home care to market care for some families and a switch from no children to at least one child with market care for others. All families will have weakly more children. This relationship is more nuanced when $\alpha < 1$. If $p > w_2 > \alpha w_2 > p'$, we clearly have the same qualitative relationship as if $\alpha = 1$. Furthermore, the same holds for if $w_2 > p > \alpha w_2 > p'$. However, if $p > w_2 > p' > \alpha w_2$, some families who previously chose home care may switch to market care and have *fewer* children. We would predict the same dynamic if the marginal choice is between market and extended family-provided care and $p > \gamma > p' > \beta$. As empirical estimates of γ are large (Anstreicher and Venator, 2024) and intuitively β may be small, there is substantial scope for this case.

Our next steps are to calibrate this model to the empirical results described above, extending our applications to include other policy designs, which may including lump-sum payments for households with a new child, annual payments to households with children, subsidized childcare, and childcare voucher programs that vary by children's age.

5 Conclusion

In this paper, we consider the question of whether costs associated with raising children are constant per child or if there are returns to scale, and if so what do those returns to scale look like. Using a variety of empirical approaches to study how parental income and time use change with the birth of a child, we find evidence for returns to scale. In particular, the time and income costs of a first child appear to be much higher than the marginal costs of a second (or third or fourth) child. Our results point to the importance of distinguishing between the intensive and extensive margins of fertility in policy design and research.

References

Aguilar-Gomez, S., E. Arceo-Gomez, and E. De la Cruz Toledo (2022). Inside the black box of child penalties: Unpaid work and household structure. *Working papar available at SSRN 3497089*.

Anstreicher, G. and J. Venator (2024). To grandmother's house we go: Childcare time transfers and female labor mobility. Working paper.

Baker, A. C., D. F. Larcker, and C. C. Wang (2022). How much should we trust staggered difference-in-differences estimates? *Journal of Financial Economics 144*(2), 370–395.

Baudin, T., D. de la Croix, and P. E. Gobbi (2015, June). Fertility and childlessness in the united states. *American Economic Review 105*(6), 1852–82.

Becker, G. S. and H. G. Lewis (1973). On the interaction between the quantity and quality of children. *Journal of Political Economy 81*(2), S279–S288.

Bhalotra, S. and D. Clarke (2019, 12). Twin birth and maternal condition. *The Review of Economics and Statistics 101*(5), 853–864.

Biro, A., S. Dieterle, and A. Steinhauer (2023). Motherhood Timing and the Child Penalty: Bounding the Returns to Delay. CEPR Discussion Papers 13732, C.E.P.R. Discussion Papers.

Clarke, D. (2018). Children and their parents: A review of fertility and causality. *Journal of Economic Surveys 32*(2), 518–540.

Doepke, M., A. Hannusch, F. Kindermann, and M. Tertilt (2023). Chapter 4 - the economics of fertility: a new era. In S. Lundberg and A. Voena (Eds.), *Handbook of the Economics of the Family, Volume 1*, Volume 1 of *Handbook of the Economics of the Family*, pp. 151–254. North-Holland.

Gardner, J. (2022). Two-stage differences in differences. *arXiv preprint arXiv:2207.05943*.

Gobbi, P. E. (2013). A model of voluntary childlessness. *Journal of Population Economics 26*(3), 963–982.

Kleven, H., C. Landais, and G. Leite-Mariante (2023). The child penalty atlas. Technical report, National Bureau of Economic Research.

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

Koopmans, P., M. van Lent, and J. Been (2024, March). Child Penalties and the Gender Gap in Home Production and the Labor Market. IZA Discussion Papers 16871, Institute of Labor Economics (IZA).

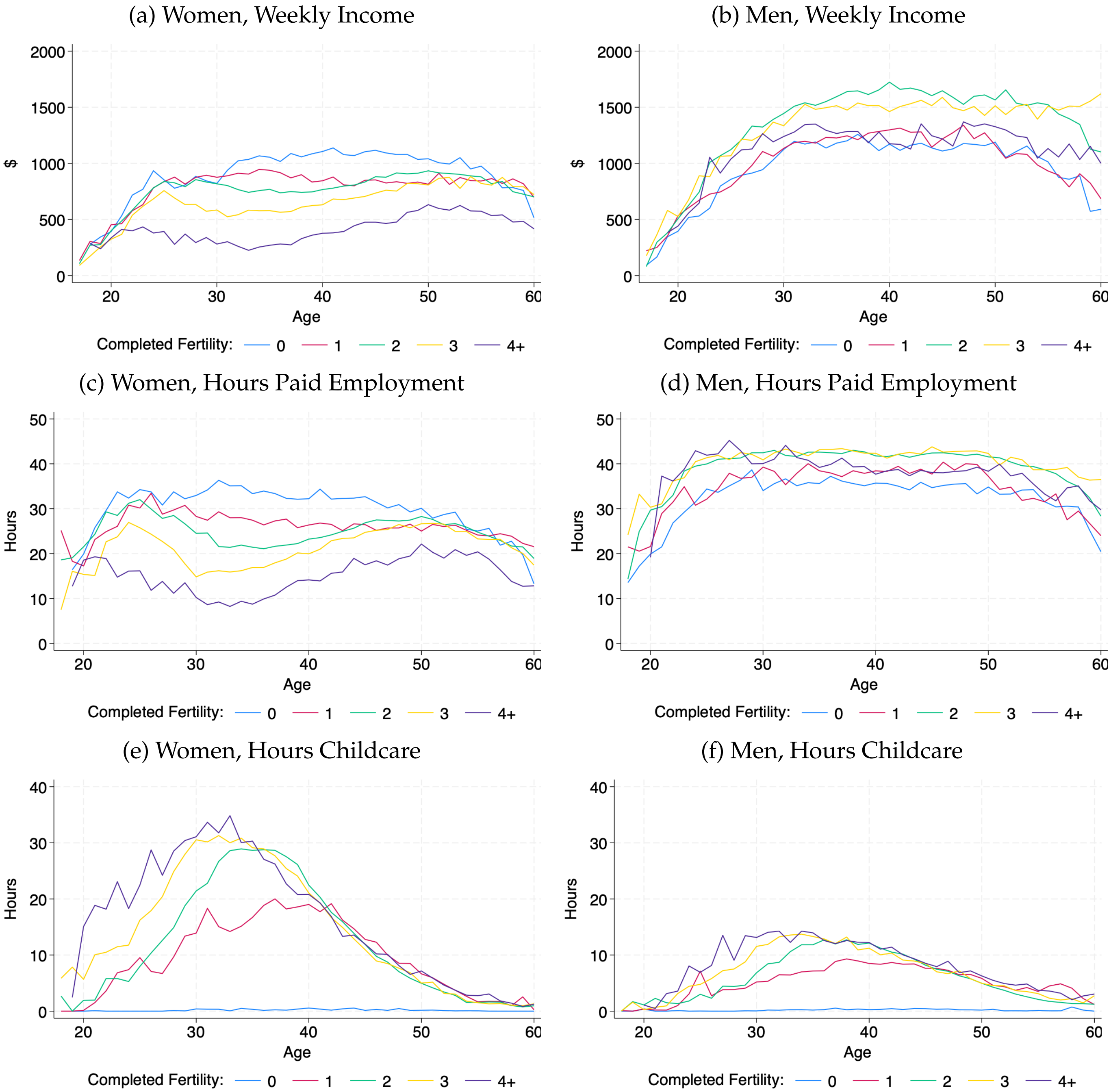
Lundborg, P., E. Plug, and A. W. Rasmussen (2017). Can women have children and a career? iv evidence from ivf treatments. *American Economic Review 107*(6), 1611–37.

Melbourne Institute of Applied Economic and Social Research (2021). The household, income and labour dynamics in australia (hilda) survey, general release 21 (waves 1-21). Melbourne Institute: Applied Economic and Social Research, University of Melbourne. https://dataverse.ada.edu.au/dataverse/hilda.

Organization for Economic Co-operation and Development (2024). Consumer price index: All items: Total for australia [auscpiallainmei]. Retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/AUSCPIALLAINMEI, February 15, 2024.

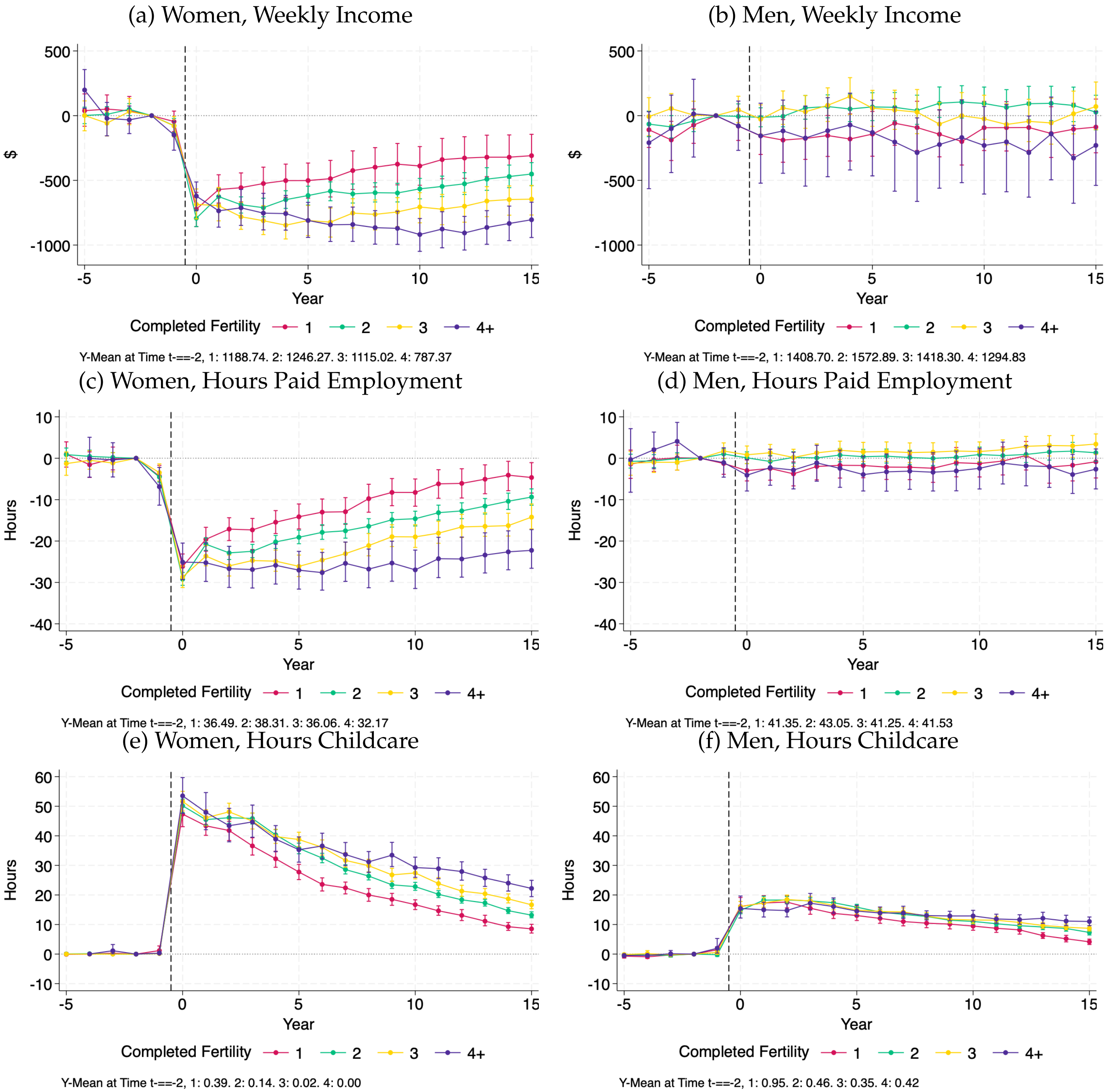
Summerfield, M., B. Garrard, M. Nesa, R. Kamath, N. Macalalad, N. Watson, R. Wilkins, and M. Wooden (2023). Hilda user manual – release 22. Melbourne Institute: Applied Economic and Social Research, University of Melbourne. https://melbourneinstitute.unimelb.edu.au/__data/assets/pdf_file/0020/4815110/HILDA-User-Manual-Release-22.0.pdf.

Figure 1: Lifecycle



Notes: This figure shows the average values of weekly income (Panels a and b); hours of paid employment (Panels c and d); and hours of childcare (Panels e and f), by age. We present results separately by predicted completed fertility of 0, 1, 2, 3, and 4 plus. Women (men) appear in Panels a, c, and e (b, d, and f). The sample is limited to people who have at least one observation in which they were at least 35 years old. Cells with fewer than 10 observation are dropped.

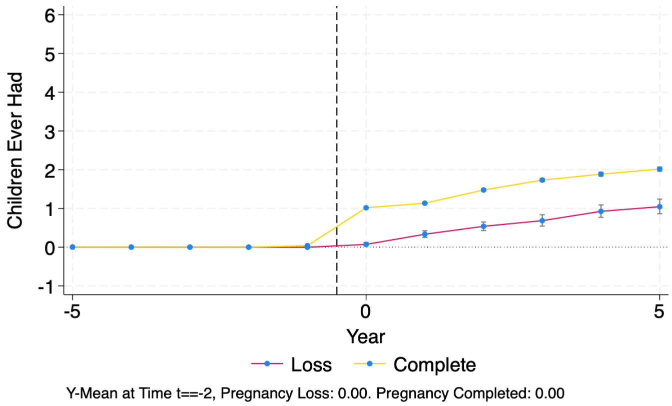
Figure 2: Event Studies



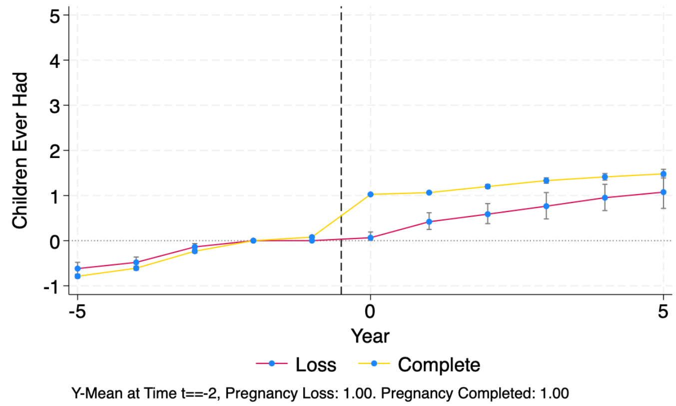
Notes: This figure presents event study estimates of the first birth for weekly income (Panels a and b); hours of paid employment (Panels c and d); and hours of childcare (Panels e and f). Female (male) estimates are presented in Panels a, c, and e (b, d, and f). Time 0 corresponds to the year the first born (computed as the age of the parent minus the age of the first-observed child). We present estimates separately by estimated completed fertility (1, 2, 3, and 4 or more). We use a two-step approach; we use the 'less conservative' event study approach: in the first step, we estimate age and year (but not person) fixed effects using never and not-yet treated observations. We do not restrict based on number of wave a person is observed. After constructing residuals by subtracting these fixed effects from the outcome variable, we regress the residuals on relative time. We construct standard errors by bootstrapping and report the 95th percentile confidence interval. We present only cells with at least 25 observations.

Figure 3: First Stage, Pregnancy Loss and Twins, Women

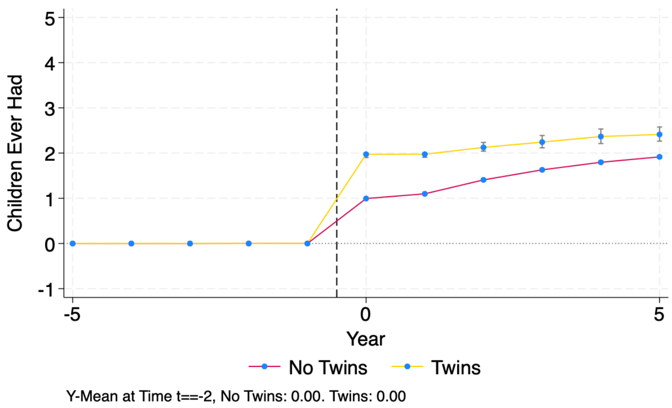
(a) Pregnancy Loss, Number of Children, Conditional on No Children



(b) Pregnancy Loss, Number of Children, Conditional on One Child

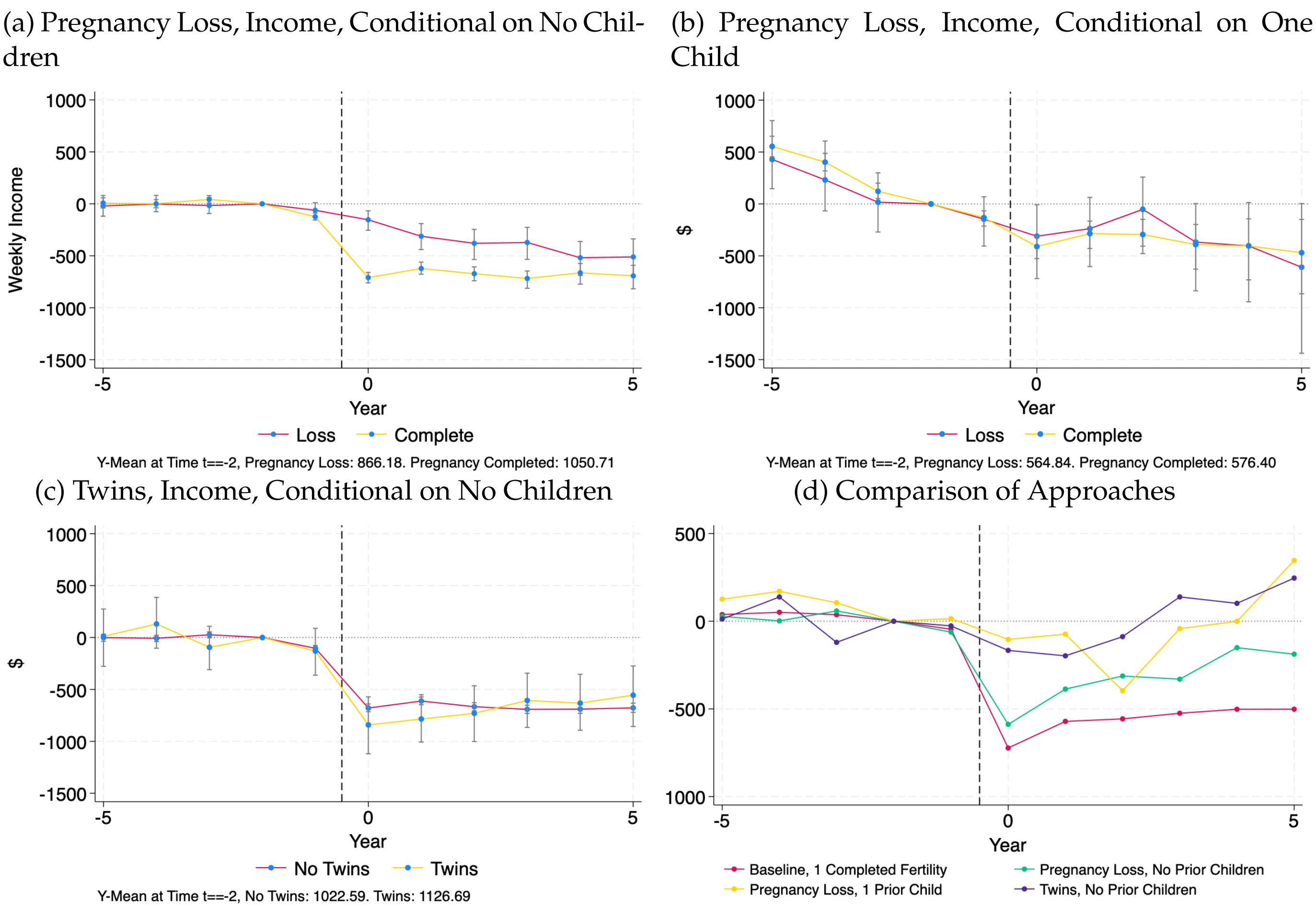


(c) Twins, Number of Children, Conditional on No Children



Notes: This figure presents event study estimates of the number of children for the first-observed pregnancy for women with 0 children (Panels a and b) and those with 1 child (Panel b). Pregnancy loss (pregnancy completion) is presented in red (yellow) for Panels a and b. Time 0 corresponds to the predicted year the next child was expected to be born (if the pregnancy were successful), which is the wave the person indicated they were pregnant if they said they were pregnant 10-12 months ago, and the subsequent wave if they said they were pregnant 0-9 months ago. We use a two-step approach here; we use the 'less conservative' event study approach: in the first step, we estimate age and year (but not person) fixed effects using observations before time 0 (or, for Panel b, observations before the expected or observed year of the first child being born); in this step, we consider only those that we assign a pregnancy completion status. After constructing residuals by subtracting these fixed effects from the outcome variable, we regress the residuals on relative time. We construct standard errors by bootstrapping and report the 95th percentile confidence interval. Panel c considers twin births. Non-twin births (twin births) are presented in red (yellow). Time 0 corresponds to the year the first born (computed as the age of the parent minus the age of the first-observed child). We use a two-step approach here; we use the 'less conservative' event study approach: in the first step, we estimate age and year (but not person) fixed effects using never and not-yet treated observations. We do not restrict based on number of wave a person is observed. After constructing residuals by subtracting these fixed effects from the outcome variable, we regress the residuals on relative time. We construct standard errors by bootstrapping and report the 95th percentile confidence interval. We present only cells with at least 25 (10) observations for pregnancy loss (twins).

Figure 4: Income, Pregnancy Loss and Twins, Women



(a) Pregnancy Loss, Income, Conditional on No Children

(b) Pregnancy Loss, Income, Conditional on One Child

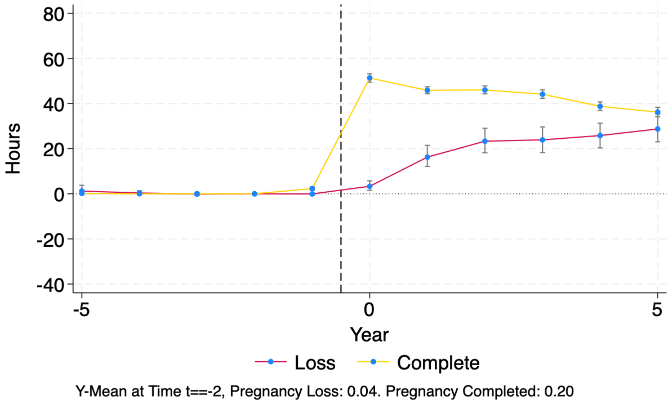
(c) Twins, Income, Conditional on No Children

(d) Comparison of Approaches

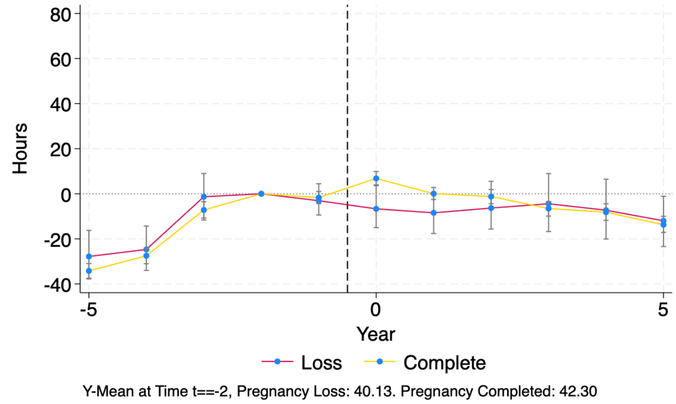
Notes: This figure presents event study estimates of income for the first-observed pregnancy for women with 0 children (Panels a and b) and those with 1 child (Panel b). Pregnancy loss (pregnancy completion) is presented in red (yellow) for Panels a and b. Time 0 corresponds to the predicted year the next child was expected to be born (if the pregnancy were successful), which is the wave the person indicated they were pregnant if they said they were pregnant 10-12 months ago, and the subsequent wave if they said they were pregnant 0-9 months ago. We use a two-step approach here; we use the 'less conservative' event study approach: in the first step, we estimate age and year (but not person) fixed effects using observations before time 0 (or, for Panel b, observations before the expected or observed year of the first child being born); in this step, we consider only those that we assign a pregnancy completion status. After constructing residuals by subtracting these fixed effects from the outcome variable, we regress the residuals on relative time. We construct standard errors by bootstrapping and report the 95th percentile confidence interval. Panel c considers twin births. Non-twin births (twin births) are presented in red (yellow). Time 0 corresponds to the year the first born (computed as the age of the parent minus the age of the first-observed child). We use a two-step approach here; we use the 'less conservative' event study approach: in the first step, we estimate age and year (but not person) fixed effects using never and not-yet treated observations. We do not restrict based on number of wave a person is observed. After constructing residuals by subtracting these fixed effects from the outcome variable, we regress the residuals on relative time. We construct standard errors by bootstrapping and report the 95th percentile confidence interval. In Panel d, we compare approaches. The 'Baseline, 1 Completed Fertility' line is the '1 Completed Fertility' line for this outcome for women in Figure 2. The 'Pregnancy Loss, No Prior Children' (and 'Pregnancy Loss, No Prior Children') line is constructed by subtracting the 'loss' estimate from the 'complete' estimate in Panels b (c) of this table and then dividing (for years 0 through 5) by the difference in children ever had between the complete and loss groups in Figure 3; the reason we do not scale the years before year 0 is that these are values are not always 0, but are very small due to measurement error, and dividing by them creates nonsensical scaled estimates. The 'Twins, No Prior Children' group is constructed similarly, scaling by the difference in Panel c of Figure 3. We present only cells with at least 25 (10) observations for pregnancy loss (twins).

Figure 5: Childcare Hours, Pregnancy Loss and Twins, Women

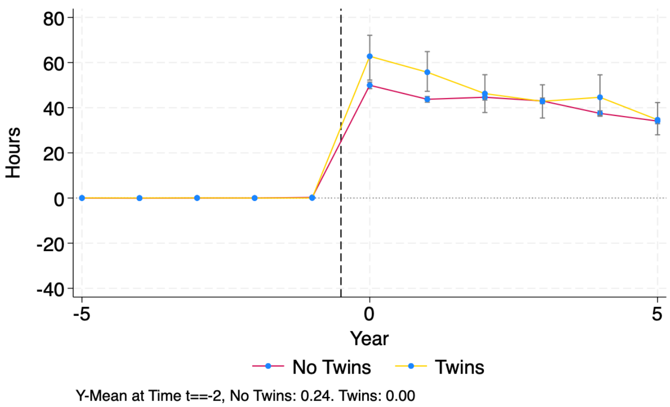
(a) Pregnancy Loss, Childcare Hours, Conditional on No Children



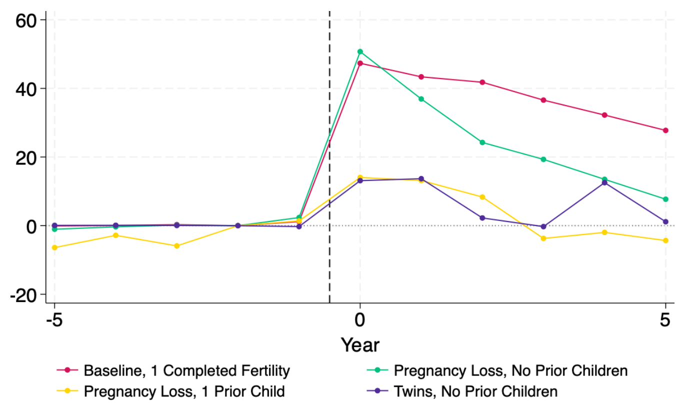
(b) Pregnancy Loss, Childcare Hours, Conditional on One Child



(c) Twins, Childcare Hours, Conditional on No Children

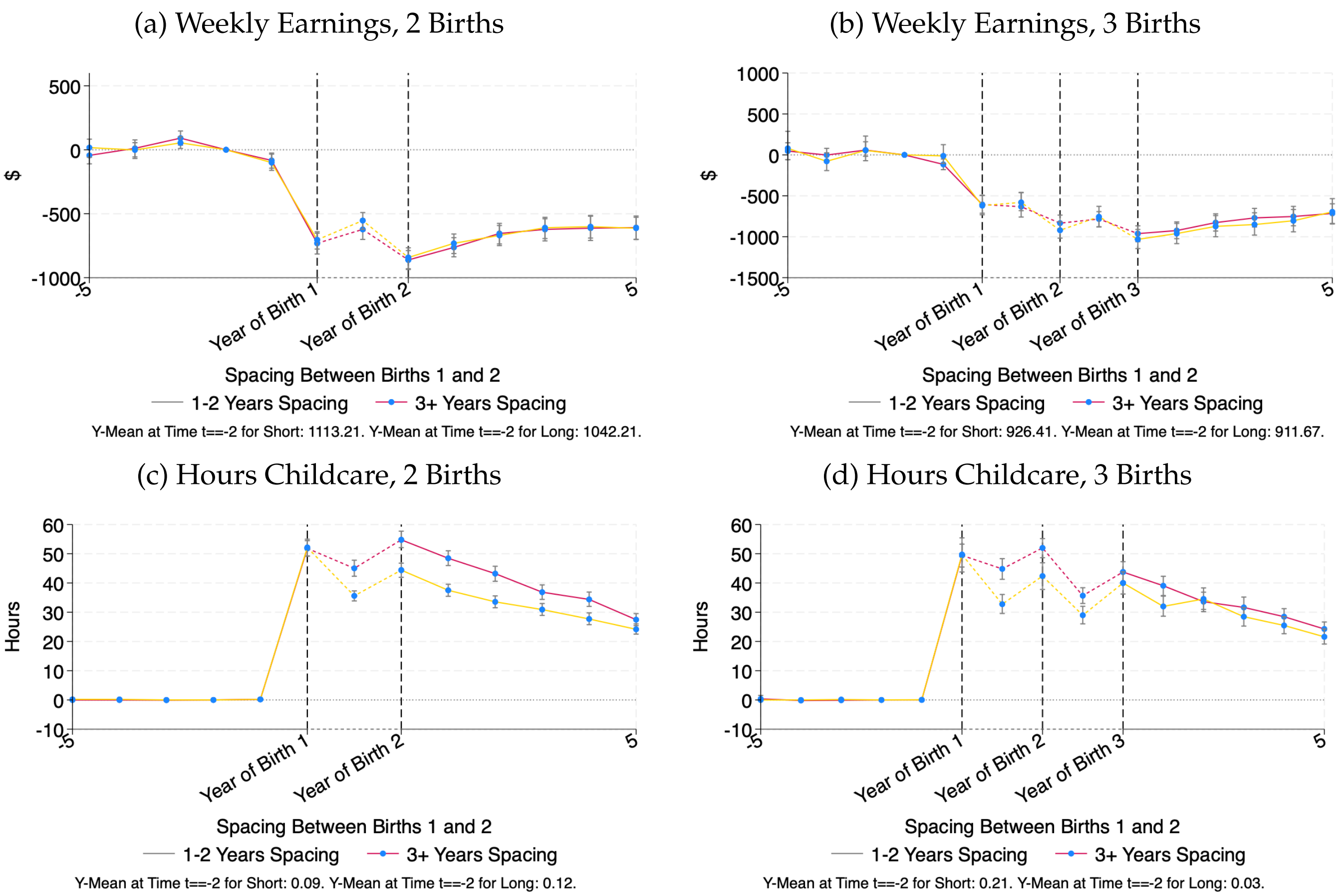


(d) Comparison of Approaches



Notes: This figure presents event study estimates of childcare hours for the first-observed pregnancy for women with 0 children (Panels a and b) and those with 1 child (Panel b). Pregnancy loss (pregnancy completion) is presented in red (yellow) for Panels a and b. Time 0 corresponds to the predicted year the next child was expected to be born (if the pregnancy were successful), which is the wave the person indicated they were pregnant if they said they were pregnant 10-12 months ago, and the subsequent wave if they said they were pregnant 0-9 months ago. We use a two-step approach here; we use the 'less conservative' event study approach: in the first step, we estimate age and year (but not person) fixed effects using observations before time 0 (or, for Panel b, observations before the expected or observed year of the first child being born); in this step, we consider only those that we assign a pregnancy completion status. After constructing residuals by subtracting these fixed effects from the outcome variable, we regress the residuals on relative time. We construct standard errors by bootstrapping and report the 95th percentile confidence interval. Panel c considers twin births. Non-twin births (twin births) are presented in red (yellow). Time 0 corresponds to the year the first born (computed as the age of the parent minus the age of the first-observed child). We use a two-step approach here; we use the 'less conservative' event study approach: in the first step, we estimate age and year (but not person) fixed effects using never and not-yet treated observations. We do not restrict based on number of wave a person is observed. After constructing residuals by subtracting these fixed effects from the outcome variable, we regress the residuals on relative time. We construct standard errors by bootstrapping and report the 95th percentile confidence interval. In Panel d, we compare approaches. The 'Baseline, 1 Completed Fertility' line is the '1 Completed Fertility' line for this outcome for women in Figure 2. The 'Pregnancy Loss, No Prior Children' (and 'Pregnancy Loss, No Prior Children') line is constructed by subtracting the 'loss' estimate from the 'complete' estimate in Panels b (c) of this table and then dividing (for years 0 through 5) by the difference in children ever had between the complete and loss groups in Figure 3; the reason we do not scale the years before year 0 is that these are values are not always 0, but are very small due to measurement error, and dividing by them creates nonsensical scaled estimates. The 'Twins, No Prior Children' group is constructed similarly, scaling by the difference in Panel c of Figure 3. We present only cells with at least 25 (10) observations for pregnancy loss (twins).

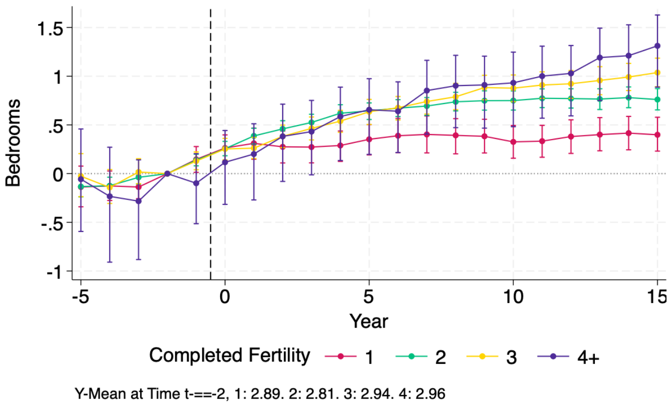
Figure 6: Event Study, Higher Order Births, Women, Birth Spacing, 2 and 3 Births



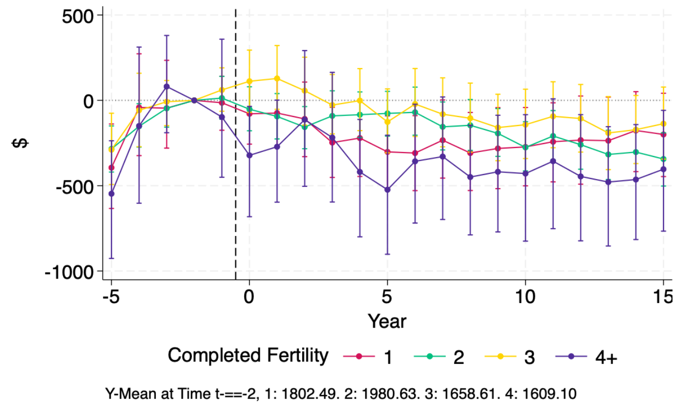
Notes: This figure presents event study estimates for women for weekly earnings (Panels a and b) and hours of childcare (Panels c and d), for those with complete fertility of 2 (Panels a and c) and 3 (Panels b and d). Short spacings (1 and 2 years) between births 1 and 2 are shown in red, while longer spacings (3+ years) are shown in yellow. Zero year spacings are not included. The first five time periods are the five years before the year of birth 1. The estimate after year of birth 1 and before year of birth 2 are all intermediate years, grouped together (and similar for the years between births 2 and 3 for Panels b and d). The five years after the last birth (birth 2 for Panels a and c, and birth 3 for Panels b and d) are also shown. We limit the sample to those for whom we observe ever having two (Panels a and c) or three (Panels b and d) children. We also drop cases where the relative age variable for a later birth is greater than the value for an earlier birth. We use a two-step approach here; we use the 'less conservative' event study approach: in the first step, we estimate age and year (but not person) fixed effects using never and not-yet treated observations (not-yet treated is defined in relation to the first birth). We do not restrict based on number of wave a person is observed. After constructing residuals by subtracting these fixed effects from the outcome variable, we regress the residuals on relative time. We construct standard errors by bootstrapping and report the 95th percentile confidence interval. We present only cells with at least 25 observations.

Figure 7: Select Consumption and Related Outcomes

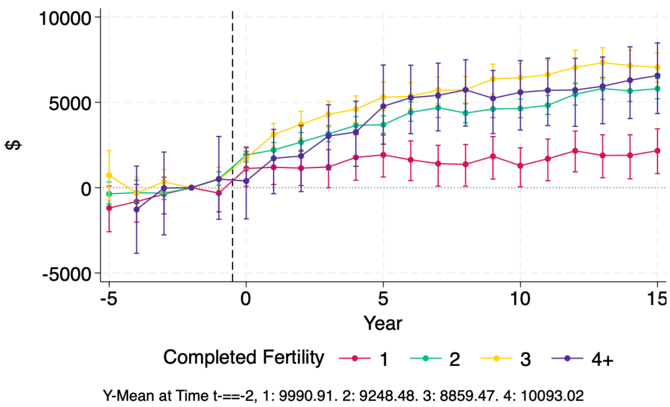
(a) Number of Bedrooms



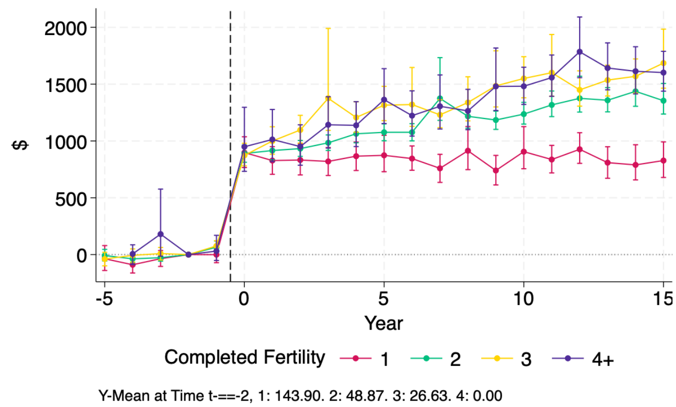
(b) Rent + Mortgage Payments



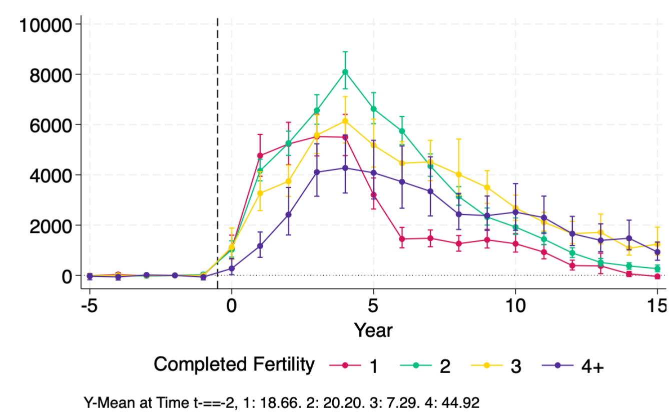
(c) Groceries



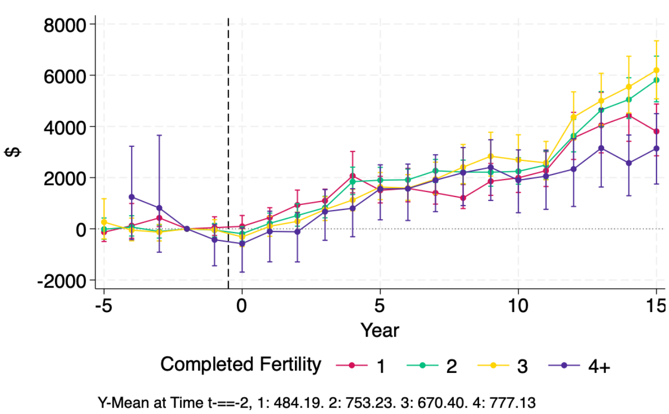
(d) Children's clothes



(e) Childcare

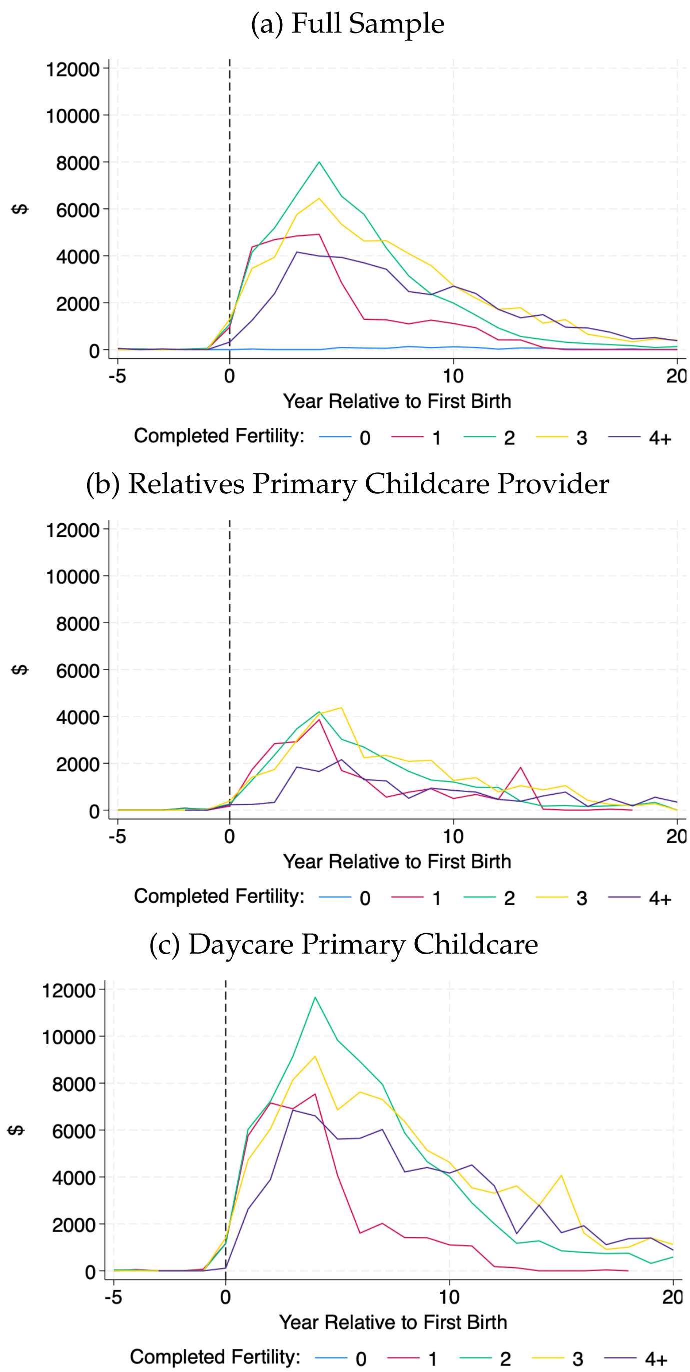


(f) Education



Notes: This figure presents event study estimates of the first birth for the number of bedrooms (Panel a), rent plus mortgage (Panel b), groceries (Panel c), children's clothing (Panel d), childcare (Panel e), and education (Panel f). Rent plus mortgage is monthly, while the remaining categories are yearly. The sample is limited to survey answers from women, but the questions ask about spending for all members of the household. Time 0 corresponds to the year the first born (computed as the age of the parent minus the age of the first-observed child). We present estimates separately by estimated completed fertility (1, 2, 3, and 4 or more). We use a two-step approach; we use the 'less conservative' event study approach: in the first step, we estimate age and year (but not person) fixed effects using never and not-yet treated observations. We do not restrict based on number of wave a person is observed. After constructing residuals by subtracting these fixed effects from the outcome variable, we regress the residuals on relative time. We construct standard errors by bootstrapping and report the 95th percentile confidence interval. We present only cells with at least 25 observations, the exceptions being that for groceries, childcare, and education, we present only cells with at least 10 observations.

Figure 8: Annual Childcare Cost, Lifecycle, Full Sample, Age Relative to Age at First Birth



Notes: This figure shows, for women, the average values by by age relative to the age at first birth of the imputed total annual childcare cost. We present results separately by predicted completed fertility of 0, 1, 2, 3, and 4 plus. The sample is limited to women, but the annual childcare cost variable is at the household level. The sample is Panel b is limited to people for whom relatives (grandparent or other relative who does or does not live with the respondent) are the primary providers of childcare for not-yet school age children; for each person-year with a positive number of hours in at least one of the relatives, daycare, and other categories, we determine the most common strategy (allowing for ties); we then classify a person's status based on the most common status of their person-years, again allowing for ties. The sample in Panel c is limited to people for whom daycare is the primary provider of childcare; this sample is selected in a similar fashion to that in Panel b. Age at first birth is calculated as one's age minus the age of the first observed child in the data; we predict age at first birth for those missing this variable. The sample is limited to people who have at least one observation in which they were at least 35 years old. Cells with fewer than ten observation are dropped.

Table 1: Summary Statistics

Variable	
Age	35.78
	(5.06)
Female	0.53
	(0.50)
Urban	0.85
	(0.35)
Born in Australia	0.76
	(0.42)
Single	0.24
	(0.42)
Bachelors or above	0.29
	(0.45)
Weekly income	953.49
	(992.37)
Predicted Completed Fertility	2.08
	(1.20)
N	9703

Notes: This table shows means and standard deviations, where each observation is a person (not a person-year). The sample is limited to individuals with at least one observation when they were 30 years or older. The statistics correspond to the first observation after age 30. All variable except age, income, and predicted completed fertility are binary. Predicted completed fertility, and is topcoded at 4. The number of observations is slightly smaller for the born in Australia, single, and bachelors variables due to missing observations.

Table 2: Fertility Distribution

	N	Pct.	Cum. Pct.
0	535	13.8	13.8
1	490	12.6	26.5
2	1426	36.8	63.3
3	862	22.2	85.5
4	347	9.0	94.5
5	121	3.1	97.6
6	52	1.3	98.9
7	23	0.6	99.5
8+	19	0.5	100.0
Total	3875	100.0	

Notes: This table shows the distribution of fertility for women. Completed fertility of more than 8 is combined with 8. After limiting the sample to age 40 and older, we define completed fertility as the maximum value of the number of the number of children ever had. This is somewhat different than predicted fertility, which incorporates expectations of future fertility, but the two are highly correlated given the age 40 restriction.

Table 3: Income, Consumption, and Time Use by Completed Fertility and Partnership Status

Panel A: In Partnership

	Women					Men				
Completed Fertility	0	1	2	3	4+	0	1	2	3	4+
Income										
Income	986.5	845.5	804.4	657.5	408.3	1216.1	1195.4	1458.3	1381.0	1198.0
Consumption										
Own Childcare	0.2	23.3	31.8	31.7	25.4	1.5	17.8	31.5	31.4	29.5
Education	14.7	30.3	47.2	54.2	41.7	10.8	29.8	45.6	49.0	42.3
Child-specific Non-Childcare (Clothing)	1.7	11.1	15.4	18.1	21.4	2.4	10.2	14.1	17.3	20.3
Non-Child-Specific Non-Childcare	1048.4	1116.4	1212.2	1216.6	1137.4	1066.2	1043.2	1209.7	1207.4	1156.2
Total of Above	1065.2	1183.3	1310.4	1324.8	1228.2	1080.6	1102.5	1304.3	1309.0	1249.3
Time Use										
Work	35.9	30.7	28.8	23.9	16.4	41.6	42.3	46.1	46.1	43.8
Household Errands, Housework, Outdoor Tasks	15.9	20.6	21.7	25.3	29.7	12.7	13.7	13.3	13.8	14.1
Own Childcare	0.2	13.7	16.1	19.3	22.7	0.3	7.0	7.7	8.6	9.9
Other Childcare, Caring for Adults, Volunteer	2.2	2.1	1.8	2.3	3.8	2.2	1.4	1.2	1.3	1.8
Sleep	52.2	50.6	50.4	49.7	47.6	50.6	48.5	49.8	48.9	48.7
Untracked (168 Minus Above)	61.8	45.8	43.9	43.1	47.0	57.4	51.0	46.5	45.6	49.3

Panel B: Single

	Women					Men				
Completed Fertility	0	1	2	3	4+	0	1	2	3	4+
Income										
Income	910.8	794.9	639.4	478.1	274.2	921.5	871.9	1058.4	1025.3	647.1
Consumption										
Own Childcare	1.9	8.8	12.2	19.0	26.8	1.0	2.2	5.8	7.6	9.9
Education	12.7	18.4	24.4	20.7	14.4	11.7	16.9	22.0	13.1	13.8
Child-specific Non-Childcare (Clothing)	1.4	9.4	12.9	18.4	22.9	2.4	5.1	7.6	8.6	10.3
Non-Child-Specific Non-Childcare	854.9	868.1	894.1	847.6	824.0	808.9	841.6	920.6	988.4	728.5
Total of Above	871.2	904.3	944.9	906.0	889.6	822.3	865.7	956.8	1017.0	765.7
Time Use										
Work	34.1	30.5	25.7	18.6	13.8	35.3	36.7	37.6	38.1	26.8
Household Errands, Housework, Outdoor Tasks	12.0	17.1	19.7	26.4	29.0	10.6	13.2	13.3	12.0	15.5
Own Childcare	0.2	8.3	11.9	17.0	18.3	0.1	2.0	4.4	5.9	9.0
Other Childcare, Caring for Adults, Volunteer	2.0	3.0	2.8	3.7	5.5	1.5	2.3	1.9	1.7	3.0
Sleep	53.1	50.6	49.2	47.9	46.6	51.0	51.2	49.9	49.3	48.8
Untracked (168 Minus Above)	65.0	56.0	57.0	57.3	58.6	68.3	62.4	57.8	61.1	60.6

Notes: This table shows the average weekly values of variables by men and women, and by predicted completed fertility, for ages 20–50. To construct the values in this table, we first compute the average of the variable for a given age. We then take the average of these averages. Income and time use are originally at the individual-level, while consumption is household consumption. All values are at or are converted to be at approximately the weekly level. Income is originally reported at the weekly level. Most of the consumption variables are reported at the yearly level; we divide these by 52.18. The remainder of the consumption variables are reported at the monthly level; we divide these by 4.35. The time use variables are reported at the weekly level. Most of the consumption variables are not observed before waves 5 or 6. Sleep is observed only in certain waves; we do not observe it for ages 20 and 21. Because untracked includes sleep, it has far fewer observations than other time use categories. We consider only people that we observe for at least one wave when they are age 35 or above. Partnership status includes legally married and de facto married. Single includes separated, divorced, widowed, and never married and not de facto. We create this variable as follows: 1) if applicable, we use the wave in which the person's first child is observed; 2) if we never observe them with a child, we use the first-observed wave when they are 30 or older; 3) if we do not observe them at age 30 or older, we use the oldest-observed wave. For the non-child-specific, non-childcare category, we group the following together: first mortgage, second mortgage, groceries, alcohol, cigarettes and tobacco, public transport and taxis, meals eaten out, motor vehicle fuel, men's clothing and footwear, women's clothing and footwear, telephone rent and calls, internet charges, private health insurance, other insurances, fees paid to health practitioner, medicines, prescriptions, and pharmaceuticals, electricity, gas bills, and other heating fuel, repairs, renovation, and maintenance to home, and motor vehicle repairs and maintenance.

Table 4: Pregnancy Loss Balance Table, Conditional on No Children

Panel A: Variables measured at time -2	Women		Men	
	Loss	Complete	Loss	Complete
Age	25.97	27.20	28.34	29.23
	(7.21)	(5.31)	(6.86)	(5.57)
Urban	0.87	0.87	0.90	0.89
	(0.34)	(0.33)	(0.31)	(0.32)
Born in Australia	0.90	0.86	0.86	0.83
	(0.31)	(0.34)	(0.35)	(0.38)
Single	0.50	0.17	0.40	0.16
	(0.50)	(0.38)	(0.49)	(0.37)
Bachelors or above	0.34	0.43	0.27	0.35
	(0.48)	(0.50)	(0.45)	(0.48)
Weekly income	866.18	1050.71	1055.03	1380.63
	(680.17)	(699.88)	(834.72)	(940.76)
N	163	914	107	829

Panel B: Variables measures at (or around) age 20	Women		Men	
	Loss	Complete	Loss	Complete
Age	23.74	23.88	24.81	25.44
	(5.41)	(4.55)	(6.42)	(5.32)
Urban	0.94	0.91	0.93	0.89
	(0.24)	(0.29)	(0.26)	(0.32)
Born in Australia	0.90	0.86	0.86	0.83
	(0.31)	(0.34)	(0.35)	(0.38)
Single	0.59	0.41	0.54	0.42
	(0.49)	(0.49)	(0.50)	(0.49)
Bachelors or above	0.24	0.31	0.17	0.25
	(0.43)	(0.46)	(0.38)	(0.44)
Weekly income	716.01	764.24	772.07	1041.86
	(549.29)	(600.84)	(643.03)	(805.55)
N	163	914	107	829

Notes: This table shows means and standard deviations for samples of pregnancy loss at first birth and completed pregnancy at first (observed) pregnancy, separately for men and women. Panel A is limited to observations at relative (to first observed pregnancy) time -2, and each variable is measured during this wave. Panel B uses the same sample of individuals in Panel A, but variables are instead measured at the first observed wave at or after age 20 (or, for individuals never observed at or after age 20, the latest wave before age 20).

Table 5: Pregnancy Loss Balance Table, Conditional on One Child

Panel A: Variables measured at time -2				
	Women		Men	
	Loss	Complete	Loss	Complete
Age	31.73	29.72	33.09	31.73
	(6.71)	(5.09)	(6.50)	(5.23)
Urban	0.93	0.86	0.93	0.85
	(0.25)	(0.35)	(0.26)	(0.35)
Born in Australia	0.74	0.82	0.74	0.82
	(0.44)	(0.38)	(0.44)	(0.38)
Single	0.24	0.09	0.14	0.05
	(0.43)	(0.28)	(0.35)	(0.21)
Bachelors or above	0.41	0.42	0.26	0.36
	(0.49)	(0.49)	(0.44)	(0.48)
Weekly income	580.32	594.58	1293.43	1493.36
	(610.96)	(744.33)	(995.16)	(1029.56)
N	74	560	43	470

Panel B: Variables measures at (or around) age 20				
	Women		Men	
	Loss	Complete	Loss	Complete
Age	26.96	24.92	28.58	26.58
	(6.61)	(4.92)	(7.08)	(5.41)
Urban	0.96	0.90	0.91	0.90
	(0.20)	(0.30)	(0.29)	(0.30)
Born in Australia	0.74	0.82	0.74	0.82
	(0.44)	(0.38)	(0.44)	(0.38)
Single	0.36	0.33	0.21	0.31
	(0.48)	(0.47)	(0.41)	(0.46)
Bachelors or above	0.30	0.33	0.23	0.29
	(0.46)	(0.47)	(0.43)	(0.46)
Weekly income	533.39	675.51	1078.39	1098.43
	(579.74)	(612.41)	(968.00)	(819.31)
N	74	560	43	470

Notes: This table shows means and standard deviations for samples of pregnancy loss at first birth and completed pregnancy at next (observed) pregnancy for those with one child, separately for men and women. Panel A is limited to observations at relative (to next observed pregnancy) time -2, and each variable is measured during this wave. Panel B uses the same sample of individuals in Panel A, but variables are instead measured at the first observed wave at or after age 20 (or, for individuals never observed at or after age 20, the latest wave before age 20).

Table 6: Twins Balance Table

Panel A: Variables measured at time -2	Women		Men	
	No Twins	Twins	No Twins	Twins
Age	26.85	30.59	28.64	32.24
	(5.63)	(5.49)	(5.93)	(6.15)
Urban	0.86	0.76	0.87	0.90
	(0.34)	(0.44)	(0.33)	(0.31)
Born in Australia	0.86	0.90	0.84	0.83
	(0.34)	(0.31)	(0.37)	(0.38)
Single	0.21	0.07	0.21	0.07
	(0.41)	(0.26)	(0.41)	(0.26)
Bachelors or above	0.40	0.59	0.31	0.48
	(0.49)	(0.50)	(0.46)	(0.51)
Weekly income	1022.59	1126.69	1311.38	1412.71
	(771.68)	(723.36)	(984.13)	(615.64)
N	1805	29	1622	29

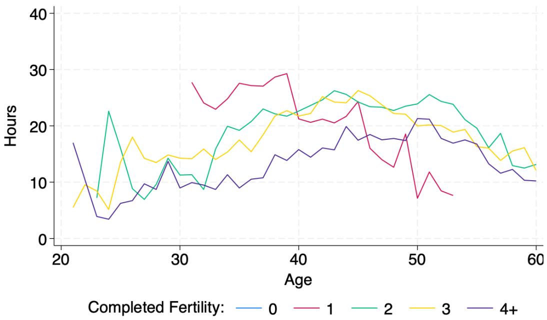
Panel B: Variables measures at (or around) age 20	Women		Men	
	No Twins	Twins	No Twins	Twins
Age	23.57	25.90	24.77	27.21
	(4.55)	(5.61)	(5.27)	(5.94)
Urban	0.89	0.83	0.89	0.90
	(0.31)	(0.38)	(0.31)	(0.31)
Born in Australia	0.86	0.90	0.84	0.83
	(0.34)	(0.31)	(0.37)	(0.38)
Single	0.43	0.31	0.45	0.41
	(0.49)	(0.47)	(0.50)	(0.50)
Bachelors or above	0.28	0.38	0.22	0.38
	(0.45)	(0.49)	(0.42)	(0.49)
Weekly income	732.49	762.58	962.86	1037.89
	(605.44)	(535.78)	(889.42)	(793.72)
N	1805	29	1622	29

Notes: This table shows means and standard deviations for samples of no twins at first birth and twins at first birth, separately for men and women. Panel A is limited to observations at relative (to first child) time -2, and each variable is measured during this wave. Panel B uses the same sample of individuals in Panel A, but variables are instead measured at the first observed wave at or after age 20 (or, for individuals never observed at or after age 20, the latest wave before age 20). For the no twins samples, some variables have a smaller number of observations than indicated due to missing data.

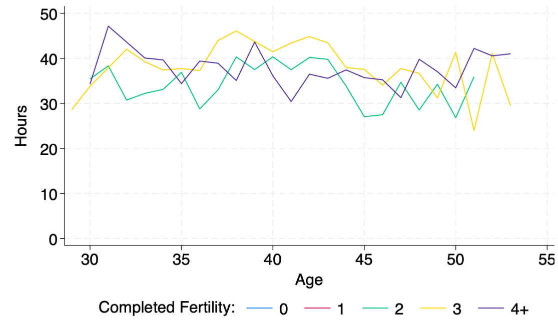
A Appendix Figures and Tables

Figure A.1: Hours Employment, Lifecycle, by Age at First Birth

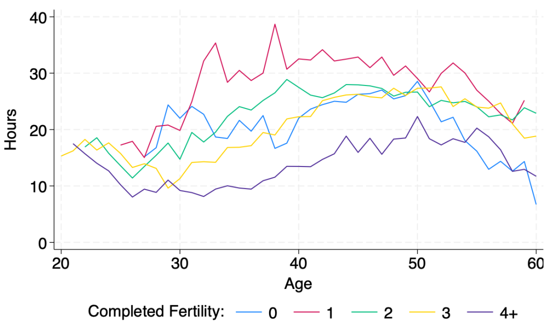
(a) Women, Age at First Birth 16–20



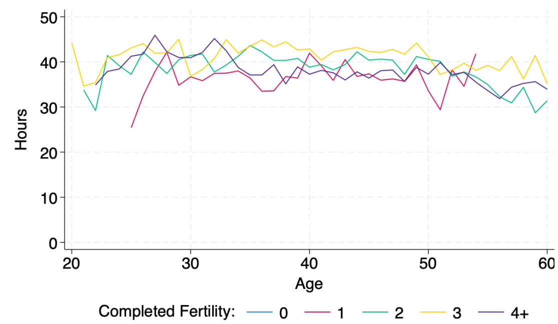
(b) Men, Age at First Birth 16–20



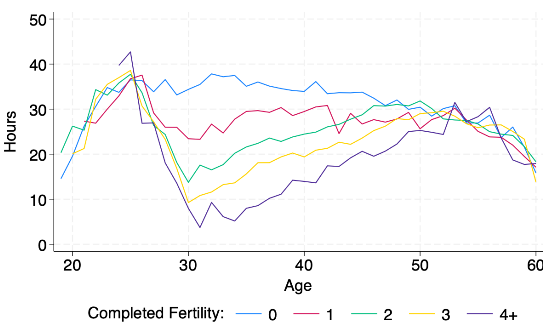
(c) Women, Age at First Birth 21–25



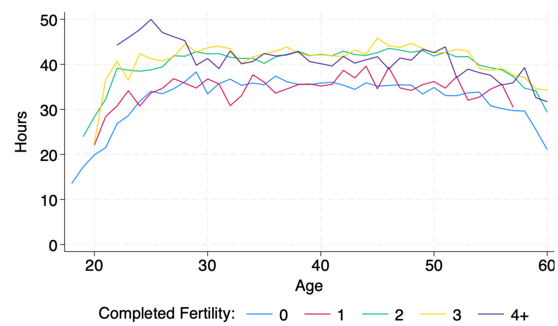
(d) Men, Age at First Birth 21–25



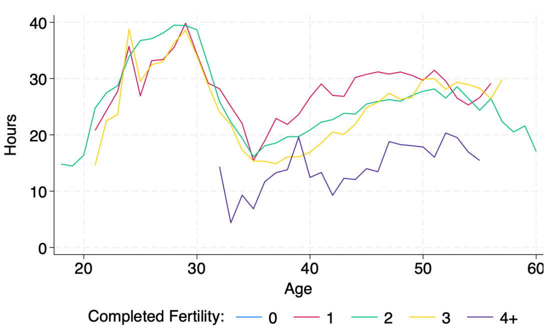
(e) Women, Age at First Birth 26–30



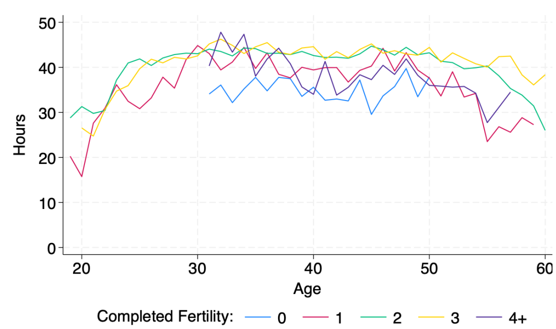
(f) Men, Age at First Birth 26–30



(g) Women, Age at First Birth 31–35

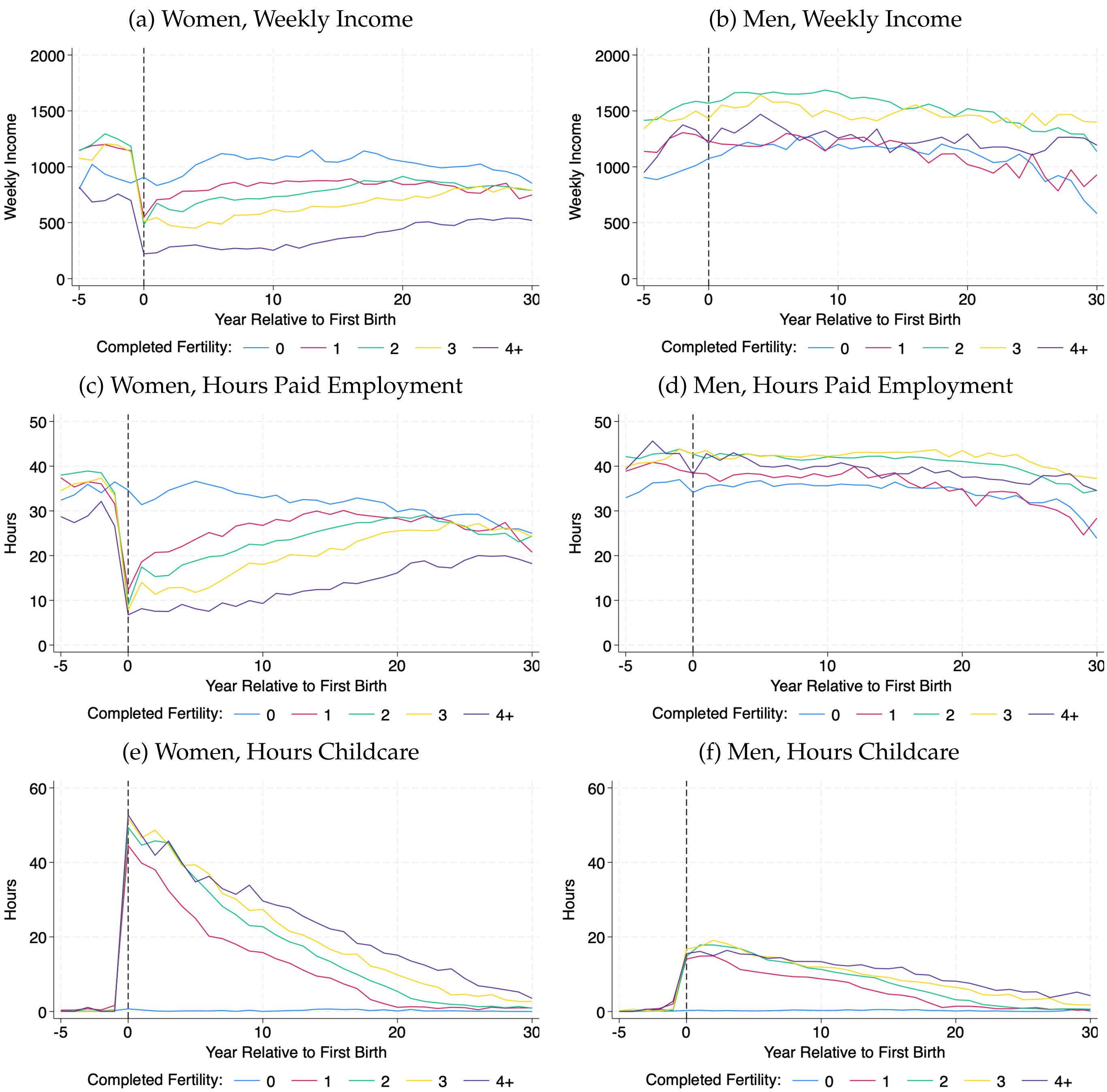


(h) Men, Age at First Birth 31–35



Notes: This figure shows the average values of hours of employment by age. We present results separately by predicted completed fertility of 0, 1, 2, 3, and 4 plus. Women (men) appear in Panels a, c, e, and g (b, d, f, and h). Panels a and b are limited to individuals who were 16–20 years old at the birth of their first child. Panels c and d are for ages 21–25; Panels e and f are for ages 26–30; and Panels g and h are for ages 31–35. Age at first birth is calculated as one's age minus the age of the first observed child in the data. If we do not observe this variable, we predict it. The sample is limited to people who have at least one observation in which they were at least 35 years old. Cells with fewer than 10 observation are dropped.

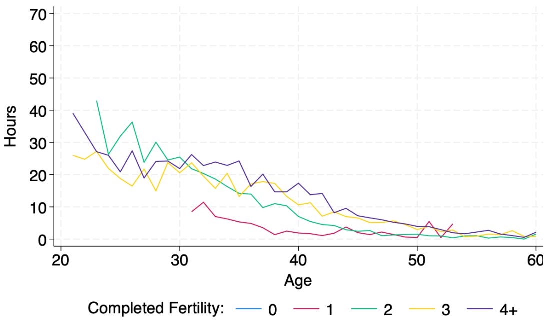
Figure A.2: Lifecycle, Relative to Year of First Birth



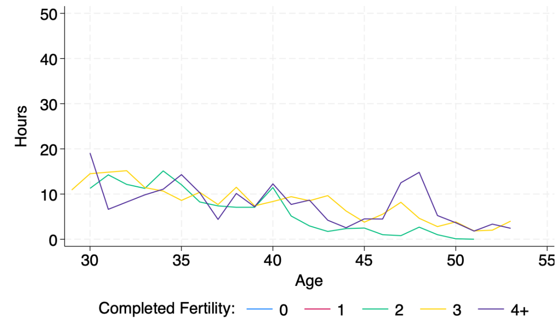
Notes: This figure shows the average values of weekly income (Panels a and b); hours of paid employment (Panels c and d); and hours of childcare (Panels e and f), by age relative to the age at first birth. We present results separately by predicted completed fertility of 0, 1, 2, 3, and 4 plus. Women (men) appear in Panels a, c, and e (b, d, and f). Age at first birth is calculated as one's age minus the age of the first observed child in the data; we predict age at first birth for those missing this variable. The sample is limited to people who have at least one observation in which they were at least 35 years old. Cells with fewer than 10 observation are dropped.

Figure A.3: Childcare, Lifecycle, by Age at First Birth

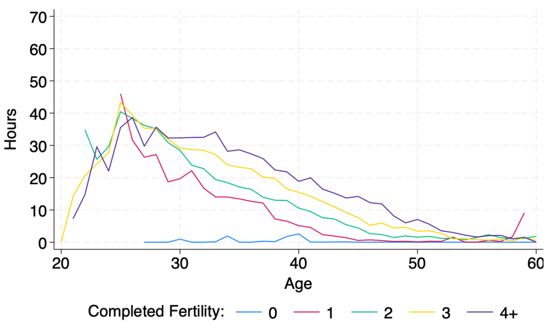
(a) Women, Age at First Birth 16–20



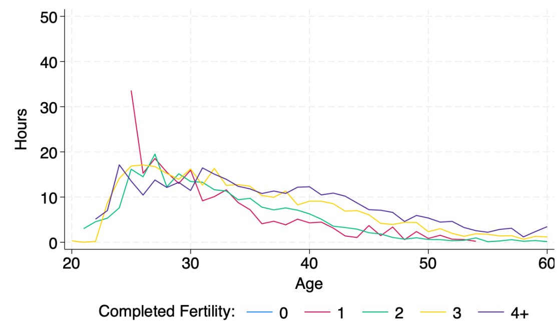
(b) Men, Age at First Birth 16–20



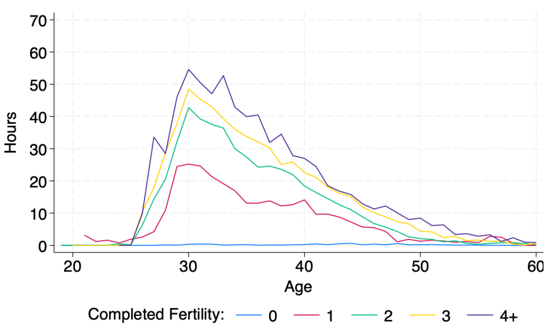
(c) Women, Age at First Birth 21–25



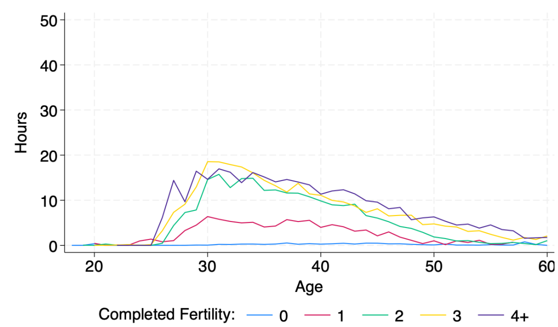
(d) Men, Age at First Birth 21–25



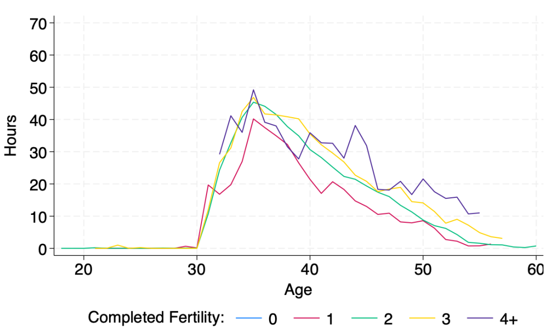
(e) Women, Age at First Birth 26–30



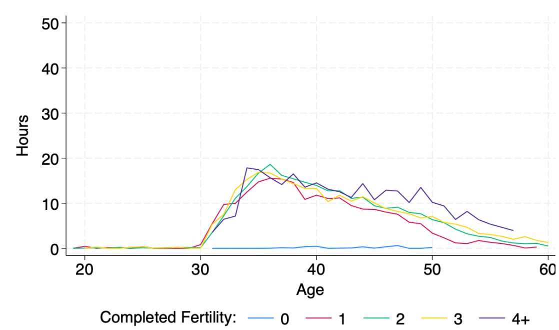
(f) Men, Age at First Birth 26–30



(g) Women, Age at First Birth 31–35

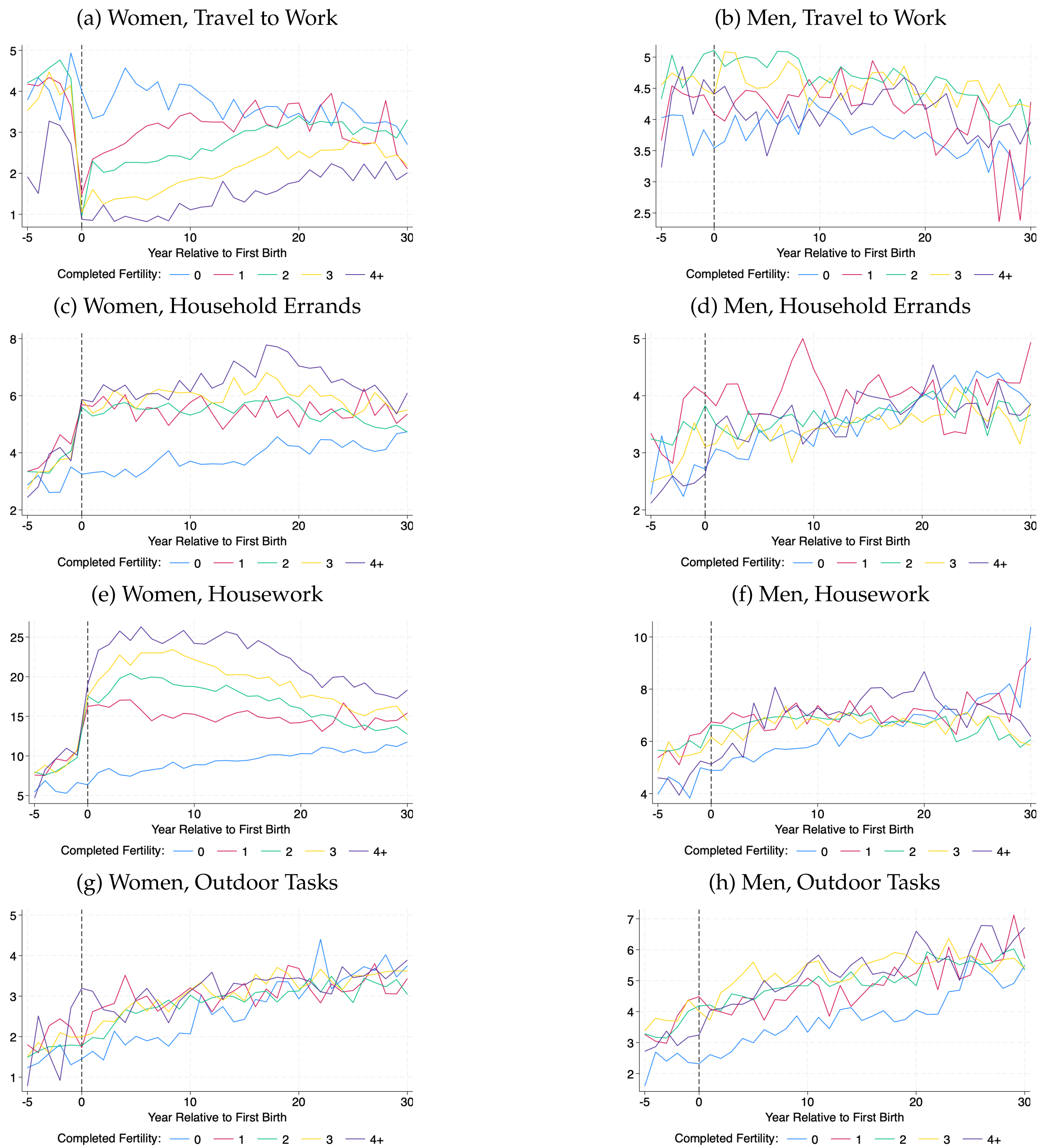


(h) Men, Age at First Birth 31–35



Notes: This figure shows the average values of hours of childcare by age, separately by age at first birth. We present results separately by predicted completed fertility of 0, 1, 2, 3, and 4 plus. Women (men) appear in Panels a, c, e, and g (b, d, f, and h). Panels a and b are limited to individuals who were 16–20 years old at the birth of their first child. Panels c and d are for ages 21–25; Panels e and f are for ages 26–30; and Panels g and h are for ages 31–35. Age at first birth is calculated as one's age minus the age of the oldest child in the data. If we do not observe this variable, we predict it. The sample is limited to people who have at least one observation in which they were at least 35 years old. Cells with fewer than 10 observation are dropped.

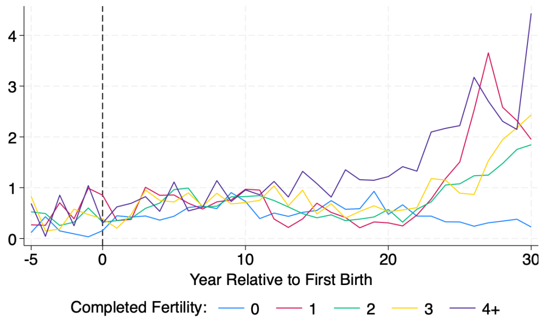
Figure A.4: Other Time Use Outcomes 1, Lifecycle, Full Sample, Age Relative to Age at First Birth



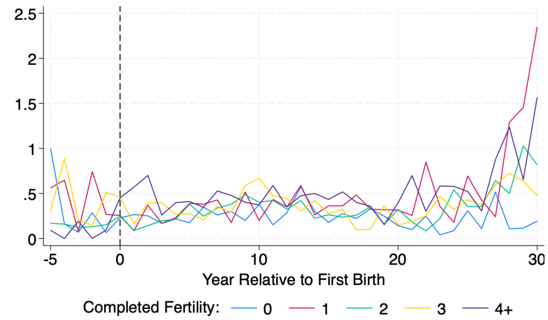
Notes: This figure shows the average values of hours of traveling to work (Panels a and b), household errands (Panels c and d), housework (Panels e and f), and outdoor tasks (Panels g and h) by age relative to the age at first birth. We present results separately by predicted completed fertility of 0, 1, 2, 3, and 4 plus. Women (men) appear in Panels a, c, e, and g (b, d, f, and h). Age at first birth is calculated as one's age minus the age of the first observed child in the data; we predict age at first birth for those missing this variable. The sample is limited to people who have at least one observation in which they were at least 35 years old. Cells with fewer than 10 observation are dropped.

Figure A.5: Other Time Use Outcomes 2, Lifecycle, Full Sample, Age Relative to Age at First Birth

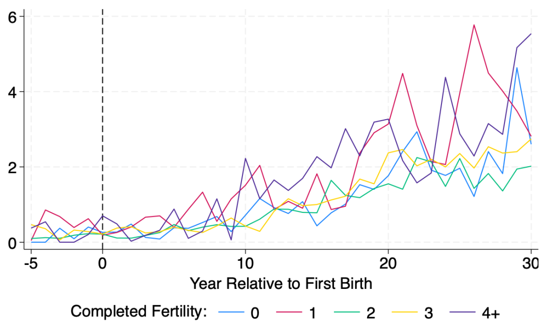
(a) Women, Caring for Others' Children



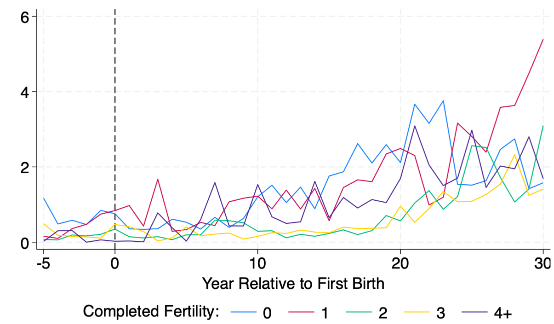
(b) Men, Caring for Others' Children



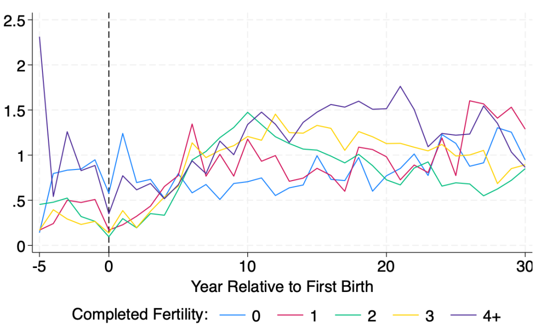
(c) Women, Caring for Adults



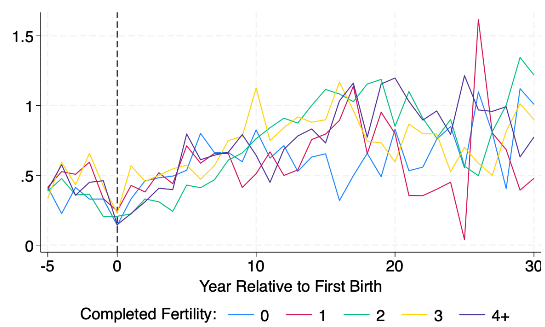
(d) Men, Caring for Adults



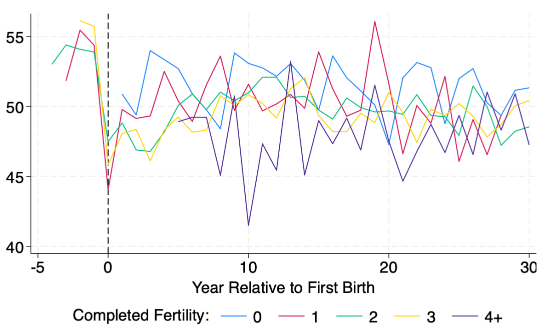
(e) Women, Volunteering



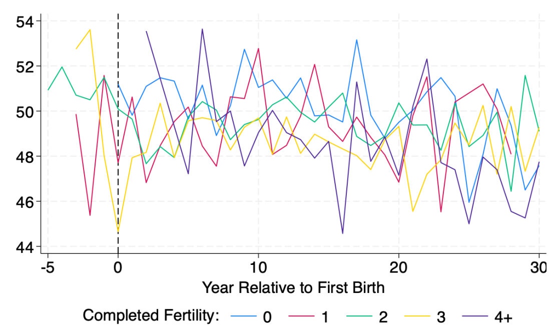
(f) Men, Volunteering



(g) Women, Sleeping

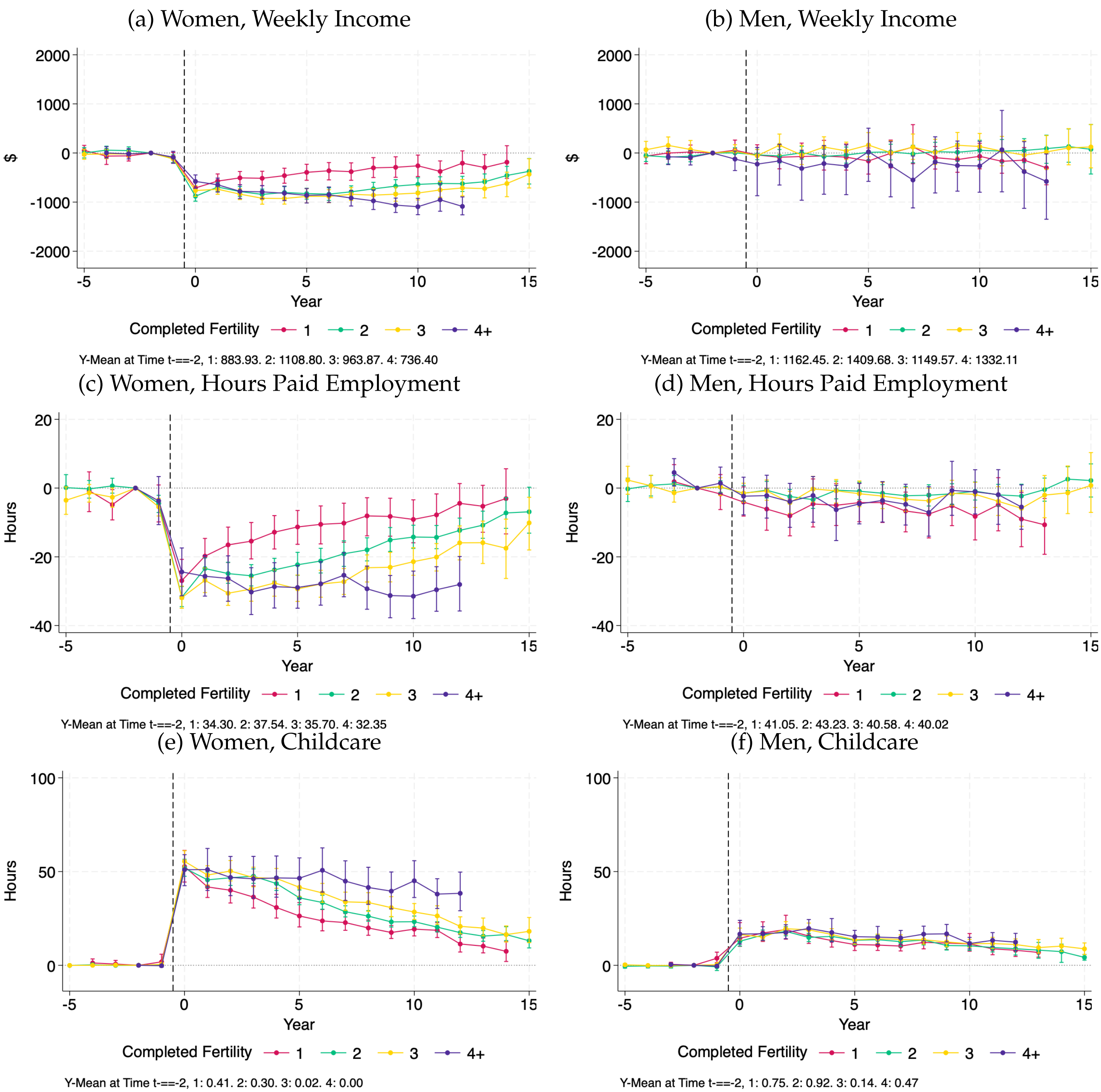


(h) Men, Sleeping



Notes: This figure shows the average values of hours of caring for others' children (Panels a and b), caring for adults (Panels c and d), volunteering (Panels e and f), and sleeping (Panels g and h) by age relative to the age at first birth. Sleeping is measured much less frequently than the other variables. We present results separately by predicted completed fertility of 0, 1, 2, 3, and 4 plus. Women (men) appear in Panels a, c, e, and g (b, d, f, and h). Age at first birth is calculated as one's age minus the age of the first observed child in the data; we predict age at first birth for those missing this variable. The sample is limited to people who have at least one observation in which they were at least 35 years old. Cells with fewer than 10 observation are dropped.

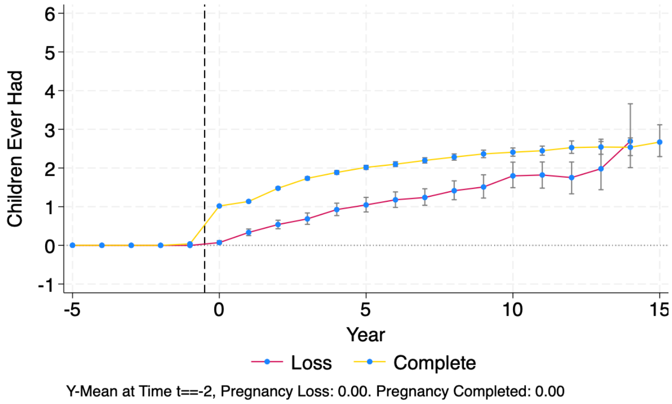
Figure A.6: Event Studies, Conservative



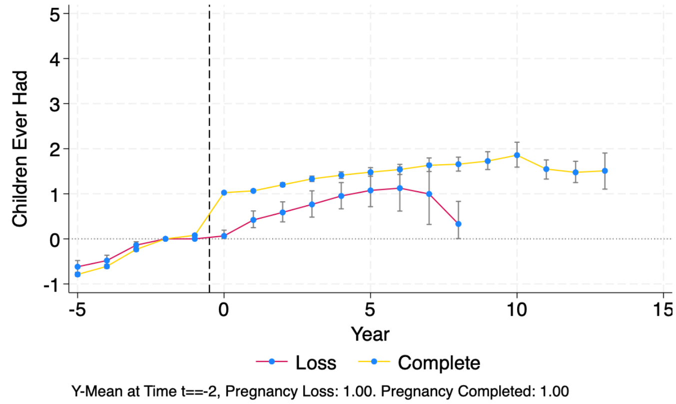
Notes: This figure presents event study estimates of the first birth for weekly income (Panels a and b); hours of paid employment (Panels c and d); and hours of childcare (Panels e and f). Female (male) estimates are presented in Panels a, c, and e (b, d, and f). Time 0 corresponds to the year the first born (computed as the age of the parent minus the age of the first-observed child). We present estimates separately by estimated completed fertility (1, 2, 3, and 4 or more). We use a two-step approach; we use the 'conservative' event study approach: In the first step, we estimate person, age, and year fixed effects using never and not-yet treated observations. After constructing residuals by subtracting these fixed effects from the outcome variable, we regress the residuals on relative time. The (treated) sample is limited to those for whom we observe at least two observations between time -5 and -1, and at least one observation between time 10 and 15. We construct standard errors by bootstrapping and report the 95th percentile confidence interval. We present only cells with at least 25 observations.

Figure A.7: First Stage, Pregnancy Loss and Twins, Women, Longer Post Period

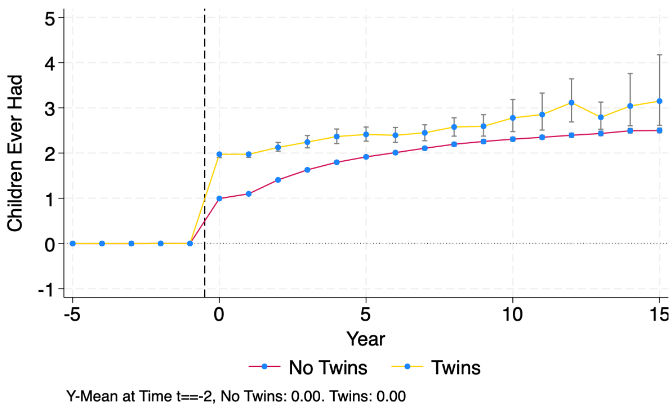
(a) Pregnancy Loss, Number of Children, Conditional on No Children



(b) Pregnancy Loss, Number of Children, Conditional on One Child



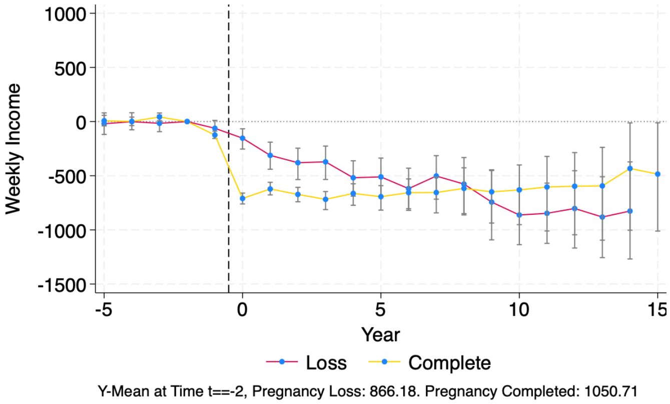
(c) Twins, Number of Children, Conditional on No Children



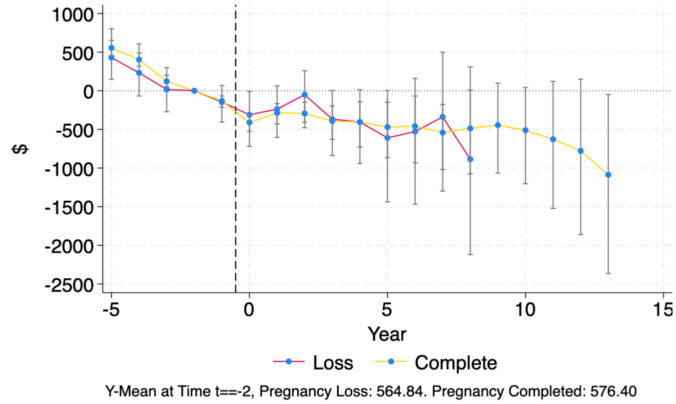
Notes: This figure presents event study estimates of the number of children for the first-observed pregnancy for women with 0 children (Panels a and b) and those with 1 child (Panel b). Pregnancy loss (pregnancy completion) is presented in red (yellow) for Panels a and b. Time 0 corresponds to the predicted year the next child was expected to be born (if the pregnancy were successful), which is the wave the person indicated they were pregnant if they said they were pregnant 10-12 months ago, and the subsequent wave if they said they were pregnant 0-9 months ago. We use a two-step approach here; we use the 'less conservative' event study approach: in the first step, we estimate age and year (but not person) fixed effects using observations before time 0 (or, for Panel b, observations before the expected or observed year of the first child being born); in this step, we consider only those that we assign a pregnancy completion status. After constructing residuals by subtracting these fixed effects from the outcome variable, we regress the residuals on relative time. We construct standard errors by bootstrapping and report the 95th percentile confidence interval. Panel c considers twin births. Non-twin births (twin births) are presented in red (yellow). Time 0 corresponds to the year the first born (computed as the age of the parent minus the age of the first-observed child). We use a two-step approach here; we use the 'less conservative' event study approach: in the first step, we estimate age and year (but not person) fixed effects using never and not-yet treated observations. We do not restrict based on number of wave a person is observed. After constructing residuals by subtracting these fixed effects from the outcome variable, we regress the residuals on relative time. We construct standard errors by bootstrapping and report the 95th percentile confidence interval. We present only cells with at least 25 (10) observations for pregnancy loss (twins).

Figure A.8: Income, Pregnancy Loss and Twins, Women, Longer Post Period

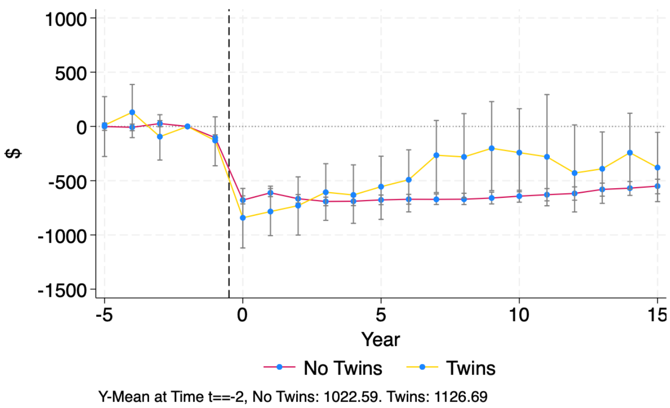
(a) Pregnancy Loss, Income, Conditional on No Children



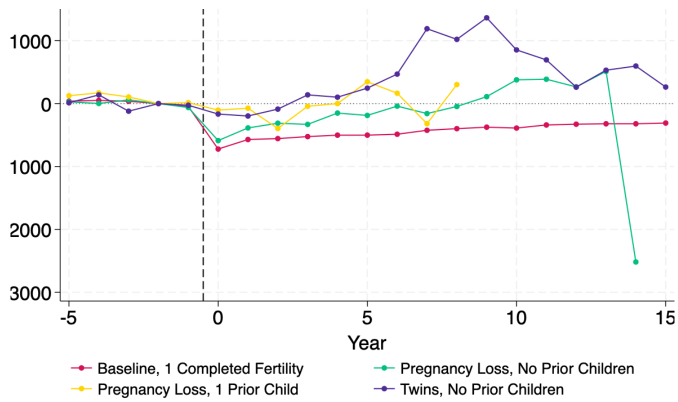
(b) Pregnancy Loss, Income, Conditional on One Child



(c) Twins, Income, Conditional on No Children



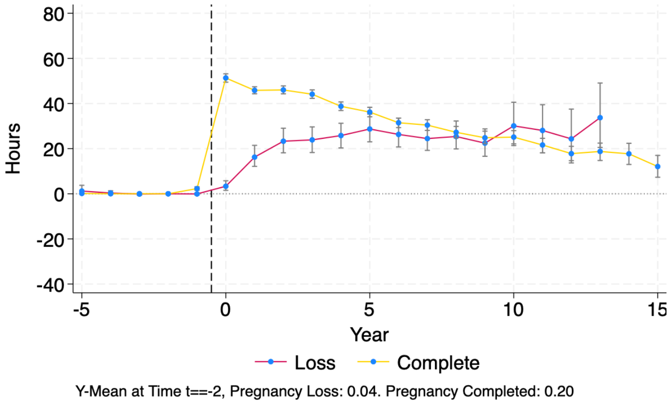
(d) Comparison of Approaches



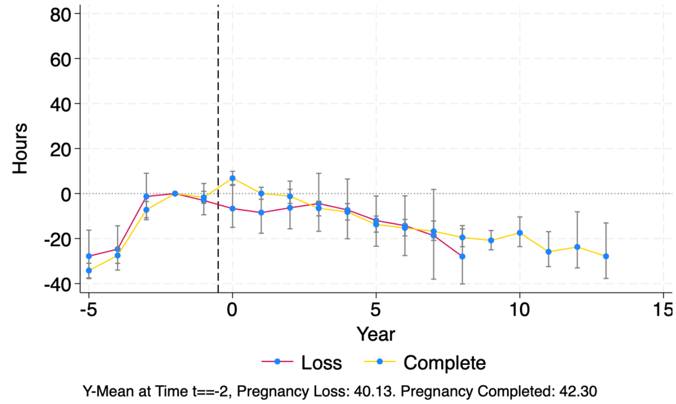
Notes: This figure presents event study estimates of income for the first-observed pregnancy for women with 0 children (Panels a and b) and those with 1 child (Panel b). Pregnancy loss (pregnancy completion) is presented in red (yellow) for Panels a and b. Time 0 corresponds to the predicted year the next child was expected to be born (if the pregnancy were successful), which is the wave the person indicated they were pregnant if they said they were pregnant 10-12 months ago, and the subsequent wave if they said they were pregnant 0-9 months ago. We use a two-step approach here; we use the 'less conservative' event study approach: in the first step, we estimate age and year (but not person) fixed effects using observations before time 0 (or, for Panel b, observations before the expected or observed year of the first child being born); in this step, we consider only those that we assign a pregnancy completion status. After constructing residuals by subtracting these fixed effects from the outcome variable, we regress the residuals on relative time. We construct standard errors by bootstrapping and report the 95th percentile confidence interval. Panel c considers twin births. Non-twin births (twin births) are presented in red (yellow). Time 0 corresponds to the year the first born (computed as the age of the parent minus the age of the first-observed child). We use a two-step approach here; we use the 'less conservative' event study approach: in the first step, we estimate age and year (but not person) fixed effects using never and not-yet treated observations. We do not restrict based on number of wave a person is observed. After constructing residuals by subtracting these fixed effects from the outcome variable, we regress the residuals on relative time. We construct standard errors by bootstrapping and report the 95th percentile confidence interval. In Panel d, we compare approaches. The 'Baseline, 1 Completed Fertility' line is the '1 Completed Fertility' line for this outcome for women in Figure 2. The 'Pregnancy Loss, No Prior Children' (and 'Pregnancy Loss, No Prior Children') line is constructed by subtracting the 'loss' estimate from the 'complete' estimate in Panels b (c) of this table and then dividing (for years 0 through 5) by the difference in children ever had between the complete and loss groups in Figure 3; the reason we do not scale the years before year 0 is that these are values are not always 0, but are very small due to measurement error, and dividing by them creates nonsensical scaled estimates. The 'Twins, No Prior Children' group is constructed similarly, scaling by the difference in Panel c of Figure 3. We present only cells with at least 25 (10) observations for pregnancy loss (twins).

Figure A.9: Childcare Hours, Pregnancy Loss and Twins, Women, Longer Post Period

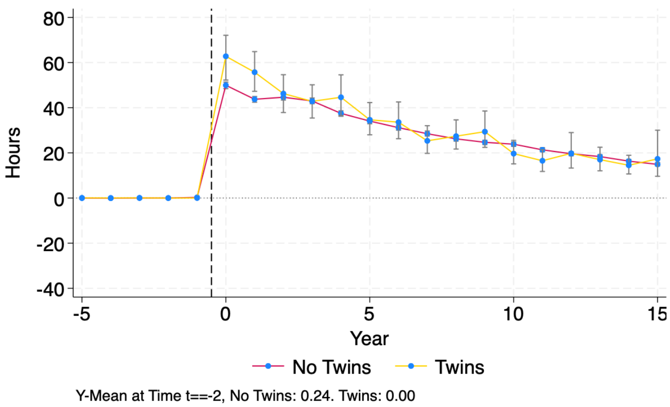
(a) Pregnancy Loss, Childcare Hours, Conditional on No Children



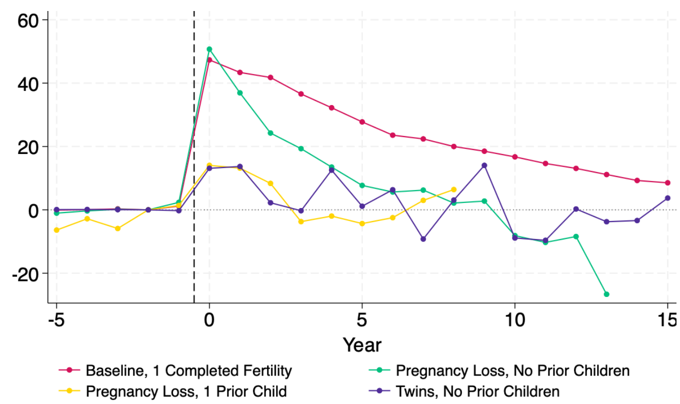
(b) Pregnancy Loss, Childcare Hours, Conditional on One Child



(c) Twins, Childcare Hours, Conditional on No Children



(d) Comparison of Approaches



Notes: This figure presents event study estimates of childcare hours for the first-observed pregnancy for women with 0 children (Panels a and b) and those with 1 child (Panel b). Pregnancy loss (pregnancy completion) is presented in red (yellow) for Panels a and b. Time 0 corresponds to the predicted year the next child was expected to be born (if the pregnancy were successful), which is the wave the person indicated they were pregnant if they said they were pregnant 10-12 months ago, and the subsequent wave if they said they were pregnant 0-9 months ago. We use a two-step approach here; we use the 'less conservative' event study approach: in the first step, we estimate age and year (but not person) fixed effects using observations before time 0 (or, for Panel b, observations before the expected or observed year of the first child being born); in this step, we consider only those that we assign a pregnancy completion status. After constructing residuals by subtracting these fixed effects from the outcome variable, we regress the residuals on relative time. We construct standard errors by bootstrapping and report the 95th percentile confidence interval. Panel c considers twin births. Non-twin births (twin births) are presented in red (yellow). Time 0 corresponds to the year the first born (computed as the age of the parent minus the age of the first-observed child). We use a two-step approach here; we use the 'less conservative' event study approach: in the first step, we estimate age and year (but not person) fixed effects using never and not-yet treated observations. We do not restrict based on number of wave a person is observed. After constructing residuals by subtracting these fixed effects from the outcome variable, we regress the residuals on relative time. We construct standard errors by bootstrapping and report the 95th percentile confidence interval. In Panel d, we compare approaches. The 'Baseline, 1 Completed Fertility' line is the '1 Completed Fertility' line for this outcome for women in Figure 2. The 'Pregnancy Loss, No Prior Children' (and 'Pregnancy Loss, No Prior Children') line is constructed by subtracting the 'loss' estimate from the 'complete' estimate in Panels b (c) of this table and then dividing (for years 0 through 5) by the difference in children ever had between the complete and loss groups in Figure 3; the reason we do not scale the years before year 0 is that these are values are not always 0, but are very small due to measurement error, and dividing by them creates nonsensical scaled estimates. The 'Twins, No Prior Children' group is constructed similarly, scaling by the difference in Panel c of Figure 3. We present only cells with at least 25 (10) observations for pregnancy loss (twins).

Table A.1: Time Use by Completed Fertility

	Women	Women w/ Kids	Women w/o Kids	Men	Men w/ Kids	Men w/o Kids
Paid Employment	23.6	17.8	31.0	37.7	40.7	34.9
	(18.3)	(17.3)	(16.9)	(18.0)	(17.2)	(18.3)
Travel to Work	2.8	2.0	3.7	4.2	4.5	3.9
	(3.6)	(3.1)	(3.9)	(4.2)	(4.3)	(4.0)
Household Errands	4.7	5.8	3.3	3.2	3.5	2.9
	(5.2)	(5.8)	(3.7)	(3.7)	(4.1)	(3.3)
Housework	14.0	19.1	7.6	5.8	6.7	4.9
	(12.5)	(13.5)	(7.0)	(5.8)	(6.6)	(4.9)
Outdoor Tasks	2.3	2.8	1.5	3.7	4.8	2.7
	(4.0)	(4.3)	(3.4)	(5.3)	(5.7)	(4.8)
Caring Own Kids	15.7	27.9	0.3	6.6	13.3	0.3
	(23.0)	(24.5)	(2.7)	(11.1)	(12.6)	(2.7)
Caring Other Kids	0.7	0.9	0.6	0.4	0.4	0.3
	(3.9)	(4.4)	(3.3)	(2.4)	(2.5)	(2.2)
Volunteer	0.7	0.8	0.6	0.6	0.6	0.6
	(2.6)	(2.6)	(2.7)	(2.6)	(2.5)	(2.7)
Caring Adults	0.8	0.9	0.6	0.5	0.4	0.5
	(6.3)	(6.8)	(5.5)	(5.1)	(4.8)	(5.3)

Notes: This table shows means and standard deviations of various time categories by predicted completed fertility. The sample is limited to ages 25-45. Having children ('w/ kids') is defined as having (reported) children under 18. "Paid Employment" through "Caring Adults" are asked in every wave, while the final four categories are asked only sometimes. All variables are hours per week.

Table A.2: Number of Years Between First and Second Child

	N	Pct.	Cum. Pct.
0	90	2.0	2.0
1	623	13.6	15.5
2	1616	35.2	50.7
3	1013	22.0	72.7
4	473	10.3	83.0
5	266	5.8	88.8
6	141	3.1	91.9
7	96	2.1	94.0
8	72	1.6	95.5
9	49	1.1	96.6
10+	157	3.4	100.0
Total	4596	100.0	

Notes: This table shows the distribution of years between the first and second child. The variable is constructed by subtracting the age of the second oldest from the age of the oldest in the first observation for which both variables are nonmissing. Spacings of more than 10 are combined with 10.

Table A.3: Percentage of People Using Type of Childcare

When parent is: Age of Kids: Type of Day:	Working School-age Term-time	Working School-age Holidays	Working Not Yet School-age -	Not Working School-age -	Not Working Not Yet School-age -
Me or partner	43.1	63.1	33.1		
	(49.5)	(48.3)	(47.1)	()	()
Brother or sister	12.5	10.9	0.9	16.2	3.1
	(33.0)	(31.2)	(9.2)	(36.8)	(17.5)
Child looks after self	16.6	11.1			
	(37.2)	(31.4)	()	()	()
Child comes to my workplace	3.1	3.7			
	(17.4)	(19.0)	()	()	()
Friend or neighbor comes to our home	2.8	2.0	1.8	5.5	5.2
	(16.4)	(13.8)	(13.4)	(22.8)	(22.1)
Friend or neighbor in their home	10.2	11.3	4.1	11.8	5.9
	(30.3)	(31.6)	(19.9)	(32.3)	(23.5)
Paid sitter or nanny	4.0	3.2	5.5	6.9	6.4
	(19.7)	(17.5)	(22.8)	(25.4)	(24.5)
Family day care	2.7	2.6	18.3	1.7	11.1
	(16.3)	(16.0)	(38.7)	(13.1)	(31.4)
Other 1	1.2	1.0	1.2	1.8	3.4
	(10.9)	(9.8)	(10.8)	(13.1)	(18.1)
Other 2	0.1	0.1	0.0	0.0	0.0
	(2.7)	(2.4)	(1.8)	(2.2)	(2.1)
Other parent not living in HH/ex-partner	0.2	0.6	0.1		
	(4.9)	(8.0)	(3.2)	()	()
Not applicable - boarding school	0.2				
	(4.0)	()	()	()	()
Not answered	0.2	0.2	0.2	0.6	0.1
	(4.2)	(4.7)	(4.9)	(7.9)	(3.0)
Child's grandparent who lives with us	2.4	2.5	2.8	2.9	3.8
	(15.2)	(15.7)	(16.6)	(16.7)	(19.2)
Child's grandparent who lives elsewhere	24.6	34.3	32.1	25.1	33.0
	(43.1)	(47.5)	(46.7)	(43.4)	(47.0)
Other relative who lives with us	0.8	0.9	0.8	2.7	2.5
	(9.0)	(9.2)	(9.1)	(16.2)	(15.6)
Other relative who lives elsewhere	9.4	14.0	9.6	19.8	18.8
	(29.2)	(34.7)	(29.5)	(39.8)	(39.1)
Formal outside of school hours care	29.0			5.0	
	(45.4)	()	()	(21.8)	()
Vacation care		21.6			
	()	(41.2)	()	()	()
Long day care center at workplace			8.1		
	()	()	(27.3)	()	()
Private or community long day care center			36.4	1.2	18.6
	()	()	(48.1)	(11.0)	(38.9)
Kindergarten / pre-school			14.4		12.5
	()	()	(35.1)	()	(33.0)
None				23.8	7.8
	()	()	()	(42.6)	(26.8)

Notes: This table shows the percentage of people who use various types of childcare. Column 1 refers to school-age children during the school term while the parent is working; Column 2 refers to school-age children during holidays while the parent is working; Column 3 refers to not-yet-school-age children while the parent is working; Column 4 refers to school-age children children while the parent is not working; and Column 5 refers to not-yet-school-age children while the parent is not working. Not all variables are asked in each column. We consider only people that we observe for at least one wave when they are age 30 or above, and only observations for people age 18 and older. For a given category, we consider only observations the question was (validly) answered.

B Data Appendix

Sample Years

For the analysis, we exclude 2020 and 2021 to avoid complications from the Covid-19 pandemic. However, when constructing most variables, we use these years. This does not matter for most variables, but allows us to obtain a slightly-better measure of, for example, completed fertility.[23]

Age at Birth of Child and Years Relative to Birth of Child

We define the parent's age at first birth as one's own age minus the age of one's oldest child, using the values from the wave with the first-observed wave containing child age. For the age of the child, we rely on the self-reported age of children variables.[24] These are reported as of the time of the survey, which is not always administered at the same time each year; this as opposed to being reported as of June 30th, as parental (respondent) age is reported. Using this variable—as opposed to other variables such as the ages of children in the household questionnaire, which are pegged to June 30th—allows us to observe age for both children in the HILDA survey and those not in the survey, who are most commonly nonresident children. Age of children is reported for both resident and nonresident children; we sort the children by age and assume that the oldest age corresponds to the oldest child.

We calculate years relative to birth of the first child by subtracting parent's age at first birth from parent's age. We also compute age at second, third, and fourth births—as well as years relative to these births—in a similar manner to above. For instance, we compute age at second birth using the first observation in which we observe the age of the second oldest child.[25]

If a person reports a child (first, second, third, or fourth birth) in wave X who is age 1, but do not report having a child in wave X-1, we recode the years relative to birth variable to be 0 in wave X.[26][27]

We drop people a) whose first child appears as being age X years or older (where X≥2) in

[23]This also leads to nonmissing age relative to birth of child observations for children born in 2020 and 2021.

[24]In a small number of cases, age is not reported. These cases will thus not be used to construct this variable.

[25]Constructing these variables this way can lead to, for a given wave, years relative to a later born being greater than years relative to an earlier born. In event studies in which we show effects for multiple births on the same plot, we exclude these cases. This can occur if there is inconsistent reporting of children across waves.

[26]Similarly, if, for example, the person reports having a child age A=2 in wave Y, does not report having a child in wave Y-2, and has a missing observation for wave Y-1, we recode the years relative to birth variable to be 0 in wave Y-2. We take similar steps for A up to 10.

[27]If we recode in this manner, we do not make a similar adjustment to age at first birth, discussed below.

wave Y; and 2) do not list having a child in waves Y-X through Y-1. For example, if a person reports having a first child age 5 in wave 9, but not report having a child in waves 4 through 8, we drop this person from the sample. In analysis specifically involving higher-order births, we also drop relevant cases like the above for higher-order births.

Measurement error can occur for several other reasons, which can include: A) If the respondent reports incorrectly. They sometimes report certain children in some waves, but not in others. B) If there are deaths of the children of the respondent. For instance, if the true firstborn child died before the respondent's first survey, we would treat their secondborn as the firstborn. While there is some information on death of children, we do not attempt to account for this issue. Overall, though, we do not think that any of these issues will substantively influence the results.

We drop from the sample individuals who do not have a value for age at first birth (they are always missing age for a child) *and* who are ever classified as having at least one child. This prevents them from entering into the never treated group of the event studies. We do not drop observations that, for example, report having a child age 0 but whose total children ever had variable is 0 in this wave.

Predicting Age at First Birth

We predict age at first birth for those who are missing this variable, either because they will never have children or because they have not yet had children as of their latest survey wave. This exercise is imprecise, but is still useful in that it allows us to show values for those with 0 completed fertility in graphs that are centered around age at first birth. To perform the prediction, we restrict the sample to those whom we observe as being at least 35 years old in a survey wave and who are born in 1935 or later,[28] and regressing age at first birth on home state fixed effects (measured in earliest observed wave; 8 values), highest level of education (measured in latest observed wave; 3 values), country of birth (3 values), year of birth, and year of birth squared. We estimate men and women separately. We then use the resulting coefficients to predict age at first birth for those that are missing age at first birth.[29] We note that if we compare the distribution of predicted

[28]Different types of biases are introduced with the values of these variables. Restricting the sample to those who are at least 35 years old gives us more years of support to estimate the year of birth and year of birth squared variables, but people who have their first child after the age of 35 will not enter into the regression. We also note that limiting to 1935 and after likely produces a better fit of year of birth for later years of birth at the expense of a worse fit for those with very early years of birth (who will later be excluded from the sample).

[29]In the first step of the prediction, we drop the tiny number of people missing at least one of these variables. There are fewer than 10 cases that do not have an observed value of age at first birth and who are missing at least one of the predictor variables. Because these are so uncommon, instead of doing something more complicated we simply assign

vs. actual values for those who are not missing the actual value, former distribution is much less dispersed than the latter.

Birth Spacing

For the main graphs in which we split by birth spacing, we construct the number of years between the first two births by subtracting the self-reported age of the second oldest child from the self-reported age of the oldest child in the first wave in which both appear. We compared this method to one in which we subtract the relative age of the second birth from the relative age of the first birth. The methods agree for about 95% of individuals. One reason for the disagreement is inconsistent reporting of children. For example a person reporting a particular child in some waves and not in others; this issue leads to negative values for birth spacing.

In the appendix graphs where we show the event study for the second birth, separately by the two most common birth spacings, we instead define birth spacing by the alternate method of subtracting relative age of the second birth from the relative age of the first birth. After restricting to those with a birth spacing of two (three) years, this method guarantees that there will be exactly one (two) years between births.

Completed Fertility

We take the following steps to increase the accuracy of the completed fertility variable, which appears in graphs in which we split by completed fertility. First, we consider only people who are at least 35 years old in at least one survey wave, noting that there is a trade-off between sample size at early ages and measurement accuracy. Second, for those who are at least X years old (43 for women and 51 for men) in at least one survey wave, we define completed fertility as the maximum value across waves of the total children ever had variable.[30] 43 (51) is the age at which approximately 99% of children have been born for women (men), based on a tabulation of a person's age when the person has a child at age 0.[31]

them the predicted age at first birth of the gender-specific median actual age at first birth among those 40 years and older.

[30]For 1.2% of people, the last-observed value of the total children ever had variable is lower than the maximum of this variable across all waves. We proceed using the maximum version. Mismatch is more common among men. Mismatch also occurs disproportionately more among those who have had a death of a child—it appears that some people exclude these children, even though they should be included.

[31]Using lower values than 43 and 51 would likely slightly improve the number of people correctly classified, but will also lead to more people being classified in a smaller bin of completed fertility, such as being classified into 1 children ever had than 2 children ever had.

Third, for people who are not at least X years old in at least one survey wave (they are between 35 and X-1 in their last survey wave), we define completed fertility as the last-observed value of expected fertility, which we define as the sum of the children they have already ever had (using the maximum across waves) and the number of additional children they intend to have (call this latter variable *Intend* for short); we provide additional explanation of the *Intend* variable.[32][33] In waves 5, 8, 11, 15, 19, respondents completed an additional module on fertility. They were first asked how many additional children they (or their partner) intend to have (Intend). (This question was not asked if they indicated that they or their partner was sterilized; we code Intend as 0. This question also was not asked in waves 5 and 8 if the individual stated that there was a physical or health reason that made it more difficult to have children; we code Intend as missing.) Waves other than 5, 8, 11, 15, and 19 did not include the fertility module. Respondents in these waves were first asked on a scale of 0–10 how likely it is that they will have additional child(ren). If they answered 0–5, they were not asked how many additional children they intend to have, and so it is not obvious how to best handle these cases. To provide guidance, among those age 35–44 who answered the fertility module, we perform a cross tab of number of additional children they intend to have and how likely it is that they will have additional child(ren). We find that 28% of people age 35–44 who answered with a likelihood of 3 intended to have a positive number of children, and that 56% and 69% of those who indicated 4 and 5 do so. The modal number of additional children for 4 is 0, and for 5 is 1. While some will be misclassified, we assign 0 intended children to those who answered 0–4, and 1 to those who answered 5.

Fourth, for the relatively few people who were not assigned a value of the variable based on the first three cases, we assign them their maximum across waves number of children they have had.

There are several sources of measurement error when constructing this variable, including relying on the person's forecast of how many children they have had and intend to have and the aforementioned censoring issue,[34] but it is our sense that because we restrict the sample to those

[32]If there is a higher value of the number of children they have ever had in a year after their last observed wave of expected fertility, we use this value.

[33]We use the total children ever had variable, which is asked in some waves and is updated by HILDA in other waves. If we construct a variable which is the sum of the number of children for which the person individually provides ages, we find that this variable matches the number of children ever had variable in 97.4% of observations, with the nonmatches sometimes having higher values and sometimes having lower values. If we add the number of children who have died since the last interview variable (which may be measured with error), the match is 98.7%.

[34]Another issue is when predicted completed fertility as assigned above is 0 but the person reports the age of a child in at least one wave (but the value for the number of children had is 0). This happens for only 19 individuals and we do not drop these cases.

that we observe as at least 35 years, which is nearer to the end of when most people have children, the majority of people will be classified correctly.

Twins

We identify parents of twins using two methods. The first method uses the twin identifier variable to find twins and to associate these twins with their parents. To do so, we keep all individuals with the twin identifier variable, which associates an individual with their twin (who is also in HILDA). We then link these children to their parents using the father and mother link variables.

The second method defines siblings of the same age whose parents say they have twins as twins and all other children who are not the same age or whose parents respond that they don't have twins as not twins. This survey question is only asked of parents about their children younger than 15 years old. It is not asked before wave 18.

We then combine the answers from both methods, considering someone to be the parent of a twin if at least one of the two methods indicates them as being so. While it appears to us that the above approach is likely to capture the majority of parents of twins, it is possible that this it will miss some.

Pregnancy Loss

To identify pregnancies and track whether or not they were completed, we use responses to questions about pregnancy (of self or partner) and birth (or adoption) over the past 12 months from adjacent survey waves. We also adjust for the number of months elapsed between waves and the reported presence of a child age 0 or 1 in either wave. The main pregnancy and birth questions come from a series of questions about major life events in the past 12 months in the self-completed questionnaire (SCQ). Respondents answer yes or no as to whether the event happened in their life in the past 12 months. If they respond yes, they are also instructed to indicate whether the event "happened or started" 0 to 3 months ago, 4 to 6 months ago, 7 to 9 months ago, or 10 to 12 months ago. As the survey is not always administered exactly 12 months apart, we interpret these answers adjusting for the time between waves. Waves 9 and beyond include the completion date of the SCQ. Around 90% of SCQ responses are within two weeks of the household survey. We assign the date of a wave using the date of the SCQ where available and the date of the household survey otherwise, and then subtract the date of the first wave from the date of the second wave, adding 365 days. We then construct months apart, where, for example, 11 months apart is between -30 and 0 days, month 12 is between 1 and 30 days, month 13 is between 31 and 60 days, and so on.

We create two samples: conditional on already having no children and conditional on already having one child who is at least two years old. In the description that follows, we focus on the former sample, but the latter sample follows a similar process.

We construct our sample of completed and not completed pregnancies as follows. We start with the sample of people who reported that they were pregnant (note we use this terminology as shorthand for if they or their partner were pregnant) in the past 12 months by selecting one or more time ranges for how many months ago the pregnancy "happened or started" (0-3, 4-6, 7-9, 10-12). We refer to this survey wave as wave 1.

To attempt to only include those with no prior children, we additionally include only individuals (with no children in prior waves) for whom any of the following hold: 1) in wave 1, the respondent reports ever having zero children and no children are observed; 2) in wave 1, the respondent reports getting pregnant 7-9 months ago, giving birth 0-3 months ago, ever having one child, and having one child observed; or 3) in wave 1, the respondent reports getting pregnant 10-12 months ago, giving birth 0-3 or 4-6 months ago, ever having one child, and having one child observed. Conditions 2 and 3 allow for the possibility that the pregnancy and birth happened in the same wave. For these conditions, we also allow for twins, where we use the twins variable described elsewhere.

We only consider a person's first-observed pregnancy. It is rare, but it happens that, within a wave, an individual reports getting pregnant for more than one range. In these cases, we consider only the first instance.

We additionally apply the following restrictions:

- We require non-missing responses in wave 1 to the month ranges (0-3, 4-6, 7-9, 10-12) of having given birth (or not) in the last 12 months (note, we use this terminology as shorthand for the question of whether they or their partner gave birth or adopted a child in the past 12 months).

- They must have responded to the next survey wave within 8 to 15 months of wave 1. We refer to this next survey wave as wave 2.
- Next we require non-missing responses in wave 2 to the month ranges (0-3, 4-6, 7-9, 10-12) of having given birth (or not) in the last 12 months.
  - If they answer yes to the birth question in either wave 1 or wave 2, the month range they select must not be implausible (e.g., indicative of answering the pregnancy question as the end rather than start of pregnancy or other respondent error). See below for how we delineate plausible and implausible ranges.
  - For every combination of pregnancy range reported in wave 1 and number of months between waves, we designate some ranges in each wave as plausible or implausible to contain a birth. Our guiding principle is to consider 8 to 10 months after a pregnancy began as a plausible birth range for that pregnancy. However, to allow for several weeks of uncertainty in our measure of time between waves (and because we group each set of 30 days into a month, but surveys, pregnancies and births happen within a month), as well as varying interpretations by survey respondents of whether "X months ago" includes or excludes the current month, in practice, end dates spanning 7 to 11 months post pregnancy start are allowed. Note that because responses regarding the timing of pregnancy and birth are ranges of three or four months, applying these rules results in a broad definition of feasible birth ranges. For example, someone who has 12 months between waves and reports a pregnancy starting 4 to 6 months ago in wave 1 will have plausible birth ranges in wave 2 of 4-6, 7-9, and 10-12 months ago. We consider a birth range plausible if at least one year in the birth range is plausible for one year in the pregnancy range.
    - A range is designated as implausible in wave 1 if it ends fewer than 8 months after the pregnancy range begins. This approach will exclude very premature births. For example, if someone reports getting pregnant 7-9 months ago and giving birth 4-6 months ago, we exclude them. Note that many of the implausible births in wave 1 are of the same month range as the pregnancy. While these answers could be re-

porting a birth from a prior pregnancy, they likely largely reflect some respondents marking the end instead of the start of pregnancy.

    · Birth ranges 0-3, and 4-6 are considered plausible in wave 1 if pregnancy is reported as starting 10-12 months ago. Birth range 0-3 in wave 1 is plausible if pregnancy started 7-9 months ago. All other birth ranges in wave 1 are implausible.

  * A birth range is plausible in wave 2 if it starts 7 or more months after the pregnancy range starts and ends 11 or fewer months after the pregnancy range ends. Note this range depends in part on how many months elapsed between surveys. This corresponds to a plausible birth range of 8 to 10 months after pregnancy starts plus an additional month on either side to allow for imprecision in our measurement of the time elapsed between surveys and survey respondents' differing interpretation of how they count the current month when counting months back.

    · For example, plausible wave 2 birth ranges for a pregnancy reported in wave 1 as starting 4-6 months ago are 0-3, 4-6, and 7-9 if 8 or 9 months elapsed between waves, 4-6, 7-9, and 10-12 if 10, 11, or 12 months elapse between waves, and 7-9 and 10-12 if 13, 14, or 15 months elapse between waves.

  * We consider birth ranges in wave 2 implausible for the completion of the pregnancy reported in wave 1 if they are earlier (less time elapsing since the reported start of pregnancy in wave 1) than the plausible ranges in wave 2.

– Birth ranges in wave 2 that are later (more time elapsing since the reported start of pregnancy in wave 1) than the plausible wave 2 ranges are not categorized as plausible (for the initial pregnancy) or implausible (indicating response error). Instead, they are consistent with a subsequent pregnancy completing after the initial pregnancy was completed or lost.

– We also identify cases in which the pregnancy could have plausibly resulted in an unreported birth between waves. We drop individuals who report a child age 0 (or 1) in wave 2, who was not present in wave 1 if they neither report a birth in wave 2 nor are among the cases we identify plausible for unreported birth between waves.

  * For example, it is plausible that a pregnancy reported in wave 1 as starting 7-9 (or 10-12) months ago was completed after wave 1 but more than 12 months prior to

wave 2 if 11 to 15 months elapsed between waves. (Note we allow for as few as 11 months to elapse in this case because there may be a few week difference between that measure of time between main survey waves and when the SCQ containing this question was completed, and it is being generous on how people can interpret the how many months ago question.) Similarly, it is plausible that a pregnancy reported in wave 1 as starting 4-6 months ago was completed after wave 1 but more than 12 months prior to wave 2 if 12 to 15 months elapsed between waves, again being generous.

Combining these sample definitions and restrictions, we are left with respondents who report a pregnancy in one wave and respond in the subsequent wave with sufficient information for us to categorize the pregnancy as (likely) completed (i.e., the pregnancy was followed by a birth within a timeframe that includes an implied 8, 9, or 10 months of gestation) or (likely) not completed (i.e., the pregnancy is not followed by a birth within 10 months of implied gestation. Note that a small portion of cases we categorize as not completed do have a child born in wave 2. These are cases where timing of birth is consistent with the original pregnancy not being carried to term but being followed relatively quickly by a subsequent pregnancy which is completed.

We classify 12.9% of births for pregnancy 1 as not complete, and 9.7% of births for pregnancy 2 as not complete.

Other Variables

We use the imputed version of the financial variables, such as weekly income.

Weekly income is "Current weekly gross wages & salary - all jobs [imputed]."

We define the working a positive number of hours variable based on the time use employment variable, coding it as '1' if the time use variable is positive. We do not make use of other similar variables, such as an employment indicator.

Marriage Status

We construct three marriage status around the birth of the first child variables: 1) married before relative age -2 (meaning two waves before wave of birth of the first child); 2) single between relative age -2 and 2; and 3) get married between relative age -2 and 2.

For the first variable, married before relative age -2, we consider the latest observe wave between waves -5 and -3, and use the marriage status of this wave.

For the second variable, single between relative age -2 and 2, we consider all waves between -2 and 2, and consider the person to be single if they are not married in each observed wave.

For the third variable, get married between relative age -2 and 2, we consider five cases. In the first case, we code them as getting married if we observe them as not married in wave -3 and then as married in a wave between waves -2 and 2. In the second case, we consider them to not have gotten married if they are single in all observed waves between -2 and 2. In the third case, we consider them to have gotten married if they are not married in a wave between -2 and 1 and get married in a later wave up to wave 2. In the fourth case, we do not consider them to have gotten married if they are married in all observed waves between -2 and 2 (unless they are not married in -2; see case 1). In the fifth case, we consider them to not have gotten married if they do not fall into a case above. Typically, this would be for those who are married and then get divorced.

It is possible, though unlikely, that we will miss situations due to an individual skipping wave(s) of the survey. For instance, if they are single, get married in waves that we do not observe, and then get remarried, we will still code them as single. We also only rely on the current marital status variable and do not use another variable that asks them to indicate if they got married in the last 12 months.

Most people will have a ‘1‘ for only one of the three variable, though in rare cases, someone will not have any ‘1s‘ (such as those who are not married before -2 and are married in [-2,2], but we don’t observe -3) or will have two ‘1s‘ (such as those who are 1) married at -3 (or possibly -4 or -5) and then unmarried in [-2,2], or 2) married in -3 (or possibly -4 or -5) and then not married and then married again in [-2, 2]).

Age in Sample Restriction

After cleaning the data as described above, we exclude all individuals who do not appear in at least one survey wave when they are age 45 or younger.