Journal of Population Economics (2026) 39:44
https://doi.org/10.1007/s00148-026-01189-5

ORIGINAL PAPER

Motherhood timing and the child penalty

Anikó Bíró[1] · Steven Dieterle[2] · Andreas Steinhauer[2,3,4]

Received: 14 July 2025 / Accepted: 29 June 2026

Abstract

Age at first birth has increased over time, which is likely to create many benefits in the labour market. Applying a number of alternative estimation approaches to administrative data from Austria, we estimate small to moderate effects of delaying motherhood on earnings, employment, and attachment to pre-birth firms in the medium run. However, even the largest of these effects are small relative to the average age effects, implying that delay leads to only small changes in child penalties. Much of the career return to later first birth comes from delaying the child penalty rather than changing how careers respond to children.

Keywords Female earnings · Fertility timing · Child penalty · Imperfect instrument

JEL Classification J13 · J31 · C26

1 Introduction

There is a growing literature documenting the presence of the child penalty—the large and persistent percentage drop in earnings for women relative to men after the birth of the first child—across developed countries (Kleven et al. 2019b, a). The child penalty has been shown to be the key contributor to the remaining gender inequality in the labour market (Goldin 2014; Blau and Kahn 2017; Kleven et al. 2019b). Therefore, understanding the mechanisms that shape child penalties is an important step to being

Responsible editor: Terra McKinnish

✉ Anikó Bíró
biro.aniko@krtk.elte.hu

Steven Dieterle
steven.dieterle@ed.ac.uk

Andreas Steinhauer
andreas.steinhauer@ed.ac.uk

1 ELTE Centre for Economic and Regional Studies, Budapest, Hungary

2 The University of Edinburgh, Scotland, UK

3 CEPR, London, UK

4 ROCKWOOL Foundation Berlin, Berlin, Germany

Published online: 22 July 2026

able to address this inequality. A separate literature has estimated large medium-run labour market returns for women from delaying first birth (Miller 2011; Bratti and Cavalli 2014).[1] It is, therefore, natural to ask whether delaying motherhood can substantially reduce the severity or persistence of the child penalty experienced by women. Indeed, theoretical models with endogenous fertility timing suggest that the size of the penalty might be affected by delay through a number of mechanisms, such as finding a better job match before birth, resulting in greater labour force attachment after birth (Erosa et al. 2002), or accumulating human capital which lowers the depreciation rate while on leave (Adda et al. 2017).

Studying the role that motherhood timing can play in changing child penalties is made challenging by the fact that fertility timing decisions are endogenous, so that OLS estimators of the returns to delay are likely biased upwards as they will capture the real effects of delay along with unobserved drivers of earnings (e.g., ability or preferences for work) that are correlated with preferences for later first birth. Nevertheless, we start our analysis using administrative data from Austria by comparing earnings and employment relative to the timing of first birth throughout the career for women who have their first child at different ages. While this naive approach recovers moderately large level returns, the implications for the child penalty are much smaller. For example, we estimate that each year of delay is associated with 4.6% higher earnings 5 years after first birth by OLS, but starting from an average penalty of 67%, this implies a reduction of only 1.0 percentage point. These OLS estimates for Austria are comparable to estimates from the USA, using data from the Panel Study of Income Dynamics (PSID). For the USA, we estimate that each year of delay is associated with 7.7% higher earnings 5 years after first birth, but starting from an average penalty of 46%, this implies a reduction of 1.5 percentage points.

Intuitively, the higher earnings 5 years after first birth from delay partially captures the effect of being 1 year older (reflecting progression along an upward-sloping age-earnings profile)—what we refer to as the mechanical effect of delay. The child penalty is calculated as a percentage of the counterfactual childless earnings, but the counterfactual earnings are also higher after a year of delay from these age effects. These mechanisms imply that, even when using an estimator that is likely to over-estimate the returns to delay, there is no substantial change to the shape of the child penalty.

In order to check the robustness of the conclusion that delay has at most a small effect on the child penalty, we make several refinements to our estimation approach that account for the main concerns with the OLS estimator, using detailed labour market and health indicators from Austrian administrative data. First, to address endogenous birth timing, we build off the literature that studies the medium-run returns to delay using pregnancy loss as an instrument for first birth timing (Miller 2011; Bratti and Cavalli 2014). While pregnancy loss does shift first birth to later ages, there are several problems with using it as an instrument, including the potential for negative selection

[1] Note that this is distinct from the literature on how the access to contraceptives or abortion affected labour market outcomes by changing the environment in which women were making human capital and labour supply decisions (Goldin et al. 2002; Bailey 2006; Bailey et al. 2012; Mølland 2016). Here, the focus is on the labour market returns to delay in a setting in which women have the opportunity to plan motherhood—and may have made human capital investment decisions accordingly.

into suffering a loss and the direct negative health effects of loss on labour market outcomes that would violate the necessary exclusion restriction. We build on the common arguments in the literature that the OLS estimator of returns to delay will be biased upward and the IV estimator may be biased downward (Miller 2011), and refine our estimates by using the approach for imperfect instruments from Nevo and Rosen (2012), lowering the estimated returns from the OLS estimates.

Overall, our goal is to estimate bounds on the average return to delay for the population of women we would observe with a pregnancy in our data, including those with both the most and least to gain from delay. Both the empirical and theoretical literature on fertility timing highlights reasons why OLS estimates of the return to delaying motherhood are likely upward biased (Erosa et al. 2002; Miller 2011; Doepke et al. 2022). Therefore, the OLS estimate will include both the causal effects of delay—potentially capturing fully anticipated fertility timing, accompanied by major educational, occupational, and job-sorting decisions made earlier—for this population along with the bias from positive selection into later fertility. By contrast, the pregnancy loss instrument shifts birth timing after key career investments and fertility plans have already been made, thereby reducing the potential impact of delay captured by the instrument—further reinforcing the lower bound interpretation for the effects of delay among the population of women who we would observe with a pregnancy. In this sense, the OLS and IV estimates can be viewed as providing upper and lower bounds, respectively, on the causal effects of delaying motherhood for the population of interest.

To check the validity of our estimates, we explore how interpreting the OLS and IV estimates as different weighted averages of heterogeneous treatment effects might threaten our results. We find evidence that the difference in weights applied by the two estimators is relatively small, implying that differences in the average biases for the two estimators are likely much more important. We also use Inverse Probability Weighting (IPW) to correct for the fact that not all women who suffer a pregnancy loss eventually have a child (over 20% in our data), which if left uncorrected results in a sample of women with a recorded pregnancy loss and subsequent birth that is positively selected in terms of pre-pregnancy earnings, labour force attachment, and health. While our re-weighting approach is based on a selection on observables assumption, this is a relaxation of the implicit assumption of selection being completely random underlying the prior work.

After applying these corrections, our estimates on the earnings return 5 years after first birth to 1 year of delay are between −0.08 and 3.55%. The fact that our corrected estimates are smaller than the OLS estimate reinforces the conclusion that delay does not induce a large change in the child penalty, with an implied change of between −0.6 and 0.7 percentage points. Therefore, even at the upper bound, the implied impact of delaying first birth by 10 years, the observed increase over the last 40 years, would be to reduce the child penalty 5 years after birth by 7 percentage points relative to a penalty of 67%, implying a sizeable remaining penalty. Further, the key result that delay has only small effects on the child penalty holds across subsamples defined by age at labour market entry (a proxy for education), and pre-birth occupation (white or blue collar), suggesting that the baseline results are not masking important returns for some groups. This potentially surprising lack of heterogeneity is due to the fact that while our estimates indicate higher level returns for those with later labour market entry, and in white-collar jobs pre-birth, this is relative to higher mechanical age effects of delay for these groups as they face steeper average age-earnings profiles, which implies higher counterfactual childless earnings as well.

Given that the child penalty a mother faces may change as their children age, it is also possible that age at first birth could impact the child penalty differentially as time passes from first birth. Therefore, we also estimate effects of delay on earnings for each year in a 16-year window around the first birth (5 years before, the year of birth, and 10 years after). Doing so, we see consistently small effects on the child penalty after birth but large effects before birth. The positive effects of delay on earnings before birth reflect large returns to staying childless a year longer. Although we find evidence that child penalties are not drastically altered by delay, the fact that we see large mechanical effects of delay before birth means that delay is still important for gender inequality over the entire career by delaying the realization of the child penalty. We quantify this effect by focusing on the cohorts of Austrian women born between 1957 and 1962, whom we observe from labour market entry to retirement age. We estimate an OLS return to a 1-year delay of 2.4% higher career earnings. For the same cohorts, we estimate a mechanical return of 1.4% higher career earnings from simply delaying a fixed penalty. This suggests that even if the majority of the effect of delay comes from delaying a fixed penalty, delay is still an important contributor to gender inequality over the career. For example, as the average age at first birth in Austria rose from 23 in the 1970s to 29 in the 2010s, our estimates imply that this shift is associated with about a 14% increase in women's average career earnings. However, most of the return to delay comes from postponing the timing of the child penalty rather than reducing its size.

Our results are closely related to those from Miller (2011), who uses the NLSY79 to estimate increases per 1 year of motherhood delay in cumulative earnings of 8.78%, average wages of 3.06%, and hours worked of 5.72% using a set of three fertility shock instruments indicating miscarriage, pregnancy while using contraceptives, and the time lapsed between wanting children and conceiving. The much smaller estimated effects for our key career outcomes are predominately driven by the fact that Miller (2011) measures earnings up to age 34 for women who have their first child between the ages of 21 and 33. This approach censors the child penalty for those with later ages at first birth. By contrast, our cumulative outcomes are measured up to the expected retirement age of 55 (see Sect. 6.3 and Appendix E for a discussion of this censoring issue).

Our results can also be compared to two recent studies that rely on local shocks to identify the effects of unplanned or unwanted childbirth. Gallen et al. (2023) find that an unplanned birth for those with a Long-Acting Reversible Contraceptive (LARC) prescription that fails reduces average earnings in the 6 years after childbirth by 32.1% for women aged 22 to 27 and 13.6% for women aged 28 to 44, implying an 18.5 percentage point smaller penalty for older mothers. Using our estimates of a 1-year birth delay, a 10 to 15 year difference in age at first birth would reduce the child penalty by 11.8 to 17.8%, broadly consistent with their findings, though differences likely reflect population selection, labour market context, and treatment definition. Londoño-Vélez and Saravia (2025) study women seeking an abortion in Colombia, where legal abortions require judge approval, and use judge gender as an instrument. Their treatment differs from a 1-year birth delay studied here and instead captures the difference in outcomes from being denied an abortion. Their estimated employment reduction of 10.6 percentage points from a baseline of 19.4% exceeds our estimated increase of 0.39 to 1.21% from a 1-year delay relative to a baseline employment of 71.4%, reflecting differences in population, context, treatment studied, and type of shock. Importantly, in both cases the local shocks—failure of a LARC and being denied an abortion—likely identify the effects for a select subset of the population that has actively sought to avoid having children that we might expect to have the most to gain from avoiding childbirth (i.e., have larger treatment effects from having children). In contrast, our estimates bound the returns for the wider population of would-be mothers.

2 Background and conceptual framework

2.1 Background

Austria is an ideal context to study the contribution of motherhood timing to gender inequality—one in which the policy environment was relatively stable over the period of our main analysis.[2] Figure 1 shows annual earnings in Austria (from our administrative data) by age for women across different ages of first birth with a comparison to the average earnings of men. Several patterns stand out. Earnings before birth are very similar between women and men (in line with Bertrand et al. 2010; Goldin 2014), which means delay leads to much higher earnings before birth. We see the earnings of women diverge from men after the birth of the first child. It follows that one of the key mechanisms driving the positive relationship between age at first birth and cumulative earnings may be by delaying the realization of the child penalty. We also note that earnings later in life tend to be positively associated with later age at first birth due to either smaller initial penalties or less persistent penalties—opening the possibility that delay changes how a woman's career responds to the birth of the first child. For instance, this faster recovery may reflect differences in pre-birth human capital or job match quality due to delay (Erosa et al. 2002; Herr 2016; Adda et al. 2017). Indeed, in Appendix Table A5, we show that there is a clear positive relation between education level, a proxy for pre-birth human capital, and age at first birth. These statistics show that the median age at first birth increases from 25.25 among those with only mandatory schooling to 31.5 among those with a university degree, with a clear step up in all quantiles for every increase in education.

[2] Austria is characterized by a generous maternity leave system. Since 1990, women were entitled to return to their pre-birth job within 2 years after birth—this policy did not change during the years we study. The maximum duration of maternity cash benefits was 18 months between July 1996 and July 2000, then it was increased to 30 months. Cash benefits consist of a flat payment indexed to inflation corresponding to about 40% of net pre-birth earnings. Until 2000, the cash benefits were conditional on 16–26 weeks of employment (depending on the mother's age and the parity of birth) the year before the childbirth; after that, the employment condition was abolished. In line with the generosity of maternity leave, the proportion of children aged under 3 using formal child-care arrangements was only 4% in 1998 (OECD 2001) and 12.5% in 2010 (OECD 2022). See Lalive et al. (2014) and Kleven et al. (2024) for further details on the institutional background.

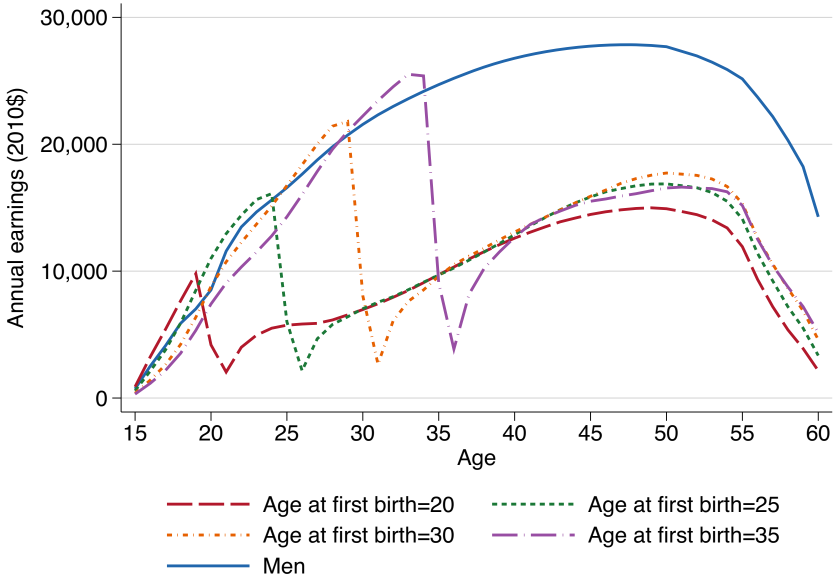

Fig. 1 Earnings and age at first birth. *Notes:* figure plots annual social security earnings by age converted to 2010 euro (CPI) and then to USD using the ECB reference exchange rate. The sample for this figure comprises every Austrian observed in the social security data (all of Austria) between 1972 and 2017. Years without earnings are included as zeroes. Years after the person has died are coded as missing. All Austrian men are included. We assign age at first birth based on year of the first observed birth spell minus year of birth and separately plot annual earnings by age for ages at first birth 20, 25, 30, and 35. We correct for topcoding using a simple Pareto distribution estimated by age and gender on income tax data (1994–2012)

Panel (a) of Appendix Fig. A7 shows average daily earnings (annual earnings divided by days of employment) by age. The pattern is similar to that we observe in Fig. 1 for annual earnings. Panel (b) shows the growth in daily earnings, and we set the growth to zero if earnings are missing (zero) in a given year. Apart from differences at the youngest ages, we do not see clear differences in wage growth by age at first birth. For annual earnings, we present the analogous figure for the USA in Appendix Fig. A1 based on the PSID data, showing patterns that are broadly similar to what we observe in the Austrian data.

2.2 Delay, child penalties, and cumulative earnings

While it is well established that motherhood is associated with a major decline in earnings and, therefore, serves as one of the main contributors to gender inequality in labour outcomes,[3] we know much less about how motherhood timing affects the motherhood earnings penalty. Here, we cast this in the "child penalty" framework of Kleven et al. (2019b) by expressing the labour market cost of motherhood as the

[3] See e.g., Goldin (2014); Blau and Kahn (2017); Kleven et al. (2019b) for important contributions and summaries of this literature. Kleven et al. (2019a) estimate the earnings decline associated with motherhood for many developed countries.

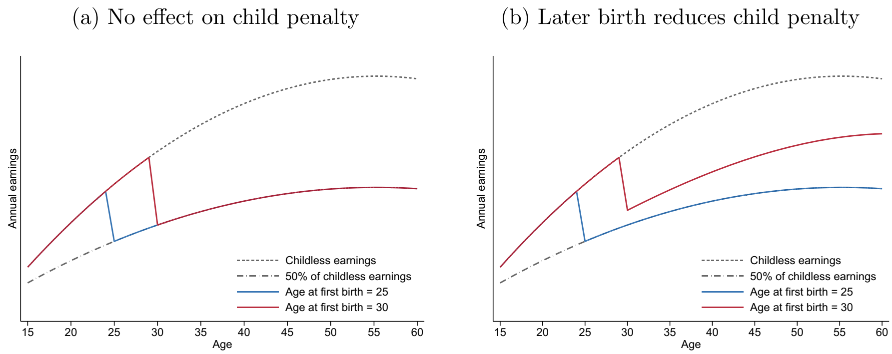

Fig. 2 Effects of birth timing on career earnings. *Notes:* for this figure, we simulate hypothetical earnings trajectories over the career for different ages at first birth. Panel (**a**) shows earnings over the career for two women who gave birth to their first child at ages 25 and 30 under the assumption that the child penalty is 50% for both of them. Panel (**b**) shows the same, but the woman with first child at 30 now faces a smaller penalty: 40% initially and then decreasing slightly with time after birth

earnings loss relative to counterfactual earnings without children. From this starting point, we are interested in two main questions. First, does motherhood timing affect the child penalty, i.e., does delaying motherhood reduce the size and persistence of the relative earnings loss after childbirth? Second, how do changes to the child penalty feed into the effect of delay on cumulative earnings over the career?

We start with a stylized example. Figure 2 depicts hypothetical age-earnings profiles for childless women and those with a first child at age 25 or 30 under different assumptions. In Panel (a), we show a case in which delay has no effect on the shape of the child penalty as it is set at a constant 50% of childless earnings, while Panel (b) plots a scenario where delay leads to a smaller penalty in terms of both the initial drop in earnings and the persistence. Delaying first birth in Panel (a) simply postpones the child penalty—moving along the childless earnings profile for longer before switching to the lower earnings profile with children—what Miller (2011) calls the "mommy track". Note, however, that even though delay has no effect on the shape of the child penalty in this case, regressions of earnings at time t relative to first birth on age at first birth will identify positive returns since women that delay will be further along the age-earnings profile. For example, in Panel (a) a comparison of earnings 5 years after first birth for those with a first child at age 30 versus 25 will simply reflect the higher earnings at age 35 versus 30 along the same age-earnings profile with children—what we refer to as mechanical age effects.[4] Importantly, this difference in earnings levels for the two women corresponds to no change in the child penalty (assumed to be a constant 50% of childless earnings in Panel (a)) as the higher age 35 earnings with a child are relative to the higher age 35 earnings without a child. Finally, note that while the penalty is unaffected by delay in Panel (a), cumulative earnings over the career (the area under each age-earnings profile) will be higher for the woman with a first child at age 30 by delaying the realization of the child penalty. That is, delay could

[4] Importantly, in regressions of earnings at time t relative to first birth on age at first birth, age at first birth and age would be perfectly collinear.

have no impact on the penalty, but still have implications for overall gender inequality in earnings.

In Panel (b) where delay reduces the child penalty with a smaller initial drop in earnings and faster relative earnings growth after birth (see Erosa et al. 2002; Adda et al. 2017) a comparison of earnings 5 years after first birth will capture the age effects (moving along the age-earnings profile for the first birth at 25 profile from age 30 to age 35 earnings) and the shift in the the child penalty from delay (the vertical difference between the profiles for age at first birth of 30 versus 25). It is important to distinguish between the mechanical age effects of delay and real changes to the earnings profile with children in order to understand the role delay can play. Once again, in terms of the child penalty, the earnings 5 years after first birth for later first births are relative to higher childless earnings. Now delaying motherhood increases cumulative earnings by not only delaying the child penalty as in Panel (a), but by also reducing the child penalty compared to Panel (a).

3 Data

We use administrative records from the Austrian Social Security Database (ASSD), which is a matched firm-worker data set covering the labour market history and some demographic characteristics of workers in dependent employment in Austria.[5] The ASSD covers the period from 1972 to 2017. We observe labour market status on a daily basis and annual earnings for worker-firm pairs as well as basic demographic characteristics (age and gender) and the birth dates of children for women. Further details related to the ASSD are provided by Zweimüller et al. (2009).

In order to construct the instrumental variable capturing pregnancy loss and to investigate selection into motherhood, we link the labour market data to health insurance records from the regional statutory health insurance fund of Upper Austria ("Oberösterreichische Gebietskrankenkasse"). Upper Austria is one of the nine states of Austria with around 1.4 million inhabitants, a sixth of the Austrian population. Health insurance is mandatory, and every Austrian resident is assigned to one of 25 "funds" depending on place of residence and occupation.[6] The time span of the health insurance records is 1st January 1998–1st October 2007. The data provide individual records of all events covered by the fund. In particular, health-care utilization associated with any kind of treatment by GPs or specialists, prescriptions, or hospital admissions is reported with quarterly accuracy. Hospital admissions include detailed diagnosis codes according to the International Classification of Diseases and Related Health Problems (ICD-9 or ICD-10), and prescription records include information based on the Anatomical, Therapeutic and Chemical (ATC) classification system.

From the matched data set, we extract a sample of women we observe giving birth to their first child or who suffer a pregnancy loss without preceding birth between 1st

[5] It does not cover the self-employed, civil servants, and some special industries (farmers, miners). See Zweimüller et al. (2009) for details.

[6] Self-employed workers, civil servants, and those employed in some public utilities (e.g., Austrian Federal Railways OEBB) or specialized industries (e.g., farmers) are insured by different funds. The data and institutional setting are described in more detail in Kuhn et al. (2009) and Hackl et al. (2015).

January 1999 and 1st October 2007.[7] We proceed in four steps. First, we extract all births we observe in either the health insurance (1998–2007) or ASSD data (1972–2017) along with all pregnancy losses in the health data (1998–2007). For each woman, we keep only the first observed "event", which is either a birth or a pregnancy loss. Second, we restrict the sample to first events in 1999–2007. This allows us to use the ASSD data to track labour market outcomes for all women and observe a subsequent birth for those who suffer a loss for a minimum of 10 years and up to a maximum of 18 years after the first event. We further restrict the sample as follows. We drop women younger than 18 or older than 40 at the time of the event—which leaves those born between 1959 and 1989. We also drop women for whom the first event takes place before labour market entry and those who have no health insurance records prior to the event. We define labour market entry as the year in which a woman first earns at least €2500 (in real terms, CPI 2000).[8] Therefore, we exclude women with very weak attachment to the labour force for the entire period before the first pregnancy. These restrictions help ensure that we include women that are likely to have their health information recorded in the insurance data. In this, we are also following Herr (2016) who suggests to consider women who work before having children separately from those that have children before entering the labour force, as experience, rather than age itself, is the important driver for the potential returns to delaying motherhood.[9]

Our main sample consists of 52,069 women for whom we either observe a first birth or pregnancy loss without preceding birth between 1999 and 2007. In this sample, the

[7] Our indicator of pregnancy loss is based on ICD-9 and ICD-10 codes. A woman is defined to suffer a pregnancy loss in a given period if an ICD code referring to miscarriage or stillbirth is recorded in that period. Using the ICD-10 terminology, the following diagnoses are included in our definition of pregnancy loss: ectopic pregnancy; molar pregnancy; other abnormal products of conception; spontaneous abortion; other abortion; unspecified abortion; maternal care for intrauterine death; fetal death of unspecified cause. Very early pregnancy losses (within a few weeks after conception) are common and often unnoticed; thus, these cannot be included in our pregnancy loss definition. Induced abortion is not considered as part of pregnancy loss. We do not include events occurring during the first year of our health data, 1998, to ensure we have at least 1 year of pre-event health data.

[8] The cut-off serves to exclude seasonal part-time jobs for women still enrolled in high school and university. We conducted robustness checks varying the cut-off between 500 and 5000, which has little effect on our results. We include dummies for age at labour market entry in all regressions below. To verify that this entry measure is reasonable, we use data on the highest educational achievement at the time of giving birth that is available for a subset of women in our sample. Using our definition of labour market entry matches expectations based on educational achievement well. For instance, women with professional apprenticeships, which are paid and follow the dual system of on-the-job training and formal education, have a median age of labour market entry of 16. Their median earnings in our sample are 1901 at age 15, 5235 at age 16, and 6697 at age 17. Earnings breakdowns for other education levels, available upon request, further support the use of the labour market entry variable as constructed.

[9] Our original sample consists of 72,248 women observed in the Upper Austrian administrative records who experienced either a birth or a pregnancy loss between 1999 and 2007. The sample restrictions affect the data as follows: 10,671 of the women have no recorded health care utilization prior to the first event; 2199 are younger than 18 or older than 40; and 9165 had not entered the labour market before the first event. Taken together, these three restrictions reduce the sample size from 72,248 to 52,069. We have conducted robustness checks on the main unweighted and weighted OLS and IV estimates using the full sample. While the estimated magnitudes differ somewhat, none of our conclusions are affected by imposing the restrictions. The corresponding results are available upon request.

first event is a live birth for 46,437 women and a pregnancy loss for 5632 (10.8%).[10] Among women whose first event is a pregnancy loss, 4561 (81%) are observed with a live birth at some point after the pregnancy loss and before the end of the ASSD data window (31st December 2017).

We construct several labour market and health-care utilization variables for the women in our sample. While the ASSD provides daily information on employment, health-care utilization is recorded only at quarterly frequency. Therefore, we aggregate variables from both data sets to either annual frequency or measured as of reference days in the middle of each quarter (15th of the second month of the quarter). Our main labour market outcomes measure conditions before and after first birth, as well as the cumulative outcomes over the entire sample period. From the ASSD we take annual earnings, which are top-coded at the maximum contribution base that is adjusted yearly (affecting roughly the top 2% of female workers), daily employment status, and the firm a woman is working for on any given day. If a woman does not work in a given year or on a given day, we assign zero earnings and employment. We deflate all earnings using the CPI (base year 2000). Using the daily spell information on employment and earnings, we construct daily earnings on quarter reference days, days worked up to a certain quarter (experience), whether a woman is employed on the quarter reference date, and an indicator for whether this employment is in a white-collar job (the only available measure of occupation).

To shed light on selection into giving birth after pregnancy loss and the direct health impacts from pregnancy loss, we make use of additional data from the health insurance fund. As a measure of utilization, we extract the total health expenditure covered by the insurance fund per quarter. Additionally, we construct several mental health indicators. Based on the recorded ICD9 and ICD10 codes associated with admission to a hospital, we generate an indicator for having ever been previously diagnosed with depression in each quarter and a separate indicator for a broader set of mental health conditions (including depression). Similarly, we use the ATC codes in the prescription records to generate variables capturing if a psychiatric drug or antidepressant was ever previously prescribed.

In some cases, we do not require women to be matched to the Upper Austria health data and consider the sample of women in all of Austria. Importantly, this allows us to relax the restriction that the first event occurs in the window of the health data so that we can measure labour market outcomes over the entire career for a sizable sample of women. In particular, we use the sample of women in the ASSD data born between 1957 and 1962, whom we observe earnings over the entire 15–55 age range. This better coverage of career outcomes comes at the cost of not being able to implement some of our estimation approaches that rely on the health data.

[10] The incidence of pregnancy loss in our sample is comparable to the incidence reported by Miller (2011) based on US survey data (14%) and Bratti and Cavalli (2014) based on Italian survey data (12.8 loss per 100 live births, thus roughly 11.3% of pregnancies). Using administrative data from Sweden, Karimi (2014) reports a lower incidence of miscarriage (around 5%).

4 OLS estimates of the return to delay

We begin by estimating the descriptive relationship between age at first birth and subsequent labour market outcomes. Following the literature, we acknowledge that OLS estimates likely capture selection on unobserved factors; specifically, women with higher unobserved earnings potential may delay motherhood to reduce the size of the penalty. Consequently, we interpret our OLS estimates as an upper bound on the causal effect of delay. As noted in the theoretical literature, the timing of birth derives from the career-fertility trade-off, where women with higher returns to experience opt for later births.[11]

We establish this descriptive upper bound to motivate our core conceptual exercise: decomposing the estimated returns to delay into a mechanical ageing effect and a genuine career adjustment effect. Our main specification simply regresses labour market outcome variables on age at first birth plus control variables. For a given event time t (e.g., 5 years after first birth), we estimate:

$$y_i = \beta A_i + \mathbf{x}_i \gamma + u_i, \tag{1}$$

where A_i is age at first birth of woman i, and $\mathbf{x}$ contains a constant and a set of control variables (cohort indicators and age at labour market entry). The resulting $\hat{\beta}^{OLS}$ captures the effect of delaying first birth on a given outcome, such as annual earnings in a specific year after birth. For example, Table 1 shows that annual earnings 5 years after first birth increase by €307.6 per year of delay. Because women are mechanically 1 year older if they delay first birth by 1 year, we expect earnings to rise simply due to age and experience effects, unrelated to deeper career adjustments.

Ideally, we would capture these differential career adjustments naturally by estimating changes in the relative child penalty. However, we cannot separately identify pure experience effects from delay effects in a standard event study specification, such as the child penalty regression (Kleven et al. 2019b), because they are mechanically related (see Thakral and Tô 2025 for a comprehensive discussion). To address this identification challenge, we compare the level effects of delay on labour market outcomes (our $\hat{\beta}^{OLS}$) to the simulated effect of simply delaying a fixed relative child penalty.

[11] Miller (2011) notes the possibility of selection into later motherhood by those with higher unobserved earnings potential in order to reduce the size of the penalty and of "unobserved social group factors" generating correlations between fertility and labour market outcomes. From a theoretical perspective, Doepke et al. (2022) propose a model (following Elizabeth et al. 2002), in which the timing of birth derives from the career-fertility tradeoff. In this model, women with lower preferences for children and those with higher returns to experience will opt for later births and invest in a career instead. The model also implies that women with high fertility preferences will select into jobs with flatter career paths. In the model of Erosa et al. (2002), women with lower tenure due to experiencing more exogenous job separations early in the career are more likely to have children early; however, these women would have faced lower earnings at any age at first birth than those who delay due to their slower accumulation of general human capital from these same job separations. While there are stories that would lead to a negative relationship between earnings and later first birth for some women—such as negative health shocks that delay fertility and lower earnings—the related literature emphasizes the positive selection into later first birth, suggesting it is the first-order concern that is likely to hold on average across all women.

Table 1 OLS effect of age at first birth on outcomes 5 years after first birth

Dependent variable	Mean 5 year after first birth (1)	OLS effect of 1-year delay (2)	Mechanical effect of 1-year delay (3)	Child penalty (4)	Effect on child penalty [95% CI] (5)
Annual earnings (€)	6680.5	307.6	112.1	−0.670	0.010
		(16.1)			[0.008, 0.011]
in %		*0.046*			
Participation	0.714	0.009	0.002	−0.289	0.007
		(0.001)			[0.006, 0.009]
in %		*0.012*			

Notes: annual earnings are total labour earnings per calendar year including zeroes. Participation is set to one if an individual ever worked in a calendar year, zero otherwise. We calculate means 5 calendar years after first birth and the OLS effect of age at first birth on the two outcomes 5 years after first birth, controlling only for year of birth of the mother (dummies). We divide the level effect by the means to show the effect in % in italics below, with robust standard errors in parentheses. The third column shows the mechanical effect of a 1-year delay, which is calculated by estimating child penalties on the relevant sample, and then calculating what would happen if everyone delayed the average child penalty by 1 year according to Eq. 5. We show estimated child penalties 5 years after first birth in column (4) and the OLS effect from column (2) converted to an effect on the child penalty in column (5). For this, we simply subtract the mechanical effect from the OLS effect and then divide by estimated counterfactual earnings. As a reading example for annual earnings: each year of delaying first birth is associated with a 1 percentage point smaller child penalty 5 years after first birth. The 95% CIs on those child penalty effects are shown in square brackets. See Sect. 4 for details on the estimation of the mechanical effect and the child penalty

Our technical implementation proceeds as follows. As in Kleven et al. (2019b), we estimate the following earnings profile regression:

$$Y_{ist} = \sum_{j \neq -1} \alpha_j \cdot I[j = t] + \sum_k \delta_k \cdot I[k = age_{is}] + \sum_y \gamma_y \cdot I[y = s] + \nu_{ist}. \quad (2)$$

Here, Y_{ist} are annual earnings for woman i in calendar year s at event time t (i.e., t years after first birth) for t in $[-5, 10]$. We include event time effects (α_j) as well as age (δ_k) and year (γ_y) effects to flexibly control for general age and business cycle effects. Denote predicted earnings from Eq. 2 as $\hat{Y}_{ist}$ and the mean earnings profile around first birth as $E[\hat{Y}_{ist}|t]$. Kleven et al. (2019b) define the child penalty at event time t as the earnings impact at event time t relative to counterfactual earnings without children. Denote predicted counterfactual earnings (setting $\alpha_j = 0$) as $\widetilde{Y}_{ist}$. The mean estimated child penalty based on Eq. 2 can be expressed as[12]:

$$P_t = \frac{E\left[\hat{Y}_{ist} \middle| t\right] - E\left[\widetilde{Y}_{ist} \middle| t\right]}{E\left[\widetilde{Y}_{ist} \middle| t\right]} = \frac{\alpha_t}{E\left[\widetilde{Y}_{ist} \middle| t\right]} \quad (3)$$

[12] As shown by Kleven et al. (2024), the child penalty for men is trivially small in Austria; hence, we ignore taking the difference between women and men when defining child penalties.

We derive the mechanical effect of delaying first birth by calculating mean predicted earnings from a 1-year shift in the age profile while holding the penalty fixed. From Eq. 3, we can write mean predicted earnings as:

$$E\left[\hat{Y}_{ist}\middle|\, t\right] = E\left[\widetilde{Y}_{ist}\middle|\, t\right](1+P_t) = E\left[\widetilde{Y}_{ist}\,(1+P_t)\middle|\, t\right]. \tag{4}$$

Therefore, we define the predicted earnings due to mechanically delaying the penalty as $\ddot{Y}^{+1}_{ist} = \widetilde{Y}^{+1}_{ist}\,(1+P_t)$—where $\widetilde{Y}^{+1}_{ist}$ are the counterfactual earnings with a 1-year delay derived by replacing the observed age with $age_{is}+1$ (i.e., moving along the childless profile from $\widetilde{Y}_{ist}$ to $\widetilde{Y}^{+1}_{ist}$ before suffering the average penalty). From here, we define the mechanical effect of delay on the earnings profile as:

$$Mechanical\ Return = E\left[\ddot{Y}^{+1}_{ist}\middle|\, t\right] - E\left[\hat{Y}_{ist}\middle|\, t\right]. \tag{5}$$

For example, consider the child penalty in annual earnings 5 years after first birth: −67% in our Austrian sample. This means mothers on average have 67% lower earnings 5 years after birth compared to having stayed childless and moved along the childless age-earnings profile. If they delayed first birth from age 30 to age 31, then earnings 5 years after first birth would mechanically increase by $67\% \times \widetilde{Y}^{age=36} - 67\% \times \widetilde{Y}^{age=35}$.

By comparing the OLS level effects against these mechanical returns, we treat the common component as the mechanical effect and the difference as the extent to which delay genuinely alters the dynamic relative penalty. In Table 1, we see that the average mechanical return to 1 year of delay in our Austrian sample is €112.1. Comparing this to the total OLS return of €307.6 implies that over a third of the estimated return to delay comes from the mechanical age effects. The remaining portion suggests that slightly less than two-thirds could be due to actual changes in the child penalty or bias of the OLS estimator.

In order to quantify the importance that a change in earnings of this magnitude would have on the child penalty itself, we standardize the €307.6 increase by dividing it by the counterfactual childless earnings. In Table 1, we see that the estimated return to a year of delay reduces the child penalty by 1.0 percentage points. Relative to an average penalty of −67%, this represents a small change in the shape of the penalty due to delay.

Similarly, the estimated return to a year of delay would reduce the child penalty in employment by 0.7 percentage points. We get very similar results based on our US sample, shown in Appendix Table A1. In the USA, the estimated return to a year of delay would reduce the child penalty by 1.5 percentage points, again a small change in the shape of the penalty due to delay, relative to an average penalty of −46%. The estimated return to a year of delay would reduce the child penalty in employment by only 0.1 percentage points in the USA.

In sum, even when likely overestimating the returns to delay, we find evidence that delay does not greatly change the child penalty. This finding holds both for Austria and the USA. In the remaining empirical analysis, we use the Austrian samples, building on the richness of the Austrian administrative data.

5 Empirical approach with instrumental variables

The previous OLS estimates of the return to delay are likely to be too large, capturing endogenous fertility timing decisions. We address this endogeneity in order to see how much smaller the impact of delay on the child penalty could be once we account for the bias. We begin by considering our simplest specification used to estimate the relationship between outcome y_i and age at first birth A_i shown in Eq. 1 above.[13] Consistent estimation by OLS is based on the following set of moment conditions: $E\left[(A, \mathbf{x})' u\right] = \mathbf{0}$. As discussed, we worry that A_i is positively correlated with the unobservables driving earnings ($\rho_{Au} > 0$) so that $E[Au] > 0$, so that the estimated returns to motherhood delay will be biased upwards.

Therefore, we follow the literature that uses pregnancy loss as an instrument for motherhood timing (Miller 2011; Karimi 2014; Bratti and Cavalli 2014).[14] Here, we define a "no loss" instrument z_i which equals zero if first pregnancy ends in loss and equals one if first pregnancy results in live birth.[15] Consistent IV estimation would require $E\left[(z, \mathbf{x})' u\right] = \mathbf{0}$. There are several concerns with the validity of this instrument. First, pregnancy loss may not be fully random but related to behavioural characteristics (smoking, drinking, among others) that have been found to be correlated with labour market outcomes (see Miller 2011, although Markussen and Strøm 2022 argue that the impact of behavioural factors on the miscarriage risk is small). Second, pregnancy loss may negatively impact mental health (Lok and Neugebauer 2007). Given the literature connecting mental health and labour market outcomes (Ernst et al. 1998; Chatterji et al. 2007, 2011; Peng et al. 2015), this raises the possibility that pregnancy loss may affect labour market outcomes through its effect on mental health, violating the exclusion restriction (Kalsi and Liu 2021).

[13] The discussion that follows is based on the linear, homogeneous effects model in Eq. 1 where age at first birth is a continuous variable. See Appendix B for a discussion of how a heterogeneous effects interpretation of the OLS and IV estimators discussed below could alter the interpretation. Importantly, Appendix B provides evidence that the difference in weighting implied by the OLS and IV estimators is likely less important than the biases discussed here.

[14] Note that the research design used here is distinct from recent work that also uses unexpected fertility shocks but with dynamic changes in treatment assignment. For example, Gallen et al. (2023) are interested in the treatment of having a child vs not having a child, with event timing relative to receiving a Long-Acting Reversible Contraceptive (LARC). Here, the initial treatment (child/no child) depends on whether the LARC failed in the 9 months following the prescription, resulting in an unplanned pregnancy. The dynamic treatment kicks in once those who were initially in the control group (no pregnancy or child) may decide to have children. Therefore, in later event times some of those in the initial control group will become treated so that the distribution of actual treatment changes with event time. Our setting differs as we are primarily interested in whether the age at which women have their first child changes the size of earnings losses at event times set relative to first birth. Treatment in this case is the continuous Age at First Birth measure. Suffering a pregnancy loss shifts the distribution of this treatment variable to later ages at first birth; however, the initial age at first birth distributions given the instrument are not changing as we move through event time—implying there are no dynamic treatment changes when looking at earnings, say, 5 years after first birth.

[15] Note, we define the instrument as "turning on" for those that do not suffer a loss in order to simplify the discussion of the methods in Sect. 5.2. In particular, it is helpful for the direct application of the methods in Nevo and Rosen (2012) if the expected sign of the correlation between the instrument and the unobserved error term in Eq. 1 is the same as the expected sign of the correlation between the endogenous regressor—age at first birth—and the error term (see Eq. 6 in Sect. 5.2. Of course, the resulting estimates would be identical if we defined the instrument as turning on for those who suffer a loss.

The main concerns with the pregnancy loss instrument suggest the IV estimator will be biased down. Further, pregnancy loss is a local shock in the sense that it affects birth timing at a point when important career choices have already been made while planning for an expected age at first birth, so that the estimate will not capture the full returns from a completely anticipated choice to delay fertility—further suggesting downward bias in the IV estimators.[16] As descriptive evidence in support of this possibility, we estimated two multinomial logit models of educational attainment at the time of first birth (compulsory schooling, some high school, high school, professional apprenticeship, teacher's college, and college) for the subset of women in our sample whose education is recorded in the birth registry. In the first specification, we include age at first birth, indicators for the mother's birth cohort, and age at labour market entry. Age at first birth is a statistically significant predictor of educational attainment, suggesting that differences in educational choices may contribute to the observed returns to delay. In the second specification, we replace age at first birth with an indicator for pregnancy loss. This variable is not statistically significant in predicting educational attainment at first birth. This finding suggests that the returns estimated using the pregnancy loss instrument are unlikely to reflect differences in educational choices, as pregnancy loss does not appear to be systematically associated with educational attainment at the time of first birth and therefore does not predict differential selection into schooling paths. See Appendix C for details. In Sect. 5.2, we discuss how we use the approach from Nevo and Rosen (2012) to adjust the OLS estimator using the imperfect pregnancy loss instrument to recover a tighter upper bound on the returns to delay.

Finally, comparing first birth timing for women who did not suffer a loss to those who did will implicitly condition on the choice made after the biological shock to attempt to have a child. As shown in Appendix Fig. A8, most women in our data have children within 3 or 4 years after suffering a pregnancy loss. However, about 20 percent of the women whose first observed pregnancy ends in a loss are never observed having a child up to 15 years after the loss and therefore "fall out" of the natural experiment. Since not every woman who suffered a loss eventually gives birth, this selection would lead to bias even if pregnancy loss itself were randomly occurring as long as the incidence of attempting another pregnancy occurs non-randomly—a reasonable possibility in this setting. Therefore, ignoring the women who select out of the sample after suffering a loss threatens the validity of the thought experiment of comparing similar women (in terms of initial choice of time of first birth) who became pregnant at the same time but where some were delayed by the biological shock. We therefore implement a probability weighting approach, described in Sect. 5.3, to account for this selection.

[16] Most instruments considered in the literature operate through "local" shocks to the motherhood timing decision that are likely to delay first birth. One exception is Gallen et al. (2023) that use the unexpected failure of IUD contraception to study women who have the age at first birth forward, rather than delayed, by a shock.

5.1 IV bias evidence

In this section, we provide descriptive evidence that suggests biases in the IV estimates using the pregnancy loss instrument. To do so, it is helpful to distinguish between two distinct requirements underpinning the IV moment condition: (1) instrument independence—that the incidence of pregnancy loss occurs as-if at random and is, therefore, unrelated to the labour market returns to delay and (2) the exclusions restriction—that the pregnancy loss only affects labour market outcomes through the effect on age at first birth,[17]

Using our labour market and health care use data, we define three groups of women. *Group 1* are women for whom the first observed event is a live birth. *Group 2* are women for whom the first observed event is a pregnancy loss, and we observe a subsequent live birth (up to 2017). *Group 3* are women for whom the first observed event is a pregnancy loss, and we do not observe a subsequent live birth (up to 2017). Appendix Table A6 shows descriptive statistics for the three groups. We see similar ages at first event for groups 1 and 2 at around 27.3 years old, while group 3 is older on average. Group 2's average age at first birth of 29.6 implies an average delay of 2.3 years due to pregnancy loss. Total number of births over our data period are also similar at around 1.92 on average among women whose first event is a live birth, and 1.85 among women in group 2. When we compare summary measures of labour market outcomes over the entire data window (1972–2017), we see generally higher average daily earnings and more experience among groups 2 and 3. In terms of total health-care utilization over the health data window (1998–2007), we see higher utilization among groups 2 and 3 compared to group 1, although only the former is statistically significant.

To look for evidence of possible violations of independence and the exclusion restriction, we apply an event history type approach around the time of pregnancy. Figure 3 displays the incidence of a birth (first stage), average daily earnings, and mental health diagnoses by quarter for different subsets of the women in our sample both before and after their first observed pregnancy.[18] Panel (a) of the figure shows that even 5 years after the first event, still less than 80% of women had a live birth among those whose first event was a pregnancy loss. Panel (c) plots average daily earnings on quarterly reference dates—recording zeros for those not in employment—separately for women whose first observed pregnancy ends with a live birth (group 1) or a pregnancy loss (groups 2 and 3). Here we see very similar trends in daily earnings before pregnancy, followed by a sudden drop for the live birth group at the time of the first birth and a slower decline in average earnings for the pregnancy loss group that corresponds with the timing of the subsequent birth for this group (see Appendix

[17] While the distinction between independence and the exclusion restriction is not clearly seen in the moment condition based on Eq. 1 it is helpful to separate these two when discussing the intuition behind the research design (Angrist and Pischke, 2009, p. 153).

[18] Note that across the groups, events relevant to labour market and health outcomes will occur before the due date. For the pregnancy loss samples, the actual pregnancy loss will occur during the three quarters preceding the (approximate) due date. For the live birth group, many will go on maternity leave before the quarter of birth—i.e., a woman who gives birth in early April will have the birth recorded in quarter two but may begin maternity leave in February or March during quarter one. For the health outcomes, it is important to note that we might expect all three groups to have more frequent doctor visits while pregnant, thereby increasing the probability of a diagnosis or prescription.

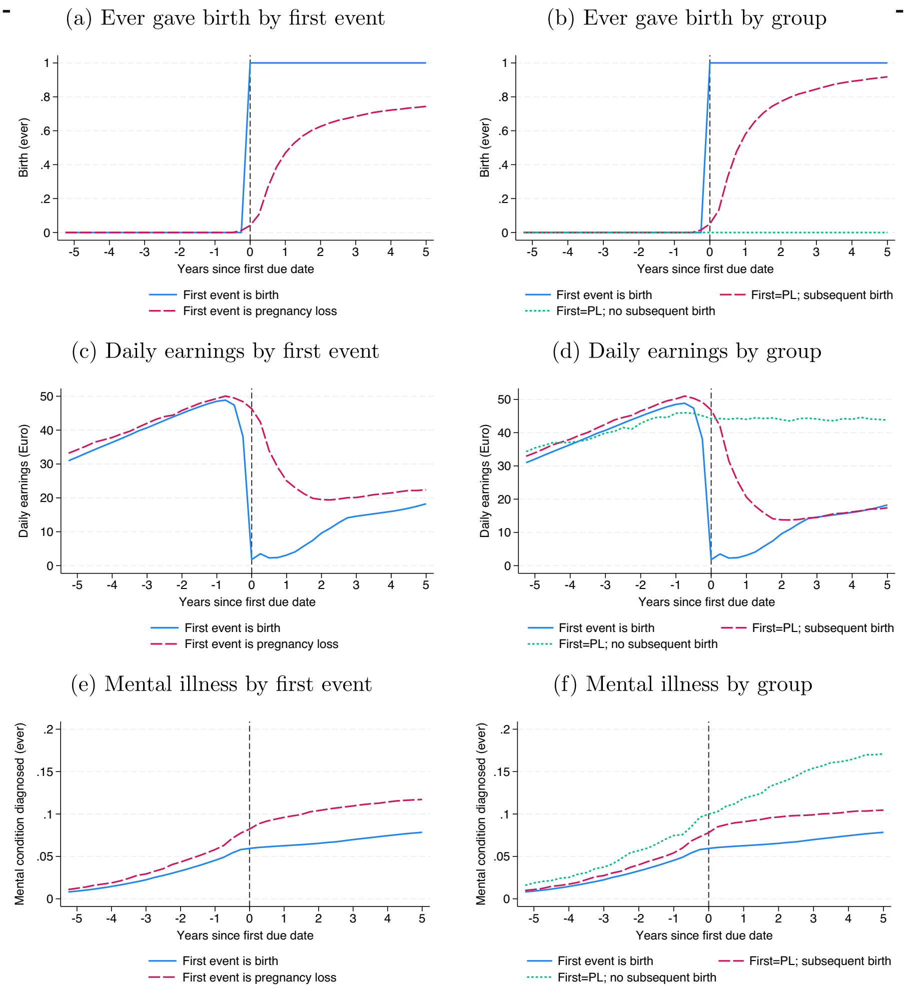

Fig. 3 Births, earnings and mental illness by years to first due date. *Notes:* Panels (**a**) and (**b**): probability of ever having given birth (first stage). Panels (**c**) and (**d**): daily earnings (in 2000 euro; quarterly on 15th of the second month; zeroes included). Panels (**e**) and (**f**): probability of ever having been diagnosed with a mental or behavioural condition (e.g., depression) up to quarter t in the health insurance data. We plot averages by time to first due date splitting the sample by whether the first observed event in the health and social security data is a birth or pregnancy loss (Panels (**a**), (**c**), (**e**)), and further whether the pregnancy loss is eventually followed by a birth up to 2017 (Panels (**b**), (**d**), (**f**)). The first due date is set to the date of live birth for women whose first event is a live birth and approximated by two quarters after the first pregnancy loss otherwise. "First=PL" means first observed event is pregnancy loss

Fig. A8). Taken on its own, Panel (c) is encouraging for the IV research design given the similar earnings before pregnancy. That is, it appears that suffering a pregnancy loss is more-or-less unrelated to factors driving earnings before pregnancy, which is suggestive evidence in support of instrument independence. However, in Panel (a), the pregnancy loss group includes those women that we never observe with a birth (group 3).

In Panels (b), (d), and (f), we further divide the pregnancy loss group into those that we observe with a birth (up to 31st December 2017) and those that we do not. Among women who suffer a pregnancy loss and are observed with a subsequent birth, more than 90% have given birth within 5 years after the loss. In Panel (d), we note that there are now clear differences in daily earnings across the three groups before pregnancy. The women who suffer a pregnancy loss and are observed with a subsequent birth have the highest pre-pregnancy earnings on average, followed by the live birth group. The women who suffer a pregnancy loss but are not observed with a subsequent birth have the lowest pre-pregnancy daily earnings. Importantly, this is the group that is excluded from the analysis when considering only those seen with a child.[19] While not an issue for the independence assumption (e.g., this would not impact the validity of the reduced form relationship between pregnancy loss and labour market outcomes), this post-loss selection could still lead to problems for the IV research design to identify the returns to delaying age at first birth. If these higher pre-pregnancy earnings for group 2 compared to group 1 reflect unobserved ability or other factors that lead to consistently higher earnings, then the comparison between the two groups in cumulative and later career outcomes underlying the pregnancy loss IV approach may overestimate the returns to delay.

Figure 3 also provides descriptive evidence that pregnancy loss may have negative effects on labour market outcomes. We see declines in earnings post pregnancy loss even for those women who do not have a subsequent child—suggesting possible violations of the exclusion restriction in this setting.

Panels (e) and (f) of Fig. 3 provide similar plots for whether an individual has ever been diagnosed with a mental health condition by time t. In Panel (e), we see that the mental health diagnoses are more likely before pregnancy for the pregnancy loss group (groups 2 and 3). This is in line with the evidence in the medical literature that prenatal depression (both treated and untreated) is associated with adverse fetal outcomes and also with increased risk of pregnancy loss (Bonari et al. 2004; Field et al. 2006; Almeida et al. 2016). While this higher incidence of mental health diagnosis before pregnancy is not associated with lower earnings (i.e., we do not see lower pre-pregnancy earnings for the pregnancy loss group in Panel (c)), this could still lead to a violation of the independence assumption if pre-pregnancy mental health changes how labour market outcomes respond to later ages at first birth (e.g., if better mental health allows women to take advantage of further delay through more successful job search). We also see an increase in diagnoses after the pregnancy loss, consistent with the possible negative mental health effects of the loss implying a possible violation of the exclusion restriction. In Panel (f), we see that the rate of mental healthdiagnoses

[19] The fact that the weighted average of the two pregnancy loss groups—where the weights are respective sample sizes—depicted in Panel (c) is very similar to the no loss group average before pregnancy is consistent with the notion that the live birth group will be a mixture of the two types of women—those that *if* they suffered a pregnancy loss they would be observed with a subsequent child in our data and those that would not.

is the highest in the group never observed with a birth (group 3), further reinforcing the possibility that this group is negatively selected.[20]

If unobservable factors driving labour market outcomes are negatively correlated with the probability of a loss and there are negative mental health effects of pregnancy loss, then we worry that the "no loss" instrument is positively correlated with unobservables (i.e., $\rho_{zu} > 0$ as those that do not suffer a loss with $z_i = 1$ are expected to have characteristics more conducive for success in the labour market and less likely to face the associated mental health shocks.). Note that as constructed, the covariance between the "no loss" instrument and age at first birth is expected to be negative ($\sigma_{Az} < 0$)—those that do not suffer a loss with $z_i = 1$ will have an earlier age at first birth on average as they do not experience the delay associated with loss—so that, ignoring the post-loss selection issue, the IV estimate will be an underestimate of the return to delayed motherhood.[21] Alone, this suggests that the IV estimates may be biased downward. As mentioned previously, Fig. 3 Panels (d) and (f) also provide evidence of post-pregnancy loss selection into being observed having a child, with around 20 percent of the women whose first observed pregnancy ended in a loss are never observed having a child. Importantly, these women had lower earnings (Panel (d)) and worse health pre-pregnancy (Panel (f)) compared to the women who are observed giving birth after a pregnancy loss. This suggests a positive selection into the pregnancy loss sample used to estimate the effects of delaying motherhood. In isolation, this selection bias would go in the opposite direction and lead the IV estimate to be too large. We now discuss our estimation approach addressing these two issues.

5.2 IV based approaches

Given concerns over both the least squares estimates and the pregnancy loss instrument, we shift attention from point estimation to producing a range of estimated returns to delay that account for the main issues outlined above under different assumptions. To do so, we adopt the estimators from Nevo and Rosen (2012).[22] Formally, under the assumptions outlined below, the Nevo and Rosen (2012) approach yields sharp bounds on the returns to delay. While the assumptions for the sharp bound interpretation are plausible in our setting, we also view this as a robustness exercise to explore whether

[20] The patterns in Fig. 3 are further supported by Appendix Table A6 where we also compare characteristics before the event and formally test for differences across the three groups. For women in group 1, we look at 1.5 years before the first birth; for groups 2 and 3, we take into account that pregnancy loss most commonly occurs during the first 12 weeks of pregnancy and look at outcomes 1 year before the approximate first due date (about 5 quarters before the pregnancy loss event). Women with a birth after pregnancy loss are more likely to be employed and have higher daily earnings compared to women who do not suffer from a pregnancy loss, whereas we find the opposite among women who do not have a birth after pregnancy loss. Average age at labour market entry is around 18 for groups 1 and 2 and age 20 for group 3. Total health-care utilization is also higher before the event among groups 2 and 3 compared to group 1. Overall, the descriptive statistics provide suggestive evidence that delaying first birth due to pregnancy loss is associated with better labour market outcomes and that both pregnancy loss and having a child after a loss are not entirely random.

[21] To see this, note that, ignoring other covariates, $plim \beta_z^{IV} = \beta + \sigma_{zu}/\sigma_{Az} < \beta$, since $\sigma_{zu} > 0$ and $\sigma_{Az} < 0$.

[22] Demir et al. (2020) also use the approach of Nevo and Rosen (2012) when analyzing the effect of motherhood timing on breastfeeding duration.

the qualitative conclusion from Sect. 4, that delay has only a small effect on the shape of child penalties, is likely to be overturned when the bias of the OLS estimator is accounted for.

Given our setting, we can directly apply the results from Nevo and Rosen (2012) to the estimation of β in Eq. 1. We first partial out the effect of $\mathbf{x}$. Denote the residuals from regressions of y, A, and z on $\mathbf{x}$ by $\widetilde{y}$, $\widetilde{A}$, and $\widetilde{z}$, respectively. Now we can express Eq. 1 using the residualized variables and omitting $\mathbf{x}$: $\widetilde{y}_i = \beta\widetilde{A}_i + u_i$. The two key NR assumptions are[23]

$$\rho_{\widetilde{A}u}\rho_{\widetilde{z}u} \geq 0 \tag{6}$$

$$|\rho_{\widetilde{A}u}| \geq |\rho_{\widetilde{z}u}|. \tag{7}$$

The assumption in Eq. 6 states that the direction of the correlation with the unobserved component is the same for both age at first birth and the No Loss instrument. As argued above, we expect $\rho_{\widetilde{A}u}$, the correlation between age at first birth and the unobserved error term to be positive due to selection into later fertility by those with higher unobserved earnings potential (see Footnote 11) and we expect $\rho_{\widetilde{z}u}$, the correlation between the no loss instrument and the unobserved error term, to also be positive as not suffering a loss is associated with characteristics expected to improve labour market outcomes. The assumption in Eq. 7 states that the No Loss instrument is less endogenous than age at first birth. This assumption seems plausible in our setting. For instance, Miller (2011) notes that over 90% of pregnancy losses are due to anomalies affecting the development of the fetus rather than, say, behavioural factors. Further, the negative mental health effects we documented in Sect. 5.1 only occur for roughly $10-20\%$ of women suffering a pregnancy loss. Taken together, it seems reasonable that the No Loss indicator will be less endogenous than age at first birth, which is related to earnings through a host of channels including positive selection and reverse causality.

Nevo and Rosen (2012) show that if $\sigma_{\widetilde{A}\widetilde{z}} < 0$ (not suffering a loss is associated with earlier ages at first birth), $\sigma_{\widetilde{A}u} > 0$ (choices to delay fertility are positively associated with unobserved factors that increase earnings), and $\sigma_{\widetilde{z}u} > 0$ (not suffering a loss is positively associated with unobserved factors that increase earnings), then β falls between β_{IV} and β_{NR}. Here, β_{IV} is simply our IV estimator using the No Loss instrument. The NR Upper Bound, which we call the "NR Adjusted estimator" (β_{NR}), is the IV estimator using the following instrument: $v = \sigma_{\widetilde{A}}\widetilde{z} - \sigma_{\widetilde{z}}\widetilde{A}$.[24] Note that,

[23] See Nevo and Rosen (2012) for a full discussion of assumptions. Formally, the observations $(y_i, A_i, \mathbf{x}_i, z_i), i = 1, \ldots, n$ have to be stationary and weakly dependent. In a cross-sectional setting like ours, this is usually stated as the "random sampling" or i.i.d assumption. Secondly, $E[\mathbf{x}'u] = \mathbf{0}$, meaning additional controls are exogenous—just as in our OLS and IV estimators above. Thirdly, $Rank(E[(z, \mathbf{x})'(z, \mathbf{x})]) = Rank(E[(A, \mathbf{x})'(A, \mathbf{x})]) = Rank(E[(z, \mathbf{x})'(A, \mathbf{x})]) = k$ is the standard rank condition for OLS and IV estimators.

[24] Intuitively, β_{NR} would be a valid instrument if the No Loss instrument is just as endogenous as age at first birth, that is if $\rho_{\widetilde{A}u} = \rho_{\widetilde{z}u}$ and Assumption (7) holds with equality. To see this, note that

$$E[vu] = E[(\sigma_{\widetilde{A}}\widetilde{z} - \sigma_{\widetilde{z}}\widetilde{A})u] = \sigma_{\widetilde{A}}\sigma_{\widetilde{z}u} - \sigma_{\widetilde{z}}\sigma_{\widetilde{A}u} = \rho_{\widetilde{z}u}\sigma_{\widetilde{A}}\sigma_{\widetilde{z}}\sigma_u - \rho_{\widetilde{A}u}\sigma_{\widetilde{A}}\sigma_{\widetilde{z}}\sigma_u = \left(\rho_{\widetilde{A}u} - \rho_{\widetilde{z}u}\right)\left(\sigma_{\widetilde{A}}\sigma_{\widetilde{z}}\sigma_u\right)$$

Clearly, if $\rho_{\widetilde{A}u} = \rho_{\widetilde{z}u}$, then $E[vu] = 0$ yielding a valid moment condition for identifying β.

β_{NR} will be less than β_{OLS} under the NR assumptions, thereby partially reducing the upward bias present in the OLS estimator.[25] Importantly, β_{NR} should still be interpreted as an overestimate or upper bound on the returns to delay—indicating how much smaller the returns would be if pregnancy loss was just as endogenous as age at first birth (i.e., (7) holds with equality). If one is willing to take the sharp bounding interpretation under the NR assumptions, then the NR bounds given by β_{IV} and β_{NR} account for any factor that may lead to a negative correlation between pregnancy loss and labour market outcomes, including the possible negative effects of pregnancy loss on mental health and the negative selection into suffering a loss while also ruling out larger returns driven by the endogeneity of age at first birth.

5.3 Accounting for post-pregnancy loss selection

We now consider the positive selection of women observed having a child after suffering a pregnancy loss. This selection could take two forms: (1) women deciding not to or being unable to have children after a pregnancy loss or (2) delaying their next pregnancy long enough that we do not observe the birth in the window of our data. Note that (2) may be less of a concern here as we observe women for a minimum of 10 years after the pregnancy loss. We cast the problem as a missing data issue where we do not observe an age at first birth or labour market outcomes with a child for the women that have a pregnancy loss and no subsequent child and apply an Inverse Probability Weighting (IPW) approach. The following discussion draws heavily on the treatment of IPW in Wooldridge (2010).

To start, denote the vector of data for an individual by $\mathbf{w}_i = (y_i, A_i, \mathbf{x}_i, z_i)$ and let s_i be a binary indicator of selection into motherhood, where $s_i = 0$ indicates women with a pregnancy loss and no subsequent observed birth and $s_i = 1$ indicates all other women in our sample—both the no pregnancy loss women and those who suffer a pregnancy loss but are observed with a subsequent birth. Intuitively, the key requirement for the IPW approach is that we observe a set of variables, here denoted by $\mathbf{r_i}$, that are such good predictors of selection that once we condition on them the selection is effectively random with respect to a woman's labour market outcomes (y_i). Formally we must assume that $P(s_i = 1|\mathbf{w}_i, \mathbf{r}_i) = P(s_i = 1|\mathbf{r}_i)$.[26]

While this selection on observables assumption is strong, it is arguably much weaker than the assumption that selection is independent of all the observable factors, i.e., that $P(s_i = 1|\mathbf{w}_i) = P(s_i = 1)$, that has been implicitly made in prior work using the pregnancy loss instrument. Indeed, the descriptive results from Sect. 5.1 suggest that

[25] Intuitively, since the new NR instrument v is a standard deviation weighted linear combination of the no loss instrument and the endogenous age a first birth—the resulting IV estimator will be a weighted average of the negatively biased IV estimator using the no loss instrument and the positively biased OLS estimator—thereby moving closer to the the lower bound than the OLS estimator and yielding tighter bounds.

[26] See Appendix D for a discussion of the full set of assumptions required.

this implicit assumption is likely violated. In practice, $\mathbf{r_i}$ will include a rich set of pre-event labour market—effectively pre-pregnancy lags of our key outcome variables—and health variables. By including pre-pregnancy labour market variables, we can control for many of the unobservable factors that may affect post-pregnancy labour market outcomes when accounting for the post-loss selection. This is conceptually similar to the arguments in favor of using pre-treatment dependent variables when creating synthetic control groups (Abadie et al. 2010). If the selection is driven by factors that lead to generally higher earnings for some women (for example, ability, costs of effort, or preferences for work), then this assumption is quite reasonable. As we show in Appendix D, reweighting based on our rich set of pre-event information does a good job of balancing the expected daily earnings and mental health diagnoses between the pregnancy loss subsequent birth group and the entire pregnancy loss group before the loss occurs.

The main concern for our approach would be selection on unobservables that lead our β_{IV} estimates to be too high. Therefore, the main threats are cases where individuals with similar pre-pregnancy earnings and health systematically select out of our motherhood sample for reasons that are related to their hypothetical post-birth earnings being lower, on average, than the women we do observe with a subsequent birth. For instance, perhaps women who suffer a particularly large health shock with the pregnancy loss are more likely to select out of our sample. If this were the case, we might expect these women to have lower earnings if they had a subsequent child than the women that were observationally similar before the first pregnancy. This would lead our IPW β_{IV} estimates to be biased upward since the observationally equivalent women we do see with a child after a pregnancy loss would have earnings that are too high.

To provide suggestive evidence on this possibility, we check the robustness of the Probit model used to estimate the IPW (discussed in Appendix D) to including the health care utilization (expenditure covered by the insurance fund) at the time of the loss as a measure of the severity of the health shock. The estimated marginal effect of contemporaneous utilization on selection is small and positive rather than negative. Further, including contemporaneous utilization does little to improve our ability to explain the selection, with the R^2 only increasing slightly from 0.1659 to 0.1677. Finally, we note that the IPW β_{IV} estimates are always lower than the unweighted estimates (see Appendix Table A9).[27]

6 Results

We present our results on the impact of age at first birth on labour market outcomes in two steps. First, we show how delaying motherhood impacts post-birth labour market

[27] Our IPW approach aims to recover the effect of delay for the population of women who would be observed having a pregnancy in our data—those with an observed live birth or pregnancy loss. See Appendix D for a discussion of alternative weighting approaches for other populations of interest.

outcomes by focusing on the period 5 years after birth. As in Sect. 4, we use the child penalty framework to provide standardized estimates on changes in the penalty. Second, we look at the effect of delay on cumulative outcomes because even if returns only come from delaying a constant relative penalty, they can significantly affect career outcomes and are thus relevant for overall gender inequality in the labour market.

6.1 Medium run effects of birth timing

We estimate the effect of age at first birth on a few key dependent variables measured 5 years after first birth using Eq. 1. Table 2 shows the results.[28] Consider annual earnings 5 years after birth. The unweighted IV and OLS estimates imply effects on annual earnings of €70.9 to €307.6 per year of delay, corresponding to a 1.1 to 4.6% increase. With inverse probability weights (IPW) applied, the estimates shift to the left, and we estimate that 1 year of delay leads to −0.1 to 4.5% higher earnings 5 years after first birth. Based on the IV estimates, we thus cannot reject that delay has zero effect on earnings after birth.[29] The OLS estimate includes substantial returns to delay; however, the NR estimate reduces this to €237—or 3.6%—higher annual earnings per year of delay. Again, β_{NR} can be viewed as an upper bound on the return to delay—removing some of the upward bias of the OLS estimator under the NR assumption that the instrument is less endogenous than age at first birth.

Next, we decompose annual earnings into days of employment (including zero) and average earnings per day if employed, shown in the second and third blocks of Table 2. The estimated $0.9-2.5\%$ more days of employment and $0.4-1.2\%$ higher earnings if employed (considering the IPW estimates) show that both margins potentially matter for delay, with relatively more weight on days of employment over daily earnings. However, based on the IV, we cannot statistically reject that delay does not lead to more work or higher daily earnings, in line with the results on annual earnings.

Importantly, when we look at daily earnings, the NR estimate (β_{NR}) on the weighted sample is even smaller than the unweighted IV estimate. The typical view in the literature is that the (unweighted) pregnancy loss IV underestimates the return to delay while OLS overestimates. Here we see that this is not necessarily the case if one ignores the selection problem from women opting not to have children after a

[28] All outcomes are measured in the calendar year 5 years after first birth (see notes to Table 2 for variable definitions). In order to have outcome measures 5 years after first birth by the end of our data window in 2017, we re-classify 129 out of 4561 women (about 3%) in group 2 (pregnancy loss and subsequent birth) who have not given birth by the end of 2012 for this analysis. We, instead, assign these women to group 3 before estimating the IPW weights. Excluding them entirely leads to almost identical estimates. "IPW" refers to the inverse probability weights from Sect. 5.3, which account for selection out of motherhood. Including the selection variables as additional controls in Eq. 1 leads to almost identical estimates.

[29] Here, we report standard errors that ignore the first stage estimation of the probability weights, which leads to more conservative inference (i.e., larger standard errors) than if we took account of this estimation (Wooldridge 2002 and Wooldridge 2010, pages 500–502, 824–826). In our setting, the loss of precision from ignoring the first stage estimation is particularly small as only 8.9 percent of our estimation sample (the no loss, subsequent birth group) has an estimated weight—while the other probabilities are treated as known ($\hat{p}(\mathbf{r}_i) = 1$ for the no loss group).

Table 2 Effect of age at first birth on outcomes 5 years after first birth

Dependent variable	Mean	Unweighted		IPW		
		$\hat{\beta}_{\text{OLS}}$	$\hat{\beta}_{\text{IV}}$	$\hat{\beta}_{\text{OLS}}$	$\hat{\beta}_{\text{NR}}$	$\hat{\beta}_{\text{IV}}$
	(1)	(2)	(3)	(4)	(5)	(6)
Annual earnings (€)	6680.5	307.6	70.9	300.5	237.3	−5.2
		(16.1)	(69.1)	(16.6)	(22.6)	(69.6)
in %		*4.61*	*1.06*	*4.50*	*3.55*	*−0.08*
[95% CI]		*[4.13, 5.08]*	*[−0.97, 3.09]*	*[4.01, 4.98]*	*[2.89, 4.22]*	*[−2.12, 1.96]*
Days employed	206.438	5.125	2.960	5.055	4.399	1.884
		(0.305)	(1.355)	(1.339)	(0.312)	(0.425)
in %		*2.48*	*1.43*	*2.45*	*2.13*	*0.91*
[95% CI]		*[2.19, 2.77]*	*[0.15, 2.72]*	*[2.15, 2.74]*	*[1.73, 2.53]*	*[−0.36, 2.18]*
Daily earnings (€)	38.304	0.485	0.424	0.460	0.398	0.169
		(0.166)	(0.677)	(0.161)	(0.165)	(0.575)
in %		*1.27*	*1.11*	*1.20*	*1.04*	*0.44*
[95% CI]		*[0.42, 2.12]*	*[−2.36, 4.57]*	*[0.38, 2.03]*	*[0.19, 1.88]*	*[−2.50, 3.38]*
Employed	0.714	0.009	0.006	0.009	0.007	0.003
		(0.001)	(0.004)	(0.001)	(0.001)	(0.004)
in %		*1.23*	*0.80*	*1.21*	*1.04*	*0.39*
[95% CI]		*[1.00, 1.46]*	*[−0.23, 1.84]*	*[0.97, 1.45]*	*[0.72, 1.37]*	*[−0.64, 1.43]*
White-collar job	0.627	0.021	0.001	0.020	0.015	−0.003
		(0.001)	(0.004)	(0.001)	(0.001)	(0.005)
in %		*3.32*	*0.23*	*3.20*	*2.41*	*−0.53*
[95% CI]		*[3.00, 3.65]*	*[−1.18, 1.63]*	*[2.86, 3.53]*	*[1.95, 2.87]*	*[−1.95, 0.89]*
Back to pre-birth firm	0.499	0.018	0.015	0.017	0.016	0.010
		(0.001)	(0.004)	(0.001)	(0.001)	(0.004)
in %		*3.62*	*3.07*	*3.51*	*3.21*	*2.09*
[95% CI]		*[3.26, 3.99]*	*[1.44, 4.70]*	*[3.13, 3.88]*	*[2.70, 3.72]*	*[0.49, 3.69]*

Notes: this table shows the estimated return to delay on outcomes 5 years after first birth ($\hat{\beta}$) based on Eq. 1 using OLS or IV. Columns (2) and (3) show unweighted regressions, (4) to (6) report results using inverse probability weights (see Sect. 5.3). Column (5) shows the Nevo-Rosen estimates, weighting the IV and OLS estimates to give an upper-bound estimate under direction-of-bias assumptions (see Sect. 5.2). Controls in all columns: dummies for the mother's year of birth cohort and her age at labour market entry. Annual earnings are total social security earnings in the calendar year 5 years after first birth. Women who are not employed in that year are included with zero earnings. Days employed is the total number of days of dependent employment in the calendar year 5 years after first birth. To get daily earnings, we divide annual earnings by days of employment. Employed is a binary variable capturing whether the woman was employed at any point in the calendar year 5 years after first birth. White-collar job is 1 if the woman had any days of white-collar employment 5 years after first birth, 0 otherwise. Back to pre-birth firm is a binary variable capturing whether the woman worked again at her pre-birth firm (any of the 4 quarters before birth) in the 5 years after first birth. To calculate the weights and regressions for this table, we re-classify from group 2 into group 3 women whose first birth was after 2012 because we don't observe their outcomes 5 years after first birth ($n = 129$). Robust standard errors in parentheses, 95% confidence intervals in brackets pregnancy loss. Both IV and OLS, if not weighted, are too large compared to the weighted IV and NR estimators.

We estimate delay to increase the probability of working in a white-collar occupation by up to 3.2% (considering the IPW estimates), which explains part of the increase in daily and annual earnings. Delay might enable women to climb the job ladder and lock in a better job before going on (job protected) parental leave.[30] In line with this result, we find a statistically significant increase in the probability of going back to the pre-birth job within 5 years. This could capture success in climbing the ladder and locking in a better match, but could also reflect that delay facilitates moving to jobs more convenient for combining work and motherhood. Such non-pecuniary benefits could explain why the IV estimates are significantly positive for returning to the pre-birth job while earnings are not affected on average.[31]

Under the NR assumptions, columns 4 and 5 of Table 2 represent bounds on the returns to delay that account for many plausible sources of endogeneity for both the pregnancy loss instrument and age at first birth. The IV estimates are consistent with delay having no causal effect on labour market outcomes, which is different from previous papers (e.g., Miller, 2011) that found significant effects using the pregnancy loss instrument. Part of the reason for this difference could be that our loss variable is based on medical diagnoses by doctors, whereas previous papers used retrospective survey information. We could thus be capturing more severe losses, and e.g., negative mental health effects could be more pronounced among our sample of women. Our NR-adjusted estimates are generally comparable to previous papers and consistent with delay affecting outcomes by strengthening labour market attachment, finding a better job match, and lowering human capital depreciation, the channels analysed by e.g., Erosa et al. (2002) and Adda et al. (2017).

In Table 3 we compare the OLS, IPW NR, and IPW IV estimates of the level effects (columns 1, 2, and 3 reproduced from Table 2) to the mechanical effects of delaying a fixed penalty (column 4). In column 5, we show the average child penalty 5 years after first birth for reference. Since the level effects estimated by IV are already small in magnitude and not statistically different from zero, we focus our discussion on the nontrivial level effects of the estimated β_{NR} to gauge how important delay could be for shifting the child penalty once we have accounted for some of the upward bias in the OLS estimator. Consider annual earnings: the level effect of delay is €237.3 (3.6%)

[30] See Lalive et al. (2014) for details. Although the duration during which mothers received child benefits was changed several times after 1990, job protection was held constant at 2 years.

[31] Appendix Table A8 includes several robustness checks for the main results. First, excluding age at labour market entry dummies does not affect the NR or IV estimates much. The labour market entry age dummies serve both as a proxy for education and as a control for switching from industries that are not covered by the Upper Austria health insurance fund (e.g., public servants) later in the career. Therefore, the robustness to its inclusion is important for ensuring that we are not confounding education decisions with birth timing in our main results. Second, including the pre-birth variables we use in our probit selection equation does lead to significantly smaller OLS estimates; however, the IV estimates are barely affected. Conceptually, these variables control for some—but, likely not all—of the endogeneity of birth timing for the OLS-based approaches, lowering the estimated returns. The relative robustness of the IV-based approaches to including these additional pre-birth controls is evidence that the selection on observables is less strong for the pregnancy loss instrument than for age at first birth. Importantly, across both robustness checks the estimated returns fall within a similar range to our main estimates, thereby supporting the conclusions drawn.

per year of delay, while the average mechanical effect from moving along the age-earnings profile would lead to an increase of €106.9. That is, nearly half the effect of delay can be attributed to moving along the age profile. This leaves room for a small decrease in the child penalty due to delay. To quantify how big this effect on the penalty would be, we subtract the mechanical effect from the estimated effect and then divide by counterfactual earnings. This gives us effects in terms of child penalty units, shown in the last three columns. For annual earnings, we estimate that the penalty would decrease by 0.7 percentage points (about 1%) per year of delay. Considering the level effect of 3.6% on earnings after birth based on the NR estimate, the result reported in the second-to-last column of Table 3 shows that the effects on child penalties are small. (see column 5 of Table 2)—which is likely larger than the true causal effect because of bias due to e.g., unobserved heterogeneity. The implied effects on the child penalty of the IV estimates are even smaller (see the last column of Table 3), reinforcing the conclusion that there is little evidence that delay drastically changes the child penalty.[32]

In Fig. 4, we show what these estimates imply for pre- and post-birth predicted earnings. The figure plots predicted earnings from the child penalty regression and the predicted earnings based on the different estimates for the return to delay. We plot this for the median age at first birth (27) and fixing the year to 2003. Next, we add the lower and upper bound predicted earnings when delaying from age 27 to age 28 (NR IV and NR OLS, respectively). Finally, we also add the predicted earnings when delaying the average child penalty by 1 year. The figure illustrates that the pre-birth effects of delay are large because the child penalty hits later. We return to this when we look at the effect of delay on total earnings below. After birth, however, effects are small over the entire 10-year period whether we look at the lower or upper bound or just delaying the average penalty.

In Appendix Table A7, we repeat the exercise shown in Table 3 for different event times. The child penalty estimates for our sample clearly show strong dynamic effects. The maximum earnings reduction is in the first year after birth, when most mothers are on parental leave, and recovers to about −0.54 in year 10 for this sample. With one exception—2 years after birth, when most mothers are still on parental leave—the estimated bounds on the effects on the child penalty, i.e., how much a 1-year delay would change the relative earnings loss from motherhood, include zero. Based on our conservative estimates, we cannot rule out the possibility of no effect. Examining the full trajectory across 0 to 10 years after childbirth reinforces that delaying motherhood

[32] As discussed in Thakral and Tô (2025), the child penalty regression imposes the strong restriction that there are no changing birth cohort effects. We assess the robustness of our results to this assumption by widening the event-time window from −10 to +10 and imposing a linear restriction that the slope between −10 and −5 is flat. In this alternative specification, we include birth cohort dummies and then recover the same −5 to 10 child penalty estimates as in the standard specification. In the Austrian case, adding cohort dummies in this way produces a somewhat larger child penalty (in absolute terms) at each event time after 0; however, our main results are only slightly different when we rely on the cohort specification instead of the baseline regression. When we standardize the level effects of delay into child penalty effects (last three columns), the OLS estimate decreases from 0.1 to 0.09%, the upper bound becomes 0.6% rather than 0.7% under the standard penalty, and the lower bound is −0.4% instead of −0.6%. Overall, although incorporating cohort effects changes the penalty itself, it does not substantially alter our estimated effects of delay.

Table 3 Effects on child penalties: 5 years after first birth

Dependent variable	Level effects			Mechanical effect of 1-year delay	Child penalty	Effects on child penalty (relative to mechanical return)		
	$\hat{\beta}_{\text{OLS}}$	$\hat{\beta}_{\text{NR}}^{IPW}$	$\hat{\beta}_{\text{IV}}^{IPW}$			$\hat{\beta}_{\text{OLS}}$	$\hat{\beta}_{\text{NR}}^{IPW}$	$\hat{\beta}_{\text{IV}}^{IPW}$
Annual earnings (€)	307.6	237.3	−5.2	106.9	−0.669	0.010	0.007	−0.006
	(16.1)	(22.6)	(69.6)			(0.001)	(0.001)	(0.003)
Days employed	5.125	4.399	1.884	1.428	−0.386	0.011	0.009	0.001
	(0.305)	(0.425)	(1.339)			(0.001)	(0.001)	(0.004)
Daily earnings (€)	0.485	0.398	0.169	0.744	−0.403	−0.004	−0.005	−0.009
	(0.166)	(0.165)	(0.575)			(0.003)	(0.003)	(0.009)
Employed	0.009	0.007	0.003	0.001	−0.287	0.007	0.006	0.001
	(0.001)	(0.001)	(0.004)			(0.001)	(0.001)	(0.004)
White-collar job	0.021	0.015	−0.003	0.008	−0.221	0.016	0.009	−0.014
	(0.001)	(0.001)	(0.005)			(0.001)	(0.002)	(0.006)

Notes: see notes to Table 2 for explanations of variables and methods. This table shows the OLS, Nevo-Rosen (NR), and IV estimates. The Nevo-Rosen and IV estimates are from the specifications using IPW. The last three columns show the coefficients minus the mechanical effect divided by counterfactual earnings. This translates the level effects into effects on the child penalty over and above what we would expect to see from mechanically being 1 year older when delaying first birth. Robust standard errors in parentheses

does not seem to substantially alter the shape of the child penalty. The gains we see in career earnings thus mainly come from delaying the incidence of the penalty itself, rather than a reduction in its magnitude.

Decomposing annual earnings into days employed and earnings per day, Table 3 shows that along both margins the mechanical effect of delay plays an important role. About a third of the estimated NR effect on days employed is likely due to just being 1 year older, whereas all of the effect on daily earnings seems to come from age, although this coefficient is somewhat imprecisely estimated. This is not the case when we look at whether women are employed at any point in the calendar year 5 years after birth. The mechanical effect is only 0.001, or one-seventh of the estimated NR estimate. We lack precision to do a precise decomposition, but the evidence is at least consistent with the mechanical effect of delay operating mainly through earnings if employed, and not employment itself. In terms of white-collar employment, the mechanical effect is again close to half the estimated NR adjusted effect. Occupational upgrading is one important channel through which earnings could be higher due to delay, and the majority of the estimated effect of 1.5 percentage points is driven by moving along the average age profile.

6.2 Heterogeneity in the returns to delay

The effects on labour market outcomes shown in Tables 2 and 3 average the heterogeneous effects of delay across women. The returns to delay for women who have little

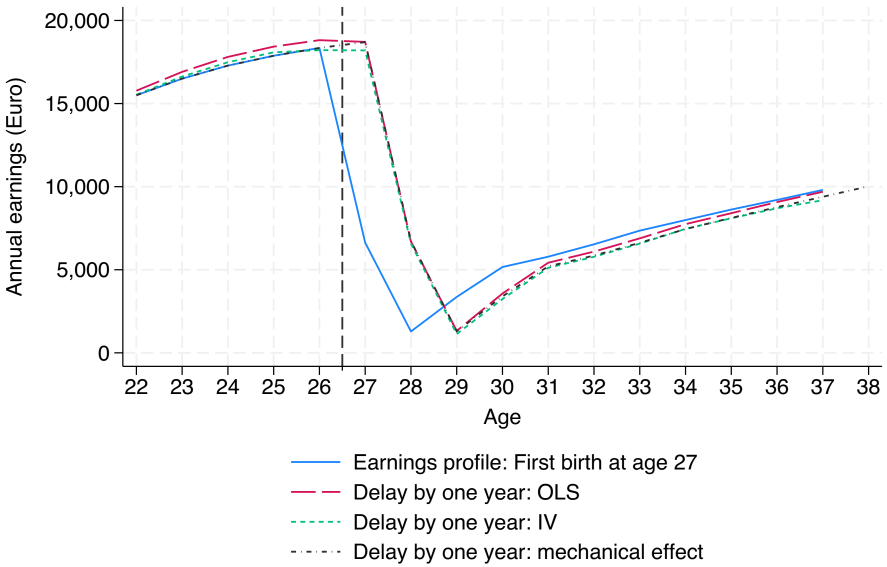

Fig. 4 Earnings effect of delaying first birth. *Notes:* the earnings profile (blue line) traces out $E\left[\hat{Y}\middle|t\right]$, in this case predicted annual earnings, after estimation of Eq. 2 for the median age at first birth (27) and fixing the median calendar year (2003). Zeroes (not employed) are included in the regression. Three separate lines then trace out the predicted effects on the level of earnings if first birth is delayed from age 27 to age 28: The lower and upper bounds on the return to delay (NR IV and NR OLS, respectively) and the mechanical effect from delaying the average penalty. See estimates and notes in Table 3 or Appendix Table A7 for the full dynamic effects

education or in low-skill jobs might be lower, having made decisions consistent with the "mommy track". Depending on the size of this group, it could mean that the average effect of delay looks small relative to the child penalty, even if there are substantial changes to child penalties for subgroups. To shed light on this, we split our sample into subgroups according to age at labour market entry (a proxy for education),[33] and whether women worked in a blue or white collar occupation before birth. Results are shown in Table 4.

The IV estimates continue to be relatively small and insignificant across specifications. We do not find evidence for significant earnings returns in any subgroups, using the pregnancy loss instrument. However, the IPW NR estimates reveal interesting heterogeneity. Consider age at labour market entry. In the Austrian education system, entering the labour market at age 17 or above is strongly correlated with being on the academic high school track, instead of doing an apprenticeship. The return to delay for those who enter the labour market later is €330.7 per year of delay. This corresponds to a 4.1% increase in annual earnings 5 years after first birth, compared to 3.7% for those who enter early. These are sizable level estimates but partly capture the effect of being mechanically older due to delaying first birth.

[33] We do not analyse heterogeneity by education because although education is observed in the birth register, it is missing for 30% of mothers.

Table 4 Heterogeneity in the effect of delay on earnings 5 years after first birth

Sample	Level effects			Mechanical effect of 1-year delay	Child penalty	Effect on child penalty (rel. to mechanical effect)		
	$\hat{\beta}_{\text{OLS}}$	$\hat{\beta}^{IPW}_{\text{NR}}$	$\hat{\beta}^{IPW}_{\text{IV}}$			$\hat{\beta}_{\text{OLS}}$	$\hat{\beta}^{IPW}_{\text{NR}}$	$\hat{\beta}^{IPW}_{\text{IV}}$
A. By age at labour market entry								
Age first job ≤ 17	239.5	207.0	96.0	61.2	−0.698	0.010	0.008	0.002
	(17.9)	(25.3)	(77.4)			(0.001)	(0.001)	(0.004)
Age first job > 17	460.6	330.7	−121.0	147.8	−0.631	0.014	0.008	−0.012
	(29.7)	(40.3)	(124.4)			(0.001)	(0.002)	(0.006)
B. By occupation before birth								
Blue collar	10.4	21.8	27.4	12.4	−0.619	−0.000	0.001	0.001
	(24.4)	(33.8)	(98.2)			(0.002)	(0.003)	(0.008)
White collar	325.3	257.6	24.0	133.3	−0.672	0.009	0.006	−0.005
	(19.4)	(27.5)	(85.7)			(0.001)	(0.001)	(0.004)

To gauge how much of the return to delay could simply be mechanical, we repeat the exercise we proposed above and calculate child penalties for the two groups to estimate the mechanical effect of delay. Results are shown in columns 4 and 5 of Table 4. Child penalties are estimated to be about 7 percentage points smaller for those who enter after age 17. The smaller penalty and a steeper childless earnings profile (not shown) combined lead to a larger mechanical effect of delay of €147.8 against €61.2 among those who enter before 17. The implied estimates on child penalties net of the mechanical effect are shown in the last three columns. Among both early and late entrants, we find child penalties only decrease by 0.8 percentage points for each year of delay (based on the IPW NR estimates). Even if we consider the OLS estimates (which are likely upward biased), the estimated decrease in the child penalties for each year of delay is less than 1.5 percentage points in both groups. The effects are similarly small if we split the sample by pre-birth occupation.[34]

Overall, the heterogeneity analysis reveals that even among subgroups the effects of delay on earnings after birth are relatively small when compared against the mechanical age effects. Across subsamples, between a third and a half of the level effects of delay are likely just due to being older and more experienced when delaying first birth.

6.3 Motherhood timing and women's careers

We now turn to the effect of motherhood timing on cumulative labour market outcomes. Even though we found small effects of delay on the relative earnings losses after birth,

[34] Even the level effects for blue-collar workers are small and insignificant across all estimators—suggesting little return to delay for this group that may reflect differences in how human capital develops in these occupations. Further, the selection into later first birth appears to operate differently for this group, with OLS estimates that are smaller than the IV (though quantitatively very similar).

delay can still have substantial effects on cumulative earnings because delaying first birth means more time on the childless earnings profile. That is, rather than simply looking at how delay shifts the earnings profile at each age as in Fig. 4, we now consider how delay changes the area under the average age-earnings profile over the entire career. Our goal is to quantify the effects of delay and compare them to the effects of mechanically delaying a fixed penalty. Unfortunately, the fact that the health data we use to measure the pregnancy loss instrument is only available between 1998 and 2007 means that we do not observe the entire career for most women in our main sample. Therefore, if we use our main sample, there will be censoring of lifetime earnings across women from different cohorts and with different ages at first birth. Importantly, this censoring will lead to estimated returns to delay that are too high.[35]

Instead, we define labour market careers to be between age 15 and 55 and leverage the social security data for all of Austria. For women born between 1957 and 1962, we observe earnings over the entire 15–55 age range by the end of the ASSD data in 2017.[36] While we cannot use the IV-based approaches with this sample, we can easily estimate the career returns to delay by OLS. Based on the same arguments as before, we expect the OLS estimator of the returns to delay on career outcomes to be biased upward.

To estimate how much of the return to delay is potentially from delaying a fixed penalty, we use a similar methodology as above to calculate the "mechanical return" of delay. We estimate long-run child penalties and use them to simulate earnings for the women born between 1957 and 1962 depending on when they had their first child. For example, if a woman born in 1960 had a child at age 30, we assign to her the counterfactual average childless earnings up to age 29 and then the average earnings after imposing the fixed (at each event time) child penalty from age 30 onward. Importantly, she will get the same relative penalty as a woman who had her first child, e.g., at age 25. This makes sure we are only comparing the effects on total earnings of delaying a fixed penalty.[37] After this assignment, we calculate total simulated earnings up to age 55 and regress them on age at first birth.

Table 5 provides the OLS estimates of career returns to delay for several labour market outcomes from regressing the actual cumulative outcome and the mechanical simulated cumulative outcome on age at first birth—effectively estimating how the area under the age-outcome profile changes linearly with age at first birth. When we

[35] As an example, consider the women born in 1977 who we observe with earnings up to age 40—for those that have a child at age 25 we will capture 15 years of the child penalty in the cumulative outcome measures while for those that have a child at age 39, we will capture only 1 year of the child penalty—clearly overstating the career returns to delay given that pre-birth earnings track childless women closely. This is an issue in the related work on motherhood timing, with cumulative outcomes often measured well before retirement and in some cases very close to the first birth. For instance, in Miller (2011), cumulative outcomes are measured at age 34 for women who have a first birth up to age 33.

[36] This ignores the retirement dimension since women in Austria start to retire once they reach age 55 (cf Fig. 1).

[37] For this exercise, we estimate child penalties using the data for all of Austria and regressing the outcome on age and year dummies, as well as event time dummies relative to first birth. We estimate this for the years 1972–2017 and limit event times from −5 to 20 years after first birth. The event time coefficients are then divided by the predicted outcome without event time coefficients. We then impute the predicted outcome from these child penalty regressions for the women in our sample. This means they all get the same relative penalty irrespective of when they have their first child.

Table 5 Effect of age at first birth on cumulative outcomes

Dependent variable	$\hat{\beta}_{\mathrm{OLS}}$	Mechanical effect of 1-year delay
Effect of delaying first birth by 1 year in % of baseline		
Total career earnings (€)	2.35	1.39
	(0.03)	
Total career days employed	1.23	0.74
	(0.02)	
Average daily earnings over career (€)	1.49	1.19
	(0.02)	
Total career days in white-collar job	3.29	1.71
	(0.04)	

Notes: each row shows separate regressions for cumulative outcomes for all-Austria sample of mothers born 1957–1962. For annual earnings, days employed, and days in a white-collar job, we sum the annual variables over the career from age 15 to age 55 (when early retirement starts to be available for some). For daily earnings, we calculate average daily earnings (annual earnings divided by days employed per calendar year) over ages 15 to 55. $\hat{\beta}_{OLS}$ is the estimate from a regression of the cumulative outcome on age at first birth with year of birth and age at labour market entry dummies. The mechanical effects use the cumulative variables from estimated child penalties based on cohorts born 1957–1967. For example, we estimate the earnings penalty, predict earnings at every age depending on a woman's age at first birth, sum the predicted earnings over the career, and regress this sum on age at first. The resulting estimate captures the mechanical effect of delaying a fixed penalty since the penalty is the same for all women in relative terms, but applied at different ages at first birth. Robust standard errors in parentheses

consider all earnings up to age 55, the mechanical return is 1.4% compared to an OLS return of 2.4%. The majority of the return to delay, even based on the OLS estimates where selection is likely an important driver, comes from delaying the child penalty and not altering the relative size of the penalty. We also see modest increases in other labour market outcomes—including spending 3.3% more days in white-collar jobs for each year of delay, with a mechanical effect roughly half that size at 1.7%.

Table 5 shows career returns to delay for a cohort of women whose entire pre-retirement labour market outcomes are observed. Interestingly, the estimated returns are substantially smaller than previous results from cumulating outcomes over a shorter age range (e.g., Miller 2011). One main reason that the estimates from the full Austrian sample are much smaller is censoring. In Appendix E, we show cumulative results for the health data sample, where we also have censoring. We show that the estimated returns using the censored career outcomes are substantially larger than those reported in Table 5. When censored to the same ages, the estimated mechanical and OLS returns for all of Austria match the IV and NR adjusted estimates, respectively, for the Upper Austria health data sample very closely. If we impose the NR assumptions and interpret the IV and NR adjusted estimates from the health data sample as bounds on the return to delay, this close match between the two samples and approaches suggests that the mechanical and OLS estimates in Table 5 may serve as reasonable bounds for the career returns to delay.[38]

7 Conclusion

We explore the potential labour market returns to delaying motherhood for women. Using administrative social security and health data from Austria, we are able to document these returns across a wide range of outcomes: earnings, employment, and experience as well as other measures of career outcomes such as their attachment to the pre-birth firm and whether employment is in a white-collar occupation.

Starting with OLS estimates, we find that in Austria, an extra year of delay is associated with 4.6% higher annual earnings, while the child penalty is reduced only by one percentage point with a year of delay. This finding is comparable to our estimates based on data from the USA.

Given the likely upward bias in the OLS estimates, using the rich administrative data from Austria, we use pregnancy loss as an imperfect instrument for age at first birth. Our approach takes account of the possible negative health effects of pregnancy loss, the negative selection into suffering a loss, the fact that the pregnancy loss-induced delay occurs after other key career investments are made, and the positive selection into the sample of mothers for those that suffer a loss.

Our results suggest that much of the observed relationship between age at first birth and labour market outcomes could be due to selection. However, when we focus on the period before and after first birth, we find that even our OLS estimates imply that delay leads to only small changes to how earnings respond to the birth of the first child. Before birth, we see larger effects on earnings; however, both before and after birth our estimates are similar to the mechanical returns we might expect from simply moving along the age-earnings profile an extra year before suffering the average child penalty. Therefore, our results are consistent with much of the cumulative returns to delay coming from the mechanical effect of delaying the realization of the child penalty rather than altering how a woman's career responds to having children.

In terms of gender inequality over the career, our estimates imply that delay is still an important factor, despite the relatively small estimates. Ages at first birth have increased substantially over the last 50 years, and delaying motherhood from around age 23 (1970s Austria) to age 29 (2010s) would imply that changes to birth timing are responsible for an increase of around 14% in average career earnings for women. Although we do not focus on cross-sectional inequality here, our results also imply lower gender inequality in any given cross section since a mechanical delay of the child penalty means more women earnings childless wages at any given point in time.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00148-026-01189-5.

[38] Indeed, the mechanical effect could be interpreted as a lower bound under the assumption that delay does not increase the size of the child penalty (that is, delay does not lower earnings post-birth) since it is calculated under the assumption the child penalty stays fixed—while, as we have argued, we suspect OLS to overestimate the returns.

Acknowledgements We thank editor Terra McKinnish and two anonymous reviewers for helpful comments. We are very grateful to Josef Zweimüller for providing access to the administrative data from the Austrian Social Security Database (ASSD) held by Statistics Austria along with data from the Upper Austria health insurance fund—the Oberösterreichische Gebietskrankenkasse (OÖGKK) at the University of Zurich's "Science Cloud Server" for secure data processing and answering many questions along the way. Anikó Bíró was supported by the National Research, Development and Innovation Office of Hungary (grant number: OTKA K 146309).

Funding Open access funding provided by ELTE Centre for Economic and Regional Studies.

Data Availability The primary data source for our analysis is the confidential administrative Austrian Social Security Database (ASSD), which we accessed and analyzed on a secure server hosted by the University of Zurich (sciencecloud) and supervised by the team of Prof Dr. Josef Zweimüller. The version of the ASSD used in this project contains social security[Pleaseinsertintopreamble]relevant information on employment histories and earnings for Austrian private-sector workers from 1972 to 2017. The data were provided with anonymized individual identifiers that allow individuals to be consistently tracked over time. We also use data from the health insurance fund of Upper Austria (Oberösterreichische Gebietskrankenkasse (OÖGKK)) that were matched to the ASSD data by a third party: amsbg, Hofmühlegasse 3-5, A-1060 Vienna. This is an independent company instructed by the ministry of labour (Arbeitsministerium) to conduct the matching so that we as researchers are not able to identify actual people.
To gain access to the confidential registry data (the "Arbeitsmarktdatenbank" data) used in our main analysis, an application must be submitted to the ASSD through the following link: https://arbeitsmarktdatenbank.at/ Interested researchers should apply for access to the ASSD data as well as the ability to link additional datasets through pseudonymized identifiers. The health insurance data, linked through pseudonymized identifiers, are from the health insurance fund of Upper Austria [Pleaseinsertintopreamble] the Oberösterreichische Gebietskrankenkasse (OÖGKK). https://www.gesundheitskasse.at/cdscontent/?contentid=10007.867331&portal=oegkportal
We also use data from the PSID (Panel Study of Income Dynamics). The PSID data is available upon registration here: https://simba.isr.umich.edu/data/data.aspx

References

Abadie A, Diamond A, Hainmueller J (2010) Synthetic control methods for comparative case studies: estimating the effect of California's tobacco control program. J Am Stat Assoc 105(490):493–505
Adda J, Dustmann C, Stevens K (2017) The career costs of children. J Polit Econ 125(2):293–337
Almeida ND, Basso O, Abrahamowicz M, Gagnon R, Tamblyn R (2016) Risk of miscarriage in women receiving antidepressants in early pregnancy, correcting for induced abortions. Epidemiology 27(4):538–546
Angrist J, Pischke J-S (2009) Mostly harmless econometrics: an empiricists' companion. Princeton Univerisity Press, Princeton, NJ
Bailey MJ (2006) More power to the pill: the impact of contraceptive freedom on women's life cycle labor supply. Q J Econ 121(1):289–320
Bailey MJ, Hershbein B, Miller AR (2012) The opt-in revolution? Contraception and the gender gap in wages. Am Econ J Appl Econ 4(3):225–254

Berndt ER, Finkelstein SN, Greenberg PE, Howland RH, Keith A, Rush AJ, Russell J, Keller MB (1998) Workplace performance effects from chronic depression and its treatment. J Health Econ 17(5):511–535

Bertrand M, Goldin C, Katz LF (2010) Dynamics of the gender gap for young professionals in the financial and corporate sectors. Am Econ J Appl Econ 2(3):228–255

Blau FD, Kahn LM (2017) The gender wage gap: extent, trends, and explanations. Journal of Economic Literature 55 (3):789–865. ISSN 0022-0515

Bonari L, Pinto N, Ahn E, Einarson A, Steiner M, Koren G (2004) Perinatal risks of untreated depression during pregnancy. Can J Psychiat 49(11):726–735

Bratti M, Cavalli L (2014) Delayed first birth and new mothers' labor market outcomes: evidence from biological fertility shocks. Eur J Popul 30(1):35–63

Caucutt EM, Guner N, Knowles J (2002) Why do women wait? Matching, wage inequality, and the incentives for fertility delay. Rev Econ Dyn 5(4):815–855

Chatterji P, Alegria M, Mingshan L, Takeuchi D (2007) Psychiatric disorders and labor market outcomes: evidence from the National Latino and Asian American Study. Health Econ 16(10):1069–1090

Chatterji P, Alegria M, Takeuchi D (2011) Psychiatric disorders and labor market outcomes: evidence from the national comorbidity survey-replication. J Health Econ 30(5):858–868

Demir F, Ghosh P, Liu Z (2020) Effects of motherhood timing, breastmilk substitutes and education on the duration of breastfeeding: evidence from Egypt. World Dev 133:105014

Doepke M, Hannusch A, Kindermann F, Tertilt M (2022) The economics of fertility: a new era. National Bureau of Economic Research WP 29948

Erosa A, Fuster L, Restuccia D (2002) Fertility decisions and gender differences in labor turnover, employment, and wages. Rev Econ Dyn 5(4):856–891

Field T, Diego M, Hernandez-Reif M (2006) Prenatal depression effects on the fetus and newborn: a review. Infant Behav Dev 29(3):445–455

Gallen Y, Schrøter Joensen J, Johansen ER, Veramendi GF (2023) The labor market returns to delaying pregnancy. Available at SSRN 4554407:

Goldin C (2014) A grand gender convergence: its last chapter. American Economic Review 104(4):1091–1119

Goldin C, Lawrence F, Katz, (2002) The power of the pill: oral contraceptives and women's career and marriage decisions. J Polit Econ 110(4):730–770

Hackl F, Hummer M, Pruckner GJ (2015) Old boys' network in general practitioners' referral behavior? J Health Econ 43:56–73

Herr JL (2016) Measuring the effect of the timing of first birth. Journal of Population Economics 29(1)

Kalsi P, Liu MY (2021) Pregnancy loss and female labor market outcomes. Available at SSRN 3829769

Karimi A (2014) Effects of the timing of births on women's earnings: Evidence from a natural experiment. Working Paper, IFAU-Institute for Evaluation of Labour Market and Education Policy

Kleven H, Landais C, Posch J, Steinhauer A, Zweimüller J (2019a) Child penalties across countries: evidence and explanations. American Economic Association Papers and Proceedings 109:122–126

Kleven H, Landais C, Posch J, Steinhauer A, Zweimüller J (2024) Do family policies reduce gender inequality? Evidence from 60 years of policy experimentation. Am Econ J Econ Pol 16(2):110–149

Kleven H, Landais C, Søgaard JE (2019b) Children and gender inequality: evidence from Denmark. Am Econ J Appl Econ 11(4):181–209

Kuhn A, Lalive R, Zweimüller J (2009) The public health costs of job loss. J Health Econ 28(6):1099–1115

Lalive R, Schlosser A, Steinhauer A, Zweimüller J (2014) Parental leave and mothers' careers: the relative importance of job protection and cash benefits. Rev Econ Stud 81(1):219–265

Lok IH, Neugebauer R (2007) Psychological morbidity following miscarriage. Best Practice & Research Clinical Obstetrics & Gynaecology 21(2):229–247

Londoño-Vélez J, Saravia E (2025) The impact of being denied a wanted abortion on women and their children. Q J Econ 140(2):1061–1110

Markussen S, Strøm M (2022) Children and labor market outcomes: separating the effects of the first three children. J Popul Econ 35(1):135–167

Miller AR (2011) The effects of motherhood timing on career path. J Popul Econ 24(3):1071–1100

Mølland E (2016) Benefits from delay? The effect of abortion availability on young women and their children. Labour Econ 43:6–28

Nevo A, Rosen AM (2012) Identification with imperfect instruments. Rev Econ Stat 94(3):659–671

OECD (2001) OECD Employment Outlook Paris. https://www.oecd-ilibrary.org/content/publication/empl_outlook-2001-en

OECD (2022) OECD Family Database. https://www.oecd.org/els/family/database.htm

Peng L, Meyerhoefer CD, Zuvekas SH (2015) The short-term effect of depressive symptoms on labor market outcomes. Health Econ 25(10):1223–1238

Thakral N, Tô LT (2025) The child penalty and an age-old problem. mimeo

Wooldridge JM (2010) Econometric analysis of cross section and panel data. MIT press

Wooldridge JM (2002) Inverse probability weighted m-estimators for sample selection, attrition, and stratification. Port Econ J 1(2):117–139

Zweimüller J, Winter-Ebmer R, Lalive R, Kuhn A, Ruf O, Wuellrich JP, Büchi S (2009) The Austrian social security database (ASSD). https://doi.org/10.2139/ssrn.1399350