DISCUSSION PAPER SERIES

DP13732

MOTHERHOOD TIMING AND THE CHILD PENALTY: BOUNDING THE RETURNS TO DELAY

Anikó Bíró, Steven Dieterle and Andreas Steinhauer

LABOUR ECONOMICS

ISSN 0265-8003

MOTHERHOOD TIMING AND THE CHILD PENALTY: BOUNDING THE RETURNS TO DELAY

Anikó Bíró, Steven Dieterle and Andreas Steinhauer

Discussion Paper DP13732
Published 13 May 2019
Submitted 13 May 2019

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

MOTHERHOOD TIMING AND THE CHILD PENALTY: BOUNDING THE RETURNS TO DELAY

Abstract

We use administrative data from Austria to analyze labor market returns to delaying motherhood. We exploit delays due to pregnancy loss to provide bounds on the returns that account for imperfect instruments and selection into the sample for mothers that suffer a loss. Our results suggest small effects of delay on earnings, employment, and firm quality--- in contrast with the prior literature. The lower bounds suggest little difference in earnings trajectories around the first birth. This raises the possibility that much of the return may come from delaying the "child penalty" rather than changing how the career responds to children.

JEL Classification: J13, J31, C26

Keywords: Female Earnings, fertility timing, imperfect instrument

Anikó Bíró - biro.aniko@krtk.mta.hu
University of Edinburgh and Centre for Economic and Regional Studies of the Hungarian Academy of Sciences

Steven Dieterle - steven.dieterle@ed.ac.uk
University of Edinburgh

Andreas Steinhauer - andreas.steinhauer@ed.ac.uk
University of Edinburgh and CEPR

Acknowledgements
We are very grateful to Josef Zweimüller for providing access to the administrative data at the University of Zurich's "Science Cloud Server" for secure data processing and answering many questions along the way. We would also like to thank seminar participants at the University of Arizona, University of Lausanne, Central European University (Budapest, Hungary), Michigan State University, University of Zurich and participants at the Society of Labor Economists Annual Meetings, the European Society for Population Economics Annual Conference and Swiss Economists Abroad End-Of-Year conference for helpful comments on earlier versions of this paper.

Motherhood Timing and the Child Penalty: Bounding the Returns to Delay*

Anikó Bíró[†], Steven Dieterle[‡], Andreas Steinhauer[§]

May 9, 2019

Abstract

We use administrative data from Austria to analyze labor market returns to delaying motherhood. We exploit delays due to pregnancy loss to provide bounds on the returns that account for imperfect instruments and selection into the sample for mothers that suffer a loss. Our results suggest small effects of delay on earnings, employment, and firm quality—in contrast with the prior literature. The lower bounds suggest little difference in earnings trajectories around the first birth. This raises the possibility that much of the return may come from delaying the "child penalty" rather than changing how the career responds to children.

Keywords: female earnings, fertility timing, imperfect instrument

JEL Classification: J13, J31, C26

*We are very grateful to Josef Zweimüller for providing access to the administrative data at the University of Zurich's "Science Cloud Server" for secure data processing and answering many questions along the way. We would also like to thank seminar participants at the University of Arizona, University of Lausanne, Central European University (Budapest, Hungary), Michigan State University, University of Zurich and participants at the Society of Labor Economists Annual Meetings, the European Society for Population Economics Annual Conference and Swiss Economists Abroad End-Of-Year conference for helpful comments on earlier versions of this paper.

†The University of Edinburgh, and Health and Population "Momentum" Research Group at the Institute of Economics, Centre for Economic and Regional Studies of the Hungarian Academy of Sciences

‡The University of Edinburgh

§The University of Edinburgh and CEPR

1 Introduction

The introduction of birth control in the 1960s gave women more control over birth timing which reduced uncertainty when making human capital investments and labor supply decisions (Goldin and Katz, 2002; Bailey, 2006; Bailey et al., 2012). This increased control led to a trend in women delaying motherhood to later ages (Caucutt et al., 2002). Many of the benefits of delay are likely to come in the form of labor market returns which is consistent with the fact that the trends in fertility timing coincided with substantial increases in labor force participation and wages for women (Juhn and Potter, 2006; Fogli and Veldkamp, 2011; Black and Juhn, 2000, among others). Age at first birth and labor outcomes for women are also strongly related in the cross-section (Caucutt et al., 2002).

In this paper we use detailed administrative worker-firm matched social security data linked to health records from Austria to analyze the relationship between delaying motherhood and labor market outcomes for women in detail. The administrative data allows us to pay particular attention to how delay affects outcomes at different points in the career— before, around, and after the first birth— as well as considering the cumulative effect over these periods — improving our understanding of how the returns to dely may devlop over the career. Delaying motherhood may improve labor market outcomes for women by reducing interruptions at a key point in the career. For instance, it may allow women to take advantage of the steeper wage profile early in their careers by accumulating more human capital or finding a better job match, resulting in greater attachment to the labor force or lower turnover after birth (Herr, 2016; Erosa et al., 2002).

To illustrate the potential returns to delaying motherhood, Figure 1 provides descriptive evidence based on our administrative data from Austria on earnings and age at first birth for the cohort of women born in 1970. We depict mean annual earnings by age for women who have their first child at different ages— including earnings for men and childless women for comparison. Before the first birth, the age-earnings profiles fall between childless women and men, increasing year-by-year. At the first birth, we see the motherhood gap realized with a sudden fall in mean earnings that never recovers to the pre-birth trend. 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" (Kleven et al., 2019). Importantly, we also tend to see steeper slopes after first birth for women who delay longer, implying a faster recovery to a higher level of earnings. Therefore, delay may also increase cumulative earnings by changing how a woman's career responds to the birth of the first child, thereby reducing the size of the child penalty.

This has led many researchers to consider the role fertility timing may play in shaping female labor market outcomes (Miller, 2011; Karimi, 2014; Bratti and Cavalli, 2014; Hotz et al., 1997; Ashcraft et al., 2013; Buckles, 2008; Cristia, 2008; Herr, 2016).[1] However, the observed relationship between age at first birth and labor market outcomes may also be driven by selection— either due to unobserved heterogeneity (preferences for fertility correlated with ability) or reverse causality (delayed fertility for women who anticipate the largest gain from delay). On balance, this has led to concerns that OLS estimates of the return to delay may be too large.

One popular approach has been to exploit delays in first birth driven by suffering a pregnancy loss (Miller, 2011; Karimi, 2014; Bratti and Cavalli, 2014).[2] Conceptually, this approach captures the thought experiment of following a group of women who chose to have children at a particular age and then compare the outcomes for those that had a child at that age to those who were delayed to a later age due to a pregnancy loss. Using prior pregnancy loss as an instrument for first birth timing has yielded mixed results, but often suggests large, positive effects of delaying first birth on labor force participation, hours worked, cumulative earnings, and long-run wages (Miller, 2011; Bratti and Cavalli, 2014).

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

[2]We use the term pregnancy loss to capture both miscarriage (early pregnancy loss) and stillbirth (later pregnancy loss). It is an unintentional loss of a pregnancy and does not include induced abortions.

Here, we use detailed health insurance and worker-firm matched social security administrative data from Austria to consider three distinct issues in the interpretation and estimation of the returns to delaying motherhood using a pregnancy loss based instrument. In line with the previous literature (Miller, 2011; Bratti and Cavalli, 2014; Karimi, 2014), we note the possibility of negative selection into pregnancy loss and the potential negative mental health effects of pregnancy loss that may violate the necessary exclusion restriction for instrumental variables (IV) estimation. Unlike the previous literature, our detailed, administrative employment and health data allow us to directly document both sources of bias. For instance, we find noticeable increases in the diagnosis of mental health conditions and psychological medication prescriptions for women that suffer a pregnancy loss. If worse mental health has a detrimental impact on labor supply, then we would expect the instrumental variable estimates to be biased down. This has lead researchers to argue that the large estimated effects are possibly a lower bound on the true effects (Miller, 2011; Bratti and Cavalli, 2014).

The second issue to arise using the pregnancy loss instrument is that in a given sample not all women who suffer a pregnancy loss will be observed having a child. Women may choose not to or be unable to have children following a pregnancy loss, or they may choose to delay a subsequent pregnancy long enough to not show up within the window of the available data. As shown in Figure 2, most women in our data have children within three or four 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 fifteen years after the loss and therefore "fall out" of the natural experiment. Even if pregnancy loss was a randomly occurring event, only women who suffer a pregnancy loss have the possibility of "selecting out" of the sample of mothers.

To see how this presents problems for the IV estimates, consider there being two types of women— those that *if* they suffer a pregnancy loss they will be observed having a subsequent child and those that will not. The group of women who do not suffer a loss will be a mixture of these two latent types while the pregnancy loss group used for estimation will only include the type that respond to a pregnancy loss in such a way that we observe a subsequent birth. If these two types of women differ in important ways, then the women we observe being delayed to a later age at first birth by a pregnancy loss may not be comparable to the group of women that were not delayed by a loss. In our setting, the women who we do observe with a subsequent birth following a loss seem to be positively selected with higher earnings and participation before pregnancy. This positive selection into the pregnancy loss sample used for estimation leads to an upward bias in the instrumental variable estimates— working against the lower bound interpretation of the IV estimates discussed above.

Finally, we also note that estimates based on a pregnancy loss instrument may not capture all of the policy relevant effects of delay. For many women, the induced variation in timing from pregnancy loss occurs after key human capital (education, training, etc.) and job search investments (search intensity, occupational choices, etc.) have been made in conjunction with the initial "optimal" fertility choice. If some of these investments are naturally front-loaded at the beginning of the career, women may be constrained from fully "re-optimizing" these investments in response to the pregnancy loss induced delay. This would lead to lower observed returns to delay than for a similar length of delay that could be anticipated earlier— further reinforcing the idea that the IV approach may underestimate the returns to delay.

Considering these issues together, we opt to move from point estimating the (average) returns to delay as in the previous literature, to estimating plausible bounds on the returns. To account for the negative selection into and the negative mental health effects of pregnancy loss we apply the approach from Nevo and Rosen (2012) to obtain two sided bounds when using an imperfect instrument. Rather than requiring pregnancy loss to be uncorrelated with unobservables driving labor market outcomes, the Nevo and Rosen bounds require the weaker assumption that pregnancy loss is less correlated with the unobservables than age at first birth is. In order to correct for the post pregnancy loss selection, we use detailed pre-pregnancy labor market and health information to predict post-pregnancy loss selection and use inverse probability weighting (IPW) to adjust for this selection.

Consistent with the motivating evidence in Figure 1, we estimate large and tight bounds for the effect of a one year delay on total pre-birth earnings of between 15.6 to 16.9 percent, on maximum pre-birth daily earnings of $3.1 - 4.7$ percent, and on highest pre-birth firm quality (measured by firm level average pay) of $3.2 - 3.9$ percent— suggesting large effects from delaying the realization of the child penalty. However, when estimating event history profiles around the time of first birth, our selection-corrected lower bounds suggest no difference in pre-birth earnings growth or post-birth earnings recovery for those who delay first birth— with similar results for other outcomes. This raises the possibility that the faster recovery highlighted in Figure 1 may be entirely due to unobserved differences for those who choose to delay motherhood and that differences in pre-birth human capital or job match quality may play little role in driving the realized returns. Taken together, our lower bound results suggest that most of the returns to delay over this period may come from delaying the realization of the child penalty, rather than changing how the mother's career responds to having children.

To summarize the overall effect of motherhood delay, we also consider several cumulative measures of labor market outcomes over our observation period (1972-2017). Our selection-corrected bounds suggest a $4.3 - 5.4$ percent increase in cumulative earnings for each year of delay. We also find small but positive bounds of a $1.3 - 1.8$ percent increase in total days of employment and $0.6 - 1$ percent increase in maximum firm quality (measured by average pay) for each year of delay.

Finally, we note that across our analysis, accounting for the post-loss selection into the sample is important for our conclusions. For instance, when considering the total observed earnings over our data, the IPW selection-corrected bounds ($4.3 - 5.4$ percent) fall below the uncorrected bounds ($5.5 - 5.8$ percent). This highlights the possibility that prior estimates based on a pregnancy loss instrument may be too large. For instance, using data from the United States, Miller (2011) estimates a 9 percent increase in earnings and 3 percent higher average wages between age 21 and 34 for a one year delay. The fact that the selection corrected bounds are wider than the uncorrected magnifies the importance of selection issues driving the OLS estimates as well.

2 Conceptual Framework and Background

2.1 Fertility Timing and Estimating Returns to Delay

To help motivate what follows, consider a stylized discussion of fertility timing choices. Suppose women face a choice between different combinations of age at first birth (A1B) and total earnings over the career, where we assume that utility is increasing and concave in total earnings and time with children. Earlier birth means more time with children over the lifetime. Assume that an initial fertility timing choice is made early in ones career and is— possibly— made jointly with decisions over human capital investments. Each woman effectively faces a locus of A1B-earnings combinations at this point. Suppose that this locus is upward sloping with higher total earnings associated with later motherhood as depicted in Figure 3 by the solid black line labeled L_1 (see Figure A.2 in the Appendix for descriptive evidence from Austria that is consistent with an upward sloping locus). Note that each point on the locus reflects the total earnings associated with that age at first birth choice and may reflect very different simultaneous choices of other factors or inputs into a woman's earnings. For instance, if motherhood greatly reduces the realized returns to education post-birth, then the same woman may choose different levels of education conditional on a choice of age at first birth— since delay increases the time over which the returns are realized. Given this locus of potential A1B-earnings combinations, the optimal fertility timing choice will be made by selecting a point on the locus that maximizes utility— such as point A on the figure where an individual's indifference curve (U_1) reflecting preferences over total earnings and time with children is tangent to the locus.

Note that estimating the cumulative earnings return to delaying motherhood requires estimating the slope of the A1B-earnings locus— or some average of the slopes of the loci faced by women in our target population. If women face the same locus— or randomly face different loci— and make different choices due to random differences in preferences, then we could easily approximate the slope by comparing women with different A1B choices. This is depicted by the dashed line connecting points A and B in the figure, where A and B reflect the fertility choices for two different women facing the same locus. However, the key concern in this literature is that differences in A1B are not driven by random differences in preferences among women facing similar options. Instead, it is suspected that women's preferences may be correlated with other factors driving their earnings. As an example, suppose that women differ in their ability and that higher ability allows women to earn more for any choice of A1B. Higher ability women will, therefore, face a higher locus— depicted by the red locus L_2 in Figure 3. If higher ability women also have stronger preferences for total earnings (flatter indifference curves $\widetilde{U}_2$ in the figure)[3], then they may choose a later A1B when faced with the higher locus— depicted by point C. Rather than estimating the intended slope of the line AB, the OLS estimate of returns to delay will confound the returns to delay with differences in ability, overestimating the returns to delay by tracing out line AC.

Faced with these concerns, several papers have used pregnancy loss as a source of variation in first birth timing to help identify the effects of delaying motherhood (Miller, 2011; Bratti and Cavalli, 2014; Karimi, 2014; Hotz et al., 1997, 2005; Ashcraft et al., 2013; Fletcher and Wolfe, 2009).[4] Miller (2011) uses three biological fertility shocks (miscarriage, conception while using contraception, elapsed time from first conception attempt to first birth) to estimate the effect of motherhood delay on earnings, wage rate, and working hours. She finds that one year delay results in 9% higher earnings, 3% higher wages and 6% increase in

[3]Note that this could also reflect preferences for the nonpecuniary aspects of having a career— not simply earnings/consumption.

[4]Many of these papers have considered the consequences of teenage childbearing (Hotz et al., 1997, 2005; Ashcraft et al., 2013; Fletcher and Wolfe, 2009). Here, we focus on estimating the returns to delay for adult women and exclude teen pregnancies from our analysis.

work hours on average over the age range from 21 to 34. Bratti and Cavalli (2014) apply a similar empirical strategy as Miller (2011), using miscarriage and stillbirth as an instrument for the timing of first birth, and estimate positive effects of motherhood delay on labour market participation and weekly working hours for a sample of Italian mothers. In contrast, Karimi (2014) finds negative effects on wages and income for women from Sweden using a pregnancy loss instrument for first birth timing.

Conceptually, the goal of using the pregnancy loss instrument is to trace out the slope of the earnings-A1B locus by comparing women who faced similar options and made similar initial fertility choices— and therefore are similar on average across unobservables, but whose realized age at first birth differs due to pregnancy loss induced delay. However, as noted above, the initial set of earnings-A1B choices reflect the outcome of several choices made simultaneously (or even sequentially). If we consider the original locus as presenting the unconstrained set of choices allowing for simultaneous changes to human capital investments (education, job search effort, occupation choices, etc.), then a concern with using the pregnancy loss instrument is that many of these choices will have been made in conjunction with the initial A1B choice and will be fixed or costly to alter at the time that the loss induced delay is realized. Therefore, at the time of the pregnancy loss, a woman may face a constrained set of A1B-earnings choices. This is depicted by the blue locus $\widetilde{L}_1$ in the figure that has a weakly lower slope after point A than the original unconstrained locus. Faced with this new set of choices, a woman is forced to choose a later A1B- here depicted by point D. Therefore, the pregnancy loss IV may approximate the slope of the line AD— not capturing the returns to delay that might occur if the delay was anticipated at the time other human capital choices were made— underestimating the unconstrained returns to delay.[5]

[5] As descriptive evidence in support of this possibility, we estimated two separate multinomial logit models of the educational level (compulsory, some high school, high school, professional apprenticeship, teacher's college, and college) at the time of first birth for a subset of women in our sample that had their education recorded in the birth registry. In the first multinomial logit, we included age at first birth, indicators for mother's birth cohort, and age at entry into the labor market. Age at first birth is found to be a statistically significant predictor of education level suggesting different education choices could play a role in the observed returns to delay. However, when we replace age at first birth with a pregnancy loss indicator, this indicator is not found to be a statistically significant predictor of education level at first birth. This suggests that

2.2 Instrument Validity

The use of prior pregnancy loss as an instrument for first birth timing also presents concerns over instrument validity. In particular, there are concerns that pregnancy loss is not fully random and can be related to behavioral characteristics (smoking, drinking, among others) that can also influence later labor market outcomes. Miller (2011) tries to control for these factors by including self-reported indicators of drug use, alcohol consumption, and smoking during pregnancy. Bratti and Cavalli (2014) and Karimi (2014), on the other hand, check the balance of covariates across pre-motherhood characteristics to look for differences between women who have and have not suffered a pregnancy loss.

In addition to possible negative selection into suffering a pregnancy loss, Miller (2011) and Karimi (2014) note the possibility of it negatively impacting mental health. In a review of the psychology literature, Lok and Neugebauer (2007) report that the ratio of women having elevated levels of depressive symptoms or being diagnosed with major depressive disorders following miscarriage varies between $10-55\%$. Given the literature connecting mental health and labor market outcomes (Berndt et al., 1998; Chatterji et al., 2007, 2011; Peng et al., 2015), this raises the possibility that pregnancy loss may affect labor market outcomes through its effect on mental health, violating the exclusion restriction necessary for IV estimation. Importantly, both of these factors— pregnancy loss being related to risky behaviors and the potential negative mental health effects of pregnancy loss— would be expected to be negatively correlated with labor market outcomes leading the IV estimates to be biased down.

Here, we raise an additional concern based on a sample selection issue that arises when studying first birth timing for adult women. Namely, when focusing on the timing of first birth for mothers our sample will necessarily only include women observed having children *at some point*.[6] As Figure 2 showed, about 20 percent of the women we observe suffering a

the returns estimated using the pregnancy loss instrument may not capture the effect of making different education choices. See Appendix B for details.

[6] In studying teenage pregnancy, Ashcraft et al. (2013) point out that pregnant teenagers who have induced

pregnancy loss in our sample are never observed with a child up to fifteen years after the loss, and are, therefore, not included in a regression of labor market outcomes on age at first birth. This selection may be driven by two different responses to pregnancy loss. First, women may be unable or choose not to have children following a pregnancy loss. Another possibility is that some women choose to delay a subsequent pregnancy longer than others, so that this pregnancy falls outside the time frame covered by a given data set. In section 4 we will consider the potential implications of this selection for estimating the returns to motherhood delay using a pregnancy loss based instrument.

3 Data

We use administrative records from the Austrian Social Security Database (ASSD) which is a matched firm-worker data set covering the labor market history and some demographic characteristics of workers in dependent employment in Austria.[7] The ASSD covers the period from 1972 to 2017. We observe labor market status on a daily basis and yearly 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).

We can link the ASSD records to data 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.[8] The time span of the health insurance

abortions are on average positively selected, leaving the miscarrying women to come on average from a more disadvantaged background. Pregnancy is more likely to be planned among adult women, thus there is less demand for induced abortion in our context.

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

[8]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).

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 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 January 1999 – 1st October 2007.[9] 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. 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 labor market entry and those who have no health insurance records prior to the event. We define labor market entry as the year in which a woman first earns at least 2,500 euro (in real terms, CPI 2000).[10] We, therefore, exclude women with very weak attachment to the labor force for the entire period before the first pregnancy. In this we follow Herr (2016) who suggests to consider women

[9] 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. 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 one year of pre-event health data.

[10] 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 5,000 which has little effect on our results. We include dummies for age at labor 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 labor 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 labor market entry of 16. Their median earnings in our sample are 1,901 at age 15, 5,235 at age 16, and 6,697 at age 17. Earnings breakdowns for other education levels, available upon request, further support the use of the labor market entry variable as constructed.

who work before having children separately from those that have children before entering the labor force, as experience, rather than age itself, is the important driver for the potential returns to delaying motherhood.

Our sample consists of 52,069 women for whom we either observe a first birth or pregnancy loss without preceding birth between 1999-2007. In this sample the first event is a live birth for 46,437 women and a pregnancy loss for 5,632 (10.8%). Among women whose first event is a pregnancy loss 4,561 (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 labor 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 a yearly frequency or measured as of reference days in the middle of each quarter (15th of the second month of the quarter). Our main labor 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 yearly 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. 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). Next, we make use of the matched firm-worker nature of the data set to calculate average daily earnings at the firm a woman is working for as a rough measure of one dimension of firm quality. To this end we use the universe of firm-worker pairs in Austria covered by the ASSD and estimate average daily earnings by firm on quarter reference days. We then match this quarterly "firm wage" to our sample of women in Upper Austria.

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 were ever previously prescribed.

3.1 Descriptive Statistics

We define three groups of women in our sample.

1. Group 1: Women, for whom the first observed event is a live birth.
2. Group 2: Women, for whom the first observed event is a pregnancy loss, and we observe a subsequent live birth (up to 2017).
3. Group 3: Women, for whom the first observed event is a pregnancy loss, and we do not observe a subsequent live birth (up to 2017).

Table 1 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 29.3 implies an average delay of 2.3 years due to the loss. Importantly, this reflects the average delay used by the loss instrument for identification. Total number of births over our data period are also similar at around 1.9 on average among women whose first event is a live birth, and 1.8 among women in group 2— reducing the role for differences in total parity to play in explaining any of the results derived using the loss instrument. When we compare summary measures of labor market outcomes over the entire data window (1972-2017) we see generally higher earnings, more experience, and better firm matches 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.

We also compare characteristics before the event. 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 one 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 labor 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 labor market outcomes and that both pregnancy loss and having a child after a loss are not entirely random.

4 Evidence for IV Biases

Here, we build on the descriptive statistics from the previous section to provide evidence supporting the concerns with the pregnancy loss instrument outlined in Section 2. We apply an event history type approach around the time of pregnancy for the three groups of women in order to look for evidence of the negative selection into pregnancy loss, the negative mental health effects of pregnancy loss, and for differences in post-pregnancy loss selection between groups 2 and 3.

Figure 4 displays 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. The solid vertical line marks either the quarter of birth for women who have a live birth (group 1) or the approximate due date for pregnancies that end in a loss (groups 2 and 3). We approximate due dates for women suffering a pregnancy loss by assuming the loss happens in the first quarter of the pregnancy (the first trimester).

Note that across the groups, events relevant to labor 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 women 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 doctors visits while pregnant, thereby increasing the probability of a diagnosis or prescription.

Panel (a) of Figure 4 plots average daily earnings on quarter 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 Figure 2). Taken on its own, panel (a) is encouraging for the IV research design given the similar earnings before pregnancy for the two groups. However, in panel (a), the pregnancy loss group includes those women that we never observe with a birth (group 3).

In panel (b), 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. The first thing to note is that now there are clear differences in daily earnings across our 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.[11] 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 over estimate the returns to delay.[12]

Figure 4 also provides the first descriptive evidence that pregnancy loss may have negative effects on labor market outcomes. We see declines in earnings post pregnancy loss even for those women who do not have a subsequent child. Of course, it is difficult to infer too much from this pattern given the potential selection into employment for these women and the fact that we lack a clear comparison group for how wages would have evolved in the absence of the pregnancy loss.

Panels (c) and (d) of Figure 4 provide similar plots for whether an individual has ever been diagnosed with a mental health condition by time t. In panel (c) we see that the mental health diagnoses are more likely before pregnancy for the pregnancy loss group (groups 2 and 3). We also see an increase in diagnoses after the pregnancy loss, consistent with the possible negative mental health effects of the loss. In Panel (d), we see that the higher level of diagnoses for the pregnancy loss group is predominately driven by the group never observed with a birth (group 3) further reinforcing the possibility that this group is negatively selected.[13]

[11] The fact that the weighted average of the two pregnancy loss groups— where the weights are respective sample sizes— depicted in panel (a) 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.

[12] Appendix Tables A.1 and A.2 provide additional evidence based on panel regressions controlling for year of birth, age, and age at labor market entry that support the patterns found here. Conditional on these controls, employment and log daily earnings before first due date are statistically significantly higher for group 2 and statistically significantly lower for group 3 compared to women of group 1.

[13] The differences become somewhat smaller when we condition on year of birth, age, and age at labor market entry (Tables A.1 and A.2). In fact, the difference in mental health diagnosis before first due date between groups 1 and 2 is not statistically significant. In the two years before first due date women of group 2 are slightly more likely to have ever been prescribed psychological medications, however.

5 Methodological Approach

The previous discussion motivated two key concerns with using pregnancy loss as an instrument for age at first birth. First, we saw that women who suffer a pregnancy loss had slightly worse health and labor market outcomes before pregnancy and appear to suffer a negative mental health shock post-pregnancy loss. Combined, these suggest that pregnancy loss may be negatively related to unobservables driving labor market outcomes. Alone, this suggests that the IV estimates may be biased downward, providing a "lower bound" on the effect of delaying motherhood. We also provide evidence of post-pregnancy loss selection into being observed having a child. Namely, about 20 percent of the women whose first observed pregnancy ended in a pregnancy loss are never observed having a child. Importantly, these women had lower earnings and worse health pre-pregnancy 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.

5.1 Accounting for Negative Mental Health Effects

Here, we will formalize the idea that the IV estimates using a pregnancy loss instrument may be downward biased due to the negative mental health effects and non-random occurance of pregnancy loss by applying the bounding approach with imperfect instruments from Nevo and Rosen (2012). We prefer the bounding approach outlined here to directly controlling for the observed health measures for a few reasons. First, and most importantly, the post-pregnancy loss health measures are outcomes of pregnancy loss themselves leading to the "bad control problem" of Angrist and Pischke (2009), making interpretation of the resulting estimates difficult. In addition, there may yet be further negative effects not captured by our health measures. The bounding exercise used here will account for these effects as well.

Consider our baseline specification relating earnings to age at first birth:[14]

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

Here, y_i are earnings for an individual, A_i is age at first birth, and $\mathbf{x}_i = (1, x_1, \ldots, x_{k-1})$ contains a constant and our control variables.[15] Consistent estimation by OLS is based on the following set of moment conditions: $E\left[(A, \mathbf{x})' u\right] = \mathbf{0}$. As we argued previously, there is a concern that age at first birth is positively correlated with the unobservables driving earnings ($\rho_{Au} > 0$) so that $E[Au] > 0$ generating an overestimate of the returns to motherhood delay. This motivates adopting an IV approach. To match the framework from Nevo and Rosen (2012), we will define our instrument as indicating that the first observed pregnancy ended in a live birth:

$$z_i = \begin{cases} 0 & \text{if first pregnancy ends in loss} \\ 1 & \text{if first pregnancy results in live birth} \end{cases}$$

We refer to z_i as our "No Loss" instrument.[16] For consistent IV estimates, we must assume that $E\left[(z, \mathbf{x})' u\right] = \mathbf{0}$, replacing the $E[Au] = 0$ condition used for OLS estimation with $E[zu] = 0$. As we have argued previously, we are concerned that a pregnancy loss based instrument may also be endogenous due to the negative mental health effects of pregnancy loss. Therefore, we worry that the No Loss instrument is positively correlated with unobservables ($\rho_{zu} > 0$), so that $E[zu] > 0$ as well. Finally note that as constructed, the covariance between the No Loss instrument and age at first birth is expected to be negative ($\sigma_{Az} < 0$)— not suffering a pregnancy loss is associated with an earlier age at first birth— so that, ignoring the post-loss selection issue, the IV estimate will be an underestimate of

[14]In our main analysis we will consider a range of outcomes and specifications. However, the basic intuition presented here will hold throughout.

[15]In practice, our main control variables will be mother cohort indicators, age at labor market entry indicators, and, in some specifications, mother's education.

[16]Defining the instrument this way, rather than as a pregnancy loss does not change the estimate, but allows us to directly apply the results from Nevo and Rosen (2012).

the return to delayed motherhood.[17]

Given our setting, we can directly apply the results from Nevo and Rosen (2012) to obtain sharp two-sided bounds on β. To apply the Nevo and Rosen approach in the presence of additional control variables, we will 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 equation (1) using the residualized variables and omitting $\mathbf{x}$:

$$\widetilde{y}_i = \beta \widetilde{A}_i + u_i$$

Formally, we must assume the following:[18]

NR1 The observations $(y_i, A_i, \mathbf{x}_i, z_i), i = 1, \ldots, n$ are stationary and weakly dependent

NR2 $E[\mathbf{x}'u] = \mathbf{0}$

NR3 $\rho_{\widetilde{A}u}\rho_{\widetilde{z}u} \geq 0$

NR4 $|\rho_{\widetilde{A}u}| \geq |\rho_{\widetilde{z}u}|$

NR5 $Rank(E[(z, \mathbf{x})'(z, \mathbf{x})]) = Rank(E[(A, \mathbf{x})'(A, \mathbf{x})]) = Rank(E[(z, \mathbf{x})'(A, \mathbf{x})]) = k$

The key assumptions for obtaining the bounds are NR3 and NR4.[19] NR3 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, a point we argued above. NR4 assumes that the No Loss instrument is less endogenous than age at first birth. This assumption seems quite plausible in our setting. For instance, Miller (2011) notes that over 90% of pregnancy losses

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

[18] Note that this differs from the setup in Nevo and Rosen (2012) for obtaining bounds with additional covariates. They allow for the more general case where the researcher has valid instruments for other endogenous regressors in $\mathbf{x}$. In that case the assumptions are cast in terms of unconditional correlations— i.e. before partialling out the other control variables— and set identification is based on z not $\widetilde{z}$. Nevo and Rosen (2012) note that with an imperfect instrument and only one endogenous regressor, the assumptions can be recast in terms of conditional correlations. This corresponds to using $\widetilde{y}$, $\widetilde{A}$, and $\widetilde{z}$ in our setting.

[19] See Nevo and Rosen (2012) for a full discussion of the assumptions. NR1 is standard assumption on the sampling process, as written here NR2 treats the other regressors as exogenous— just as in our OLS and IV estimators above, and NR5 includes the standard rank conditions for OLS and IV estimators.

are due to anomalies affecting the development of the fetus rather than, say, behavioral factors. Further, the negative mental health effects we documented in section 4 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.

Under assumptions NR1 to NR5, Nevo and Rosen (2012) show that if $\sigma_{\widetilde{A}\widetilde{z}} < 0$, $\sigma_{\widetilde{A}u} > 0$, and $\sigma_{\widetilde{z}u} > 0$, then β falls between β_z^{IV} and β_v^{IV}. Here, β_z^{IV} is simply our IV estimator using the No Loss instrument. We will also refer to it as the Nevo-Rosen (NR) Lower Bound. This formalizes the intuition presented in past work that β_z^{IV} would be a lower bound if pregnancy loss has detrimental mental health effects. The NR Upper Bound, β_v^{IV}, is the IV estimate using the following instrument that we will refer to as the NR instrument:

$$v = \sigma_{\widetilde{A}}\widetilde{z} - \sigma_{\widetilde{z}}\widetilde{A} \tag{2}$$

Intuitively, β_v^{IV} provides an upper bound as v 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 NR4 holds with equality. To see this, note that

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

Clearly, if $\rho_{\widetilde{A}u} = \rho_{\widetilde{z}u}$, then $E[vu] = 0$ yielding a valid moment condition for identifying β. This procedure will provide upper and lower bounds for the effects of age at first birth on labor market outcomes that are valid under weaker assumptions than those used in prior work for point identification. Importantly, the bounds directly account for any factor that may lead to a negative correlation between pregnancy loss and labor market outcomes, including the possible negative affects of pregnancy loss on mental health and the negative selection into suffering a loss.

5.2 Accounting for Post-Pregnancy Loss Selection

In this subsection, we discuss how we account for positive selection of women observed having a child after suffering a pregnancy loss. As we discussed previously, this selection could take two forms: (1) women deciding not to or being unable to have children after a pregnancy loss or (2) choosing to delay their next pregnancy sufficiently long that we do not observe the birth in the window of our data. We cast the problem as a missing data issue where we do not observe an age at first birth or labor market outcomes with a child for the women that have a pregnancy loss and no subsequent child. However, we do observe detailed pre-pregnancy loss labor market and health data for these women. Setting up the problem this way leads directly to the Inverse Probability Weighting (IPW) approach to dealing with selection issues. The following discussion draws heavily on the treatment of IPW in Wooldridge (2010).

To start, denote the vector of data for an individual by $\mathbf{w}_i = (y_i, A_i, \mathbf{x}_i, z_i)$. Let our random draw from the population be given by $(\mathbf{w}_i, s_i)$ where s_i is a binary indicator of selection, where when $s_i = 0$ we only observe part of $\mathbf{w}_i$. In our case, $s_i = 0$ indicates women with a pregnancy loss and no subsequent observed birth— those for who we do not observe y_i or A_i— 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.

Again, our aim is to estimate the parameters of equation (1). Following Wooldridge (2010), denote $g(\mathbf{w}_i)$ as any scalar function of the data where the mean exists. Note that $E\left[g(\mathbf{w}_i)\right]$ could represent any of the moments used for identification for our three estimators of interest— OLS, NR Lower Bound (IV), and NR Upper Bound.[20] Quite generally, instead of using data representative of the population average $E[g(\mathbf{w}_i)]$, we only have the selected sample to identify $E[s_i g(\mathbf{w}_i)]$. The goal is to recover the population average from the selected subsample.

[20] Recall that all three are based on a set of moment conditions that include $E[\mathbf{x}'u] = \mathbf{0}$, but simply differ in the final condition used for identification— $E[Au] = 0$, $E[zu] = 0$, or $E[vu] = 0$.

The IPW solution requires us to have some variables that we observe for everyone ($s_i = 1$ and $s_i = 0$) that are good predictors of selection. Wooldridge (2010) sets out the following assumptions for IPW:

IPW1 $\mathbf{w}_i$ is observed when $s_i = 1$

IPW2 There is a vector $\mathbf{r}_i$ such that $P(s_i = 1|\mathbf{w}_i, \mathbf{r}_i) = P(s_i = 1|\mathbf{r}_i) \equiv p(\mathbf{r}_i)$

IPW3 For all $\mathbf{r} \in \mathscr{L} \subset \mathbb{R}^J$, $p(\mathbf{r}) > 0$

IPW4 $\mathbf{r}_i$ is observed when $s_i = 1$

IPW1 simply states that we observe all of $\mathbf{w}_i$ when women are observed to have a child. In practice, $\mathbf{r}_i$ will include pre-pregnancy health and labor market information so that IPW4 clearly holds in our setting as we observe this for everyone regardless of selection. IPW3 just states that there is no set of predictors for which the probability of selecting into the sample is zero—that is, no matter what someone's pre-pregnancy loss earnings or health are they have some chance of having a subsequent child.

IPW2 is the key assumption— basically assuming that $\mathbf{r}_i$ is such a good predictor of selection that once we condition on it the selection probability is independent of the other elements of $\mathbf{w}_i$. IPW2 is essentially an ignorability or unconfoundedness assumption for selection conditional on the variables included in $\mathbf{r}_i$.

In our setting, we must assume that conditional on the pre-pregnancy earnings and health variables the selection decision is independent of post-pregnancy earnings. 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. However, one might be concerned that individuals with the same pre-pregnancy earnings and health might select out of our mother sample for reasons that are related to post-pregnancy earnings. For instance, perhaps women who suffer a particularly large health shock with the pregnancy loss are more likely to select out of our sample—by either never giving birth or delaying longer so the birth is right censored. 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 estimates to be upward biased since the observationally equivalent women we do see with a child after a pregnancy loss would have earnings that are too high. Importantly, the prior work that has ignored this selection issue has implicitly made the much stronger assumption that selection is independent of all the observable factors, i.e. that $P(s_i = 1|\mathbf{w}_i) = P(s_i = 1)$. The descriptive results from Section 4 suggest that this implicit assumption is likely violated.

The following shows how the IPW approach— weighting by the inverse of the probability of selection— helps to recover the population mean from the selected sample:

$$\begin{aligned} E\left[s_i g(\mathbf{w}_i)/p(\mathbf{r}_i)\right] =& E\left\{E\left[s_i g(\mathbf{w}_i)/p(\mathbf{r}_i)|\mathbf{w}_i, \mathbf{r}_i\right]\right\} \\ =& E\left\{E\left[s_i|\mathbf{w}_i, \mathbf{r}_i\right] g(\mathbf{w}_i)/p(\mathbf{r}_i)\right\} \\ =& E\left\{P(s_i = 1|\mathbf{w}_i, \mathbf{r}_i) g(\mathbf{w}_i)/p(\mathbf{r}_i)\right\} \\ =& E\left\{p(\mathbf{r}_i) g(\mathbf{w}_i)/p(\mathbf{r}_i)\right\} = E\left[g(\mathbf{w}_i)\right] \end{aligned}$$

Intuitively, the IPW puts more weight on people who look similar to those who have selected out of the sample—so that they are effectively standing in for the people who selected out.

In order to implement the IPW, we need to estimate the selection probabilities, $p(\mathbf{r}_i)$. To account for the fact that only women who suffer a pregnancy loss have the opportunity to select out, we write $s_i = z_i + (1 - z_i)\tau_i$ where

$$\tau_i = \begin{cases} 0 & \text{if we do not observe a birth post-pregnancy loss} \\ 1 & \text{if we do observe a birth post-pregnancy loss.} \end{cases}$$

Since we can include z_i in $\mathbf{r}_i$, we have:

$$P(s_i = 1|\mathbf{r}_i) = E(s_i|\mathbf{r}_i) = z_i + (1 - z_i)E(\tau_i|\mathbf{r}_i)$$

Therefore, we need to estimate $E(\tau_i|\mathbf{r}_i) = P(\tau_i = 1|\mathbf{r}_i) = p_\tau(\mathbf{r}_i)$, and then plug it into our expression for s_i. In practice we estimate the probabilities by probit using a rich set of pre-event labor market — effectively pre-pregnancy lags of our key outcome variables — and health variables. By including pre-pregnancy labor market variables, we can control for many of the unobservable factors that may affect post-pregnancy labor 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).

Given the probit fitted values $\hat{p}_\tau(\mathbf{r}_i)$ we can write the estimated selection probabilities as:

$$\hat{p}(\mathbf{r}_i) = z_i + (1 - z_i)\hat{p}_\tau(\mathbf{r}_i) \tag{3}$$

and use $1/\hat{p}(\mathbf{r}_i)$ as the weights in our regressions.[21] By combining the Nevo and Rosen bounding exercise with the IPW, we can estimate two-sided bounds on the effect of delaying motherhood that account for both sources of potential bias found in the prior work.

It is worth noting that we expect the IPW correction for post-pregnancy loss selection to be more important for the IV than the OLS estimates. Intuitively, this is because the pregnancy loss sample is crucial to identification in the former while it makes up a small portion of the total sample in the latter. In Appendix A, we provide a stylized example to formalize this intuition. Finally, note that the Nevo and Rosen upper bound estimate, β_v^{IV}, is a weighted average of the original IV and OLS estimates. Therefore, we expect the upper bound to be less sensitive to post loss selection than the lower bound— which is the

[21]See Appendix C for a discussion of the IPW estimates.

IV estimate β_z^{IV}— as well.

6 Results

We begin by presenting our baseline results on the effects of age at first birth on labor market outcomes. In each case we estimate the following specification.

$$y_i = \beta A_i + \mathbf{x}_i \gamma + u_i \tag{4}$$

where y_i is one of several labor market outcomes discussed below. A_i is age at first birth and $\mathbf{x}_i$ includes dummies for year of birth and dummies for age at labor market entry. Since we only include women in the sample who we observe working before the first event (and control for age at labor market entry) age at first birth is equivalent to measuring birth timing relative to labor market entry, the measure suggested by Herr (2016).

As the discussion in Section 2 highlighted, delaying first birth could be the result of women optimizing over the life cycle and considering expected career earnings as one important variable. Delay could lead to higher pre-birth earnings since women might benefit from the higher childless earnings for a larger part of their career. It could also lead to higher earnings after birth if delay allows them to find better job matches and invest in general and job-specific human capital. In terms of relative importance Figure 1 highlighted the possibility that a large part of the returns to delaying motherhood may come from delaying the realization of the child penalty. To shed light on the relative importance of the different benefits of delay, we look at outcomes before birth, after birth, and cumulative outcomes over the entire period covered by our social security data. The evidence suggests that much of the labor market returns to delay are, indeed, concentrated before birth. This suggests that a conservative estimate of the return to delay is well approximated by the return to experience before birth. To make this point more directly, we also present earnings profiles in an event study framework where we account for the returns to experience while flexibly controlling for age.

6.1 Medium-run: Five Years After Birth

Table 2 presents results for outcomes 5 years after first birth providing medium-run measures of labor market outcomes and accumulated human capital.[22] The left and right panels display our selection-corrected OLS and IV estimates and corresponding Nevo-Rosen bounds, using IPW weights and unweighted, respectively. According to the OLS estimates each year of delay is associated with large increases in yearly and daily earnings (11.9% and 6.0%, respectively), yearly days employed (5.1 days per year), the probability of being employed (1.4 percentage points) or working in a white collar job (2.0 percentage points) relative to a blue-collar job or not working at all. With the pregnancy loss instrument, the IV estimates are much smaller for all outcomes. We cannot reject at the 5% level that earnings or the probability of being employed or working in a white-collar job 5 years after first birth are not affected by delay. We can marginally reject zero for yearly days employed. Unlike the prior literature that tends to find positive effects on medium-run wage outcomes (Miller, 2011; Herr, 2016) our conservative lower bounds raise the possibility that delay may have little affect on post-birth outcomes.[23] The weighted IV estimates are smaller than the unweighted estimates for all outcomes — therefore not accounting for post-loss selection into motherhood leads to lower bound estimates that are too high. This effect is even more pronounced when we look at cumulative outcomes below.

The previous measures reflect career outcomes that capture different aspects of human capital accumulation that might benefit from delaying motherhood. However, we could also imagine women benefiting in terms of other non-earnings based job characteristics. For ex-

[22] In order to have outcome measures five years after first birth by the end of our data window in 2017, we re-classify 129 out of 4,561 women (about 3%) in group 2 (pregnancy loss and subsequent birth) who have not given birth by the end of 2012 for this analysis only. We, instead, assign these women to group 3 before estimating the IPW weights. Excluding them entirely leads to almost identical estimates.

[23] We explore the robustness of these results to excluding the age at labor market entry dummies, which are potentially related to pregnancy loss through educational choice and adverse health in Table A.3. The OLS estimates are somewhat larger when excluding the age at labor market entry dummies while the IV estimates are barely affected.

ample, delay may allow them to find a job match that is more convenient for combining work and motherhood after birth at a similar or even lower wage. This compensating differential is difficult to measure directly in administrative data. Instead, we look at a proxy measure: whether women return to their pre-birth firm within 5 years after birth. In Austria during the years we study, women are entitled to return to their pre-birth job within two years after birth.[24] This job protection can be seen as an insurance policy that allows women to look for better jobs after birth or return to the same job should the search prove unsuccessful. Locking in a better job match before birth thus has the direct benefit of holding a better insurance policy, not just in terms of wage but also other amenities like hours flexibility or a short commute. A higher likelihood of returning to the pre-birth job could reflect a better job match, even if it does not translate into higher wages. That said, we need to be careful in interpreting any changes in the probability of returning to the same employer since it could also be a sign of a failed search for a better job. When we consider the probability that the mother returns to the pre-birth firm within 5 years after birth, we find a positive and statistically significant effect of delay of 1.2 percentage points per year of delay for the weighted IV estimate. The OLS estimate is again larger (1.8 pctp) suggesting that selection operates in a similar way as for earnings and is consistent with delay allowing women to find a better match – even along non-wage dimensions — if they delay birth.

6.2 Pre-birth Outcomes

While the selection-corrected lower bound confidence intervals for many of the post-birth outcomes included zero, these reflect only part of the return to delay. According to the descriptive evidence in Figure 1 a large component of the return to delay could simply be that women benefit from higher earnings before birth along an experience profile, and delaying birth thus allows them to enjoy these higher earnings for longer. Here, we document these returns by considering how age at first birth affects pre-birth labor market outcomes.

[24]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 two years.

Table 3 presents the weighted and unweighted NR bounds for two earnings outcomes, log total pre-birth earnings in the social security data (starting in 1972 when the oldest women in the sample were 13 years old) and highest observed daily earnings before birth. We find large positive effects on pre-birth earnings with tight bounds. Our weighted NR bounds reflect returns of 15.6%-16.9% for total earnings and 3.1%-4.7% for highest daily earnings per year of delay. Both effects are consistent with women moving along an upward-sloping earnings profile before first birth.

Similarly, we find large positive effects on total days of employment with weighted bounds of 10.2%-11.0% per year of delay. The lower bound corresponds to about 7 months more experience per year of delay. Exploiting the fact that we have firm identifiers we also estimate the effect of delay on highest pre-birth firm-level earnings as a rough proxy for firm quality. For this outcome we construct a "firm wage" calculating the average firm-level daily earnings on reference days for the universe of private-sector firms in Austria and take logs. Our NR bounds for highest firm daily earnings imply an increase of 3.2%-3.9% per year of delay. The lower bound is similar to the highest daily earnings observed at the worker level which is consistent with the idea that women climb a job ladder before birth and achieve higher earnings in part by moving to better paying firms. We also consider the probability that a woman ever works in a white-collar job before birth. These jobs are on average better paid and could allow women to better combine work and family. Our weighted bounds for the effect of delay on the probability of white-collar employment are 0.9-2.0 percentage points per year of delay.

It is worth noting the contrast between the strong, positive effects of delay captured by the tightly estimated bounds for pre-birth outcomes and the wider bounds that include very small or even negligible effects on post-birth outcomes from the previous section. First, the wider bounds for post-birth outcomes likely reflect the larger scope for bias in the OLS estimates when considering post-birth rather than pre-birth measures. Simply put, concerns over unobserved heterogeneity driving both the fertility timing choice and labor market outcomes are more important when considering what happens after birth — consistent with the similar trajectories before birth for women with different age at first birth choices in Figure 1. Second, the fact that small post-birth effects are included in the bounds may reflect recent evidence that the child penalty is particularly large and persistent in Austria (Kleven et al., 2019). In particular, Kleven et al. (2019) find that a substantial share of women stay at home until the youngest child starts attending kindergarten at age 3 and those who return often reduce their hours of work — consistent with the idea of a "mommy track" (Miller, 2011). Our lower bound results suggest that the severity of the child penalty may not be considerably moderated by delaying first birth given the institutional context. Our event study analysis in Section 6.4 will further explore the persistence of such effects up to ten years after birth.

6.3 Cumulative Outcomes

The previous sections helped characterize different aspects of the returns to delay by separately considering outcomes before and after first birth. Here, we summarize the combined effect of age at first birth over the period covered by our data. Table 4 shows post-loss-selection-corrected IPW and unweighted NR bounds for cumulative outcomes. Our weighted bounds for the effect of age at first birth on total earnings ever observed in our data window (1972-2017) are 4.3%-5.4% per year of delay. The lower bound combines the positive pre-birth effects and the negligible or even negative effects on post-birth outcomes uncovered previously.

When considering cumulative outcomes up to 2017, it is important to note that this will measure outcomes at different points in the career for women from different cohorts. In addition, while we have at least ten years of labor market data after the first event (live birth or loss recorded between 1999-2007), the first birth for the small subset women who suffer a loss and experience the longest delays (see Figure 2) may occur close to the end of our data window. This may partially censor the initial child penalty, leading to larger positive effects of delay than if the full penalty had been observed for these women. Ideally one would observe total career earnings, but due to the structure of our data we do not observe many women up to retirement age. Note that this is also an issue in related work, with many papers considering outcomes only up to a certain age—for example age 34 in Miller (2011) and an average age of 42 in Herr (2016). We experimented with different cumulative outcomes such as total earnings up to 20 years after labor market entry, which can be calculated for all women in our sample, but does exacerbate the problem of censoring some of the child penalty for the women who delay the longest. As a result, this leads to slightly larger estimated effects.[25] On balance we decided to incorporate the information we have in the data up to 2017, when the median age of our sample is 41 and the oldest women are 58 years old.

Our bounds imply that each year of delaying first birth leads to 1.3%-1.8% more days of employment and an 0.9-2.3% increase in highest daily earnings. The latter is about half of what we estimated before birth. Additionally, each year of delay leads to 0.6%-1.1% higher average firm-level daily earnings and a percentage point increase of 0-0.7 for the probability to ever work in a white-collar job. We cannot reject at the 5% level that delay has no effect on the probability to ever work in a white-collar job.

Finally, we note that the correction for post-loss selection into the sample of mothers proves particularly important for these cumulative outcome measures, consistent with non-random selection into motherhood after suffering a pregnancy loss. For example, the selection corrected IPW upper bound for the effect of delay on total earnings up to 2017 falls below the uncorrected lower bound. Given the argument that the IV estimate may be downward biased by the negative mental health effects of pregnancy loss, the fact that we find a much smaller effect when accounting for post-pregnancy loss selection is non-trivial. This also

[25]Based on the evidence in Section 6.4 that the short and medium-run returns to delay may be quite small, if the long-run effects later in a women's career are also negligible, then stopping at 20 years after labor market entry may overstate the overall returns to delay since delay means a larger part of the child penalty falls outside this window. In contrast, if differences by age at first birth develop in the long-run, then returns within 20 years after entry understate the true returns.

suggests that some of the results found in the prior literature that did not account for this selection may have been too large. Further, once we account for the post-loss selection by IPW, the contrast between the upper and lower bound estimates becomes larger, suggesting that, unlike the unweighted case, there may be a fair degree of selection driving the OLS results.

6.4 Event History Around First Birth

The previous bound estimates suggest positive effects of delaying motherhood both before birth and cumulatively over a substantial portion of the career— however, the lower bounds for outcomes in the medium run after the first birth could not rule out a small or even no effect of delay. If the lower bounds are correct, this suggests that delay mainly operates through allowing women to move along a steeper earnings profile before birth for longer. To make this point more directly, we move from estimating cross-sectional regressions to regressions tracing out the earnings profile for women with different ages at first birth — where we will also consider how controlling for age effects impacts these profiles. Specifically, we start by estimating the following event history type regressions by OLS and 2SLS using the IPWs for annual earnings relative to first birth:

$$y_{it} = \alpha + \gamma A1B_i + \sum_{t=-5}^{10} [\beta_t D_{it} + \delta_t D_{it} \times A1B_i] + \mathbf{Cohort_i}\eta + u_{it} \quad (5)$$

where

t measures time relative to first birth

y_{it} is annual earnings at time t

$A1B_i$ is the mother's age at first birth

D_{it} is a dummy variable indicating earnings measured at time t

$\mathbf{Cohort_i}$ is a vector of mother's year of birth cohort dummies.

Following the IV procedure from the previous section, when estimating by 2SLS we use a pregnancy loss indicator as well as interactions between the pregnancy loss dummy with the event time dummy variables as instruments for $A1B_i$ and the $D_{it} \times A1B_i$ interactions. Note that equation (5) allows for very flexible estimation of the baseline earnings profile from five years before to ten years after first birth with separate indicators for each time period. Given that the instruments used in the 2SLS estimation are based on the single indicator for suffering a pregnancy loss, the effect of age at first birth on the relative time specific earnings is constrained to be linear at each age. However, it does allow the effect of age at first birth to differ by the relative timing around the event (i.e. the effect of $A1B$ is allowed to be different one year before the event compared to one year after).

In Figure 5 we use $\hat{\beta}_t$ and $\hat{\delta}_t$ to generate the predicted annual earnings for women with an age at first birth of 27 or 30 holding the mother's year of birth effect constant at the 1975 cohort level.[26] In panel (a), we see higher annual earnings in the years prior to first birth for the profile based on an age at first birth of 30 compared to an age at first birth of 27 when estimating the delay effects by OLS. Both predictions suggest similar earnings right after the birth likely due to most women in our sample being out of the labor force directly after birth regardless of their age. We do see a steeper slope post-birth for the $A1B = 30$ prediction which— if correct— would suggest a faster recovery for women who delay, which is consistent with the motivating descriptive evidence in Figure 1. Moving to the IV based predictions in panel (b), we see a smaller difference in earnings prior to birth than in panel (a). Also in contrast to panel (a), we see very similar slopes after birth for both the age 27 and 30 predictions. Appealing to the bounding idea from the previous section,[27] this

[26] We use ages 27 and 30 for predicted earnings as these age correspond to the average age at first birth in the no loss and pregnancy loss groups.

[27] We avoid a strict bounding interpretation here as it would require the NR assumptions to hold at every time period relative to the first birth. For instance, while it may be reasonable for age at first birth to be positively correlated with unobservables driving a cumulative earnings measure it may not be at every point around the birth. Perhaps women who choose to delay longer also have higher savings— or a partner with higher income— allowing them to stay out of the labor force longer after a birth. This could lead to a negative correlation between age at first birth and the unobservables driving earnings one year after birth which would violate assumption NR3.

suggests that the faster recovery observed in panel (a) and in Figure 1 may be driven by the selection issues outlined previously, rather than reflecting an actual effect of delay on the post-birth profile.

Importantly, the higher annual earnings prior to first birth for those who delay may result from several different factors. First, annual earnings for women who delay will be measured at older ages and will necessarily be higher as long as earnings are increasing in age for childless women. Women may also use the time to increase effort in searching for better job matches or investing in general or job-specific human capital. Such choices may lead to earnings growth above what would be expected based on the typical childless earnings-age profile and they might systematically change how their careers respond post-birth.

To understand the extent that the improved earnings before birth can be explained by simply moving along the typical childless profile longer, we would like to control for age-specific earning effects in the event history regressions and see if there is any additional effect of age at first birth conditional on age. However, differences in age at first birth will be collinear with the set of age and time to first birth indicators.[28] Therefore, we consider reduced form event history regressions replacing $A1B_i$ in equation (5) with the pregnancy

[28] Consider the age 29 ($Age_{29} = 1$) earnings for the women one year away from first birth ($D_{i,-1} = 1$)—they will all have the same age at first birth of 30.

loss instrument.

$$y_{it} = \alpha + \gamma Loss_i + \sum_{t=-5}^{10} [\beta_t D_{it} + \delta_j D_{it} \times Loss_i] + \mathbf{Cohort_i}\eta + \mathbf{Age}_{it}\gamma + u_{it} \quad (6)$$

where

t measures time relative to first birth

y_{it} is annual earnings at time t

$Loss_i$ is an indicator for suffering a pregnancy loss at the first observed event

D_{it} is a dummy variable indicating earnings measured at time t

$\mathbf{Cohort_i}$ is a vector of mother's year of birth cohort dummies

$\mathbf{Age}_{it}$ is a vector of age dummies.

The pregnancy loss indicator will not suffer from the same collinearity problem as age at first birth, allowing us to identify differences in earnings for those who suffer a loss— and experience the accompanied delay— conditional on both age and event time dummies.

The reduced form profiles in panel (c) are similar to the IV profiles by design[29]— showing an annual earnings gap before birth and overlapping profiles after. However, when we control for age effects in panel (d) we see the pre-birth differences for the loss and no loss groups disappear. This is consistent with delay allowing women to move along the childless earnings-age profile without showing an additional effect of delay over this period. We do see a slightly larger drop after birth for the loss group once we control for age, however this difference disappears over time. Across panels (b), (c), and (d)— estimates reflecting our lower-bound on the event history profiles— we see little evidence of a difference in profiles after birth for those who delay. This suggests that the mechanisms that are expected to work through post-birth earnings— such as those found in the model of fertility and job search in Erosa

[29]We chose to generate predicted yearly earnings at ages 27 and 30 based on the estimates of equation (5) to allow for a straightforward comparison to the predictions based on the reduced form equation (6) since the mean age at first birth rounded to the nearest whole number in our no loss and loss groups are ages 27 and 30, respectively.

et al. (2002) that would lead to higher average earnings through lower turnover post-birth for those who delay longer (higher human capital, both general and job-specific, and better job matches)— are potentially not strong drivers of women's earnings here.

6.5 Interpreting the Bounds

While the lower bound estimates raise the possiblity that the key mechanism driving returns to delay comes from delaying the child penalty, it is important to note that the upper bound estimates are consistent with larger medium-run effects that could be due to real differences in human capital or job match after the first birth. Ultimately, the true effects lie somewhere in between depending on the relative endogeneity of the instrument— which can be measured by the ratio of the correlation between the no-loss instrument and the unobservables to the correlation between age a first birth and the unobservables, denoted $\lambda = \rho_{\tilde{z}u}/\rho_{\tilde{A}u}$. The bounds focus attention on a best case in which the instrument is exogenous ($\lambda = 0$) and a worst case in which it is just as endogenous as age at first birth ($\lambda = 1$). To provide additional context for interpreting the reported bounds, we can derive the associated effect that would be consistent with any choice of λ — denoted $\beta(\lambda)$. Importantly, $\beta(\lambda)$ is not linear in λ, rather, from equation (7) in Nevo and Rosen (2012) we can write

$$\beta(\lambda) = \frac{\sigma_{\tilde{A}}\sigma_{\tilde{z}\tilde{y}} - \lambda\sigma_{\tilde{z}}\sigma_{\tilde{A}\tilde{y}}}{\sigma_{\tilde{A}}\sigma_{\tilde{A}\tilde{z}} - \lambda\sigma_{\tilde{z}}\sigma^2_{\tilde{A}}} \tag{7}$$

Figure 6 displays $\beta(\lambda)$ for $0 \leq \lambda \leq 1$ for two of our key outcome measures: log daily earnings five years after first birth and total earning up to 2017. In both cases, the relationship between the appropriate estimate and the degree of endogeneity of the instrument is concave — so that small deviations from assuming an exogenous instrument initially results in larger increases to the implied consistent estimate. In this case, even a moderate level of endogeneity ($\lambda = 0.2$) is associated a considerably larger effect of delay on daily earnings post-birth (3.39%) than the lower bound (1.39%) would suggest. This pattern provides helpful context for the careful interpretation of the bounds presented in this paper.

7 Conclusion

We consider the potential labor market returns to delaying motherhood for women. Using administrative social security and health data from Austria, we are able to document these returns in terms of earnings, employment, and the quality of the firm (measured by average pay) for an individual. Our estimates are based on using pregnancy loss as an instrument for age at first birth where we modify the empirical approach to account for several concerns with the pregnancy loss instrument. This leads us to estimating plausible bounds for the returns to delay that account for 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. Throughout, we find accounting for the positive selection into the pregnancy loss sample to be particularly important— with corrected estimates consistently smaller than the uncorrected ones.

While the previous literature tends to find large positive effects of delay, we see little difference in our lower bound earnings trajectories before and after the birth of the first child. Instead, the largest returns seem to develop before the first birth due to delaying the realization of the child penalty. Together, our results raise the possibility that the cumulative returns to delay may be mostly driven by the mechanical effect of delaying the realization of the child penalty rather than altering how a woman's career responds. This suggests that many of the mechanisms discussed in the literature— increased attachment due to higher human capital or job match quality— for the improved outcomes for those who delay may not necessarily hold on average.

References

Abadie, A., A. Diamond, and J. Hainmueller (2010). Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program. *Journal of the American Statistical Association 105*(490), 493–505.

Angrist, J. and J. Pischke (2009). *Mostly Harmless Econometrics: An Empiricists' Companion*. Princeton, NJ: Princeton Univerisity Press.

Ashcraft, A., I. Fernández-Val, and K. Lang (2013). The consequences of teenage childbearing: Consistent estimates when abortion makes miscarriage non-random. *The Economic Journal 123*(571), 875–905.

Bailey, M. J. (2006). More power to the pill: the impact of contraceptive freedom on women's life cycle labor supply. *The Quarterly Journal of Economics*, 289–320.

Bailey, M. J., B. Hershbein, and A. R. Miller (2012). The opt-in revolution? Contraception and the gender gap in wages. *American Economic Journal: Applied Economics 4*(3), 225–254.

Berndt, E. R., S. N. Finkelstein, P. E. Greenberg, R. H. Howland, A. Keith, A. J. Rush, J. Russell, and M. B. Keller (1998). Workplace performance effects from chronic depression and its treatment. *Journal of Health Economics 17*(5), 511–535.

Black, S. E. and C. Juhn (2000). The rise of female professionals: Are women responding to skill demand? *The American Economic Review 90*(2), 450–455.

Bratti, M. and L. Cavalli (2014). Delayed first birth and new mothers' labor market outcomes: Evidence from biological fertility shocks. *European Journal of Population 30*(1), 35–63.

Buckles, K. (2008). Understanding the returns to delayed childbearing for working women. *The American Economic Review 98*(2), 403–407.

Caucutt, E. M., N. Guner, and J. Knowles (2002). Why do women wait? Matching, wage inequality, and the incentives for fertility delay. *Review of Economic Dynamics 5*(4), 815–855.

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

Chatterji, P., M. Alegria, and D. Takeuchi (2011). Psychiatric disorders and labor market outcomes: Evidence from the national comorbidity survey-replication. *Journal of Health Economics 30*(5), 858–868.

Cristia, J. P. (2008). The effect of a first child on female labor supply evidence from women seeking fertility services. *Journal of Human Resources 43*(3), 487–510.

Erosa, A., L. Fuster, and D. Restuccia (2002). Fertility decisions and gender differences in labor turnover, employment, and wages. *Review of Economic Dynamics 5*(4), 856–891.

Fletcher, J. M. and B. L. Wolfe (2009). Education and labor market consequences of teenage childbearing evidence using the timing of pregnancy outcomes and community fixed effects. *Journal of Human Resources 44*(2), 303–325.

Fogli, A. and L. Veldkamp (2011). Nature or nurture? Learning and the geography of female labor force participation. *Econometrica 79*(4), 1103–1138.

Goldin, C. and L. F. Katz (2002). The power of the pill: Oral contraceptives and women's career and marriage decisions. *Journal of Political Economy 110*(4).

Hackl, F., M. Hummer, and G. J. Pruckner (2015, September). Old boys' network in general practitioners' referral behavior? *Journal of Health Economics 43*, 56–73.

Herr, J. L. (2016). Measuring the effect of the timing of first birth. *Journal of Population Economics 29*(1).

Hotz, V. J., S. W. McElroy, and S. G. Sanders (2005). Teenage childbearing and its life cycle consequences exploiting a natural experiment. *Journal of Human Resources 40*(3), 683–715.

Hotz, V. J., C. H. Mullin, and S. G. Sanders (1997). Bounding causal effects using data from a contaminated natural experiment: Analysing the effects of teenage childbearing. *The Review of Economic Studies 64*(4), 575–603.

Juhn, C. and S. Potter (2006). Changes in labor force participation in the United States. *The Journal of Economic Perspectives 20*(3), 27–46.

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., C. Landais, J. Posch, A. Steinhauer, and J. Zweimüller (2019). Child penalties across countries: Evidence and explanations. *forthcoming American Economic Association Papers and Proceedings*.

Kleven, H., C. Landais, and J. E. Søgaard (2019). Children and gender inequality: Evidence from Denmark. *forthcoming American Economic Journal: Applied Economics*.

Kuhn, A., R. Lalive, and J. Zweimüller (2009, December). The public health costs of job loss. *Journal of Health Economics 28*(6), 1099–1115.

Lalive, R., A. Schlosser, A. Steinhauer, and J. Zweimüller (2014, 09). Parental Leave and Mothers' Careers: The Relative Importance of Job Protection and Cash Benefits. *The Review of Economic Studies 81*(1), 219–265.

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

Miller, A. R. (2011). The effects of motherhood timing on career path. *Journal of Population Economics 24*(3), 1071–1100.

Nevo, A. and A. M. Rosen (2012). Identification with imperfect instruments. *Review of Economics and Statistics 94*(3), 659–671.

Peng, L., C. D. Meyerhoefer, and S. H. Zuvekas (2015). The short-term effect of depressive symptoms on labor market outcomes. *Health Economics*.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

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

A Implications of Post-Pregnancy Loss Selection for OLS and IV

To highlight the importance of the observed selection out of the motherhood sample after a pregnancy loss for our OLS and IV estimates of the labor market returns to delayed motherhood, we appeal to a very stylized example. We will consider estimating the effect of delay on a labor market outcome y_i. For exposition we consider measuring delay as a binary indicator for delaying first birth, so that our treatment indicator, instrument, and selection indicator can be written as:

$$A_i = \begin{cases} 0 & \text{if an early first birth} \\ 1 & \text{if a delayed first birth} \end{cases}$$

$$z_i = \begin{cases} 0 & \text{if first pregnancy ends in loss} \\ 1 & \text{if first pregnancy results in live birth} \end{cases}$$

$$\tau_i = \begin{cases} 0 & \text{if we do not observe a birth post-pregnancy loss} \\ 1 & \text{if we do observe a birth post-pregnancy loss} \end{cases}$$

For exposition we will assume that pregnancy loss forces a delay— a shift from $A_i = 0$ to $A_i = 1$. In this simple set up our OLS and IV estimators of the return to delay will be the following Wald estimators:

$$\beta_{OLS} = E\left[y_i|A_i = 1\right] - E\left[y_i|A_i = 0\right] \tag{8}$$

$$\begin{aligned} \beta_{IV} &= \frac{E\left[y_i|z_i = 1\right] - E\left[y_i|z_i = 0\right]}{E\left[A_i|z_i = 1\right] - E\left[A_i|z_i = 0\right]} \\ &= E\left[y_i|z_i = 0\right] - E\left[y_i|z_i = 1\right] \end{aligned} \tag{9}$$

Note, the simplification for β_{IV} follows from the fact that under the assumption that pregnancy loss forces a delay the denominator of the Wald IV estimator is $E\left[A_i|z_i=1\right] - E\left[A_i|z_i=0\right] = 0 - 1 = -1$. Next note that the first term in the OLS estimate represents the expected outcomes for women who delayed motherhood and can be rewritten as a weighted average of outcomes for those who chose to delay and those that were delayed by a pregnancy loss:

$$E\left[y_i|A_i=1\right] = Pr\left(z_i=1\right)E\left[y_i|A_i=1,z_i=1\right] + Pr\left(z_i=0\right)E\left[y_i|A_i=1,z_i=0\right]$$

Here, those with $A_i = 1$ and $z_i = 1$ are women who chose to delay birth and did not suffer a pregnancy loss. We can then express the expected outcome for the pregnancy loss group, $z_i = 1$, as a weighted average of the two selection types:

$$\begin{aligned} E\left[y_i|A_i=1\right] = {} & Pr\left(z_i=1\right)E\left[y_i|A_i=1,z_i=1\right] \\ & + Pr\left(z_i=0\right)\left\{Pr(\tau=1)E\left[y_i|A_i=1,z_i=0,\tau_i=1\right]\right. \\ & \left. +Pr(\tau=0)E\left[y_i|A_i=1,z_i=0,\tau_i=0\right]\right\} \end{aligned}$$

Note that by ignoring the selection issue, we effectively assume $E\left[y_i|A_i=1,z_i=0,\tau_i=1\right] = E\left[y_i|A_i=1,z_i=0,\tau_i=0\right]$. Denote the estimator under this assumption by $\tilde{\beta}_{OLS}$. In this simplified setting, the second term in the OLS estimate reflecting the average outcomes for those who have a child early, $E\left[y_i|A_i=0\right]$, includes only those that do not suffer a pregnancy loss and will, therefore, be the same for both estimators. The difference between the two can then be written as:

$$\begin{aligned} & \tilde{\beta}_{OLS} - \beta_{OLS} = \\ & \quad Pr\left(z_i=0\right)Pr(\tau=0)\left\{E\left[y_i|A_i=1,z_i=0,\tau_i=1\right] - E\left[y_i|A_i=1,z_i=0,\tau_i=0\right]\right\} \\ & \quad = Pr\left(z_i=0\right)Pr(\tau=0)\left\{E\left[y_i|z_i=0,\tau_i=1\right] - E\left[y_i|z_i=0,\tau_i=0\right]\right\} \end{aligned}$$

The importance for the OLS estimate therefore depends on how strong the selection on earnings is and the prevalence of pregnancy loss and post-pregnancy loss selection. In our sample, $Pr\left(z_i = 0\right) = 0.11; Pr(\tau = 0) = 0.19$ so that $Pr\left(z_i = 0\right) Pr(\tau = 0) = 0.021$.

In contrast, for the IV estimator, we can rewrite the first term representing the expected outcomes for those that suffer a pregnancy loss as a weighted average of the two selection types:

$$E\left[y_i|z_i = 0\right] = Pr(\tau = 1)E[y_i|z_i = 0, \tau_i = 1] + Pr(\tau = 0)E[y_i|z_i = 0, \tau_i = 0]$$

Again, ignoring selection effectively assumes that $E[y_i|z_i = 0, \tau_i = 1] = E[y_i|z_i = 0, \tau_i = 0]$. If we denote the IV estimator under this assumption as $\tilde{\beta}_{IV}$ then the difference in the IV estimate can be written as:

$$\begin{aligned}&\tilde{\beta}_{IV} - \beta_{IV} = \\ &\quad Pr\left(\tau = 0\right)\left\{E\left[y_i|z_i = 0, \tau_i = 1\right] - E\left[y_i|z_i = 0, \tau_i = 0\right]\right\}\end{aligned}$$

Here the same selection term is multiplied by $Pr(\tau = 0) = 0.19$ instead of $Pr\left(z_i = 0\right) Pr(\tau = 0) = 0.021$, implying that the post-pregnancy loss selection issue will have a larger impact on the IV estimate than the OLS.

B Education Level at First Birth

The mother's education information is recorded in the birth registry at the time of birth under the following categories: compulsory, some high school, high school, professional apprenticeship, teacher's college, and college. We thus only have information about the highest education achieved for women who ever give birth (groups 1 and 2). Table A.4 shows shares of education levels for women whose first event we observe is a live birth, and those whose first observed event is a pregnancy loss, followed eventually by a live birth.[30] We focus on the subsample of women with non-missing education information (about 75% of the main sample).

Importantly, since the education level is recorded at the time of the birth of a child, it reflects the investments in education made up to that point. This allows us to consider whether delaying motherhood may provide an opportunity to invest in human capital before having a child— at which point the opportunity costs of further investment may be higher. Given the categorical nature of the education variable, we model the probability of having each education level at the birth of the child by a multinomial logit (MNL):

$$P(E = j|A1B, x) = \frac{exp(A1B_i\gamma_j + \mathbf{x}_i\beta_j)}{1 + \sum_{h=1}^{J} exp(A1B_i\gamma_h + \mathbf{x}_i\beta_h)} \tag{10}$$

where $j = 1, ...J$ indicate the education level and $\mathbf{x}_i$ includes mother's cohort indicators (rounded to 5 year bins) and the age at entry to the labor market. We set the compulsory education level as the base category with the other categories generally reflecting higher educational investments and present estimated odds ratios for the age at first birth in Table A.5. We also present the joint test of whether the coefficients on age a first birth are equal to zero for all education levels ($\gamma_j = 0 \ \forall j$)—- or effectively whether the odds ratio relative to compulsory schooling is equal to one for each.

In column (1), we see that age at first birth is a significant predictor of education level— both for each category relative to compulsory school and jointly across category— with higher age at first birth associated with an increased probability of attaining higher educational credentials (odds ratio grater than one). While this relationship may be driven by a host of factors, it is consistent with human capital accumulation playing a role in the returns to delaying motherhood. In contrast, when we replace age at first birth with an indicator for suffering a pregnancy loss in column (2), the pregnancy loss indicator is not a statisti-

[30] Unfortunately, data quality is an issue with many missing values: 23.8% among group 1 and 43.6% among group 2. The birth registry data ends in 2007, which explains the higher share of missings among women whose first birth is delayed by pregnancy loss.

cally significant predictor of education level. This suggests that the sudden, surprise delay induced by the pregnancy loss is not likely to increase educational investments— indeed, the estimated odds ratios relative to compulsory schooling are all very close to or less than one. Therefore, using the pregnancy loss instrument for identification of delay effects may not capture an important mechanism for delay to improve labor market outcomes.

C Inverse Probability Weighting Details

We model the probability of being observed with a birth after a pregnancy loss, $p_\tau(\mathbf{r}_i)$, as a probit where $\mathbf{r}_i$ includes:

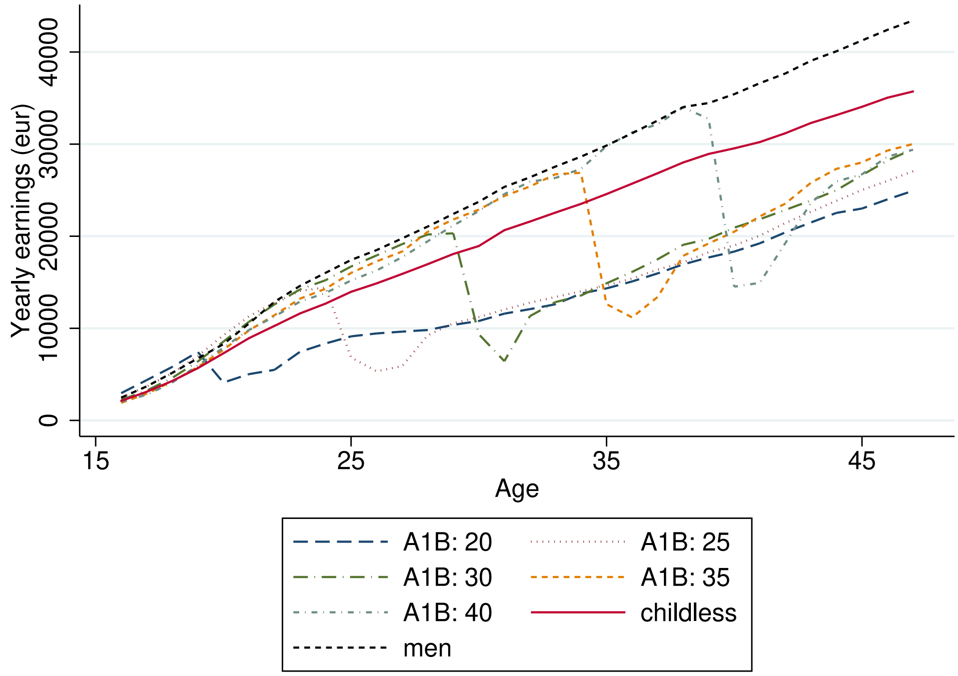
1. Indicators for the mother's birth cohort and year of the pregnancy loss

2. Pre-pregnancy annual histories: employment, earnings, health care utilization, depression diagnosis and medication use, psychological diagnosis and medication use

We then use the estimated probabilities, $\widehat{p_\tau}(\mathbf{r}_i)$ to generate our IPWs as $1/\hat{p}(\mathbf{r}_i)$ where $\hat{p}(\mathbf{r}_i) = z_i + (1 - z_i)\hat{p}_\tau(\mathbf{r}_i)$.

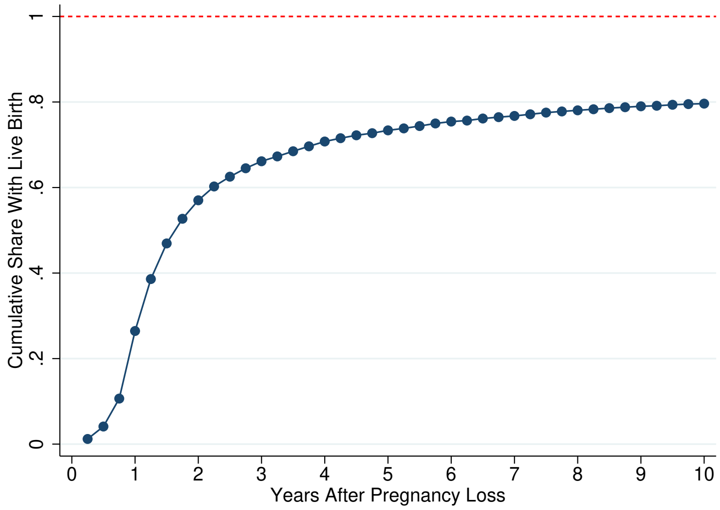
Figure A.1 provides evidence on how reweighting by the IPW may help account for the positive selection into the pregnancy loss sample used for estimation. Panel (a) presents four separate series of log earnings in the five years before the first pregnancy for women who suffer a pregnancy loss. Again we see the positive selection into the estimation sample with the unweighted earnings being higher for the subsequent birth sample than for the entire pregnancy loss group. Importantly, when we reweight the subsequent birth sample using the IPW, the series overlaps the pre-pregnancy earnings for the entire pregnancy loss group. Panel (b) shows a similar pattern for the mental health diagnoses before pregnancy. This suggests that weighting the subsequent birth group better reflects the experience of those that suffer a pregnancy loss— including those that will fall out of the natural experiment— thereby better matching the thought experiment behind the pregnancy loss IV approach.

Figure 1: Earnings and Age at First Birth (a1b): Women Born in Austria in 1970



Notes: This figure plots yearly social security earnings by age for women and men born in 1970. We classify this cohort of women into age at first birth groups by the first age at which we observe a birth in the social security data (birth events only available for women). For women classified as childless we do not observe a birth up to and including 2017. E.g. the blue dashed line (A1B: 20) includes all women born in 1970 who had their first child in 1990. Earnings are topcoded (see Section 3) for about 2% of employed women in the ASSD. We calculate age at first birth by year at which we observe the first live birth in the social security data minus year of birth. In this figure we include all of Austria.

Figure 2: Cumulative Share of First Births After Pregnancy Loss



Notes: The sample in this figure is all women who we observe with a pregnancy loss in 1999-2007 (see Section 3). We plot the share of this group who is ever observed with a live birth by year after the pregnancy loss event. The data include births up to and including 2017.

Figure 3: Fertility timing choices – conceptual framework

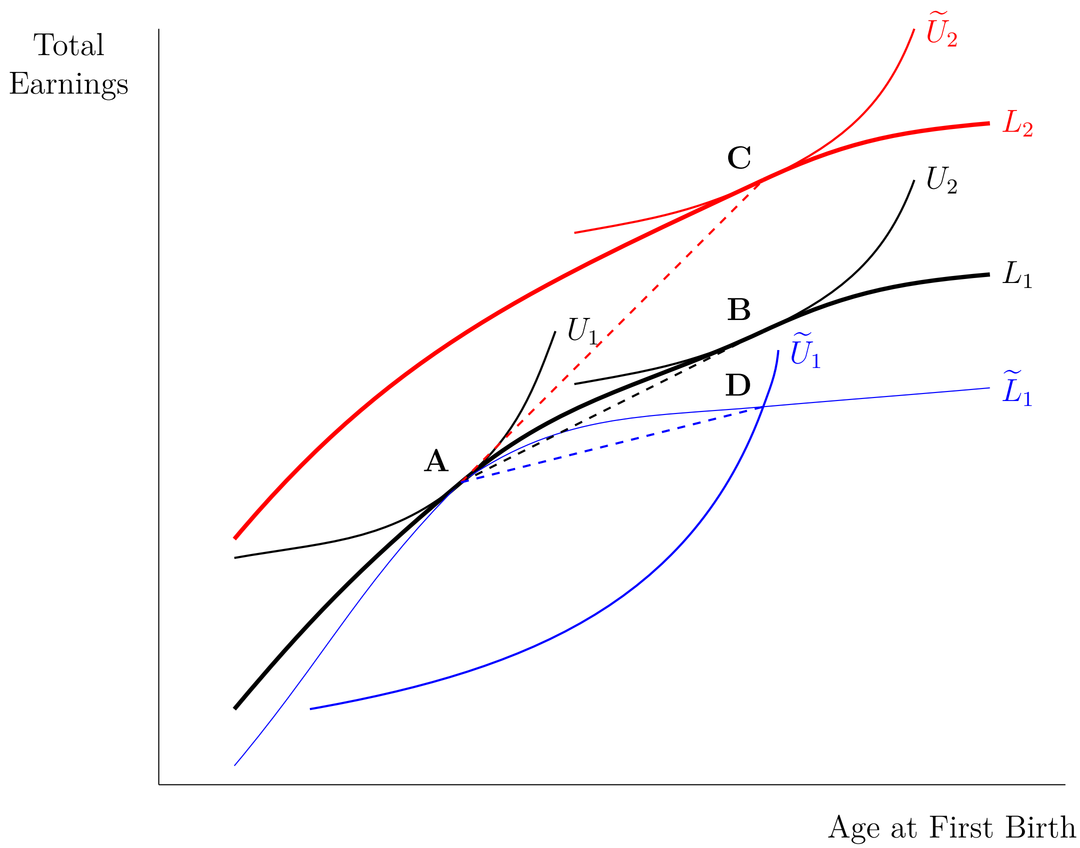
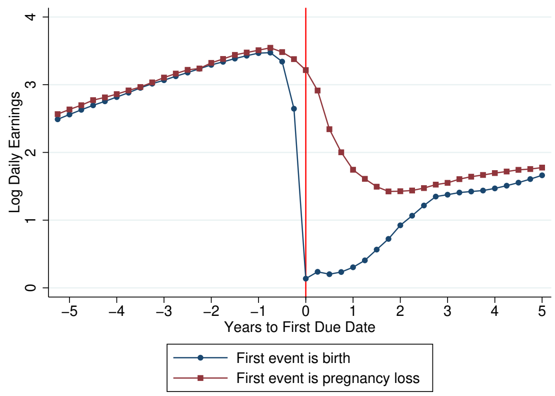
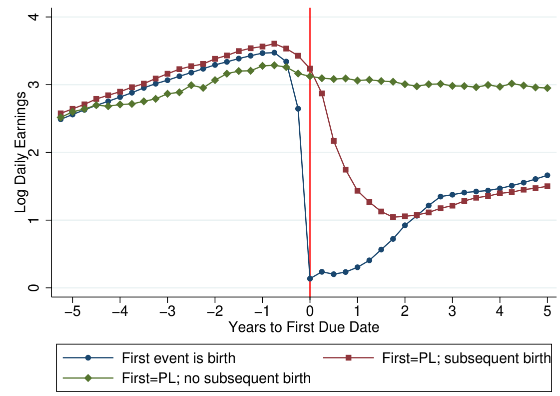


Figure 4: Earnings and Mental Illness by Years to First Due Date

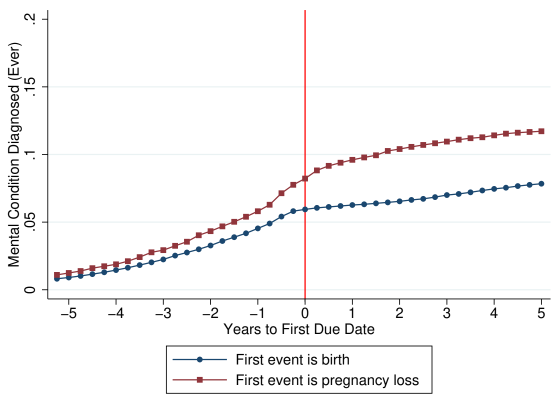
(a) Daily earnings by first event



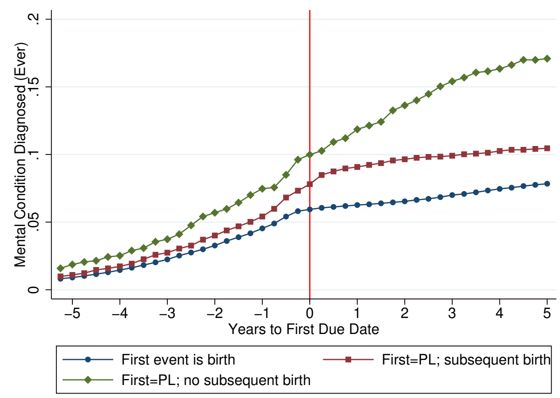
(b) Daily earnings by group



(c) Mental illness by first event

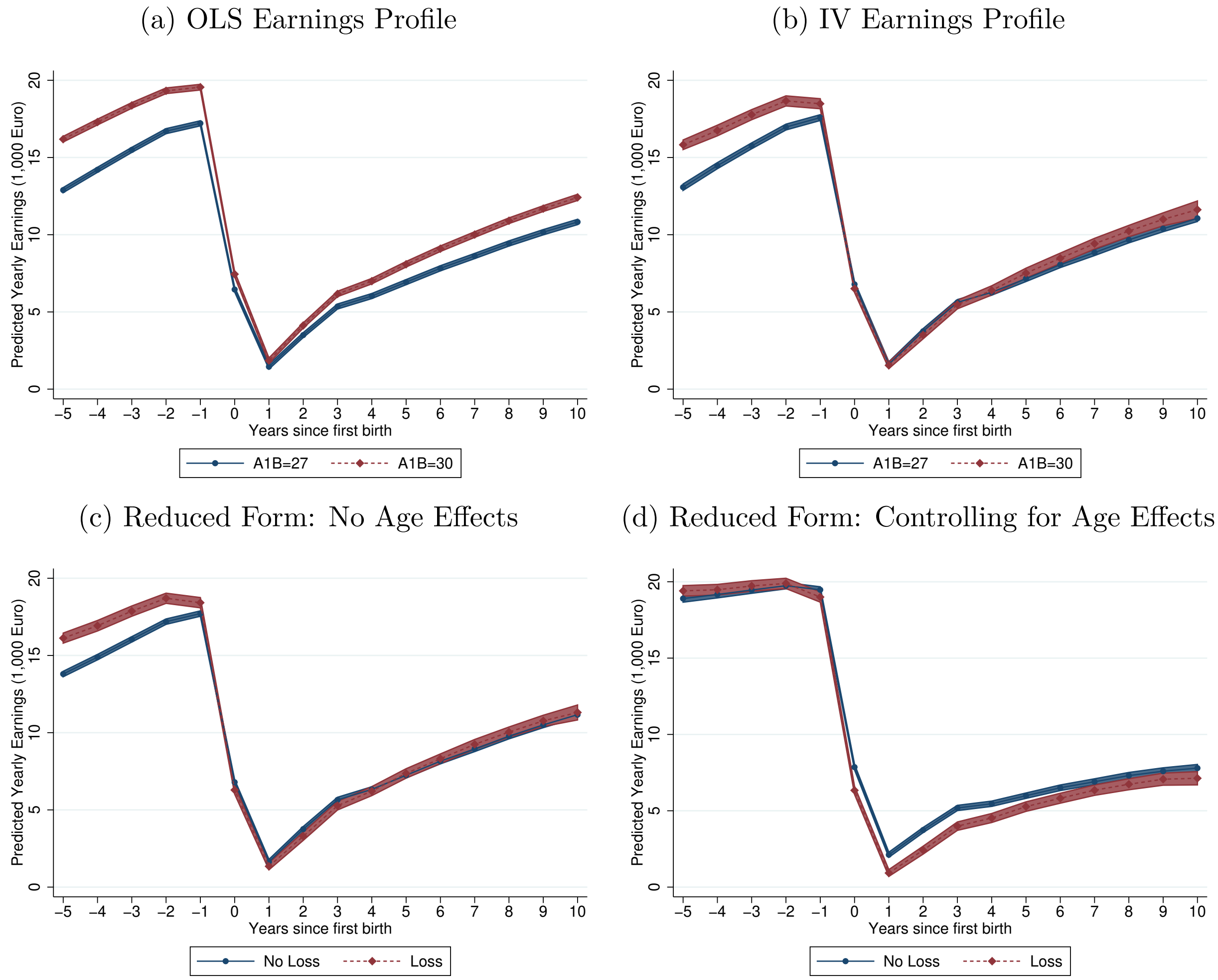


(d) Mental illness by group



Notes: 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 first pregnancy loss otherwise. Log daily earnings in real (CPI base 2000) euro, as of the middle of the reference quarter (e.g. 15th February for Q1) and set to zero if the woman is not employed. Ever diagnosed mental condition is set to one if a woman has ever been diagnosed with mental or behavioural conditions up to time t. "First=PL" means first observed event is pregnancy loss.

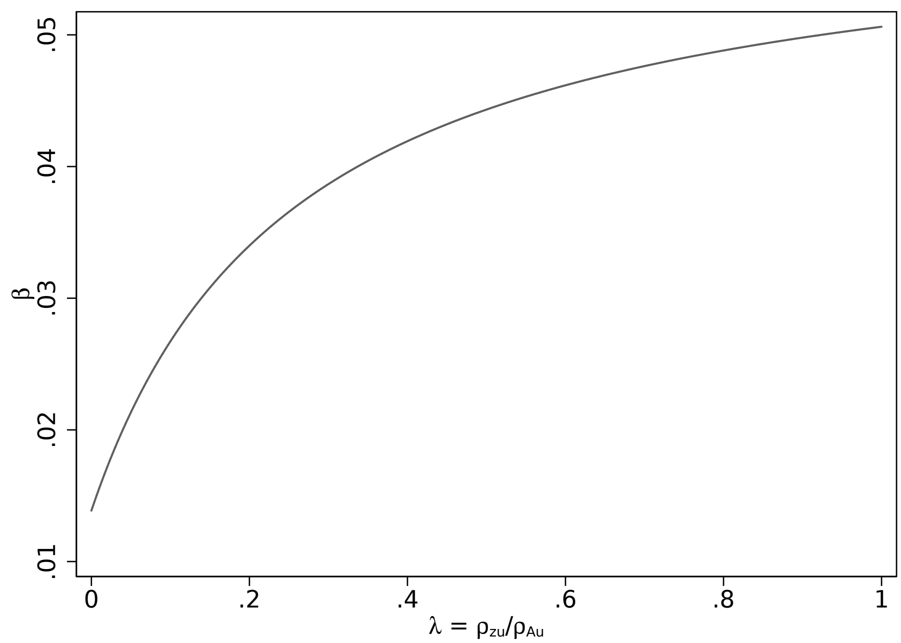
Figure 5: Estimated Earnings Profiles Relative to First Birth



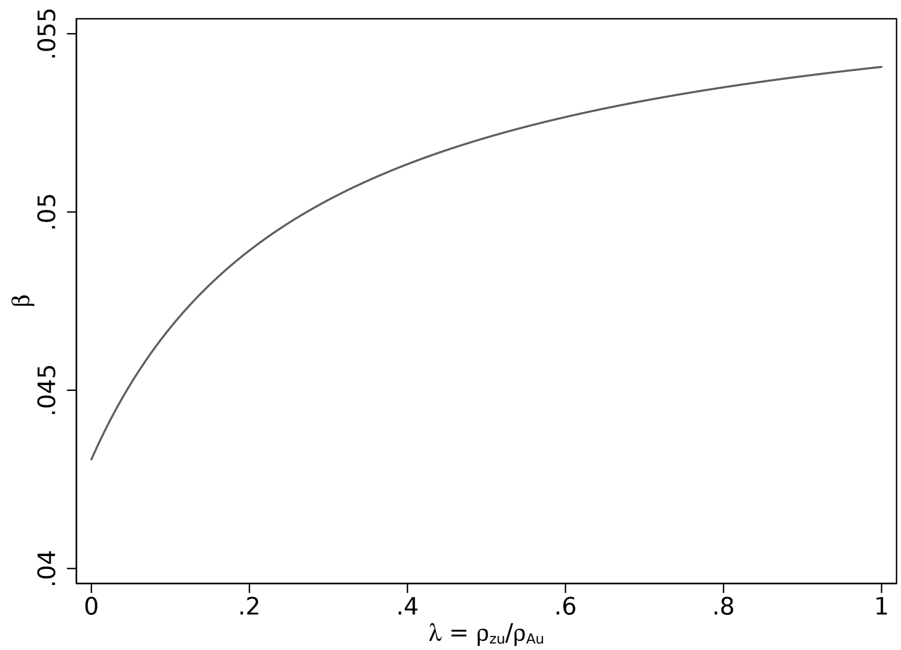
Notes: This figure plots predicted earnings setting year of birth to 1975 and age at first birth to 27 and 30 in panels (a) and (b). In panels (c) and (d) we predict for year of birth in 1975 and age 30. Shaded areas are 95% CIs. All regressions include dummies for year of birth and dummies for age at labor market entry.

Figure 6: Return to Delay Estimate as a Function of the Relative Endogeneity of the No-loss Instrument

(a) Log Daily Earnings Five Years Post-birth



(b) Log Total Earnings



Notes: Each panel depicts the estimated return to delay associated with an instrument that would be valid for different values of the endogeneity of the no-loss instrument relative to age at first birth, denoted λ, given by: $\beta(\lambda) = \frac{\sigma_{\tilde{A}}\sigma_{\tilde{z}\tilde{y}} - \lambda\sigma_{\tilde{z}}\sigma_{\tilde{A}\tilde{y}}}{\sigma_{\tilde{A}}\sigma_{\tilde{A}\tilde{z}} - \lambda\sigma_{\tilde{z}}\sigma^2_{\tilde{A}}}$. The range for β is the lower and upper bounds found in the main estimation tables.

Table 1: Descriptive Statistics

	Means and std devs			Differences and std errs		
		First event is pregnancy loss				
	First event is live birth (G1)	Subsequent birth (G2)	No subsequent birth (G3)	G1-G2	G1-G3	G2-G3
Age at first event	27.3317	27.3169	31.7386	0.0148	−4.4069	−4.4217
	(4.7456)	(5.0598)	(5.7376)	(0.0741)	(0.1474)***	(0.1764)***
Age at first birth	27.3317	29.5612		−2.2295		
	(4.7456)	(5.2224)		(0.0743)***		
Total number of births	1.9152	1.8459		0.0693		
	(0.8018)	(0.7809)		(0.0124)***		
Log total earnings	12.1615	12.2702	12.6025	−0.1087	−0.4409	−0.3322
	(0.8002)	(0.7458)	(0.9371)	(0.0123)***	(0.0248)***	(0.0267)***
Log total experience	8.6296	8.6617	8.8148	−0.0321	−0.1853	−0.1532
	(0.5276)	(0.4994)	(0.6588)	(0.0081)***	(0.0164)***	(0.0181)***
Log average daily earnings	3.6964	3.7400	3.8846	−0.0437	−0.1882	−0.1446
	(0.3417)	(0.3240)	(0.3683)	(0.0053)***	(0.0106)***	(0.0113)***
Log highest firm daily earnings	4.4361	4.4519	4.4731	−0.0157	−0.0370	−0.0212
	(0.2995)	(0.2954)	(0.3136)	(0.0046)***	(0.0093)***	(0.0102)**
Employed before event	0.8701	0.8891	0.8329	−0.0189	0.0373	0.0562
	(0.3361)	(0.3141)	(0.3733)	(0.0052)***	(0.0104)***	(0.0111)***
Log daily earnings before event	3.3379	3.4317	3.1627	−0.0938	0.1752	0.2690
	(1.4484)	(1.3651)	(1.5973)	(0.0224)***	(0.0449)***	(0.0480)***
Age at labor market entry	18.1099	18.1679	20.1298	−0.0580	−2.0199	−1.9618
	(3.6416)	(3.8087)	(5.8683)	(0.0567)	(0.1146)***	(0.1452)***
Log total healthcare utilization	8.1285	8.5106	8.1767	−0.3821	−0.0482	0.3339
	(1.2639)	(0.7038)	(1.2005)	(0.0190)***	(0.0390)	(0.0279)***
Log total healthcare utilization before event	5.1308	5.6926	5.4325	−0.5618	−0.3017	0.2602
	(2.7204)	(2.4285)	(2.8371)	(0.0418)***	(0.0842)***	(0.0853)***
Observations	46,437	4,561	1,071	50,998	47,508	5,632

Notes: Group 1—first event is live birth. Group 2—first event is pregnancy loss, subsequent live birth is observed. Group 3—first event is pregnancy loss, subsequent birth is not observed. All earnings and utilization values in real (CPI base 2000) euro.

First three columns show means and standard deviations in parentheses. Last three columns show differences in means between the groups and standard errors of the differences in parentheses. Significance tests against zero with t-test allowing for unequal variance. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 2: Outcomes 5 Years After First Birth

		IPW			Unweighted		
Dependent variable	Mean	OLS	IV	NR Bounds	OLS	IV	NR Bounds
Yearly earnings	6689.38	306.91	23.73	[23.73,248.44]	310.81	66.23	[66.23,265.63]
		(16.53)	(69.11)	{-111.73,292.38}	(16.07)	(69.16)	{-69.33,305.53}
Log yearly earnings	5.3916	0.1188	0.0158	[0.0158,0.0975]	0.1195	0.0295	[0.0295,0.1029]
		(0.0085)	(0.0367)	{-0.0562,0.1202}	(0.0084)	(0.0373)	{-0.0435,0.1239}
Yearly days employed	206.4375	5.1218	2.6341	[2.6341,4.6081]	5.1250	2.9596	[2.9596,4.7250]
		(0.3087)	(1.3182)	{0.0504,5.4244}	(0.3050)	(1.3551)	{0.3037,5.4870}
Daily earnings	18.3079	0.8800	-0.0240	[-0.0240,0.6933]	0.8912	0.0897	[0.0897,0.7431]
		(0.0492)	(0.1983)	{-0.4128,0.8202}	(0.0483)	(0.1990)	{-0.3003,0.8589}
Log daily earnings	1.6661	0.0602	0.0139	[0.0139,0.0506]	0.0604	0.0185	[0.0185,0.0527]
		(0.0035)	(0.0148)	{-0.0151,0.0599}	(0.0034)	(0.0151)	{-0.0110,0.0612}
Employed	0.5637	0.0143	0.0067	[0.0067,0.0127]	0.0143	0.0072	[0.0072,0.0130]
		(0.0009)	(0.0040)	{-0.0012,0.0152}	(0.0009)	(0.0041)	{-0.0009,0.0153}
White-collar job	0.3738	0.0202	0.0010	[0.0010,0.0163]	0.0207	0.0044	[0.0044,0.0177]
		(0.0009)	(0.0039)	{-0.0066,0.0187}	(0.0009)	(0.0040)	{-0.0035,0.0200}
Back to pre-birth firm	0.4988	0.0177	0.0116	[0.0116,0.0164]	0.0182	0.0153	[0.0153,0.0176]
		(0.0010)	(0.0040)	{0.0037,0.0189}	(0.0009)	(0.0041)	{0.0072,0.0200}

Notes: All regressions include dummies for year of birth and dummies for age at labor market entry. Robust standard errors in parentheses. NR bounds in brackets, 95% CI in braces. Women in group 2 with a live birth after 2012 ($n = 129$) are assigned to group 3 to have 5 years after first birth for everyone in the sample, $N = 50,869$. All earnings variables in real (CPI base 2000) euro. Yearly earnings are calendar year earnings in the year 5 years after first birth. Daily earnings and employment are as of the reference day (e.g. 15th February for the first quarter) in the quarter 5 years after the quarter in which the first birth is observed. Logs are taken of $x + 1$ to include zeroes. White-collar job is an indicator for whether we observe any days of white-collar employment in the quarter 5 years after first birth. This is zero for blue-collar jobs or non-employment. Back to pre-birth firm is an indicator taking value 1 if a woman is ever observed working for any of the firms 5 years after birth for which she worked in the year before birth, and 0 otherwise (also if never employed after birth).

Table 3: Labor Market Outcomes Before First Birth

Dependent variable	Mean	IPW OLS	IPW IV	IPW NR Bounds	Unweighted OLS	Unweighted IV	Unweighted NR Bounds
Log total earnings	11.4299	0.1721	0.1557	[0.1557,0.1685]	0.1759	0.1740	[0.1740,0.1755]
		(0.0013)	(0.0045)	{0.1468,0.1718}	(0.0012)	(0.0039)	{0.1663,0.1783}
Log highest daily earnings	4.0625	0.0519	0.0312	[0.0312,0.0474]	0.0535	0.0374	[0.0374,0.0502]
		(0.0007)	(0.0024)	{0.0265,0.0491}	(0.0006)	(0.0023)	{0.0330,0.0518}
Log avg. daily earnings	3.4646	0.0583	0.0449	[0.0449,0.0553]	0.0605	0.0565	[0.0565,0.0597]
		(0.0010)	(0.0036)	{0.0379,0.0579}	(0.0009)	(0.0032)	{0.0503,0.0619}
Log total experience	7.8318	0.1125	0.1020	[0.1020,0.1102]	0.1146	0.1111	[0.1111,0.1139]
		(0.0007)	(0.0024)	{0.0974,0.1120}	(0.0006)	(0.0021)	{0.1070,0.1154}
Log highest firm daily earnings	4.1550	0.0413	0.0322	[0.0322,0.0393]	0.0420	0.0352	[0.0352,0.0406]
		(0.0005)	(0.0020)	{0.0283,0.0407}	(0.0005)	(0.0020)	{0.0313,0.0419}
Ever white-collar job	0.7699	0.0232	0.0094	[0.0094,0.0202]	0.0238	0.0118	[0.0118,0.0214]
		(0.0007)	(0.0029)	{0.0038,0.0221}	(0.0007)	(0.0029)	{0.0061,0.0232}

Notes: All regressions include dummies for year of birth and dummies for age at labor market entry. Robust standard errors in parentheses. NR bounds in brackets, 95% CI in braces. $N = 50,998$. All earnings variables in real (CPI base 2000) euro. Total earnings is the sum of earnings ever observed in the social security data before first birth (starting in 1972). Highest daily earnings are the highest-ever observed daily earnings on quarter reference days (e.g. 15th February for the first quarter). Total experience is the number of days of employment before birth (starting in 1972). There are no zeroes in these variables due to the construction of our sample (see Section 3). Highest firm daily earnings is constructed as follows. We take the full Austrian worker-firm data and calculate average daily earnings by firm for each quarter reference day in 1972-2017. We then assign to each woman in our sample the highest ever firm daily earnings we observe before first birth. Ever white-collar job is an indicator for whether we ever observe a woman working in a white-collar job before birth.

Table 4: Cumulative Outcomes

Dependent variable	Mean	IPW OLS	IPW IV	IPW NR Bounds	Unweighted OLS	Unweighted IV	Unweighted NR Bounds
Log cumulative earnings	12.1712	0.0572	0.0431	[0.0431,0.0541]	0.0589	0.0545	[0.0545,0.0580]
		(0.0012)	(0.0045)	{0.0343,0.0571}	(0.0011)	(0.0040)	{0.0466,0.0607}
Log cumulative experience	8.6325	0.0194	0.0125	[0.0125,0.0179]	0.0201	0.0171	[0.0171,0.0195]
		(0.0007)	(0.0028)	{0.0069,0.0198}	(0.0007)	(0.0026)	{0.0120,0.0212}
Log highest daily earnings	4.1706	0.0267	0.0086	[0.0086,0.0227]	0.0276	0.0130	[0.0130,0.0247]
		(0.0006)	(0.0024)	{0.0039,0.0243}	(0.0006)	(0.0023)	{0.0086,0.0261}
Log average daily earnings	3.7003	0.0270	0.0168	[0.0168,0.0248]	0.0278	0.0218	[0.0218,0.0266]
		(0.0005)	(0.0021)	{0.0127,0.0262}	(0.0005)	(0.0020)	{0.0180,0.0279}
Log highest firm daily earnings	4.4377	0.0118	0.0059	[0.0059,0.0105]	0.0122	0.0084	[0.0084,0.0114]
		(0.0005)	(0.0022)	{0.0017,0.0119}	(0.0005)	(0.0021)	{0.0043,0.0127}
Ever white-collar job	0.8406	0.0095	-0.0005	[-0.0005,0.0073]	0.0097	0.0003	[0.0003,0.0078]
		(0.0007)	(0.0027)	{-0.0057,0.0091}	(0.0007)	(0.0026)	{-0.0048,0.0094}

Notes: All regressions include dummies for year of birth and dummies for age at labor market entry. Robust standard errors in parentheses. NR bounds in brackets, 95% CI in braces. $N = 50,998$. All earnings variables in real (CPI base 2000) euro. Cumulative earnings is the sum of earnings ever observed in the social security data (1972-2017). Cumulative experience is the number of days of employment (1972-2017). Highest daily earnings are the highest-ever observed daily earnings on quarter reference days (e.g. 15th February for the first quarter). There are no zeroes in these variables due to the construction of our sample (see Section 3). Highest firm daily earnings is constructed as follows. We take the full Austrian worker-firm data and calculate average daily earnings by firm for each quarter reference day in 1972-2017. We then assign to each woman in our sample the highest ever firm daily earnings we observe. Ever white-collar job is an indicator for whether we ever observe a woman working in a white-collar job.

Figure A.1: Patterns of earnings and mental health diagnoses with and without IPW

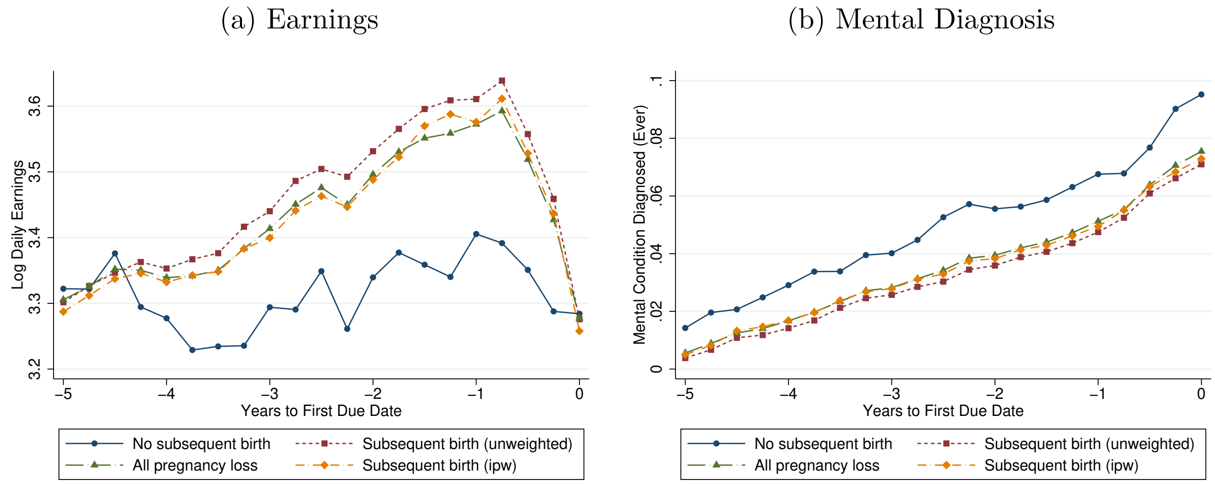
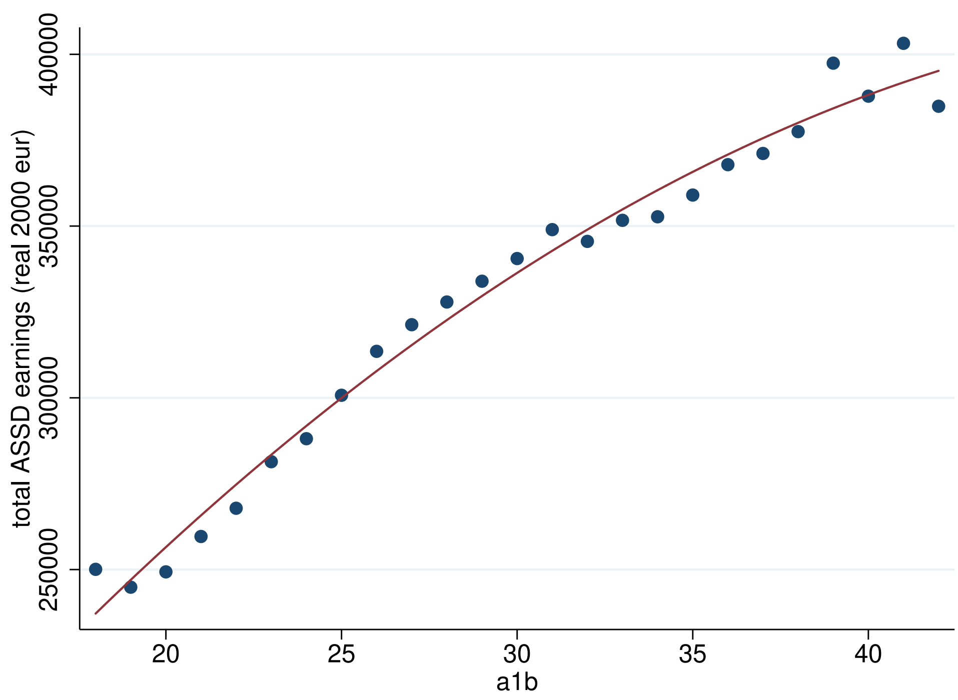


Figure A.2: Earnings and Age at First Birth: Women Born in Austria in 1970



Notes: This figure plots the sum of ever observed social security earnings by age at first birth for women born in 1970. The ASSD include data for 1972-2017. Earnings are topcoded (see Section 3) for about 2% of employed women in the ASSD. We calculate age at first birth by year at which we observe the first live birth in the social security data minus year of birth. In this figure we include all of Austria.

Table A.1: Pre-due-date profiles: All groups

Dep. var.	Employed		Log daily earnings		Ever mental diagnosis		Ever psychol. med.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
t = -4	0.0603	0.0353	0.258	0.119	0.00549	0.000745	0.00777	−0.00112
	(0.00174)***	(0.00191)***	(0.00615)***	(0.00685)***	(0.000343)***	(0.000390)*	(0.000408)***	(0.000465)**
t = -3	0.115	0.0629	0.507	0.224	0.0134	0.00306	0.0172	−0.00131
	(0.00205)***	(0.00239)***	(0.00728)***	(0.00880)***	(0.000533)***	(0.000676)***	(0.000603)***	(0.000785)*
t = -2	0.162	0.0806	0.733	0.296	0.0236	0.00677	0.0314	0.00150
	(0.00222)***	(0.00270)***	(0.00806)***	(0.0104)***	(0.000705)***	(0.000996)***	(0.000809)***	(0.00118)
t = -1	0.195	0.0826	0.907	0.306	0.0363	0.0115	0.0503	0.00637
	(0.00229)***	(0.00293)***	(0.00852)***	(0.0118)***	(0.000868)***	(0.00139)***	(0.00101)***	(0.00165)***
Group 2 × t = -4	0.0144	0.00952	0.0795	0.0497	0.00276	0.000645	0.00771	0.00421
	(0.00650)**	(0.00570)*	(0.0264)***	(0.0218)**	(0.00201)	(0.00200)	(0.00253)***	(0.00249)*
Group 2 × t = -3	0.0176	0.0119	0.0952	0.0617	0.00499	0.00255	0.00773	0.00370
	(0.00582)***	(0.00538)**	(0.0242)***	(0.0211)***	(0.00251)**	(0.00249)	(0.00292)***	(0.00288)
Group 2 × t = -2	0.0161	0.00976	0.0892	0.0517	0.00741	0.00441	0.0119	0.00682
	(0.00511)***	(0.00502)*	(0.0219)***	(0.0203)**	(0.00302)**	(0.00299)	(0.00353)***	(0.00348)*
Group 2 × t = -1	0.0151	0.0106	0.0974	0.0661	0.00878	0.00424	0.0175	0.0102
	(0.00448)***	(0.00476)**	(0.0196)***	(0.0198)***	(0.00349)**	(0.00344)	(0.00418)***	(0.00412)**
Group 3 × t = -4	−0.0472	−0.0684	−0.109	−0.290	0.0107	0.00920	0.0319	0.0203
	(0.0140)***	(0.0124)***	(0.0566)*	(0.0487)***	(0.00482)**	(0.00487)*	(0.00683)***	(0.00682)***
Group 3 × t = -3	−0.0600	−0.0817	−0.200	−0.382	0.0149	0.0134	0.0477	0.0360
	(0.0133)***	(0.0125)***	(0.0549)***	(0.0499)***	(0.00583)**	(0.00586)**	(0.00821)***	(0.00815)***
Group 3 × t = -2	−0.0520	−0.0744	−0.225	−0.410	0.0242	0.0223	0.0597	0.0469
	(0.0121)***	(0.0122)***	(0.0513)***	(0.0496)***	(0.00713)***	(0.00714)***	(0.00936)***	(0.00926)***
Group 3 × t = -1	−0.0402	−0.0610	−0.188	−0.368	0.0293	0.0259	0.0799	0.0650
	(0.0109)***	(0.0119)***	(0.0467)***	(0.0493)***	(0.00809)***	(0.00811)***	(0.0108)***	(0.0107)***
_cons	0.699	0.999	2.560	4.051	0.00907	0.00720	0.0126	0.0117
	(0.00213)***	(0.0778)***	(0.00835)***	(0.334)***	(0.000440)***	(0.00934)	(0.000517)***	(0.0126)
N	260,345	260,345	260,345	260,345	260,345	260,345	260,345	260,345
Controls	No	Yes	No	Yes	No	Yes	No	Yes

Notes: Controls are dummies for year of birth, age at labor market entry, and calendar year. Robust standard errors in parentheses, clustered by person. *** p<0.01, ** p<0.05, * p<0.1.

Table A.2: Pre-due-date profiles: Pool groups 2/3

Dep. var.	Employed		Log daily earnings		Ever mental diagnosis		Ever psychol. med.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
t = -4	0.0603	0.0355	0.258	0.120	0.00549	0.000715	0.00777	−0.00119
	(0.00174)***	(0.00191)***	(0.00615)***	(0.00685)***	(0.000343)***	(0.000390)*	(0.000408)***	(0.000465)**
t = -3	0.115	0.0633	0.507	0.226	0.0134	0.00300	0.0172	−0.00144
	(0.00205)***	(0.00239)***	(0.00728)***	(0.00880)***	(0.000533)***	(0.000676)***	(0.000603)***	(0.000786)*
t = -2	0.162	0.0811	0.733	0.299	0.0236	0.00668	0.0314	0.00131
	(0.00222)***	(0.00270)***	(0.00806)***	(0.0104)***	(0.000705)***	(0.000997)***	(0.000809)***	(0.00119)
t = -1	0.195	0.0832	0.907	0.309	0.0363	0.0113	0.0503	0.00611
	(0.00229)***	(0.00293)***	(0.00852)***	(0.0118)***	(0.000868)***	(0.00139)***	(0.00101)***	(0.00165)***
ploss × t = -4	0.00266	−0.00477	0.0436	−0.0122	0.00426	0.00218	0.0123	0.00707
	(0.00601)	(0.00529)	(0.0243)*	(0.0203)	(0.00189)**	(0.00189)	(0.00246)***	(0.00242)***
ploss × t = -3	0.00286	−0.00541	0.0390	−0.0201	0.00688	0.00452	0.0153	0.00963
	(0.00546)	(0.00507)	(0.0227)*	(0.0200)	(0.00235)***	(0.00233)*	(0.00288)***	(0.00282)***
ploss × t = -2	0.00315	−0.00573	0.0295	−0.0336	0.0106	0.00772	0.0210	0.0142
	(0.00483)	(0.00477)	(0.0207)	(0.0194)*	(0.00284)***	(0.00281)***	(0.00342)***	(0.00336)***
ploss × t = -1	0.00460	−0.00253	0.0431	−0.0139	0.0127	0.00826	0.0294	0.0204
	(0.00426)	(0.00457)	(0.0185)**	(0.0191)	(0.00326)***	(0.00323)**	(0.00402)***	(0.00395)***
_cons	0.699	0.981	2.560	3.960	0.00907	0.0104	0.0126	0.0189
	(0.00213)***	(0.0790)***	(0.00835)***	(0.341)***	(0.000440)***	(0.00886)	(0.000517)***	(0.0121)
N	260,345	260,345	260,345	260,345	260,345	260,345	260,345	260,345
Controls	No	Yes	No	Yes	No	Yes	No	Yes

Notes: Controls are dummies for year of birth, age at labor market entry, and calendar year. Robust standard errors in parentheses, clustered by person. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A.3: Outcomes 5 Years After First Birth: Robustness to Excluding Age at Labor Market Entry

Dependent variable	(1) OLS	(2) IV	(3) OLS	(4) IV
Yearly earnings	306.9131	23.7302	378.1519	10.3903
	(16.5308)***	(69.1111)	(16.4405)***	(68.9612)
Log yearly earnings	0.1188	0.0158	0.1376	0.0141
	(0.0085)***	(0.0367)	(0.0083)***	(0.0364)
Yearly days employed	5.1218	2.6341	5.1478	2.4882
	(0.3087)***	(1.3182)**	(0.3015)***	(1.3044)*
Daily earnings	0.8800	−0.0240	1.0782	−0.0573
	(0.0492)***	(0.1983)	(0.0489)***	(0.1978)
Log daily earnings	0.0602	0.0139	0.0669	0.0127
	(0.0035)***	(0.0148)	(0.0034)***	(0.0147)
Employed	0.0143	0.0067	0.0144	0.0063
	(0.0009)***	(0.0040)*	(0.0009)***	(0.0040)
White-collar job	0.0202	0.0010	0.0199	−0.0012
	(0.0009)***	(0.0039)	(0.0009)***	(0.0039)
Back to pre-birth firm	0.0177	0.0116	0.0187	0.0106
	(0.0010)***	(0.0040)***	(0.0009)***	(0.0040)***
Age at lm entry dummies	Yes	Yes	No	No

Notes: All regressions include dummies for year of birth. Robust standard errors in parentheses. $N = 50,869$. All earnings variables in real (CPI base 2000) euro. Yearly earnings are calendar year earnings in the year 5 years after first birth. Daily earnings and employment are as of the reference day (e.g. 15th February for the first quarter) in the quarter 5 years after the quarter in which the first birth is observed. Logs are taken of $x + 1$ to include zeroes. White-collar job is an indicator for whether we observe any days of white-collar employment in the quarter 5 years after first birth. This is zero for blue-collar jobs or non-employment. Back to pre-birth firm is an indicator taking value 1 if a woman is ever observed working for any of the firms 5 years after birth for which she worked in the year before birth, and 0 otherwise (also if never employed after birth).

Table A.4: Education by Group

	First event is live birth (G1)	First event is preg. loss, subsequent birth (G2)	difference
Education: primary	0.1268	0.1305	-0.0037
	(0.3328)	(0.3370)	(0.0067)
Education: apprenticeship	0.4400	0.4117	0.0284
	(0.4964)	(0.4922)	(0.0101)***
Education: some HS	0.1755	0.1884	-0.0129
	(0.3804)	(0.3911)	(0.0077)*
Education: HS	0.1529	0.1532	-0.0003
	(0.3599)	(0.3603)	(0.0073)
Education: teachers college	0.0482	0.0537	-0.0056
	(0.2141)	(0.2255)	(0.0044)
Education: College	0.0565	0.0624	-0.0059
	(0.2310)	(0.2420)	(0.0047)
Observations	31,089	2,643	33,732

Notes: Women with missing education information (23.8% among group 1 and 43.6% among group 2) excluded. Differences tested with t-test allowing for unequal variance. *** p<0.01, ** p<0.05, * p<0.1.

Table A.5: Education Level at First Birth

	Odds Ratio	
	(1)	(2)
Education Category	Age at First Birth	Pregnancy Loss
Professional Apprenticeship	1.0564***	0.9064
	(0.0089)	(0.0610)
Some High School	1.0853***	0.9853
	(0.0100)	(0.0735)
High School	1.1158***	0.9173
	(0.0105)	(0.0716)
Teacher's College	1.105***	0.9946
	(0.0141)	(0.1065)
College	1.1887***	0.8902
	(0.0155)	(0.0984)
$H_0 : \gamma_j = 0 \ \forall j$	$\chi^2_5 = 240.52$	$\chi^2_5 = 4.47$
	p-value=0.0000	p-value=0.4841

Notes: All columns include indicator variables for year of birth and age at which we first observe yearly earnings above 2,500€. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.6: Probability of Live Birth After Pregnancy Loss (IPW)

	(1)	(2)
	Probit Coef	AME
Year of birth = 1960	0.1660	0.0289
	(0.6741)	(0.1136)
Year of birth = 1961	0.4507	0.0919
	(0.6143)	(0.1091)
Year of birth = 1962	0.7722	0.1836
	(0.5891)	(0.1068)*
Year of birth = 1963	1.0984	0.2947
	(0.5749)*	(0.1018)***
Year of birth = 1964	1.0468	0.2762
	(0.5733)*	(0.0995)***
Year of birth = 1965	1.1555	0.3154
	(0.5676)**	(0.0963)***
Year of birth = 1966	1.5735	0.4711

Continued on next page

Table A.6 – continued from previous page

	(1) Probit Coef	(2) AME
	(0.5672)***	(0.0959)***
Year of birth = 1967	1.4857	0.4385
	(0.5657)***	(0.0947)***
Year of birth = 1968	1.8016	0.5534
	(0.5648)***	(0.0929)***
Year of birth = 1969	2.0222	0.6268
	(0.5657)***	(0.0925)***
Year of birth = 1970	1.9864	0.6154
	(0.5653)***	(0.0922)***
Year of birth = 1971	1.9041	0.5884
	(0.5640)***	(0.0918)***
Year of birth = 1972	2.3653	0.7238
	(0.5655)***	(0.0903)***
Year of birth = 1973	2.3633	0.7233
	(0.5659)***	(0.0905)***
Year of birth = 1974	2.4588	0.7460
	(0.5659)***	(0.0899)***
Year of birth = 1975	2.6306	0.7821
	(0.5672)***	(0.0896)***
Year of birth = 1976	2.5632	0.7687
	(0.5669)***	(0.0897)***
Year of birth = 1977	2.8453	0.8188
	(0.5696)***	(0.0891)***
Year of birth = 1978	2.8266	0.8159
	(0.5692)***	(0.0891)***
Year of birth = 1979	2.9526	0.8337
	(0.5715)***	(0.0890)***
Year of birth = 1980	2.8948	0.8259
	(0.5713)***	(0.0891)***
Year of birth = 1981	2.9192	0.8293
	(0.5738)***	(0.0893)***
Year of birth = 1982	3.1153	0.8527
	(0.5779)***	(0.0890)***
Year of birth = 1983	3.1744	0.8585
	(0.5836)***	(0.0891)***
Year of birth = 1984	3.0635	0.8471
	(0.5840)***	(0.0896)***
Year of birth = 1985	3.0600	0.8467
	(0.5866)***	(0.0898)***
Year of birth = 1986	3.2215	0.8628
	(0.6019)***	(0.0898)***
Year of birth = 1987	3.5156	0.8833

Continued on next page

Table A.6 – continued from previous page

	(1) Probit Coef	(2) AME
	(0.6327)***	(0.0891)***
Year of birth = 1988	3.3398	0.8723
	(0.6470)***	(0.0908)***
Year of first event = 2000	-0.0356	-0.0060
	(0.0995)	(0.0168)
Year of first event = 2001	-0.0339	-0.0057
	(0.1152)	(0.0194)
Year of first event = 2002	-0.2581	-0.0483
	(0.1236)**	(0.0229)**
Year of first event = 2003	-0.2824	-0.0535
	(0.1279)**	(0.0239)**
Year of first event = 2004	-0.5430	-0.1144
	(0.1261)***	(0.0254)***
Year of first event = 2005	-0.6128	-0.1326
	(0.1317)***	(0.0275)***
Year of first event = 2006	-0.7041	-0.1575
	(0.1364)***	(0.0297)***
Year of first event = 2007	-0.8946	-0.2133
	(0.1421)***	(0.0330)***
Age at lm entry <= 16	-0.0126	-0.0030
	(0.0653)	(0.0154)
Age at lm entry = 17	0.0287	0.0067
	(0.0770)	(0.0178)
Age at lm entry = 18	0.1261	0.0282
	(0.0877)	(0.0192)
Age at lm entry = 19	0.1661	0.0365
	(0.0863)*	(0.0184)**
Age at lm entry = 20	-0.0318	-0.0075
	(0.1000)	(0.0239)
Age at lm entry = 21	0.0037	0.0009
	(0.1233)	(0.0289)
Age at lm entry = 22	0.2672	0.0564
	(0.1526)*	(0.0294)*
Age at lm entry = 23	0.0153	0.0036
	(0.1428)	(0.0332)
Age at lm entry = 24	0.2315	0.0496
	(0.1555)	(0.0308)
Age at lm entry = 25	0.0821	0.0187
	(0.1490)	(0.0330)
Age at lm entry = 26	0.1436	0.0319
	(0.1605)	(0.0339)
Age at lm entry = 27	0.0638	0.0146

Continued on next page

Table A.6 – continued from previous page

	(1) Probit Coef	(2) AME
	(0.1635)	(0.0367)
Age at lm entry = 28	0.0671	0.0154
	(0.1797)	(0.0402)
Age at lm entry = 29	-0.1226	-0.0301
	(0.1843)	(0.0469)
Age at lm entry >= 30	0.1952	0.0424
	(0.1285)	(0.0263)
Employed 1y before PL	-0.4455	-0.1011
	(0.1555)***	(0.0352)***
Employed 2y before PL	0.0315	0.0071
	(0.1594)	(0.0362)
Employed 3y before PL	0.2456	0.0557
	(0.1641)	(0.0372)
Employed 4y before PL	-0.2819	-0.0640
	(0.1635)*	(0.0371)*
Employed 5y before PL	-0.2458	-0.0558
	(0.1442)*	(0.0327)*
Log daily earnings 1y before PL	0.1569	0.0356
	(0.0419)***	(0.0095)***
Log daily earnings 2y before PL	0.0573	0.0130
	(0.0487)	(0.0111)
Log daily earnings 3y before PL	-0.0737	-0.0167
	(0.0507)	(0.0115)
Log daily earnings 4y before PL	0.0589	0.0134
	(0.0515)	(0.0117)
Log daily earnings 5y before PL	0.1046	0.0237
	(0.0441)**	(0.0100)**
Log utilization 1y before PL	0.0224	0.0051
	(0.0179)	(0.0041)
Log utilization 2y before PL	-0.0397	-0.0090
	(0.0229)*	(0.0052)*
Log utilization 3y before PL	0.0240	0.0054
	(0.0244)	(0.0055)
Log utilization 4y before PL	-0.0164	-0.0037
	(0.0260)	(0.0059)
Log utilization 5y before PL	-0.0118	-0.0027
	(0.0223)	(0.0051)
Ever anti-depressants 1y before PL	-0.5079	-0.1152
	(0.3213)	(0.0729)
Ever anti-depressants 2y before PL	0.2043	0.0464
	(0.4668)	(0.1059)
Ever anti-depressants 3y before PL	0.0297	0.0067

Continued on next page

Table A.6 – continued from previous page

	(1) Probit Coef	(2) AME
	(0.5661)	(0.1284)
Ever anti-depressants 4y before PL	-0.2138	-0.0485
	(0.6293)	(0.1428)
Ever anti-depressants 5y before PL	0.8389	0.1903
	(0.6197)	(0.1405)
Ever psychol.med. 1y before PL	0.3911	0.0887
	(0.2926)	(0.0664)
Ever psychol.med. 2y before PL	-0.5739	-0.1302
	(0.4293)	(0.0974)
Ever psychol.med. 3y before PL	0.1255	0.0285
	(0.5238)	(0.1188)
Ever psychol.med. 4y before PL	0.2268	0.0515
	(0.5683)	(0.1289)
Ever psychol.med. 5y before PL	-0.6608	-0.1499
	(0.5603)	(0.1271)
Ever diagnosed depression 1y before PL	-0.1134	-0.0257
	(0.2116)	(0.0480)
Ever diagnosed depression 2y before PL	0.3191	0.0724
	(0.3148)	(0.0714)
Ever diagnosed depression 3y before PL	-0.6298	-0.1429
	(0.3629)*	(0.0823)*
Ever diagnosed depression 4y before PL	0.2668	0.0605
	(0.4327)	(0.0982)
Ever diagnosed depression 5y before PL	0.2086	0.0473
	(0.4520)	(0.1026)
Missing health data 1y before PL	0.0016	0.0004
	(0.1030)	(0.0234)
Missing health data 2y before PL	-0.0258	-0.0059
	(0.1104)	(0.0250)
Missing health data 3y before PL	-0.0601	-0.0136
	(0.1159)	(0.0263)
Missing health data 4y before PL	-0.0682	-0.0155
	(0.1172)	(0.0266)
Missing health data 5y before PL	-0.2491	-0.0565
	(0.1174)**	(0.0266)**
_cons	-1.1954	
	(0.5889)**	
N	5632	5632

Notes: The table shows probit coefficients and average marginal effects for probit of probability of having a child after pregnancy loss. Labor market and health measures from 1y-5y before pregnancy loss (PL). The health information is only available from 1998 and variables are set to zero if missing. Indicator variables control for the availability of health information relative to time of first pregnancy loss. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.