The Economics of Infertility: Evidence from Reproductive Medicine*

Sarah Bögl† Jasmin Moshfegh‡ Petra Persson§ Maria Polyakova¶

First version: March 21, 2024

This version: June 17, 2024

Abstract

As the share of births that rely on assisted reproductive technologies (ART) keeps growing, policies around infertility treatments remain ridden with controversy. We use population-wide Swedish administrative data with uniquely detailed information on individual-level use of ARTs, combined with quasi-experimental empirical methods, to characterize the rate of infertility burden, its private and public costs, and the role of insurance coverage in alleviating infertility. We estimate that one in eight women will experience primary infertility – the inability to have any child at all – over her fertile years. Our analysis reveals that persistent infertility causes a long-run deterioration of mental health and couple stability, with no long-run "protective" effects (of having no child) on earnings. Insurance coverage plays a central role in driving the demand for expensive infertility treatments (IVF). The rate of IVF initiations drops by half when treatment is not covered by health insurance. Our estimates imply that couples are willing to pay at most 33% of their annual disposable income for a course of IVF treatment that gives an about 40% chance of having a child. The response to insurance coverage is more pronounced at the lower end of the income distribution. We show that, as a result, coverage of infertility treatments determines both the total number of additional children as well as their allocation across the socio-economic spectrum.

*We are grateful for helpful comments from Kate Bundorf, Rebecca Diamond, David Dorn, Amy Finkelstein, Matthew Notowidigdo, Hannes Schwandt, Dan Zeltzer, and seminar participants at the University of Michigan, University of Zurich, UC Davis, Stanford University (× 2), the Research Institute for Industrial Economics, the Becker Friedman Health Economics Initiative Annual Conference, and the Whistler Junior Health Economics Summit. We thank Iliriana Shala at the Research Institute for Industrial Economics for excellent research assistance. We also gratefully acknowledge support from the National Science Foundation (CAREER SES-2144072, Persson), 2022 Stanford Discovery Innovation Fund (Polyakova), the National Institute on Aging (K01AG05984301, Polyakova), and the Sweden-America Foundation (Moshfegh). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Science Foundation or the National Institute on Aging. We also thank FertilityIQ for generously providing free access to their informational resource.

†Stanford University

‡Stanford University

§Stanford University and NBER

¶Stanford University and NBER

1 Introduction

The WHO estimates that as many as one in six individuals worldwide are affected by infertility, which is defined as attempting but failing to conceive a child after twelve months of trying (WHO, 2023). Since the 1980s, the technology of assisted reproduction (ART) has experienced dramatic advances.[1] Throughout its history, infertility treatments have been at the center of many ethical, demographic, and economic policy debates. Variation in policies that affect the availability and price of infertility treatments around the world is striking, reflecting many differences in how societies think about infertility and assisted reproduction. At the most fundamental level, there is disagreement about the ethics of ART procedures altogether. This has recently sharply come into the public spotlight after the 2024 Alabama Supreme Court ruling that frozen embryos should be considered children – a ruling that lead to pauses in IVF treatments in this U.S. state – as well as the 2024 vote to oppose IVF by America's largest Protestant denomination (Supreme Court of Alabama, 2023; Albert, 2024; Graham, 2024). Also in 2024, "U.S. Senate Republicans blocked legislation that would codify the right to access fertility treatments such as in vitro fertilization" (Miller, 2024). More commonly, however, reproductive assistance is available, but there is variation in whether infertility is considered a disease and its treatment a medical necessity that should be covered by health insurance, or whether it is viewed as an elective procedure that shouldn't be subsidized.[2] Because some ARTs are expensive, these technicalities may ultimately have a large impact on utilization. The recent increase in court cases and legislative discussions related to coverage of infertility treatments in the U.S. (e.g., the proposed Access to Infertility Treatment and Care Act), constitutes but one example of a public policy discussion around the affordability of ARTs.

This remarkable variation in policy and the intense public debate highlight the need for more evidence on the rate of infertility burden, its private and public costs, and the impact of public policies that affect the affordability of infertility treatment. We use population-wide Swedish administrative data with uniquely detailed information on individual-level use of ARTs, combined with empirical methods that leverage natural experiments, to begin to fill some of these gaps.

After describing the data and institutional context in Section 2, in Section 3 we begin by quantifying the rate of infertility burden in our context. One in eight women will experience primary

[1] The ability of ARTs to lead to a successful birth has improved over time after medication-assisted reproduction became available in the 1940s and in-vitro-fertilization in the 1980s (Nordlund, 2008). The first IVF child was born in 1981 in the U.S.. As the success of ARTs has improved, its price has remained high.

[2] The WHO recognized infertility as a disease in 2009 and the American Medical Association followed in 2017 (Adashi, 2018).

infertility – the difficulty of having the first child – over her fertile years, a risk on par with the lifetime risk of breast cancer (American Cancer Society, 2024). This estimate, which leverages couples' revealed willingness to have a child through drug purchase data, improves on survey-based measures of infertility prevalence that rely on recall (WHO, 2023).[3] Many infertility stories have a happy end, with nearly 75% of women bearing a child within eight years of starting their first infertility treatment of any kind. Nearly a quarter of women, however, remain infertile.

In Section 4, we proceed to quantify the private monetary and non-monetary costs of infertility. We use a sample of women who initiate infertility treatment (and hence have a revealed preference for a child), to estimate the consequences of remaining infertile on the woman's own and partner's mental health, labor market outcomes, and couple stability. We leverage two sources of quasi-random variation in the success of assisted reproductive technologies: a failure of the first conception after ART initiation (our primary specification), or a failure to conceive altogether. While many women who initiate infertility treatment eventually carry a pregnancy to term, a failure of the first conception is associated with a 33 percentage point lower long-run probability of having a child. Yet, this outcome is rarely predictable.

Our analysis reveals large and negative impacts of wanting, but being unable, to have a child. Women who do not give birth within five years after the first unsuccessful conception are 48% more likely to fill a prescription for a mental health drug than women whose first conception succeeds. We also find an increase in the use of mental health medication among their partners. Our estimates further reveal that couples who remain infertile are six percentage points more likely to get divorced in the long-run, suggesting that infertility takes a toll on couple stability. At the same time, we find *no* protective long-run impact on labor-market incomes for women, and a small decline in the long-run income of partners.

This finding can appear at odds with a large literature that has documented a negative effect of child-bearing on female earnings, which hence may be expected to imply a large positive effect of infertility on earnings.[4] The lack of a "child penalty" in our estimates likely reflects the fact that our analysis focuses on women who reveal a preference for having a child, and is consistent with an emerging body of evidence using IVF-births (Bensnes et al., 2023; Lundborg et al., 2024). Among

[3] Our definition does not capture involuntary "social" infertility that women may face if they would like to have a child, but do not have a suitable partner, or do not initiate infertility treatments due to individual ethical or religious considerations. We also do not capture cases of infertility in which the woman does not use infertility treatments for other medical reasons.

[4] Goldin (2021) provides an in-depth overview of the relationship between career and family for women over the course of the last century.

women who want a child, our results suggest that remaining childless does not have a long-run protective effect on income. One interpretation of this result is that, while the arrival of a child causes a drop in women's earnings, the longer-run toll of infertility on mental well-being among women who are involuntarily childless may have a countervailing (negative) impact.

Substantial deterioration in mental health and couple stability as a consequence of infertility suggests that policies affecting access to treatment may have a substantial private and social value. In Section 5, we explore the impact of such a policy on IVF utilization and the key demographic outcome: childbirths. We start by leveraging sharp discontinuities in insurance coverage of IVF to estimate the price sensitivity of demand for IVF. Swedish public insurance uses age ceilings that vary over time and across regions to determine who is eligible to be on the waitlist for subsidized IVF cycles. On the day a woman crosses the age cutoff, the out-of-pocket cost for the typical course of IVF jumps to roughly two monthly disposable incomes of an average couple. As wait times and the course of treatment are both uncertain, women close to the age cutoff cannot perfectly control their ability to get access to subsidized IVF before the cutoff birthday. We use this variation to estimate that 12.3 out of 23.9 per 10,000 childless women refrain from initiating treatment when they become ineligible for (nearly) free IVF cycles – a 51% decline in the rate of treatment initiations. We infer that the maximum willingness to pay for a course of IVF treatment (or a circa 40% probability of giving birth) is four months of an average couple's disposable income.

The responsiveness to price is especially pronounced at the lower end of the income spectrum, where couples are 50% more likely to opt out of IVF when insurance coverage lapses. Our estimates by income translate into a (nominal) maximum willingness to pay of circa $14,000 for couples in the lowest income decile and $20,000 for couples in the highest income decile. Notably, this willingness to pay gradient reverses when expressed in relation to disposable income. We discuss several mechanisms through which insurance alters demand. A textbook interpretation is moral hazard, which generates socially inefficient (over)utilization. Several institutional facts that we discuss in Section 5 point to the importance of liquidity constraints in our context, which, we estimate, may account for a third of the demand response at the lowest decile of the income distribution. In other words, in addition to the classic tradeoff between risk protection and moral hazard, insurance coverage of IVF may be providing an important consumption smoothing benefit (Chetty, 2008).

In the last part of the paper, we turn to a more normative discussion. Because insurance coverage has a larger impact on IVF treatments at lower income levels, the design of coverage policies ultimately affects the socio-economoic distribution of children. Our simulation of counterfactual

policies suggests that over 45% fewer children are born in the relevant population in the absence of a subsidy as compared to a universal subsidy. But the effect of subsidies varies by income: We estimate a more than three times larger decline in births in response to the subsidy age ceiling among lower-income couples. The difference in the share of "marginal" beneficiaries directly translates into differences in the marginal value of public funds (Finkelstein and Hendren, 2020; Hendren and Sprung-Keyser, 2020) that an IVF subsidy generates at different income levels. We estimate an MVPF (for a subsidy above the age ceiling) that only reflects private willingness to pay to be \$0.65 among lower-income households and \$0.76 among higher-income households. Allowing for differential social welfare weights (Saez and Stantcheva, 2016), we would get an MVPF of a pooled subsidy that exceeds \$1 if we let society's valuation of providing IVF to lower-income households be about two times their private valuation (or more).

This paper contributes to several literatures. First, we contribute to a broader literature investigating the impact of health shocks on family behavior and well-being, and the implications of those responses for policy design (see, e.g., Fadlon and Nielsen, 2019, 2021; Persson and Rossin-Slater, 2018, forthcoming). Second, our specific focus on infertility as the health shock relates the paper to an extensive literature on the causal effects of having children on maternal health and labor market outcomes.

A large body of work has examined the consequences of having a child on labor market outcomes.[5] Most closely related to our paper is a literature that studies the causal effects of having *any* child using miscarriages and/or stillbirths (Hotz et al., 1997; Hotz et al., 2005; Bratti and Cavalli, 2014) or the success of infertility treatment (Bensnes et al., 2023; Gallen et al., 2023; Lundborg et al., 2017, 2024) as an instrument for having a child.[6] We contribute to this literature by providing causal estimates of the impact of *not* having children, when children are desired (by revealed preference, individuals who undergo infertility treatment want to have children), on measures including but not limited to labor market outcomes. The extent to which the impacts of having a child are informative of the (reverse) impacts of not having a child, holding preferences for a child constant, is a priori ambiguous.

Evidence on the impact of fertility on outcomes other than those on the labor market is

[5] An extensive literature has studied "child penalties" following the birth of the first child in general and not restricted to children conceived through infertility treatments. The existence of child penalties has been documented in labor markets in a range of contexts (Andresen and Nix, 2022; Kleven et al., 2019), including in Sweden (Angelov et al., 2016).

[6] Also closely related are the recent papers examining the consequences of miscarriage per se (i.e., not restricted to individuals seeking infertility treatment) (Bütikofer et al., 2024; Kalsi and Liu, 2022; Rellstab et al., 2022).

limited, in particular as pertains to overall well-being. There is conflicting evidence on the impact of having a child on maternal well-being, with some studies suggesting detrimental effects (e.g., Rizzo et al., 2013; Margolis and Myrskylä, 2015; Glass et al., 2016; Ahammer et al., 2023), in line with the notion of the "parenthood paradox" or the "parenthood happiness gap." However, as highlighted by Cetre et al. (2016), part of these associations may reflect selection into childbearing, something that we partially control for when we condition our sample on individuals who have a preference for childbearing.[7] Conceptually more closely related to our study is a notably smaller body of work examining the associations of infertility treatments with a variety of non-labor market outcomes. Prior evidence on the long-term effects of successful/unsuccessful infertility treatments finds mixed effects on the association with couple stability (e.g., Johansson et al., 2009; Peterson et al., 2011; Kjaer et al., 2014; Martins et al., 2014) and maternal well-being (e.g., Rädestad et al., 1997; Verhaak et al., 2007; Heazell et al., 2016; van den Berg et al., 2017; Vikström et al., 2017).

Third, this paper contributes to a literature analyzing the arrival of new fertility-related health technologies, and the implications of subsidies that influence access to and affordability of ART.[8] In the context of infertility treatments, insurance coverage for IVF has been shown to be associated with increased utilization and birth rates,[9] increased educational investments (Kroeger and La Mattina, 2017; Gershoni and Low, 2021b), and changes in marriage and divorce patterns (Abramowitz, 2014, 2017; Cintina and Wu, 2019; Gershoni and Low, 2021a; Low, 2024). Less is known about the uptake of other forms of ART.[10] We contribute to this literature by exploiting discontinuities in a nationwide subsidy schedule for IVF, which limits concerns of selection, to estimate the impact of subsidies on IVF take-up as well as the willingness to pay for IVF. Our evidence on the diffusion of ART across couples of different socio-economic status also contributes new facts to the broader work on (health and non-health) technology diffusion and how it relates to socioeconomic status; see Cutler et al. (2006) for an overview.[11]

[7] A related literature documents that the postpartum period is a high-risk period for the onset of anxiety and depression in a woman's life, with estimates of the share of new mothers who experience postpartum depression and anxiety ranging from 6.5% to 20% (Mughal et al., 2018).

[8] Outside of the context of infertility treatments, see e.g. Conner et al. (2022) and Junhong (2001) for analyses of the arrival of fertility-related technologies that allow expecting individuals to learn about the characteristics of the fetus during pregnancy.

[9] See, e.g., Banks et al. (2010), Bitler and Schmidt (2006), Bitler and Schmidt (2012), Boulet et al. (2015), Buckles (2013), Bundorf et al. (2007), Crawford et al. (2016), Griffin and Panak (1998), Hamilton and McManus (2012), Hamilton et al. (2018), Henne and Bundorf (2008), Jain et al. (2002), Machado and Sanz-de Galdeano (2015), Provost et al. (2016), and Schmidt (2005).

[10] See, e.g., Kessler et al. (2013).

[11] Some examples include Glied and Lleras-Muney (2008), Jaravel (2019), Jayachandran et al. (2010), Lleras-Muney and Lichtenberg (2005), and Moshfegh (2024).

Finally, our paper contributes to a relatively small but growing body of work that provides a normative analysis of moral hazard. This includes work that evaluates the private willingness to pay and the marginal value of public funds spent on healthcare or health insurance (e.g., Finkelstein et al., 2019a,b). It also includes the work that documents the importance of liquidity sensitivity in healthcare that may observationally be identical to moral hazard, but that has a different normative interpretation (e.g., Ericson and Sydnor, 2018; Lyngse, 2020; Belchior and Gomes, 2022; Gross et al., 2022). In contrast to a large literature that estimates the magnitude of moral hazard in the context of health insurance, the normative discussion is much scarcer (Einav and Finkelstein (2018) provide an overview; also see Nyman (2003) for a discussion). The evidence we present is consistent with an importance of liquidity constraints in driving the observed behavioral response to the prices of care. This in turn suggests that insurance coverage may provide a consumption smoothing benefit in the case of expensive infertility treatments. While existing work has shown that liquidity sensitivity may cause consumers to delay healthcare consumption that is associated with small lump-sum payments (co-pays) until the receipt of income (Gross et al., 2022), our evidence suggests that liquidity sensitivity may crowd out healthcare consumption entirely, as potential children are simply not born when insurance is removed.

The rest of the paper proceeds as follows. We describe our data sources and measurement in Section 2. We present facts about the prevalence of infertility and the success rates of infertility treatments in our data in Section 3. We quantify the consequences of infertility on monetary and non-monetary well-being in Section 4. We estimate the price sensitivity, the private willingness to pay, and discuss the potential role of liquidity in Section 5. In that section, we also turn to a more normative discussion, deriving the implications of insurance coverage on the number of children born and quantifying the MVPF of subsidizing IVF. Section 6 concludes with a discussion.

2 Institutional Background, Data, and Measurement

2.1 Infertility and Infertility Treatment

Involuntary infertility was classified by the World Health Organization as a disease in 2009 (Zegers-Hochschild et al., 2009), defined as a disease of the reproductive system.[12] Infertility treatment starts with an infertility evaluation, typically after at least one year of trying to conceive. The

[12]The American Medical Association followed suit in 2017. The designation of a condition as a disease has significant implications for insurance coverage, as it changes the designation of treatment as being "medically necessary."

subsequent choice of treatment, if any, depends on the underlying cause of infertility. Often the first line of treatment is what we will refer to as low-intensity treatment (henceforth LIT). LIT can take two forms. Patients can take prescription medication that regulates ovulation, improving chances of conception. Another form of LIT is in-utero insemination, which may or may not be supported by the same medication. Thus, LIT is a non-invasive or minimally invasive method of assisted reproduction. It is generally cheap and considered well-established and safe.

If LIT is not successful or if the patient's case is not medically appropriate for LIT, a high-intensity treatment is available. High intensity treatments include various forms of in vitro fertilization (IVF). IVF involves strong ovulatory stimulation, an egg retrieval procedure, fertilization outside of the woman's body, and subsequent implantation of a fertilized embryo. IVF is a substantially more invasive and more expensive procedure.

The market price of a single cycle of IVF in Sweden in 2010 (which is roughly in the middle of our main analytic sample) was about 26,813 SEK ($3,250 in 2022 dollars).[13] Typically, a woman needs two to three cycles to arrive at one birth, for the total price of treatment corresponding to 13% to 20% of the average couple's disposable income in the relevant population (as we discuss in more detail in Section 5).

The effective out of pocket price for IVF varies dramatically depending on health insurance coverage. For primary infertility (i.e., for women who are childless and are struggling to conceive their first child), Swedish public health insurance provides coverage for three IVF cycles, but with some important restrictions. During our sample period, coverage eligibility was based on age cutoffs—a woman needed to be young enough to be eligible for insurance coverage. The age cutoffs varied across Swedish regions and over time. The lowest age cutoff was 36; see Table A1 for the cutoffs in all regions and years. When a woman is eligible for insurance coverage, the out of pocket cost of IVF is comparatively low, corresponding to at most 2,700 SEK in 2010 in co-pays for prescription drugs and office visits ($327 in 2022 USD) within the public insurance system. Even when a woman is covered by insurance financially, access to care can be challenging in practice, with significant wait times.[14]

[13] The average price of one IVF cycle is as reported by eight clinics in Sweden. Data on prices was collected for 2010 using the web.archive.org website snapshots.

[14] Typically the waiting period is between three to twelve months, but can be as long as two years depending on where in Sweden the patient lives (https://xn--fertilitetsrd-0fb.se/ivf-vantetid/; https://www.1177.se/Halland/barn--gravid/barnloshet/ivf-provrorsbefruktning/).

2.2 Data and Measurement

The backbone of our data is an extract from the Swedish Population Register of all individuals residing in Sweden from 2000 onwards. We merge this data with the Statistics Sweden's longitudinal database of individuals (LISA) from 1990 through 2021, which contains demographic and socio-economic information drawn from various administrative records (Statistics Sweden, 2019). We further link individuals to their spouses using marital records and, from 2011, also to cohabiting partners using information about shared addresses. Hereafter, we refer to a person's spouse or cohabiting partner simply as the person's partner.

To construct measures of ART utilization and health outcomes, we merge in health records from the National Board of Health and Welfare (Socialstyrelsen, 2019). For each individual, we observe the universe of prescription drug purchases made in outpatient pharmacies from July 2005 through 2019. Importantly, the records include all purchases of prescription drugs, irrespective of whether the purchase is covered by public health insurance.[15] For each purchase, we observe the name of the drug and the drug's seven-digit Anatomical Therapeutic Chemical (ATC) classification code. We also observe the universe of inpatient hospital visits and specialist outpatient visits from 2002 through 2019. For each visit, we observe the date of the visit and the diagnosis codes (ICD-10) attached to the visit. We also observe birth records that contain all live births and stillbirths in Sweden from 1985 through 2019.[16] For each birth we observe whether it is a live birth or a still birth, gestational age at birth, and the due date. We also observe if the birth resulted from a medically assisted conception.

To capture all women who initiate an ART treatment, we create a list of all drugs (ATC codes) that are used in LIT and IVF treatments, respectively. We define a woman as initiating LIT (IVF) treatment on day d if she fills an LIT (IVF) prescription for the first time in 365 days during the period from July 2006 through November 2019.[17] Appendix A provides the details. Figure A1 validates our measure. In this figure we compare the rate of ART-births as recorded in the

[15] In other words, this is not insurance claims data, but rather pharmacy transaction data for prescription drug fills.

[16] For dates prior to July 1 2008, the birth record contains alive and stillborn infants from week 28+0. Starting on July 1, 2008, the register contains alive and stillborn deliveries from week 22+0.

[17] Some drugs are only used in IVF treatments, whereas some drugs are used both in LIT and IVF treatments. Thus, to define a woman as an LIT initiator, we require that she does not take an IVF drug. To ascertain that we capture only women who undergo treatment due to infertility, we exclude women for whom we observe a cancer diagnosis in either the inpatient or specialist outpatient records in the six months following d, as they may undergo infertility treatment in preparation for their cancer treatment. We additionally exclude women for whom we observe a breast cancer diagnosis in the two years before d to avoid misclassifying cancer treatment as ART since similar drugs are used for breast cancer hormone therapy and ART.

administrative birth register (dashed line) with the rate of assisted births computed based using our ART measure (solid line). For the overlapping years of data, the two measures are very close to each other.

We identify the women who *conceive* following an ART treatment as those who either (i) appear in the birth records with a live birth or a birth of a stillborn child, or (ii) appear in the inpatient or specialist outpatient records with a primary diagnosis code indicating a miscarriage.[18] For each birth, we infer the date of conception by subtracting 280 days from the due date. For miscarriages, we observe the date of diagnosis, but not the date of conception. For individuals who initiated ART treatment within 22 weeks of the miscarriage, we take the ART initiation date to be the date of conception. For individuals who initiated ART treatment more than 22 weeks before the miscarriage, we infer the date of conception as 12 weeks earlier (thus, we assume that the miscarriage occurred on day 12+0 of pregnancy).

We use these data sources to define our key outcome variables for women of fertile age and their partners. First, we create indicator variables for any drug purchase in the following three categories: anti-anxiety, anti-depressant, and antibiotic. Appendix A lists the exact ATC codes. To capture labor market outcomes, we use LISA variables to calculate total annual work-related income, which includes wages from employers as well as income from self-employment. Finally, we create an indicator variable for divorce among those who are married.

The merged dataset further allows us to define additional variables used in the analysis including a categorical variable indicating the individual's region ("län") of residence, a categorical variable indicating the individual's highest level of completed schooling (i.e., no college, some college, completed college), and place (usually country) of birth. We also compute two measures of income ranks (within gender, calendar year, and birth cohort) that we use throughout the paper. First, we compute deciles of women's individual disposable income at age 32. We use this measure of relative income in our analysis of the willingness to pay for infertility treatment, as it captures income well before any woman approaches the insurance coverage age cutoff. Second, for all women irrespective of their ART use, we compute the woman's parental income rank taken over the two years when the woman was age 20 and 21. Appendix A provides more details of each measure.

[18] See Appendix A for diagnosis codes for a miscarriage. We do not capture conceptions ending in early miscarriages that occur at home, without any contact with the healthcare system. As ART pregnancies are usually closely monitored, this is rarer than for non-assisted pregnancies, but likely nonetheless a relevant limitation of our definition of conception.

2.3 Samples

Table 1 presents summary statistics for several analytic samples. We start by defining a sample of women who are "at-risk" for involuntary infertility. Intuitively, as we do not (a priori) know which women are childless because they struggle to conceive, we start by taking all women who are of fertile age and who do not (yet) have any child. Specifically, our at-risk sample (Column (1)) includes all person-year observations over the time period from 2006 through 2019 when the woman is (i) childless and (ii) age 16–45. This at-risk sample consists of 13.4 million observations and includes 1.8 million unique women. Column (1) reports average characteristics across all person-year observations. The typical woman who has not yet had her first child is 25 years old. Further, 22% have completed college, 9% are married (2% had an experience of a divorce), and 17% were born outside of Sweden. These women earn on average 133,540 SEK (measured in 2010 SEK) in work income (ca. $20,000 USD).

Column (2) presents characteristics for the "initiator" sample. This includes all women who are in the at-risk sample of Column (1) and who initiate either IVF or LIT treatment at some point over the time period when we observe initiations (i.e., from July 2006 through 2019). Column (2) reports summary statistics for this sample of women measured in the year before the initiation of ART, for one observation per woman. Women who initiate ART are older on average (31.1 vs. 25.4), more likely to have completed college (47% vs. 22%), more likely to be married (35% vs. 9%), and have nearly twice as much income as the typical woman who hasn't yet had her first child.

The sample in Column (3) is identical to the one in Column (2), but restricted to individuals initiating ART treatment by December 2012; we use this sample in Section 3 to describe the success rates of ART over a longer (eight years) time horizon after initiation. The characteristics of women in the two initiator samples (measured, in both cases, in the year before initiation) are similar. Columns (4) through (6) show summary statistics for samples that we will return to in detail in the subsequent parts of the paper. Column (4) shows summary statistics for the event study sample used in Section 4, which is a sub-sample of Column (2), restricted to months and years up to February 2016 (to allow for follow-up observation) and to women who conceive within three years of treatment initiation. Women in this sample are generally similar to women in the two other initiator samples. Columns (5) and (6) present summary statistics for the two samples that we use in our Regression Discontinuity (RD) analysis of Section 5. Both are sub-samples of the childless at-risk sample in (1), but restricted to observations within two years of the ART eligibility cutoff,

and Column (6) is further restricted to women who initiate LIT. The characteristics are measured at age 32, i.e., well before any woman approaches the insurance eligibility cutoff. As compared to all at-risk women in Column (1), the women close to the insurance eligibility cutoff are (mechanically) older, are somewhat more likely to have completed college and to be married, have higher income, and are more likely to have been born outside of Sweden.

3 Prevalence of Infertility and Infertility Treatments

Measuring the prevalence of infertility in the general population has traditionally been challenging, as infertility is inherently defined as an absence of an event (pregnancy) over a certain time period, making measurement complex as it requires knowing intent (WHO, 2023). As infertility treatment is not easily available (or may not be utilized by all couples) in many countries, the WHO measure of worldwide infertility relies on responses to surveys about a woman's reproductive experience. In the Swedish context, the low intensity ART treatment is cheap and readily available, and so the take-up of at least *some* infertility treatment for a woman who is experiencing infertility is less of a concern. In turn, using the occurrence of treatment in administrative healthcare records to reveal which woman is experiencing infertility has the advantage of being able to capture the full population without variation in individual recall or interpretation. With this in mind, we consider a woman in our data to have encountered primary infertility if she is childless and starts using any assisted reproductive technology (IVF or LIT) based on the prescription drug records.

There are several ways to measure how common infertility is in our context. Figure 1, panel (a), illustrates the incidence of infertility by plotting the share of childless women who initiate ART treatment, by age, using data for 2018 (the last full year observed in the data). The solid line depicts initiations of any ART treatment (LIT or IVF), and the two dashed lines show initiations of LIT and IVF, respectively. Very few women initiate any ART below age 25. The initiation rate then increases until age 38 – where it peaks at more than three percent – and subsequently falls sharply. The rates of LIT and IVF initiations are similar until age 30, after which the rate of IVF initiations is higher, reflecting the fact that many women start with LIT and subsequently initiate IVF; see Figure 2, panel (b). The sharp drop in ART initiations around age 40 is driven by a steep decline in the rate of IVF initiations.[19] There is no discontinuous drop in the rate of initiations of (readily available) LIT.

[19]This is consistent with the fade-out of IVF health insurance coverage around age 40, a fact that we explore in detail in Section 5.

Next, we use these numbers to compute "period lifetime infertility." Intuitively, akin to "period life expectancy," which reflects an individual's expected longevity under the assumption that mortality rates remain fixed at today's (or a given period's) levels, in Figure 1, panel (b), we compute the likelihood that a childless woman encounters infertility over her remaining fertile years, by age, under the assumption that the incidence of infertility and non-ART childbearing rates remain at 2018 levels. Specifically, for each age a on the x-axis, we calculate and plot $100 * P(\text{Initiating ART between } a \text{ and } 45 | \text{Childless and no history of ART at } a)$.[20] Among women who are childless at age 16 (i.e., basically all women in Sweden), the likelihood of ever encountering infertility is 12%. This likelihood then increases by age until it peaks at age 34, where more than 15% of all childless women (who haven't yet encountered infertility) will encounter infertility over their remaining fertile years. Beyond age 34, the rate of lifetime infertility falls, as we would expect if a gradually larger share of childless women at older ages are those women who prefer to remain childless.

Figure 2, panel (a), illustrates how infertility journeys resolve over an eight-year horizon after an ART initiation. Here we use the sample of ART initiators whom we can follow for eight years, i.e., the sample in Column (3) of Table 1. The figure plots the rate of childbirth as a function of time since initiation of LIT (dashed line) and IVF (solid line), respectively. Two years after initiation, the birth rate is around 50%–60%; eight years after initiation, the birth rate is around 75%.[21]

Figure 2, panel (b), illustrates the precise infertility treatment pathways over the first two years after ART initiation. 60% of women who initiate a treatment start with LIT, while 40% directly initiate a more intensive treatment. Among those who start with the less intensive treatment, 37% eventually switch to IVF. All in all, within two years of initiating any infertility treatment, 60.8% of women give birth. 25% go all the way through at least one cycle of an intensive IVF treatment and still do not have a child within two years. About 14% of women start with LIT but then appear to "drop out"—they neither switch to IVF, nor have a child, within two years. Figure A4 shows the analogous infertility pathways evaluated eight years after an ART initiation. With this longer follow-up period, a lower fraction of women ends up trying IVF and not succeeding (15%) and the

[20] Appendix B provides the details of this calculation. Appendix Figure A2 shows the underlying values for this calculation, namely the probability of giving birth and the probability of being childless and initiating ART at each age a in the 2018 cross-section.

[21] The fact that the eight-year follow-up birth rates are similar is not informative of the underlying effectiveness of each technology. First, we do not require births to be associated with any specific technology. Second, even though IVF is generally considered to be more effective than LIT, it is substantially more invasive, so women with less complex infertility cases are likely to initiate LIT as the first line of treatment. In other words, the observed pattern likely reflects sorting into treatment. Consistent with this, Figure A3 shows that women who undergo LIT tend to be younger than women who undergo IVF; age tends to be one of the main predictors of treatment success.

overall birth-rate is higher.

Taken together, the descriptive patterns in our data suggest that, at the population-level, infertility is common and the majority of women who encounter infertility as measured by ART use ultimately end up with a birth. Yet nearly a quarter of women remain infertile; we next investigate how remaining infertile affects women's and their partner's well-being in the long-run.

4 The Consequences of Persistent Infertility

4.1 Empirical Approach

Estimating the causal effects of infertility is challenging for (at least) two reasons. First, infertility is hard to observe in most data sources, as we usually do not know which childless women are involuntarily medically infertile and which are childless for other reasons. Second, even when infertility status is observed, being infertile is not randomly assigned among couples. In addition to these empirical challenges, there is a conceptual issue that has received little attention in the prior literature on the effects of children: Conditional on wanting children, the counterfactual of experiencing infertility is to experience the arrival of a child. The treatment effects of infertility (or of children) are thus only defined in relative terms, as conceptually there is no untreated group once we restrict attention to a sample of couples who want a child. Another conceptual issue is that there are several "paths" to persistent infertility: Women may remain infertile by failing to conceive at all, or by conceiving (perhaps repeatedly) but failing to carry any pregnancy to term. While the end result remains that the woman (couple) experiences persistent infertility, it is conceivable that the treatment effects of each path are distinct (and it is conceivable that they are similar).

Our data and setting allow us to make progress in addressing these challenges. To address the first empirical challenge – that infertility is hard to observe in most data sources – we zoom in onto childless women of child-bearing age (16–45) who are initiating an infertility treatment of any kind (see Appendix A for the detailed definition of treatment initiation). These women are a set of women who have revealed their preferences for having a child in the data.

Next, we turn to the second empirical challenge, that infertility is not randomly assigned. Indeed, the likelihood of infertility depends on a series of choices made by each woman (couple), including when to start trying to conceive, whether and what type of assisted reproductive technologies to use, how long to keep trying (with or without technologies), and potentially broader behaviors that may influence the likelihood of conception or the ability to carry a pregnancy to

term. At the same time, there are components of women's (in)fertility trajectories that contain elements of chance. The basic idea of our empirical strategies is to leverage these elements of chance in the context of an event study design.

Our primary specification uses a sample of women who initiate any ART (reveal a preference for a child) and conceive within three years of initiation; for women who conceive more than once, we use the first conception after ART initiation.[22] Conditional on conception, some pregnancies fail due to the biological risk of a miscarriage or stillbirth, rendering infertility persistent for some. The sporadic miscarriage risk is considered to be largely a matter of chance in the medical literature. Many women go on to conceive after experiencing a miscarriage. For some women, however, miscarriages early in the fertility journey end up uncovering longer-term infertility. The woman's "type" is not easily predictable by the patient or their physician ex ante; indeed, a medical evaluation for recurrent pregnancy loss is only advised after two or three consecutive miscarriages (e.g., Larsen et al., 2013; Melo et al., 2023; Oster and Fox, 2024).[23] While *which* woman will end up experiencing persistent infertility is hard to predict, in aggregate the failure of the first post-ART conception is predictive of infertility over a longer time horizon, as we shall see below.

We use the variation across the two subsets of women—(i) women whose first conception after ART initiation resulted in a live birth and (ii) women whose first conception resulted in a miscarriage or stillbirth ("failed")—to measure the consequences of persistent (medical) infertility on mental health, earnings, and couple stability, in the medium- and long-run. More formally, we estimate the following event-study model centered around the time of the index conception:

$$Y_{it} = \alpha_i + \sum_{\tau} \kappa_\tau D_{\tau,it} + \sum_{\tau} \sigma_\tau D_{\tau,it} * Failed_i + \gamma_t + \beta * \mathbf{X}_{it} + \epsilon_{it} \tag{1}$$

where Y_{it} is an individual-level outcome (e.g., having filled a prescription for a mental health drug), t is the absolute time (quarter or year depending on the outcome) and τ is time relative to the quarter (or year) of conception. $D_{\tau,it}$ are fixed effects for relative time, α_i denotes individual fixed

[22] We do not require any conception to coincide with the precise month in which any ART is used.

[23] According to the American College of Obstetricians and Gynecologists (ACOG), chromosomal abnormalities are a common cause of early miscarriages, largely described as "random" events (https://www.acog.org/womens-health/faqs/early-pregnancy-loss). More broadly, in the words of Oster and Fox (2024) on early-stage miscarriage, "practically without exception, miscarriage is not due to something the mother did or did not do." Later-stage pregnancy loss can be caused by chromosomal abnormalities but also by several other causes, e.g., fetal demise, preterm labor, or cervical insufficiency. These causes, too, are usually sporadic and not related to any underlying condition (Oster and Fox, 2024). Finally, some causes may involve actions taken by the mother, but coupled with an element of chance. For example, a Listeria infection can cause stillbirth, an occurrence that in principle is driven by what the pregnant woman eats; however, Listeria outbreaks are hard to predict (Oster, 2013).

effects, and γ_t denotes absolute time fixed effects. $Failed_i$ is a dummy equal to 1 if the woman had an unsuccessful pregnancy following the index conception. $\mathbf{X}_{it}$ include dummies for age at time t in two-year bins. The coefficients of interest are σ_τ that measure how the time path of the outcome relative to the time of conception differs between women whose first conception after ART initiation is successful and women whose first conception fails.

While above we cited several sources suggesting that pregnancy failure early in the fertility trajectory can often be considered to be as good as randomly assigned, our event study design relies on a weaker "parallel trends" identifying assumption. Specifically, the identifying assumption needed to interpret our event study coefficients as the causal effect of infertility relative to the desired outcome (and to the societal default) of having a child, is that women whose first conception succeeds provide an accurate counterfactual for how the outcomes of women whose first conception fails would have evolved over time in the counterfactual scenario (i.e., if these womens' pregnancies, too, had succeeded). While we cannot test this assumption directly, we can examine trends in outcomes prior to conception.

In addition to the reduced form specification in Equation (1), we also estimate a first stage. The specification is analogous to Equation (1), but the outcome variable is an indicator for having had any childbirth. By definition, there are no births in the sample during the period prior to the first conception. In the post-period, we do not restrict births to be directly related to any infertility treatments.

In our baseline event study, we compare women whose first conception after ART initiation fails to women whose first conception succeeds. In the language of potential outcomes (where infertility is the treatment and the outcome of the first conception is an instrument), women whose first conception succeeds are compliers who are not treated (by infertility). Women whose first conception fails are either never takers (women whose first conception fails, but they go on to have a child in a subsequent pregnancy) or treated compliers (women whose first conception fails and who subsequently remain childless). There are no always takers in our context, as women cannot remain infertile if their first conception succeeds. A unique feature of our environment is that if we restrict ourselves to a finite follow-up time horizon, we can empirically differentiate between treated compliers (whose first conception failed) and never takers. We define never takers as those women whose first conception fails, but they go on to have a birth within a five year follow-up period. Taking these women out of the sample allows us to do a direct comparison of compliers whose first conception succeeds and compliers whose first conception fails and who remain infertile (at least

during a five to ten year follow-up, depending on the outcome). Throughout, we also estimate the analogue of Equation 1 only on this complier sample, which recovers the treatment effect of infertility among those women whose experience with the first assisted conception determines their long-term fertility.

Our baseline event study estimates reflect the impact of infertility that involves the experience of a pregnancy loss. As mentioned above, another path to infertility involves a failure to conceive altogether. To assess the impact of infertility through this alternate path, in Appendix D we also present results from event studies that center on infertility treatment initiation, without conditioning on conception. These event studies compare women who conceive within three months of treatment initiation to women who fail to conceive in the first three months of initiation. This specification thus captures the impact of infertility through a path that involves a failure to conceive – and it effectively puts women who experience a miscarriage into the control group (as a subset of the women who conceive subsequently experience a miscarriage). Interestingly, the results from the specification that does not condition on conception are generally similar to the results from our baseline event studies. This suggests that the long-run impacts of persistent infertility are similar whether the path to this end result involves pregnancy loss or persistent non-conception.[24]

4.2 Results

Figure 3 illustrates the first stage. Panel (a) is a flow measure, which shows the share of women giving birth in each quarter relative to the time of the first conception after ART initiation. Mechanically, there are no births prior to conception. Among women whose first conception results in a live birth (gray dashed line, 80% of the sample), births are concentrated in quarter 3 (as we would expect given the modal length of pregnancy). Among women whose first conception fails (blue solid line, 20% of the sample), we observe live births only after quarter 4. There is a small mass of births in quarter 5 after the first conception, but births are generally spread across multiple quarters post the first unsuccessful conception. Panel (b) shows that some of the first conception failures result in persistent infertility. This panel plots a stock measure—the share of women who have had at least one live birth by a given quarter relative to the first conception. Mechanically, 100% of women whose first conception is successful have had a live birth by quarter 5. Among women whose first conception fails, three quarters still experience a birth in the longer-run (Table A2). In Table 2

[24]This finding is consistent with survey evidence suggesting that both pregnancy loss and the "non-event" of involuntary childlessness through failures to conceive are associated with decreased life satisfaction (Nynas et al., 2015; Rooney and Domar, 2018).

Column (1) we estimate that on average two years after the first conception and beyond (up to 20 quarters), women whose first conception fails are 33 percentage points less likely to have a child. In other words, the failure of the first conception post ART initiation predicts a substantial longer-term fertility difference.

We next estimate Equation (1) to quantify the impacts of remaining involuntarily infertile on women and their partners. For each outcome, in Figures 4 and 5, we show two types of panels. The left column of panels plots the raw data for each outcome, re-normalized to the common mean in the quarter (or year) prior to conception. As in the first stage figures, we continue to split the sample into women whose first conception results in a live birth (gray dashed lines; this is the control group, or untreated compliers) and women whose first conception fails (blue solid line; the treatment group, which is a mix of treated compliers and never takers). In addition, we add a third line that plots each outcome for the sub-sample of women whose first conception fails and who subsequently remain childless for five years (navy dotted line; treated compliers only, i.e., this is a subset of the women included in the blue solid line).

The right column of panels reports the point estimates from the event study in Equation (1). In light blue are our baseline estimates that use the full sample, i.e., that compare women whose first conception succeeds to (all) women whose first conception fails. The navy estimates instead compare women whose first conception succeeds to (only the) women whose first conception fails and who remain childless for five years. In other words, the navy estimates are restricted to treated and untreated compliers only, effectively scaling our baseline estimates by the first stage.

We first consider the effect of infertility on health, focusing especially on mental health. The first row of panels in Figure 4 and Figure 5 show results for the probability of (ever) using antidepressant and anti-anxiety medications. We generally observe relatively high rates of mental health medications at baseline—about 25% of women and 17% of partners have ever taken these medications by the quarter prior to conception (see Columns (2) and (6) of Table 2). Women (and partners) whose first conception fails and whose first conception succeeds have similar time trends in mental health drug use prior to the first conception after ART initiation. The event study figures confirm the visual lack of pre-trends in the raw data and also reveal a substantial divergence in the use of mental health medication after the first conception fails. For both women and their partners, we observe a much faster growth in the probability of using mental health prescriptions after the conception failure. Our baseline estimates in Column (2) of Table 2 suggest that two years after the first failed conception, women are four percentage points (16%) more likely to have

taken a mental health medication. For women who remain infertile, this increase is substantially larger (48%). To shed light on whether this may simply capture a higher probability of taking any medication (or a deterioration in general health), in Appendix Figure A5 we present event study estimates for antibiotics claims.[25] Our estimates suggest that infertility causes a decline in antibiotics use around the age when couples whose first conception succeeded start sending their children to daycare (which happens around age one in Sweden);[26] however, in the longer run we observe no differences in antibiotics claims between women whose first conception succeeds and women whose first conception fails.

Next, we consider the impact of persistent involuntary infertility on economic outcomes. The second row of panels in Figure 4 and Figure 5 show our estimates for women's and partners' work income. The trend in income is similar between the treatment and control groups prior to the first conception (income is growing slightly slower for both women and partners in the treated complier sample). We see no changes in income trends among complier women and partners whose first and subsequent conceptions are unsuccessful. Consistent with the prior literature on the effect of children on earnings, however, we observe a sharp drop (by 56% among women and by 7% among partners) in income coinciding with the time of birth among those women and partners whose first conception is successful. We observe a muted and delayed drop in income in the sample of women whose first conception is unsuccessful but two thirds of whom go on to have a child at a later point in time. The estimates in Table 2 indicate that remaining infertile "protects" income by 2.5% when we average over all years two to ten after the first conception. This protective effect, however, is driven by the years closer to conception and dissipates in the longer run (as is visible in Figure 4). At seven years after the first conception, women and men whose first conception was successful not only catch up in income levels, but experience a steeper income growth trajectory as compared to the couples who remain infertile. Persistent infertility thus has no "protective" long-run effect on women's incomes, while for partners, persistent infertility leads to income declines in the long-run. In other words, for women who reveal their preference for having a child (and hence who experience infertility when they don't have a child), we estimate a zero long-run child penalty on income. One interpretation of this result is that, while the arrival of a child causes a drop in earnings, the longer-run consequences of infertility on mental well-being among women who are involuntarily childless

[25] As few chronic conditions are present in the relevant age group, we focus on infections that are common at any age.

[26] This is consistent with Daysal et al. (2021), who show that preschool children bring home germs that infect younger siblings with respiratory disease.

may have a countervailing (negative) impact on their earnings. In other words, women who desire children do not avoid an earnings penalty by remaining childless.

Finally, we turn to the impact of infertility on couple stability, and more specifically, the likelihood of a divorce. The last row of panels in Figure 4 shows the raw data and the event study for the share of couples who have divorced. This panel restricts our sample to the subset of couples who were married in the year before the first conception. In both the raw data and the event studies, we observe that couples who had a failed first conception are divorcing at faster rates compared to couples who had a birth after the first conception. The difference is especially striking for couples who remain infertile. Ten years after the first failed conception, 30% of these couples have divorced, as compared to 20% among those whose first conception was successful. The event study estimate in Column (5) of Table 2 suggests that after two years, women whose first conception fails are two percentage points more likely to get divorced. The effect is six percentage points when scaled by the first stage (i.e., capturing the effect for those women who remain infertile in the long-run).

In sum, in contrast to the common notion that childbearing may lead to long-term unhappiness, drops in economic productivity, and couple instability, we find the opposite in the sample of couples who prefer to have children. Among couples who remain involuntarily infertile in the long run, mental health deteriorates, earnings remain the same (among women) or experience a slower growth (among partners), and these couples experience substantially higher rates of separations.

5 Demand for Infertility Treatments

The estimates in Section 4 reveal substantial private (non-pecuniary) costs of infertility. We next examine the nature of demand for medical technologies that treat infertility, and how health insurance may affect this demand.

Swedish public insurance (nearly fully) covers three cycles of IVF treatment. The eligibility for coverage primarily depends on age. Historically the age ceiling has varied across Swedish regions and over time. See Appendix Table A1 for the maximum age at which coverage was provided in each region-year. As patients age out of coverage, they can still access IVF services, but have to pay the full cost out of pocket.[27] While the age-based eligibility change is deterministic and foreseeable, patients cannot perfectly control the timing of their treatment around the age cutoff. First, women typically have no knowledge of their fertility (and thus the likelihood of encountering

[27]Patients can also choose to forgo the public coverage waitlist even before they age out of coverage. We include all treatment initiations in our measure of initiations, irrespective of coverage.

infertility) until they decide to try to have a child, which can be at different ages relative to the coverage ceiling. Second, due to capacity constraints, most patients have to wait for their publicly covered treatment, with wait times typically ranging from three to twelve months.[28] As a result, some women requesting to start a cycle when they are eligible for coverage may age out of insurance coverage by the time they receive their first appointment. This institutional context implies that right at the age cutoff, access to (highly) subsidized IVF becomes as good as randomly assigned. Taking advantage of this fact, we use a regression discontinuity design to estimate how much women respond to the sharp change in the price of IVF.

5.1 Graphical Evidence

Figure 6 shows the raw data. The x-axis in both panels shows the age of the woman (in months) relative to the (region-year specific) insurance eligibility cutoff. The y-axis in panel (a) shows the number of women who initiate IVF in a given (relative) month per 10,000 childless women who have not initiated IVF before. The sample includes all childless women whose age falls within five years of their (region-specific) age eligibility threshold. We observe a steady increase in the probability of initiating IVF as the woman ages, reaching over 25 per 10,000 initiations in the month just prior to the cutoff. There is no apparent anticipatory pattern. The initiation rate drops sharply at the eligibility cutoff, to circa 12 per 10,000 initiations. The initiation rate then stays flat for around two years, and gradually declines to nearly zero at five years after the cutoff, when all women are over the age of 40 and many are approaching age 45 (depending on their region's coverage cutoffs).

Figure 6, panel (b), shows that the drop in the take-up of IVF is accompanied by a proportional drop in births within two years of treatment.[29] This figure plots the number of IVF initiations in each month of age relative to the cutoff in dark blue—this is the same data as in the numerator of panel (a). In light gray, we show the number of (the same) women giving birth within two years of each relative month. At the insurance eligibility cutoff, we observe a sharp decline in the number of children born, which is proportional to the decline in treatment initiations. Appendix Figure A7 shows that there is no change in the treatment success rate (panel (a)) or treatment intensity (panel (c)) around the cutoff, suggesting that the sharp drop in childbirths is a direct consequence of the reduction in the number of IVF attempts.

[28] Source: https://xn--fertilitetsrd-0fb.se/ivf-vantetid/. Wait times can be as high as 30 months in some regions, see https://www.altinget.se/vard/artikel/stora-skillnader-i-vantetider-for-barnlosa. See also Appendix Figure A6 for our estimates of wait times between an infertility diagnosis and treatment, which can be multiple years.

[29] We include any birth, without conditioning on the birth being attributable to any specific IVF cycle.

5.2 Formal Estimates

We next convert the graphical evidence in the preceding discussion into formal regression discontinuity (RD) estimates. Our running variable is the year-month of birth of an individual in the at-risk sample of all childless women relative to the region-year specific age cutoff for IVF coverage. Our primary specification is a parametric linear RD:

$$Y_{itc} = \beta_0 + \beta_1 \mathbb{1}\left[a_{it} > A_{ct}\right] + \beta_2 \left(a_{it} - A_{ct}\right) + \beta_3 \mathbb{1}\left[a_{it} > A_{ct}\right] \times \left(a_{it} - A_{ct}\right) + \mathbf{x}'_{itc}\kappa + \epsilon_{itc} \qquad (2)$$

Y_{itc} is the outcome of interest (e.g., an indicator for initiating IVF treatment) for individual i at time t residing in region c. A_{ct} denotes the age cutoff that region c has at time t, above which there is no health insurance coverage for IVF. The variable $\mathbb{1}[a_{it} > A_{ct}]$ is an indicator for the individual's age a_{it} at time t being above the region-year specific cutoff, and hence the individual being ineligible for coverage. The coefficient on this indicator, β_1, is the main coefficient of interest; it measures how much the probability of initiating IVF treatment changes when the woman becomes ineligible for insurance coverage. $(a_{it} - A_{ct})$ is an individual's age in months centered around the cutoff. We estimate how the initiation of IVF treatment varies by age and allow for this relationship to have different slopes on the opposite sides of the cutoff. We show results with and without a vector of controls x_{itc}, which include indicators for the region of residence, calendar years, an individual's place of birth, and education categories (high school only, some college, college degree or more).

The regression model is estimated on a panel of woman-year-month observations for women within two years on either side of the cutoff (note that this is a narrower window of the running variable than what is displayed in the graphical evidence). For each woman residing in region c in year t, we calculate relative age by comparing the month-year of birth to the age cutoff A_{ct} in region c in year t. A woman of child-bearing age is in our panel in a given year-month as long as she is childless. A woman exits the panel upon childbirth or IVF initiation—whichever (if any) occurs first. In addition to our primary specification, we also estimate a "donut" RD that omits five months on each side of the cutoff. Although we do not observe anticipatory behavior in the graphical evidence, the donut RD verifies that anticipatory behavior to the left of the threshold is not driving our baseline estimates. Throughout, we cluster standard errors at the individual level.

To interpret β_1 as the effect of insurance (or equivalently of the out of pocket price) on demand for IVF, we need two ingredients. First, we need to assume that, close to the threshold, women do not have precise control over whether they get access to IVF at the subsidized price. Note

that our design explicitly allows women, both before and after the threshold, to choose to "opt out" of the waiting list and get IVF at a private clinic, paying the full price out of pocket—and this is the only option that remains for women who (as good as randomly) do not get access to IVF at the subsidized price to the right of the cutoff. While age, by definition, is both perfectly predictable and cannot be manipulated, as we have discussed above, waiting lines for treatment and the unpredictability of needing treatment introduce an element of uncertainty. Thus, women can try to influence—but cannot perfectly control—when they will be able to start an IVF procedure.[30] The second assumption that we need is that no other major transitions that could affect IVF take-up happen at exactly the cutoff age. Here, we are helped by the fact that cutoff age varies across regions and over time. Appendix Figure A8, panel (a) shows the density around the cutoff, which reveals no evidence of treatment re-timing in our baseline sample (estimates from McCrary tests (McCrary, 2008) are reported in the notes). Further, panels (b), (c), and (d) plot average individual characteristics (income, having a college degree, and being foreign-born, respectively) measured at age 32 – well before any woman in our sample approaches the threshold – again showing no evidence of manipulation at the threshold.

Table 3 shows the results of estimating Equation (2) on observations 24 months around the cutoff. In the full sample of childless women, 23 per 10,000 women initiate IVF on average in the twelve months before their insurance coverage eligibility lapses. As women age out of insurance coverage, their rate of initiations drops by 12 per 10,000 or by 51%. The estimate is not very sensitive to including controls (column (2), drop of 51%). We get a similar estimate of the decline (45%) in the "donut" specification as reported in column (3).[31] The last two columns of the table show the estimated effect (using the specification including controls) in two samples: splitting women into below versus above the median of the income distribution based on their disposable income measured at age 32 (far away from the cutoff). The price response as a share of the pre-cutoff mean is smaller as we move up in the income distribution—a fact that we consider in more detail in the next sub-section.

[30] In a newspaper article covering women who have aged out while waiting for IVF – and thus missed out on coverage – the representative of one of Sweden's regions, Region Uppsala, advises women to start at least one year in advance in order to reduce the likelihood of aging out; further, he states that there is no protocol to prioritize women who approach 40 over other patients. See https://www.aftonbladet.se/family/a/V9nyjp/ensamstaende-hann-inte-fa-assisterad-befruktning (in Swedish).

[31] As we would expect, IVF initiation rates are substantially higher in the sample of women with a history of LIT treatment. 410 per 10,000 women initiate IVF in that sample in the twelve months before the age cutoff (Table A3). This rate goes down by 167 per 10,000 (or 137 per 10,000 in the donut RD estimate), or by 41% (33% in the donut). In other words, women with an observed history of infertility treatment are somewhat less responsive to the change in insurance coverage.

5.3 Willingness to Pay

All estimates in the previous sub-section imply that the decision to initiate an IVF treatment is sensitive to the availability of insurance coverage, and in turn to the out of pocket price of IVF. In this section we convert our estimates of price response into willingness to pay terms.

Aggregate Willingness to Pay Estimate One cycle of IVF during our sample period costs circa \$3,250 (in 2022 US dollars). This implies a full price of about \$8,775 for the course of 2.7 months of IVF treatment,[32] which is the average length of treatment around the age cutoff (panel (c) of Appendix Figure A7). This is 2,584% higher than the cost of \$327 in co-pays under insurance coverage. The implied arc elasticity in the full sample for which we observe a 51.3% drop in demand amounts to -0.37. To get at the maximum willingness to pay in our sample, we extrapolate out of sample by assuming that demand for IVF is linear in price. Our estimates imply that the rate of IVF initiations would go to zero if the price of the full course of treatment went up to \$16,741. As the full course of treatment is observed to yield a 40% birth rate around the age cutoff (Panel (a) of Appendix Figure A7), this in turn implies a maximum willingness to pay for a 40% chance of having a child of \$16,741.[33]

These willingness to pay estimates are hard to interpret without the context of consumption levels in Swedish households. \$8,775 (the cost of 2.7 cycles of IVF) is the equivalent of 2.1 months of average (couple's) disposable income, or 18% of annual disposable income. The *maximum* willingness to pay for a 40% chance of having a child in the full sample is then equivalent to four monthly incomes, or a third of average annual disposable income.

Willingness to Pay by Income Naturally, the effective price of IVF in relation to a couple's income will vary substantially across the income distribution, and (given the substantial cost of the treatment) so likely will willingness to pay. Figure 7, panel (a), shows the price of 2.7 months of IVF

[32] We assume that each month in which a woman is observed purchasing an IVF-related drug corresponds to one treatment cycle.

[33] This last computation assumes that individuals' perception of the IVF success rate is accurate. IVF success rates are widely publicized by clinics and when surveyed, individuals report an IVF success rate of 47% to 59%, which is close to what we observe in different subsamples of the data (Devroe et al., 2022; Fauser et al., 2019); thus, this assumption appears reasonable in our context. Another, more nuanced, assumption in this calculation is that individuals make their decision about initiating treatment based on the full expected cost of 2.7 cycles. An alternative would be to assume that the behavior at the threshold reflects only the cost of doing one initial cycle of IVF, after which couples can decide whether to proceed based on information learned in the first cycle. As standard protocols recommend three cycles before reevaluating the chances of success, and as the average duration of treatment remains unchanged after the cutoff, we assume that couples make initiation decisions based on the expectation that 2.7 may be needed.

treatment as a multiple of couples' monthly disposable income, at different points in the income distribution.[34] At the lowest decile of the income distribution, a couple must give up 3.5 months of disposable income to pay for IVF (or about 30% of annual income); a couple in the top decile must give up only 1.4 months of disposable income (12% of annual income).

Figure 7, panel (b), plots estimates from our baseline regression discontinuity, estimated separately at each decile of the income distribution. Fitting a linear regression to the pattern of these RD coefficients, we estimate that price sensitivity, as measured by the percent change in demand at the insurance eligibility cutoff, drops by three percentage points for each income decile. These estimates translate into a maximum willingness to pay (for a 40% chance of having a child) of circa $14,000 for couples in the lowest income decile and $20,000 for couples in the highest income decile.[35] Notably, this gradient *reverses* when we consider willingness to pay in terms of relative (as opposed to absolute) income: Couples in the lowest decile of the income distribution are willing to pay at most 5.5 months of their disposable income for IVF, while couples at the top are willing to pay at most 3.1 months of their disposable income. This implies that lower-income households scale down their willingness to pay disproportionately *less* than what would have been implied by their consumption levels.

Potential Role of Liquidity There are two potential drivers of revealed willingness to pay in this context. The first—textbook—explanation would consider the full response to insurance coverage to be traditional moral hazard, reflecting changes in the consumption of medical care in response to distortions in relative prices. Within that framework, our estimates would imply that on average, about half of childless families who want a child, and are close to the age cutoff, value IVF below its cost, and thus below the value of the insurance subsidy that covers treatment at close to 100%.[36] An alternative explanation is that (at least part) of the response reflects liquidity constraints. In the presence of liquidity constraints, couples may be willing to pay for IVF treatment—but they may not be able to come up with the lump-sum of a third of their annual consumption equivalent. In that case, insurance would be providing a consumption smoothing benefit rather than (or in addition to) inducing moral hazard.

Two observations suggest that liquidity constraints are likely important in our setting. First,

[34] As we do not always observe income for partners, we define income percentiles (x-axis) based on the woman's disposable income, in order to keep the x-axis the same in panel (a) and panel (b) of Figure 7. To obtain our measure of couple income on the y-axis in panel (a), we use the sub-sample for which we observe partners.

[35] In our sample, we observe no statistically significant income gradient in the success rate (Appendix Figure A7, panel (c)). Interestingly, Groes et al. (2024) observe an education gradient in IVF success in Denmark.

[36] The valuation of IVF here is unlikely to reflect downward biased beliefs about success chances; see footnote 33.

in their analysis of Swedish asset data, Kolsrud et al. (2020) find that Swedish households almost fully rely on government programs to smooth consumption in response to health shocks, with no evidence of other smoothing mechanisms being readily available.[37] Second, in many countries, including Sweden, there has been a pronounced growth of fertility-specific funding arrangements over the last two decades.[38] A key benefit of these types of credit solutions is that they do not require a collateral, they are interest-free up to certain amounts, and they are available to individuals across the income distribution.[39] While our estimates cover the period where such plans were not yet widespread (and hence demand behavior would still reflect liquidity constraints), the emergence of the market for this financial smoothing instrument specifically for IVF implies that liquidity is an important factor driving demand.

While the exact magnitude of liquidity constraints in relation to moral hazard is hard to quantify, we can make some progress by following the approach similar in spirit to that in Chetty (2008). Consider the difference in the magnitude of the behavioral response across income levels that we have shown in Figure 7, panel (b). Conceptually, this income gradient can either reflect an income gradient in the ability to pay a high lump-sum price for IVF treatments (a liquidity effect), or an income gradient in the underlying preferences and/or ability to pay the high price of conceiving a child even if it were not lump-sum (moral hazard). It appears reasonable to interpret the response of high-income individuals in Figure 7, panel (b), as a measure of behavioral response (i.e. moral hazard) among liquidity unconstrained households. If we assume that moral hazard is constant across the income distribution, we would estimate that a third of the reduction in fertility treatment utilization among the lowest-income households can be attributed to the liquidity effect rather than traditional moral hazard.

Naturally, a constant degree of moral hazard across the income distribution is a strong assumption. It is very plausible that both innate preferences for children as well as the marginal utility for non-IVF and non-child-related consumption altogether (or "ability to pay" from lifetime income) could be lower among lower-income households. Appendix Figure A9 sheds some (suggestive) light

[37] The median individual in Kolsrud et al. (2020)'s sample of individuals who experience a health shock is 38 years old – close to the age thresholds in our RD design – and has only 77,000 SEK in total net wealth (ca. $8,500) and zero liquid assets (bank holdings) prior to the onset of the health shock. This net wealth level is nearly equal to the full cost of 2.7 cycles of IVF treatment.

[38] In Sweden, today, several IVF clinics have partnered up with Human Finans, which is one such credit institution. Human Finans was founded in 2004, and has been growing over time.

[39] At one clinic (Linne kliniken), customers are currently (at the timing of writing) able to pay interest-free installments over a period of six to twelve months for a maximum amount of SEK 50,000 (which would cover ca. two IVF cycles) through Human Finans. Amounts exceeding 50,000 SEK and up to 100,000 SEK are also possible to pay through installments, but these are not interest-free. Source: Linne kliniken. https://www.linne.se/priser/

on these alternative interpretations. It plots (for the full population) the share of women who have non-ART children, by the end of their fertile years, by income rank. The figure includes cohorts born in 1970 to 1975, so that we can observe their completed fertility by 2019. While women at the top decile of the income distribution are more likely to have a child than women at the bottom decile, this difference is small—of only two percentage points relative to the overall level of more than 80% of women having a child. This relationship between income and the probability of having a child captures both the heterogeneity in innate preferences for having a child across the income distribution, and how those preferences may be moderated by the high, but mostly not lump-sum, lifetime cost of raising a child. Swedbank (a large bank in Sweden) estimates the child-rearing cost in Sweden to be on the order of ca. $135,000 net of government transfers over the course of 18 years (Swedbank, 2024).[40] This implies that for couples who are revealing their strong preference for having a child and the willingness to take on the associated lifetime expenses, an IVF procedure would raise the overall cost of having a child by a relatively modest 6%.

We interpret Appendix Figure A9 as ultimately suggesting two things. First, there is no clear evidence of differences in innate preference for a child and/or the willingness to devote a substantial amount of lifetime income to child-related consumption across the income distribution.[41] Second, the price of IVF is modest relative to the overall cost of raising a child. This brings us back to the idea that liquidity likely plays a central role in the observed response to the price of IVF. While this evidence is certainly suggestive, it highlights the importance of understanding the degree to which IVF insurance coverage provides redistribution to individuals with higher marginal utility of income or the consumption smoothing benefit for liquidity-constrained households.[42]

5.4 Distribution of Children

The sharp drop in IVF use and births in the absence of insurance highlights how public policy can shape the number of children born, and not born, at the population level. To further inform this discussion, we next consider several policy counterfactuals. Using our regression discontinuity

[40] For a benchmark, USDA estimated the cost (without the cost of college) of raising a child born in 2015 in the US to be $233,610 for a middle-income family (U.S. Department of Agriculture, 2024). They estimate that lower-income families will spend $174,690, while families with higher incomes are expected to spend $372,210. Applying the same scaling factor to the Swedish number (assuming it reflects a median family), we would get a cost of circa $100,000 for lower-income families, which is still far more substantial than the cost of IVF.

[41] This implicitly interprets fertility as a rational economic decision. The wide availability and the low cost of family planning services in Sweden make interpretation more plausible in our institutional context.

[42] Naturally, if consumption smoothing or redistribution are an important part of what insurance does, a question arises as to whether using health insurance to address market failures in credit markets or for redistribution may in itself be a second-best policy—an important question for future work.

sample, we predict the number of children that would have been born under two different policy regimes; one in which the subsidy is universal and not tied to age, the other in which there is no subsidy.

Figure 8, panel (a) shows the number of women giving birth within two years of having initiated IVF by age relative to the subsidy cutoff (this is simply replicating the gray bars from panel (b) of Figure 6). To estimate the number of births that would have occurred in a counterfactual regime of a universal subsidy, we use data points to the left and a linear regression to form a prediction (dashed gray line). To estimate the number of births that would have occurred in a counterfactual regime of no subsidy, we perform a similar exercise using only data points to the right of the cutoff (solid line). The difference between predicted births and actual births on each side of the cutoff constitutes an estimate of "incremental" or "missing" births, respectively, under the assumption that the observed fertility responses are causal effects of the subsidy design. In panel (a), we estimate approximately 2,120 missing births under the current policy regime relative to a regime with no age-based cutoffs, and 4,013 incremental births as a result of the current subsidy relative to a regime with no subsidy. In total, 13,450 children would have been born if the subsidy were available to everyone, while 7,317 (or 46% fewer) children would have been born in the absence of subsidies.

In addition to affecting the birth rate in the overall population, panels (b) and (c) illustrate how the effects differ across childless women with above and below median income, respectively. The figures reveal that the mass of missing births is more pronounced in the below-median income population (1,285 vs 187 missing births). In relative terms, the removal of the subsidy to the right of the age threshold results in 70% of potential IVF children not being born (becoming missing) in the below-median income population, as compared to 21% in the above-median income population. Thus, in terms of IVF children born, the age cutoff policy has a substantially larger "bite" at the lower end of the income distribution, where it eliminates a much larger share of the IVF children that would have been born within two years in a counterfactual world with no insurance eligibility cutoff (i.e., if subsidies were universally available at any age). This underscores that subsidy design in this context affects not only the level, but also the ultimate distribution of births across the income spectrum.

5.5 Normative Analysis

Notably the estimates above imply that the distributional nature of subsidizing IVF depends on whether we consider couples before or after infertility information is realized. Higher-income couples have almost twice the rate of IVF births. Hence, the incidence of overall spending on a universal IVF subsidy will mostly fall onto higher-income couples. In this sense, subsidizing IVF is regressive. At the same time, conditional on a couple finding out that they are infertile, the subsidy is progressive in the sense that as a share of annual income it is larger for lower-income couples, and many of these couples would forgo having children in the absence of the subsidy.

To make this more concrete, suppose the government offers the subsidy to any infertile woman in the population. Per 10,000 eligible women, 24 women will initiate IVF every month, out of which 8.6 women will be from below-median income households and 15.4 women from above-median income households (these numbers are based on the relative means of dependent variables in Table 3, Columns (4) and (5)). For 10,000 women, the government then expects to pay ca. $205,000 per month, 64% of which accrues to above-median income households (and hence this policy is regressive).

In the marginal value of public funds (MVPF) framework (Finkelstein and Hendren, 2020; Hendren and Sprung-Keyser, 2020), a universal policy such as the one described above, which offers an IVF subsidy to any infertile woman in the population regardless of her income or age, has an aggregate MVPF of 0.8.[43] We get a somewhat smaller MVPF if instead of considering a counterfactual with and without universal subsidies, we focus on a policy of removing the age ceiling and subsidizing older women. The MVPF "to the right of the cutoff" is 0.68, which is lower because a higher share of beneficiaries is marginal to the subsidy to the right of the cutoff.

If we compute the MVPF "to the right of the cutoff" for lower and higher-income households separately, we get an MVPF of 0.76 for higher-income households and an MVPF of 0.66 for lower-income individuals.[44] The MVPF is lower for lower-income individuals because this calculation

[43]The numerator of MVPF is the full benefit to the subsidy beneficiaries. In our case, in our overall RD sample, we estimate that 61% of beneficiaries are inframarginal—the subsidy doesn't change their rate of IVF initiations (these estimates integrate the projected IVF initiations across all months −60 to 60 around the cutoff, analogously to the projections in Figure 8). For these beneficiaries a subsidy of $8,448 (full cost of 2.7 cycles of IVF net of co-pays in public insurance that beneficiaries also pay when they are covered by insurance) is a transfer worth $8,448. 39% of beneficiaries are marginal and do not initiate an IVF treatment in the absence of the subsidy. The IVF subsidy is large, hence we cannot assume that the marginal beneficiaries are indifferent. Instead, we use the Harberger approximation (Harberger, 1971; Finkelstein and Hendren, 2020) and assume that the marginal beneficiary values the subsidy (i.e., the provision of in-kind IVF treatment) at 50%, or at $4,224. Adding these together, we get an MVPF numerator of $6,800. Dividing the benefits by costs, we get an MVPF of 0.81. We get an MVPF of 0.77 if we use the response of childbirths to measure the share of marginal beneficiaries.

[44]To obtain these MVPFs, we use the income-specific shares of inframarginal versus marginal beneficiaries (49%

attributes a lower social valuation to a subsidy directed at an individual who is marginal, i.e., who ceases to use IVF in the absence of the subsidy, than to a subsidy directed at an individual who is inframarginal and values a dollar of subsidy at the full dollar (and indeed, the share of marginal beneficiaries is much higher at the lower end of the income distribution). This implicitly assumes that the private behavioral response to the subsidy (and, in particular, which share of beneficiaries is marginal) is the right normative measure of benefits from the subsidy. But as we discussed above, among lower-income households the behavioral response may in part reflect liquidity constraints or low levels of income. If so, society's willingness to pay for an IVF subsidy may be higher if it is directed at a lower-income household.[45]

Finkelstein et al. (2019a) discuss a similar point in the context of private willingness to pay for a health insurance program serving lower-income families in the US, Medicaid. As they highlight, one approach to thinking about the private versus societal MVPF involves specifying social welfare weights (Saez and Stantcheva, 2016). The pooled MVPF of a universal subsidy for IVF in our context would be circa one if the society's valuation of providing IVF to below-median income households were just above two times the private valuation.[46]

6 Concluding Discussion

As the share of children born from conceptions assisted by reproductive technologies keeps growing around the world, public policies in the space of infertility treatments remain ridden with controversy. These debates highlight a critical need for more evidence on the rate of infertility burden, its private and public costs, and the impact on infertile couples of public policies that influence the affordability of infertility treatments. This paper uses population-wide Swedish administrative

versus 69% marginal beneficiaries in the high- versus low-income population), and implement the same calculation as in footnote 43. Notably if we use the childbirth response, the difference in the share of marginal beneficiaries is much more pronounced (21% versus 70% marginal beneficiaries in the high- versus low-income population). In that case, we get MVPFs "to the right of the cutoff" of 0.65 and 0.9 among below-median versus above-median beneficiaries, respectively.

[45] Societal willingness to pay for a couple of any income level may also differ from private willingness to pay if the government internalizes a variety of externalities that may arise from subsidizing IVF. First, the government may be concerned about dynamic effects of IVF coverage on the timing of childbearing. In several contexts, it has been documented that IVF subsidies may induce delay in when women decide to try to have children (see e.g., Abramowitz, 2017; Gershoni and Low, 2021a). Second, the government may internalize positive or negative externalities from having more children overall. Both issues are beyond the scope of this paper, but are important for further understanding of optimal IVF coverage policies.

[46] We obtain this number from the following calculation: 0.64×0.76 for higher income households plus $0.36 \times 2.2 \times 0.66$ for lower-income households, where 0.64 and 0.36 reflect the shares of spending on the subsidy that accrue to beneficiaries at the high and low end of the income spectrum, respectively, and 0.76 and 0.66 are the income-specific MVPFs.

data with uniquely detailed information on individual-level use of ARTs, combined with empirical methods that leverage quasi-natural experiments, to begin to fill some of these gaps. Our analysis delivers three main take-aways.

First, involuntary primary medical infertility is very common. We estimate that one in eight women will experience primary infertility – the inability to have any child at all – over her fertile years, a risk on par with the lifetime risk of breast cancer (American Cancer Society, 2024). This estimate, which leverages couples' revealed willingness to have a child through drug purchase data, improves on existing survey-based measures of the rate of infertility burden that rely on recall (WHO, 2023), and puts the real prevalence of infertility in stark contrast to how it is often perceived in public discourse – indeed, infertility was only recognized as a disease by the American Medical Association in 2017 (Adashi, 2018).

Second, our analysis reveals large and negative impacts of wanting, but being unable, to have a child. Women who remain involuntarily childless are 48% more likely to suffer from mental health issues several years after encountering infertility, and are six percentage points more likely to get divorced. Further, persistent infertility has no "protective" long-run effect on women's incomes. Involuntary childlessness also causes a deterioration of partners' mental health and (some) longer-run income losses. In other words, to the extent that common wisdom suggests that children reap havoc on women's careers or make couples unhappy, our results offer a sharp counterpoint. One interpretation of these results is that, while the arrival of a child causes a drop in women's earnings, the longer-run toll of infertility on mental well-being among women who are involuntarily childless may have a countervailing (negative) impact on their earnings. In practice, this would suggest that women who desire children may not be able to avoid an earnings penalty by remaining childless. In a nutshell, it may be the *preference to have a child per se* that is associated with a penalty – regardless of whether the woman is able to realize her wish to have a child.

Third, we show that policies that affect the affordability of IVF have major impacts on the use of IVF treatments, especially at the lower end of the income spectrum. This, in turn, means that insurance coverage of IVF ultimately affects the distribution of children across the income spectrum: In the absence of insurance coverage, many potential IVF children are never conceived, and the share of such non-conceived children is larger at the lower end of the income spectrum. A policy that makes IVF universally affordable is thus progressive in terms of its impact on the socioeconomic distribution of children. However, as high-income individuals are more likely to use IVF, it is regressive in terms of overall spending. These disparate effects across the income distribution imply

that the ultimate marginal value of public funds devoted to a universal subsidy depends crucially on the social welfare weights placed on beneficiaries across the income distribution.

These implications for the design of government policies that affect affordability of ARTs offer a perspective that, more broadly, has the potential to be relevant for the millions of couples that are struggling with infertility around the world. Today, IVF insurance coverage varies widely across countries. In the OECD context, many public health insurance programs offer some coverage, but with many and varying limitations. The debate about insurance coverage of IVF is particularly salient in the United States. The U.S. National Survey of Family Growth estimates that 19% of currently married childless women ages 15–49 in the U.S. are experiencing infertility (Nugent and Chandra, 2024) at a given point in time. This is about 2.5 million women, a number that almost surely underestimate the national prevalence of infertility among all childless women.[47] In the absence of a federal health insurance program for this population of child-bearing age in the U.S., coverage of infertility treatments that these women have access to varies dramatically across states, employers, and income groups.[48]

As a result of limited insurance coverage, many couples experiencing infertility in the U.S. pay out-of-pocket for treatments that can cost between $40,000–$60,000 (Fertility IQ, 2024).[49] Our analysis in Section 5 suggests the this high price tag may substantially deter many couples from pursuing IVF altogether. To illustrate the magnitudes in the U.S. context: At a point in time, 4.3 million of 30–39 year old women are estimated to be childless in the US. Applying our RD estimates (grossly) out of sample, we would conclude that a full universal IVF subsidy (relative to no subsidy),

[47] Among 40 to 49 year old childless currently married women (who are close to having completed their fertility) in the US, 26.8% are considered to be experiencing infertility. Our estimates in Section 4 would imply that about 1.3 million women in the US may be experiencing the long-run cost to well-being from infertility, including a nearly 50% increase in the prevalence of mental health illness.

[48] As of 2023, 21 U.S. states and Washington D.C. have passed laws mandating insurance coverage of infertility treatments. 15 of those states include IVF coverage. There is substantial variation in what a "coverage mandate" means in practice (Kaiser Family Fundation, 2020). Across states with some coverage mandate in place, there is variation in the type of infertility service covered (e.g., diagnosis or treatments like IVF), the type of insurers affected by the mandate (.e.g, HMO, large employers plans), and the type of indication (e.g., infertility for at least a year and within certain age ranges). Individuals with low income are particularly exposed to limited coverage. As of 2020, 12 Medicaid state programs offered some coverage. These program, however, tend to be very incomplete and may cover only the prescription drug portion of the treatment, or diagnostic services, but not the treatment procedures themselves. Several states are debating whether infertility coverage should be included in their Medicaid programs.

[49] As in Sweden, credit markets for financing IVF have been growing in the U.S., suggesting that liquidity constraints are also likely to be important in the U.S. context. More than 25 different financing programs were available to customers as of 2023 (`https://resolve.org/learn/financial-resources-for-family-building/`). For instance, CCRM Fertility, a large infertility treatment provider in the U.S., has currently partners with four credit institutions, including CapexMD, Lighstream, future family, and Lending Club. The conditions vary widely across financing providers, with for instance Lighstream offering loans to customers with good to excellent credit scores, while Lending Club has no such specific requirement. Some programs are offered by the clinics, while others involve clinics partnering with credit institutions.

could generate 3.5% extra first births in this age group.[50]

Our results thus highlight the substantial effects that IVF financing policies may have on the number and distribution of children. Arriving at the *optimal* design of public policies affecting the affordability of ARTs further requires understanding whether the most effective intervention would be subsidies or loans, and in the case of subsidies whether the main goal is redistribution, in which case policy instruments other than public health insurance systems may be more efficient. We leave these important questions for future work.

References

Abramowitz, Joelle, "Turning Back the Ticking Clock: The Effect of Increased Affordability of Assisted Reproductive Technology on Women's Marriage Timing," *Journal of Population Economics*, 2014, *27* (2), 603–633.

_ , "Assisted Reproductive Technology and Women's Timing of Marriage and Childbearing," *Journal of Family and Economic Issues*, 2017, *38* (1), 100–117.

Adashi, Eli Y., "Infertility: A Disease by Any Other Name," *JAMA Forum Archive*, 02 2018, *A7* (1).

Ahammer, Alexander, Ulrich Glogowsky, Martin Halla, and Timo Hener, "The Parenthood Penalty in Mental Health: Evidence from Austria and Denmark," *IZA Discussion Paper No. 16459*, 2023.

Albert, Victoria, "Major Alabama Health System Pauses IVF Treatments After State Court Says Frozen Embryos Qualify as Children," *The Wall Street Journal*, February 21, 2024.

American Cancer Society, "Key Statistics for Breast Cancer," 2024. Accessed June 2024. https://www.cancer.org/cancer/types/breast-cancer/about/how-common-is-breast-cancer.html#:~:text=younger%20than%2045.-,Lifetime%20chance%20of%20getting%20breast%20cancer,will%20never%20have%20the%20disease.

Andresen, Martin Eckhoff and Emily Nix, "What Causes the Child Penalty? Evidence From Adopting and Same-Sex Couples," *Journal of Labor Economics*, 2022, *40* (4), 971–1004.

Angelov, Nikolay, Per Johansson, and Erica Lindahl, "Parenthood and the Gender Gap in Pay," *Journal of Labor Economics*, 2016, *34* (3), 545–579.

Banks, Nicole K, John M Norian, M Kate Bundorf, and Melinda B Henne, "Insurance Mandates, Embryo Transfer, Outcomes—The Link is Tenuous," *Fertility and Sterility*, 2010, *94* (7), 2776–2779.

Belchior, Carlos Alberto and Yara Gomes, "Liquidity Constraints, Cash Transfers and the Demand for Health Care in the Covid-19 Pandemic," *Health Economics*, 2022, *31* (11), 2369–2380.

[50]The estimates in our RD analysis imply that 144 per 10,000 would initiate IVF and give birth to 58 children per 10,000 women per year. With a full subsidy, the number of births would be double at 116 per 10,000, or about 50,000 children per year in this age group, which is 3% of all births, or 7% of all first births in this age group in the US, for 3.5% of incremental first births.

Bensnes, Simon, Ingrid Huitfeldt, and Edwin Leuven, "Reconciling Estimates of the Long-Term Earnings Effect of Fertility," Report, IZA Discussion Papers 2023.

Bitler, Marianne and Lucie Schmidt, "Health Disparities and Infertility: Impacts of State-Level Insurance Mandates," *Fertility and Sterility*, 2006, *85* (4), 858–865.

Bitler, Marianne P and Lucie Schmidt, "Utilization of Infertility Treatments: The Effects of Insurance Mandates," *Demography*, 2012, *49* (1), 125–149.

Boulet, Sheree L, Sara Crawford, Yujia Zhang, Saswati Sunderam, Bruce Cohen, Dana Bernson, Patricia McKane, Marie A Bailey, Denise J Jamieson, and Dmitry M Kissin, "Embryo Transfer Practices and Perinatal Outcomes by Insurance Mandate Status," *Fertility and Sterility*, 2015, *104* (2), 403–409.

Bratti, Massimiliano and Laura Cavalli, "Delayed First Birth and New Mothers' Labor Market Outcomes: Evidence from Biological Fertility Shocks," *European Journal of Population*, 2014, *30*, 35–63.

Buckles, Kasey S, "Infertility Insurance Mandates and Multiple Births," *Health Economics*, 2013, *22* (7), 775–789.

Bundorf, M Kate, Melinda Henne, and Laurence Baker, "Mandated Health Insurance Benefits and the Utilization and Outcomes of Infertility Treatments," *NBER Working Paper No. 12820*, 2007.

Bütikofer, Aline, Deirdre Coy, Orla Doyle, and Rita Ginja, "The Consequences of Miscarriage on Parental Investments," *IZA DP No. 16858*, 2024.

Cetre, Sophie, Andrew E Clark, and Claudia Senik, "Happy People Have Children: Choice and Self-Selection into Parenthood," *European Journal of Population*, 2016, *32* (3), 445–473.

Chetty, Raj, "Moral Hazard versus Liquidity and Optimal Unemployment Insurance," *Journal of Political Economy*, 2008, *116* (2), 173–234.

Cintina, Inna and Bingxiao Wu, "How Do State Infertility Insurance Mandates Affect Divorce?," *Contemporary Economic Policy*, 2019, *37* (3), 560–570.

Conner, Peter, Liran Einav, Amy Finkelstein, and Petra Persson, "Targeting Precision Medicine: Evidence from Prenatal Testing," *NBER Working Paper No. 30669*, 2022.

Crawford, Sara, Sheree L Boulet, Denise J Jamieson, Carol Stone, Jewel Mullen, and Dmitry M Kissin, "Assisted Reproductive Technology Use, Embryo Transfer Practices, and Birth Outcomes After Infertility Insurance Mandates: New Jersey and Connecticut," *Fertility and Sterility*, 2016, *105* (2), 347–355.

Cutler, David, Angus Deaton, and Adriana Lleras-Muney, "The Determinants of Mortality," *Journal of Economic Perspectives*, 2006, *20* (3), 97–120.

Daysal, N Meltem, Hui Ding, Maya Rossin-Slater, and Hannes Schwandt, "Germs in the Family: The Short-and Long-Term Consequences of Intra-Household Disease Spread," *NBER Working Paper No. 29524*, 2021.

Devroe, J, K Peeraer, TM D'Hooghe, J Boivin, A Laenen, J Vriens, and EAF Dancet, "Great Expectations of IVF Patients: the Role of Gender, Dispositional Optimism and Shared IVF Prognoses," *Human Reproduction*, 2022, *37* (5), 997–1006.

Einav, Liran and Amy Finkelstein, "Moral Hazard in Health Insurance: What We Know and How We Know It," *Journal of the European Economic Association*, 05 2018, *16* (4), 957–982.

Ericson, Keith Marzilli and Justin R Sydnor, "Liquidity Constraints and the Value of Insurance," *NBER Working Paper No. 24993*, 2018.

Fadlon, Itzik and Torben Heien Nielsen, "Family Health Behaviors," *American Economic Review*, September 2019, *109* (9), 3162–91.

_ and _, "Family Labor Supply Responses to Severe Health Shocks: Evidence from Danish Administrative Records," *American Economic Journal: Applied Economics*, July 2021, *13* (3), 1–30.

Fauser, Bart CJM, Jacky Boivin, Pedro N Barri, Basil C Tarlatzis, Lone Schmidt, and Rachel Levy-Toledano, "Beliefs, Attitudes and Funding of Assisted Reproductive Technology: Public Perception of Over 6,000 Respondents From 6 European Countries," *PLoS One*, 2019, *14* (1), e0211150.

Fertility IQ, "What is the Total Cost of IVF?," 2024. Accessed June 2024. https://www.fertilityiq.com/fertilityiq/ivf-in-vitro-fertilization/costs-of-ivf#what-is-the-total-cost-of-ivf.

Finkelstein, Amy and Nathaniel Hendren, "Welfare Analysis Meets Causal Inference," *Journal of Economic Perspectives*, November 2020, *34* (4), 146–67.

_ , _ , and Erzo FP Luttmer, "The Value of Medicaid: Interpreting Results from the Oregon Health Insurance Experiment," *Journal of Political Economy*, 2019, *127* (6), 2836–2874.

_ , _ , and Mark Shepard, "Subsidizing Health Insurance for Low-Income Adults: Evidence from Massachusetts," *American Economic Review*, April 2019, *109* (4), 1530–67.

Gallen, Yana, Juanna Schrøter Joensen, Eva Rye Johansen, and Gregory F Veramendi, "The Labor Market Returns to Delaying Pregnancy," *Available at SSRN 4554407*, 2023.

Gershoni, Naomi and Corinne Low, "Older Yet Fairer: How Extended Reproductive Time Horizons Reshaped Marriage Patterns in Israel," *American Economic Journal: Applied Economics*, 2021, *13* (1), 198–234.

_ and _, "The Power of Time: The Impact of Free IVF on Women's Human Capital Investments," *European Economic Review*, 2021, *133*, 1036–45.

Glass, Jennifer, Robin W Simon, and Matthew A Andersson, "Parenthood and Happiness: Effects of Work-Family Reconciliation Policies in 22 OECD Countries," *American Journal of Sociology*, 2016, *122* (3), 886–929.

Glied, Sherry and Adriana Lleras-Muney, "Technological Innovation and Inequality in Health," *Demography*, 2008, *45*, 741–761.

Goldin, Claudia, *Career and Family: Women's Century-Long Journey toward Equity*, Princeton University Press, 2021.

Graham, Ruth, "Southern Baptists Vote to Oppose Use of I.V.F.," *The New York Times*, June 12, 2024.

Griffin, Martha and William F Panak, "The Economic Cost of Infertility-Related Services: An Examination of the Massachusetts Infertility Insurance Mandate," *Fertility and Sterility*, 1998, *70* (1), 22–29.

Groes, Fane, Anna Houstecka, Daniela Iorio, and Raul Santaeulalia-Llopis, "The Unequal Battle Against Infertility: Theory and Evidence from IVF Success," 2024. Centre for Economic Policy Research.

Gross, Tal, Timothy Layton, and Daniel Prinz, "The Liquidity Sensitivity of Healthcare Consumption: Evidence from Social Security Payments," *American Economic Review: Insights*, 06 2022, *4*, 175–190.

Hamilton, Barton H and Brian McManus, "The Effects of Insurance Mandates on Choices and Outcomes in Infertility Treatment Markets," *Health Economics*, 2012, *21* (8), 994–1016.

_ , Emily Jungheim, Brian McManus, and Juan Pantano, "Health Care Access, Costs, and Treatment Dynamics: Evidence from In Vitro Fertilization," *American Economic Review*, 2018, *108* (12), 3725–77.

Harberger, Arnold C, "Three Basic Postulates for Applied Welfare Economics: an Interpretive Essay," *Journal of Economic Literature*, 1971, *9* (3), 785–797.

Heazell, Alexander EP, Dimitrios Siassakos, Hannah Blencowe, Christy Burden, Zulfiqar A Bhutta, Joanne Cacciatore, Nghia Dang, Jai Das, Vicki Flenady, and Katherine J Gold, "Stillbirths: Economic and Psychosocial Consequences," *The Lancet*, 2016, *387* (10018), 604–616.

Hendren, Nathaniel and Ben Sprung-Keyser, "A Unified Welfare Analysis of Government Policies," *Quarterly Journal of Economics*, 2020, *135* (3), 1209–1318.

Henne, Melinda B and M Kate Bundorf, "Insurance Mandates and Trends in Infertility Treatments," *Fertility and Sterility*, 2008, *89* (1), 66–73.

Hotz, V Joseph, Charles H Mullin, and Seth G Sanders, "Bounding Causal Effects Using Data from a Contaminated Natural Experiment: Analysing the Effects of Teenage Childbearing," *The Review of Economic Studies*, 1997, *64* (4), 575–603.

_ , Susan Williams McElroy, and Seth G Sanders, "Teenage Childbearing and its Life Cycle Consequences: Exploiting a Natural Experiment," *Journal of Human Resources*, 2005, *40* (3), 683–715.

Jain, Tarun, Bernard L Harlow, and Mark D Hornstein, "Insurance Coverage and Outcomes of In Vitro Fertilization," *New England Journal of Medicine*, 2002, *347* (9), 661–666.

Jaravel, Xavier, "The Unequal Gains from Product Innovations: Evidence From the US Retail Sector," *The Quarterly Journal of Economics*, 2019, *134* (2), 715–783.

Jayachandran, Seema, Adriana Lleras-Muney, and Kimberly V Smith, "Modern Medicine and the Twentieth Century Decline in Mortality: Evidence on the Impact of Sulfa Drugs," *American Economic Journal: Applied Economics*, 2010, *2* (2), 118–146.

Johansson, Marianne, Annsofie Adolfsson, Marie Berg, Jynfiaf Francis, Lars Hogström, Per Olof Janson, Jan Sogn, and Anna-Lena Hellström, "Quality of Life for Couples 4–5.5 Years After Unsuccessful IVF Treatment," *Acta Obstetricia Et Gynecologica Scandinavica*, 2009, *88* (3), 291–300.

Junhong, Chu, "Prenatal Sex Determination and Sex-Selective Abortion in Rural Central China," *Population and Development Review*, 2001, *27* (2), 259–281.

Kaiser Family Fundation, "Mandated Coverage of Infertility Treatment," 2020. Accessed May 2024. https://www.kff.org/womens-health-policy/state-indicator/infertility-coverage/?currentTimeframe=0&sortModel=%7B%22colId%22:%22Location%22,%22sort%22:%22asc%22%7D.

Kalsi, Priti and Maggie Y Liu, "Pregnancy Loss and Female Labor Market Outcomes," *Available at SSRN 3829769*, 2022.

Kessler, Lawrence M, Benjamin M Craig, Shayne M Plosker, Damon R Reed, and Gwendolyn P Quinn, "Infertility Evaluation and Treatment Among Women in the United States," *Fertility and Sterility*, 2013, *100* (4), 1025–1032. e2.

Kjaer, Trille, Vanna Albieri, Allan Jensen, Susanne K Kjaer, Christoffer Johansen, and Susanne O Dalton, "Divorce or End of Cohabitation Among Danish Women Evaluated for Fertility Problems," *Acta Obstetricia Et Gynecologica Scandinavica*, 2014, *93* (3), 269–276.

Kleven, Henrik, Camille Landais, and Jakob Egholt Søgaard, "Children and Gender Inequality: Evidence from Denmark," *American Economic Journal: Applied Economics*, October 2019, *11* (4), 181–209.

Kolsrud, Jonas, Camille Landais, and Johannes Spinnewijn, "The Value of Registry Data for Consumption Analysis: An Application to Health Shocks," *Journal of Public Economics*, 2020, *189*, 104088.

Kroeger, Sarah and Giulia La Mattina, "Assisted Reproductive Technology and Women's Choice to Pursue Professional Careers," *Journal of Population Economics*, 2017, *30* (3), 723–769.

Larsen, Elisabeth Clare, Ole Bjarne Christiansen, Astrid Marie Kolte, and Nick Macklon, "New Insights Into Mechanisms Behind Miscarriage," *BMC Medicine*, 2013, *11*, 1–10.

Lleras-Muney, Adriana and Frank R Lichtenberg, "Are the More Educated More Likely to Use New Drugs?," *Annales d'Économie et de Statistique*, 2005, pp. 671–696.

Low, Corinne, "The Human Capital–Reproductive Capital Trade-Off in Marriage Market Matching," *Journal of Political Economy*, 2024, *132* (2), 552–576.

Lundborg, Petter, Erik Plug, and Astrid Würtz Rasmussen, "Can Women Have Children and a Career? IV Evidence from IVF Treatments," *American Economic Review*, 2017, *107* (6), 1611–37.

_ , _ , and Astrid Würtz Rasmussen, "Is There Really a Child Penalty in the Long Run? New Evidence from IVF Treatments," *IZA DP No. 16959*, 2024.

Lyngse, Frederik Plesner, "Liquidity Constraints and Demand for Healthcare: Evidence from Danish Welfare Recipients," *arXiv preprint arXiv:2010.14651*, 2020.

Machado, Matilde P and Anna Sanz de Galdeano, "Coverage of Infertility Treatment and Fertility Outcomes," *SERIEs*, 2015, *6* (4), 407–439.

Margolis, Rachel and Mikko Myrskylä, "Parental Well-Being Surrounding First Birth as a Determinant of Further Parity Progression," *Demography*, 2015, *52* (4), 1147–1166.

Martins, Mariana V, Patrício Costa, Brennan D Peterson, Maria E Costa, and Lone Schmidt, "Marital Stability and Repartnering: Infertility-Related Stress Trajectories of Unsuccessful Fertility Treatment," *Fertility and Sterility*, 2014, *102* (6), 1716–1722.

McCrary, Justin, "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test," *Journal of Econometrics*, 2008, *142* (2), 698–714.

Melo, Pedro, Rima Dhillon-Smith, Md Asiful Islam, Adam Devall, and Arri Coomarasamy, "Genetic Causes of Sporadic and Recurrent Miscarriage," *Fertility and Sterility*, 2023.

Miller, Maya, "Senate G.O.P. Blocks I.V.F. Access Bill as Democrats Press for Political Edge," *The New York Times*, June 13, 2024.

Moshfegh, Jasmin, "Who Gains When Medicine Becomes More Precise? Evidence from Genomic Testing In Breast Cancer," 2024. Stanford University mimeo.

Mughal, Saba, Yusra Azhar, and Waquar Siddiqui, "Postpartum Depression," 2018. In: StatPearls. StatPearls Publishing, Treasure Island (FL); 2023. PMID: 30085612.

Nordlund, Christer, *Hormoner för Livet: Endokrinologin, Läkemedelsindustrin och Drömmen om ett Botemedel Mot Sterilitet 1930–1970*, Arkiv förlag, 2008.

Nugent, N. Colleen and Anjani Chandra, "Infertility and Impaired Fecundity in Women and Men in the United States, 2015–2019," 2024. Accessed May 2024. https://www.cdc.gov/nchs/data/nhsr/nhsr202.pdf.

Nyman, J.A., *The Theory of Demand for Health Insurance* Stanford Business Bks, Stanford University Press, 2003.

Nynas, Johnna, Puneet Narang, Murali K Kolikonda, and Steven Lippmann, "Depression and Anxiety Following Early Pregnancy Loss: Recommendations for Primary Care Providers," *The Primary Care Companion for CNS Disorders*, 2015, *17* (1), 26225.

Oster, E., *Expecting Better: Why the Conventional Pregnancy Wisdom Is Wrong–and What You Really Need to Know*, Penguin Publishing Group, 2013.

_ **and N. Fox**, *The Unexpected: Navigating Pregnancy During and After Complications* The ParentData Series, Penguin Publishing Group, 2024.

Persson, Petra and Maya Rossin-Slater, "Family Ruptures, Stress, and the Mental Health of the Next Generation," *American Economic Review*, April 2018, *108* (4-5), 1214–52.

_ **and** _ , "When Dad Can Stay Home: Fathers' Workplace Flexibility and Maternal Health," *American Economic Journal: Applied Economics*, forthcoming.

Peterson, Brennan D, Matthew Pirritano, Jessica M Block, and Lone Schmidt, "Marital Benefit and Coping Strategies in Men and Women Undergoing Unsuccessful Fertility Treatments Over a 5-Year Period," *Fertility and Sterility*, 2011, *95* (5), 1759–1763. e1.

Provost, Meredith P, Samantha M Thomas, Jason S Yeh, William W Hurd, and Jennifer L Eaton, "State Insurance Mandates and Multiple Birth Rates After In Vitro Fertilization," *Obstetrics & Gynecology*, 2016, *128* (6), 1205–1214.

Rellstab, Sara, Pieter Bakx, and Pilar Garcia-Gomez, "The Effect of a Miscarriage on Mental Health, Labour Market, and Family Outcomes," 2022. Tinbergen Institute Discussion Paper 2022-027/V.

Rizzo, Kathryn M, Holly H Schiffrin, and Miriam Liss, "Insight Into the Parenthood Paradox: Mental Health Outcomes of Intensive Mothering," *Journal of Child and Family Studies*, 2013, *22* (5), 614–620.

Rooney, Kristin L and Alice D Domar, "The Relationship Between Stress and Infertility," *Dialogues in Clinical Neuroscience*, 2018, *20* (1), 41–47.

Rädestad, Ingela, Berit Sjögren, Conny Nordin, and Gunnar Steineck, "Stillbirth and Maternal Well-Being," *Acta Obstetricia Et Gynecologica Scandinavica*, 1997, *76* (9), 849–855.

Saez, Emmanuel and Stefanie Stantcheva, "Generalized Social Marginal Welfare Weights for Optimal Tax Theory," *American Economic Review*, January 2016, *106* (1), 24–45.

Schmidt, Lucie, "Infertility Insurance Mandates and Fertility," *American Economic Review Papers and Proceedings*, May 2005, *95* (2), 204–208.

Socialstyrelsen, "Hälsodataregister - Räddar Liv och Ger Bättre Vård," 2019. Accessed August 2021. https://www.socialstyrelsen.se/globalassets/sharepoint-dokument/dokument-webb/ovrigt/halsodataregister-information-om-nyttan-med-register.pdf.

Statistics Sweden, "Longitudinal Integrated Database for Health Insurance and Labour Market Studies (LISA)," 2019. Accessed May 2019. https://www.scb.se/en/services/ordering-data-and-statistics/ordering-microdata/vilka-mikrodata-finns/longitudinella-register/longitudinal-integrated-database-for-health-insurance-and-labour-market-studies-lisa/.

Supreme Court of Alabama, "October Term, 2023-2024, Case SC-2022-0515," 2023. Accessed May 2024. https://publicportal-api.alappeals.gov/courts/68f021c4-6a44-4735-9a76-5360b2e8af13/cms/case/343D203A-B13D-463A-8176-C46E3AE4F695/docketentrydocuments/E3D95592-3CBE-4384-AFA6-063D4595AA1D.

Swedbank, "Så Mycket Kostar det att Ha Barn," 2024. Accessed June 2024. https://www.swedbank.se/privat/livet/familj/ekonomens-tips-till-blivande-foraldrar.html#:~:text=Faktum%20%C3%A4r%20att%20ett%20barn,kronor%20fram%20till%2019%2D%C3%A5rsdagen.

U.S. Department of Agriculture, "Families Projected to Spend an Average of $233 610 Raising a Child Born in 2015," 2024. Accessed June 2024. https://fns-prod.azureedge.us/sites/default/files/2016%20CRC%20Press%20Release.2.pdf.

van den Berg, Gerard J, Petter Lundborg, and Johan Vikström, "The Economics of Grief," *The Economic Journal*, 2017, *127* (604), 1794–1832.

Verhaak, Chris M, Jesper MJ Smeenk, Andrea WM Evers, Jan AM Kremer, Floris W Kraaimaat, and Didi DM Braat, "Women's Emotional Adjustment to IVF: A Systematic Review of 25 Years of Research," *Human Reproduction Update*, 2007, *13* (1), 27–36.

Vikström, Josefin, Ann Josefsson, Mats Hammar, Marie Bladh, and Gunilla Sydsjö, "Risk of Postpartum Psychosis After IVF Treatment: A Nationwide Case-Control Study," *Human Reproduction*, 2017, *32* (1), 139–146.

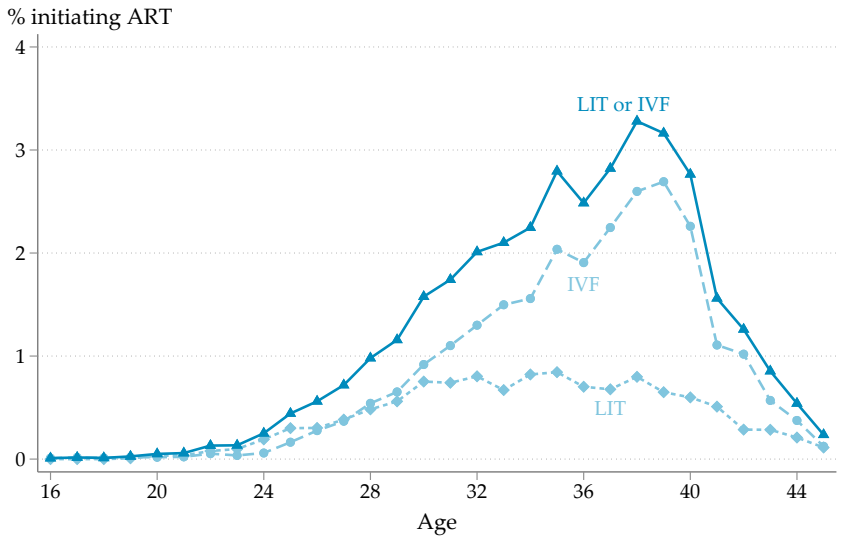
WHO, *Infertility Prevalence Estimates, 1990–2021*, Geneva: World Health Organization, 2023.

Zegers-Hochschild, Fernando, G David Adamson, Jacques De Mouzon, Osamu Ishihara, Ragaa Mansour, Karl Nygren, Elisabeth Sullivan, and Sheryl Van der Poel, "The International Committee for Monitoring Assisted Reproductive Technology (ICMART) and the World Health Organization (WHO) Revised Glossary on ART Terminology, 2009," *Human Reproduction*, 2009, *24* (11), 2683–2687.

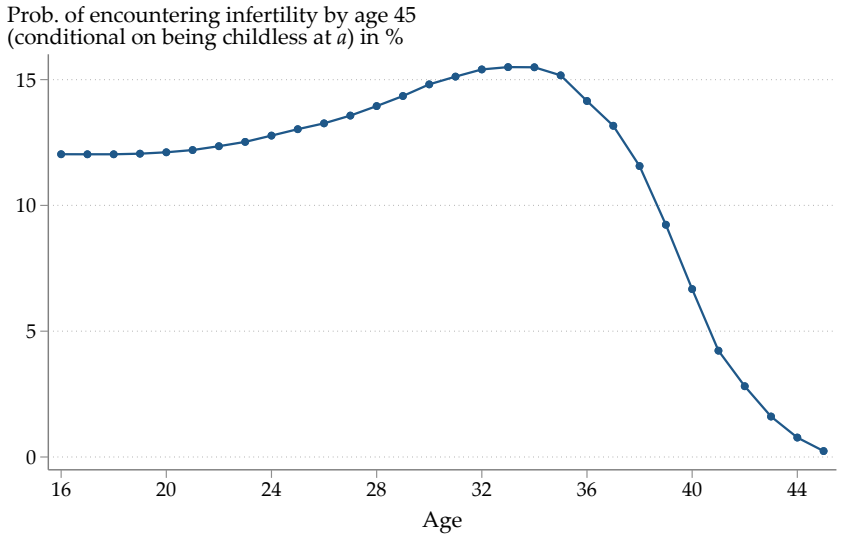
Figures and Tables

Figure 1: **Prevalence of Infertility as Measured by ART Use**

(a) Share of childless women initiating ART in 2018



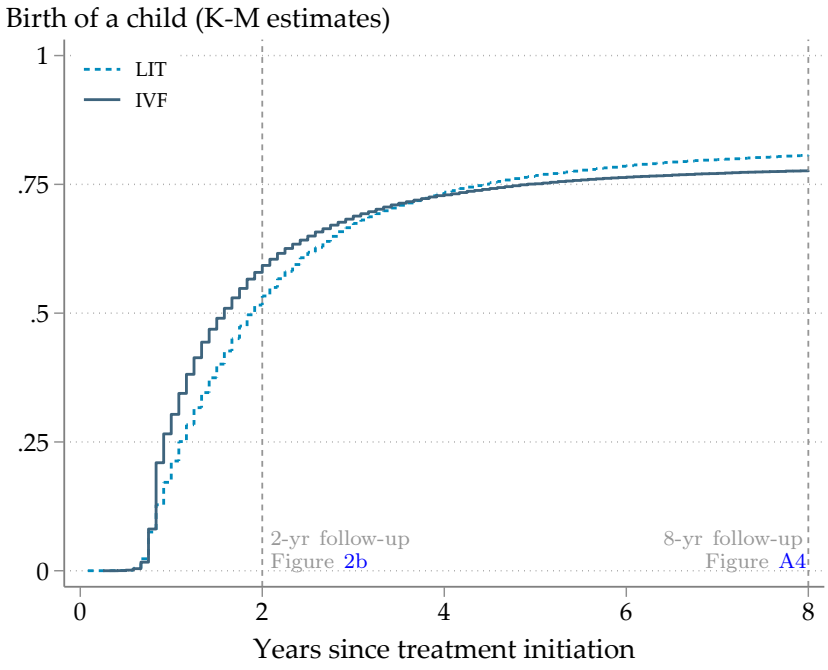
(b) Probability of encountering infertility over fertile lifetime



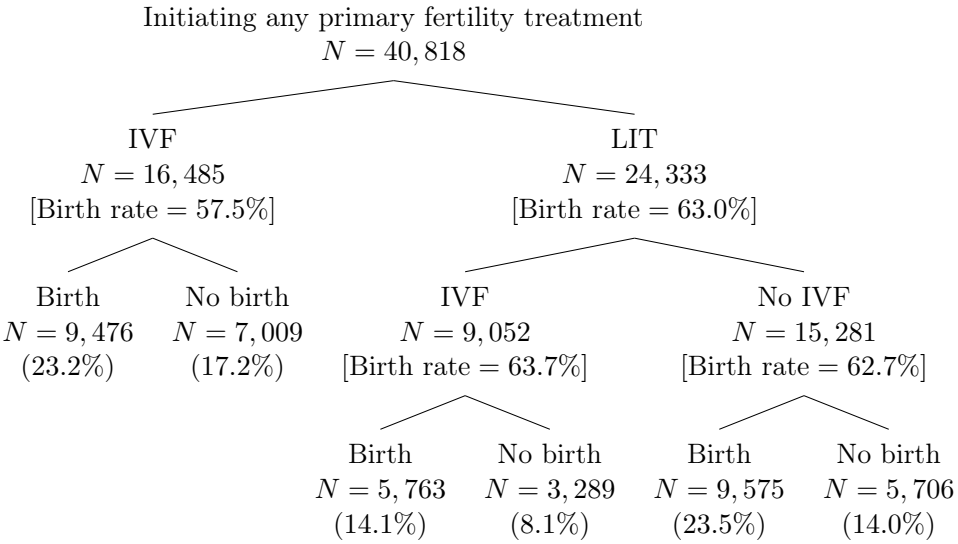
Notes: Both panels use data from 2018, the last full year observed in the data. Panel (a) shows the share of childless 16–45 year old women (as of the end of the preceding year) who initiate ART at each age. Panel (b) aggregates the cross-sectional probabilities in Panel (a) to estimate period lifetime infertility analogous to period life expectancy. For each age a on the x-axis, we calculate and plot $100 * P(\text{Initiating ART between } a \text{ and } 45 | \text{Childless and no history of ART at } a)$. We calculate this probability as the sum of the probabilities of being childless and initiating ART for each age between a and 45 divided by the probability of being childless (and having no history of ART) at a. A more detailed explanation of this calculation can be found in Appendix B. Underlying values for this calculation, namely the probability of giving birth and the probability of being childless and initiating ART at each age a in 2018 cross-section, can be found in Figure A2.

Figure 2: **Infertility Treatment Pathways**

(a) Infertility Treatment and Long-Run Birth Rate



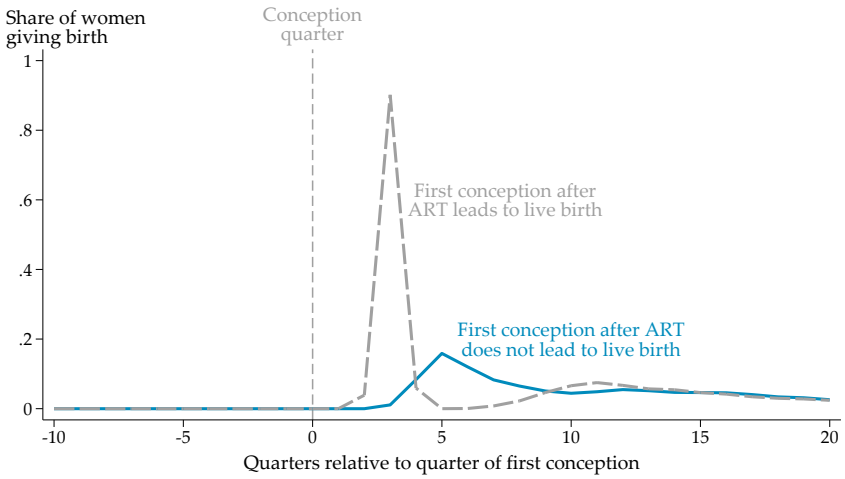
(b) Infertility Treatment Pathways at 2 Year Follow-Up



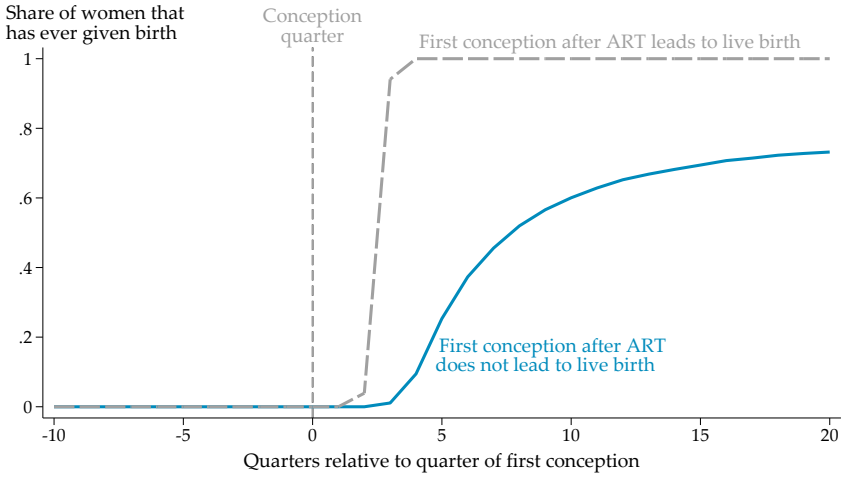
Notes: Panel (a) tracks the progression of fertility outcomes over an eight-year time horizon after the year of treatment initiation, for the universe of childless women in Sweden age 16–45 who initiate an ART treatment between July 2006 and December 2012. The characteristics of this sample of women are reported in Column (3) of Table 1. The figure plots the hazard rate (estimated as the Kaplan-Meier function without covariates) for women to have given birth to a child as a function of time since the initiation of an LIT (dashed line) or IVF (solid line) treatment. Dashed vertical lines mark two-year and eight-year follow up windows, for which Panel (b) and Figure A4 show more details on the treatment pathways and fertility outcomes, respectively. Panel (b) tracks the progression of infertility treatments over a two-year time horizon after the year of treatment initiation, for the same sample as in Panel (a). When a woman switches from LIT to IVF treatment, the two-year follow-up window is reset to allow for a two-year follow-up after IVF initiation. Fertility outcomes are evaluated at two years after the initiation of infertility treatment without requiring a birth outcome to directly follow a treatment cycle. The total birth rate within two years after treatment initiation is 60.8%. Figure A4 characterizes the same treatment pathways and fertility outcomes at eight year follow-up, at which the total birth rate is 78.9%.

Figure 3: **First Conception after ART Initiation and Subsequent Fertility**

(a) Rate of live births



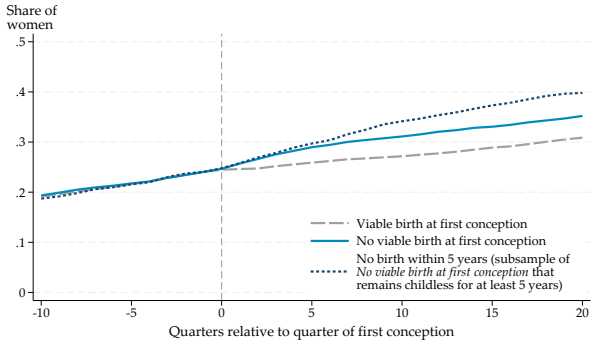
(b) Stock of live births



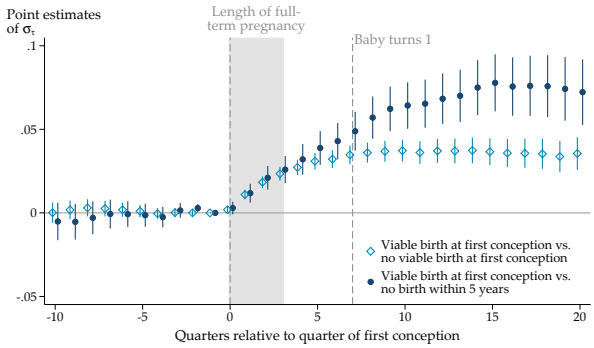
Notes: The figure shows the relationship between the first conception after a woman has initiated any ART treatment and subsequent fertility, measured as either the share of women giving birth in a given quarter relative to the first conception (y-axis in panel a), or as a share of women who have had a live birth by a given quarter (y-axis in panel b). The sample includes childless women who initiate either an LIT or an IVF treatment between July 2006 and February 2016 and who conceive within three years of treatment initiation. The characteristics of this sample of women are reported in Column (4) of Table 1. The sample is then split up by the whether the index conception leads to a live birth (dashed gray line) or not (solid blue line). The x-axis indexes time in quarters relative to the date of the index conception.

Figure 4: **First Conception after ART Initiation and Subsequent Outcomes (Women)**

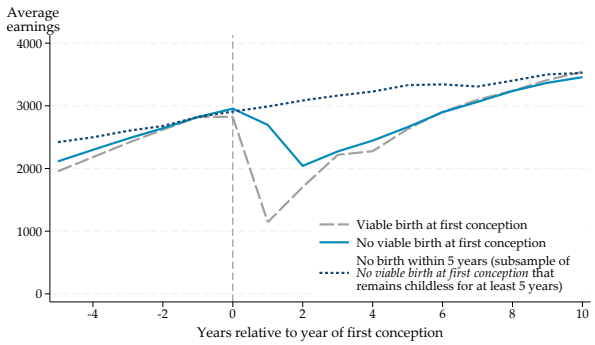
(a) Mental health R_x, raw data



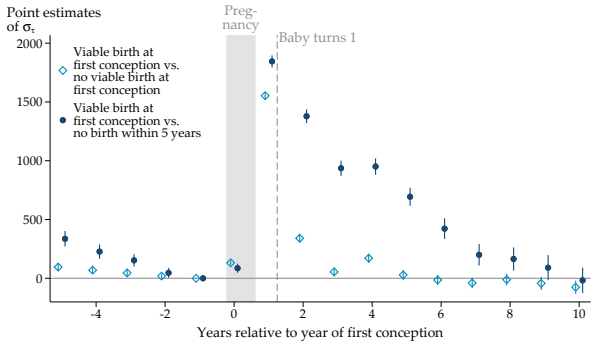
(b) Mental health R_x, event study



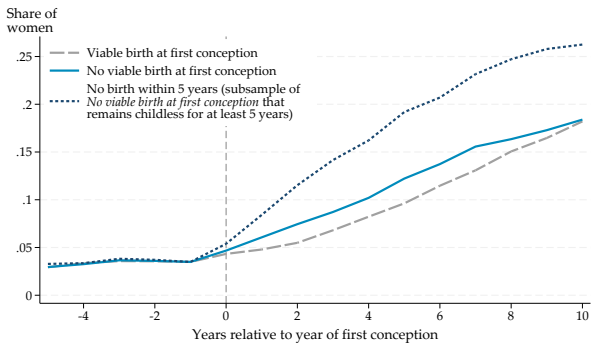
(c) Work income, raw data



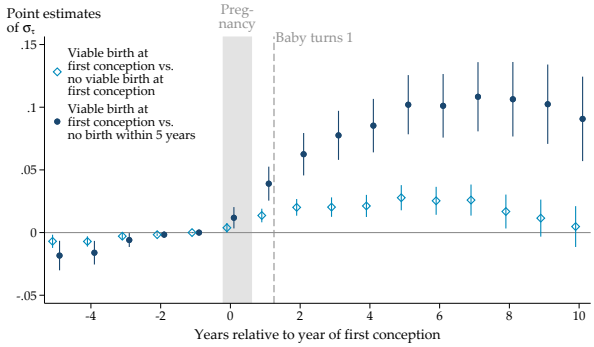
(d) Work income, event study



(e) Divorce, raw data



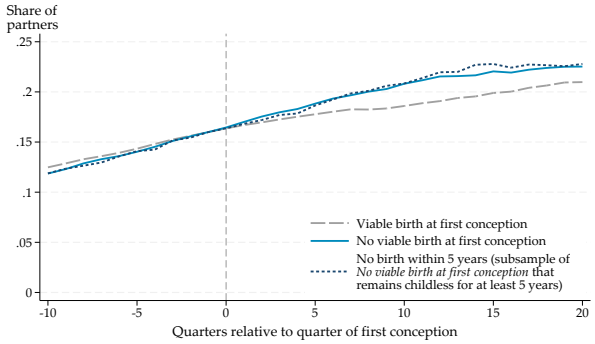
(f) Divorce, event study



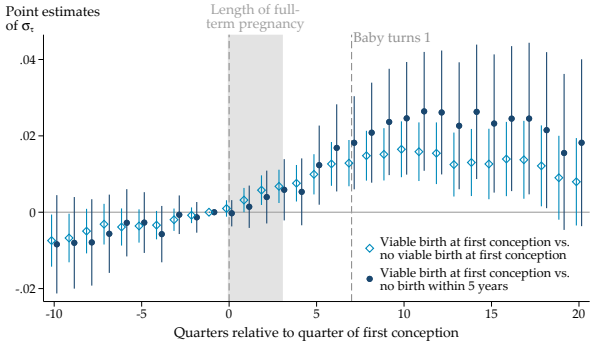
Notes: The figure shows raw data (panels on the left, normalized to the mean in time −1) and event study estimates as specified in Equation (1) (panels on the right) for the effect of having an unsuccessful first conception after ART initiation (which predicts persistent infertility, see Figure 3) on outcomes of women. The sample includes childless women who initiate LIT or IVF treatments between July 2006 and February 2016 and conceive within 3 years of treatment initiation. The characteristics of this sample of women are reported in Column (4) of Table 1. The x-axis indexes time relative to the index conception in either quarters (panels a–b) or years (panels c–f). The outcome in panels (a) and (b) is a stock measure for having a history of drug claims with ATC codes N05 (psycholeptics, commonly used to treat anxiety) or N06 (analeptics, commonly used to treat depression). In panels (c) and (d) the outcome is annual individual work income in 100 SEK. In panels (e) and (f) the outcome is the history of a divorce. Event study panels show the results of two separate event studies. The control group in both are those women whose first conception after ART initiation resulted in a live birth (long dashed gray line in the raw data). We provide estimates for two treatment groups: (i) for all women whose first conception was not successful (solid line in the raw data; light blue estimates) and (ii) for a subset of (i) where no birth happens within 5 years of that conception, i.e., those couples who remain infertile in the longer-run (short-dashed line in the raw data; dark blue estimates). Vertical lines in the event studies denote 95% confidence intervals. Standard errors are clustered at the individual level.

Figure 5: **First Conception after ART Initiation and Subsequent Outcomes (Partners)**

(a) Mental health R_x, raw data

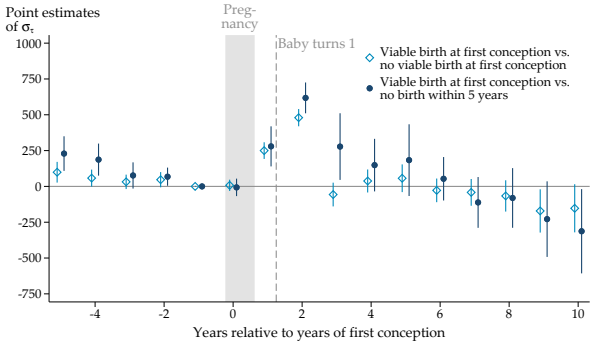


(b) Mental health R_x, event study



(c) Work income, raw data

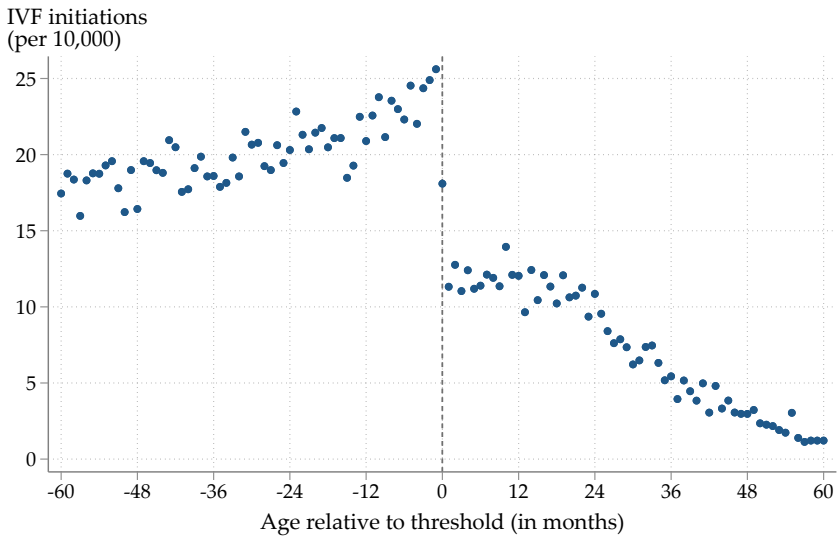
(d) Work income, event study



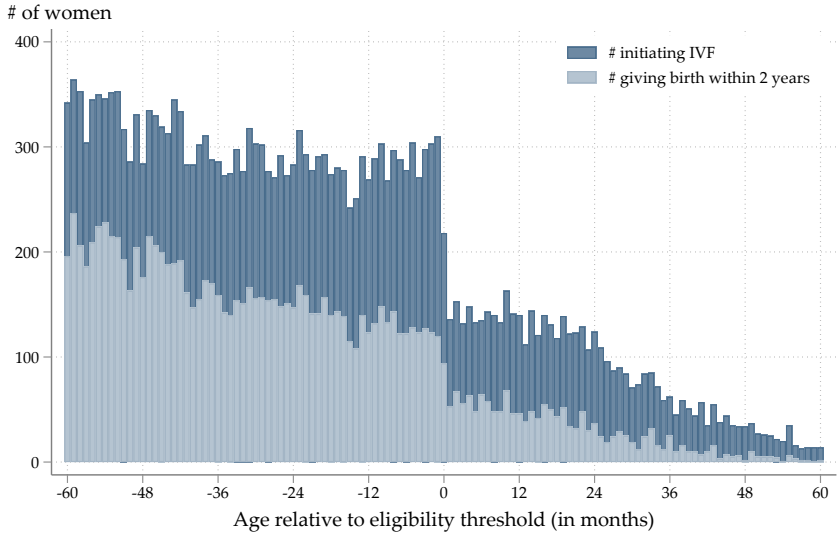
Notes: The figure shows raw data (panels on the left, normalized to the mean in time −1) and event study estimates as specified in Equation (1) (panels on the right) for the effect of having an unsuccessful first conception after ART initiation (which predicts persistent infertility, see Figure 3) on outcomes of partners. The sample includes married or co-habitating partners of childless women who initiate LIT or IVF treatments between July 2006 and February 2016 and conceive within 3 years of treatment initiation. The x-axis indexes time relative to the first conception in either quarters (panels a–b) or years (panels c–d). The outcome in panels (a) and (b) is a stock measure for having a history of drug claims with ATC codes N05 (psycholeptics, commonly used to treat anxiety) or N06 (analeptics, commonly used to treat depression). In panels (c) and (d) the outcome is annual individual work income in 100 SEK. Event study panels show the results of two separate event studies. The control group in both are partners in those couples where the woman's first conception after ART initiation resulted in a live birth (long dashed gray line in the raw data). We provide estimates for two treatment groups: (i) for all partners of women whose first conception was not successful (solid line in the raw data; light blue estimates) and (ii) for a subset of (i) where no birth happens within 5 years of that conception, i.e., those couples who remain infertile in the longer-run (short-dashed line in the raw data; dark blue estimates). Vertical lines in event studies denote 95% confidence intervals. Standard errors are clustered at the individual level.

Figure 6: **Impact of IVF Insurance Coverage**

(a) IVF initiations



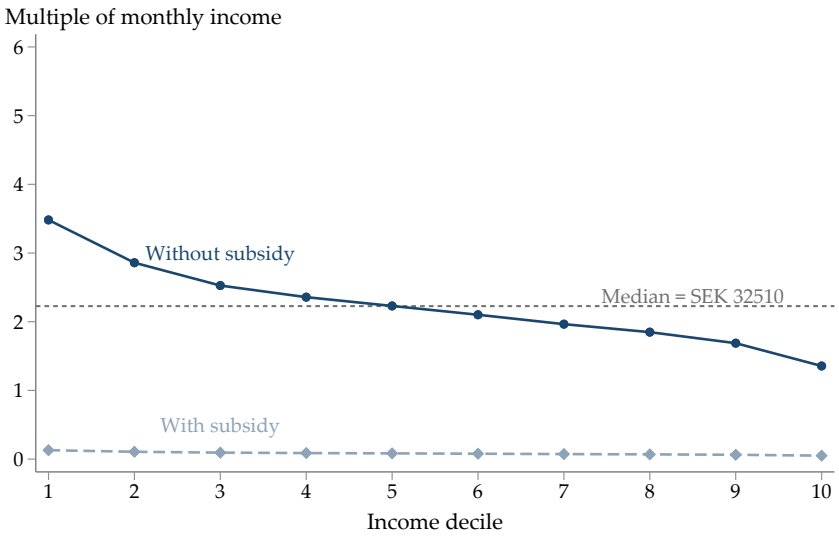
(b) Births after IVF initiation



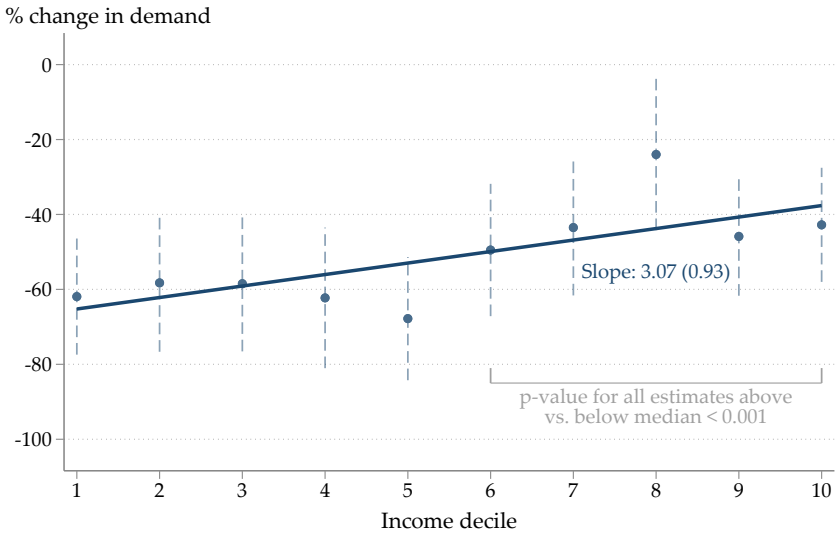
Notes: This figure illustrates the relationship between public health insurance coverage of IVF and the probability of IVF initiation (panel a) as well as the number of births (panel b). The x-axis in both panels indexes a woman's age in months relative to the insurance eligibility cutoff age, which is specific to the woman's region of residence and year. Panel (a) shows the number of women initiating IVF per 10,000 women in the sample of all childless women who are within five years of the insurance age eligibility cutoff, and who haven't initiated an IVF treatment yet by a given month. The characteristics of this sample of women are reported in Column (5) of Table 1; these women are the age-based subset of all age 16–45 childless women. In panel (b), the dark blue bars show the count of women who initiate IVF at every relative month—this is simply the numerator of the y-axis share in panel (a). The gray bars show the number of the same women who give birth within two years of the relative month on the x-axis. Women's dates of birth are recorded at the year-month level, thus in relative month 0 they can be either eligible or not, depending on their exact day of birth.

Figure 7: **Income Gradient in the Impact of IVF Insurance Coverage**

(a) Cost of 2.7 IVF Cycles in Monthly Disposable Incomes



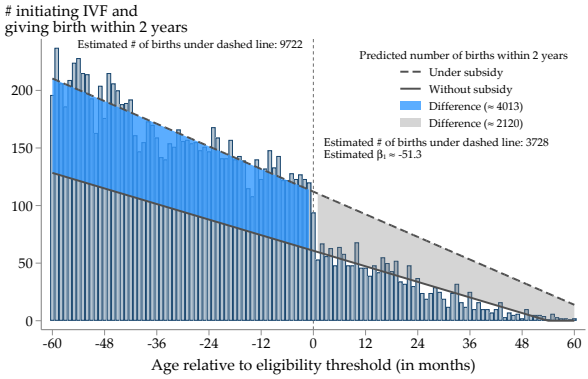
(b) Demand response by income decile



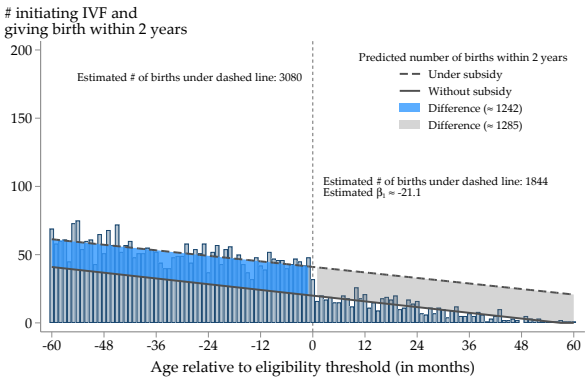
Notes: Panel (a) shows the cost of 2.7 IVF cycles with insurance coverage (dashed line) and without insurance coverage (solid line) in 2010 divided by a couple's monthly disposable income measured when the woman is age 32 (in 2010 SEK). The ratio is plotted for each decile in the woman's monthly disposable income distribution. Section 2 provides more detailed information about the cost of IVF and nature of insurance coverage. 2.7 cycles is our proxy for the number of cycles individuals expect to undergo, based on the number of months an average woman around the coverage cutoff claims IVF medication within two years of treatment initiation (see Figure A7c). Panel (b) shows regression discontinuity estimates (Equation (2)) separately for each decile in the woman's monthly disposable income distribution. For each regression discontinuity point estimate, we calculate the change in demand in % as $100 * \hat{\beta}_1$ divided by the sample mean in the 12 months before the cutoff. Vertical bars denote 95% CIs, calculated using the delta method. The p-value of a statistical test for the difference in demand response below and above median income is reported in gray in the bottom right corner. The solid line is a line of best fit for the relationship between the estimated price response and income decile; the slope and the bootstrapped standard error (in parentheses) are reported close to the regression line. The intercept of the regression line is −68.3.

Figure 8: **IVF Insurance Coverage and (Missing) Births**

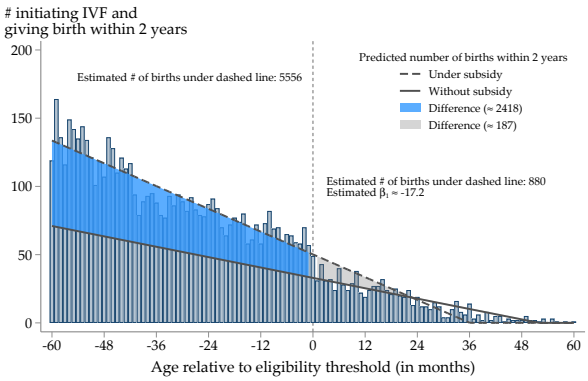
(a) All childless women



(b) Childless women with below-median income



(c) Childless women with above-median income



Notes: Panel (a) of the figure replicates a part of panel (b) in Figure 6. It plots the number of women giving birth within two years of IVF initiation in the sample of all childless women who are within five years of the insurance age eligibility cutoff and who haven't yet initiated an IVF treatment. The characteristics of this sample of women are reported in Column (5) of Table 1; these women are the age-based subset of all age 16–45 childless women. Panel (b) and (c) plot the same figure, but splitting the sample into women with below (panel b) and above (panel c) woman's median disposable income at age 32. The x-axis in each panel indexes a woman's age a in months relative to the insurance eligibility cutoff age A. The bars show the actual number of women who initiate IVF at a relative age, $a - A$, and give birth within two years of that initiation. The dashed lines are predicted numbers of births from a regression of the numbers of births on the left side of the cutoff ($a < A$) on the relative age, $a - A$. The solid lines are predicted numbers of births from a regression of the numbers of births on the right side of the cutoff ($a > A$) on $a - A$. The shaded blue and light gray areas show the difference between the two lines for $a < A$ and $a > A$, respectively. "Estimated β_1" refers to the coefficient on an indicator variable for $a > A$ from an RD-style regression, similar to Equation (2), where the number of births is the outcome.

Table 1: **Summary Statistics**

	Sample					
	(1) At-risk	(2) All initiators	(3) Initiators until 12/2012	(4) Event study	(5) RD sample	(6) LIT RD sample
Year of measurement	2012.6	2011.7	2008.1	2009.8	2005.1	2004.7
	(4.0)	(4.0)	(1.9)	(2.9)	(4.8)	(4.1)
Demographics						
Age	25.4	31.1	31.1	30.6	32.0	32.0
	(7.1)	(5.2)	(5.2)	(4.9)	(0.0)	(0.0)
No college	0.59	0.35	0.38	0.35	0.43	0.31
	(0.49)	(0.48)	(0.48)	(0.48)	(0.49)	(0.46)
Some college	0.15	0.15	0.14	0.14	0.15	0.16
	(0.36)	(0.35)	(0.35)	(0.35)	(0.35)	(0.37)
Full college (or more)	0.22	0.47	0.45	0.49	0.26	0.35
	(0.41)	(0.50)	(0.50)	(0.50)	(0.44)	(0.48)
Married	0.09	0.35	0.37	0.36	0.11	0.12
	(0.29)	(0.48)	(0.48)	(0.48)	(0.31)	(0.32)
Having ever divorced	0.02	0.05	0.05	0.04	0.05	0.06
	(0.15)	(0.22)	(0.22)	(0.20)	(0.21)	(0.24)
Work income (in 100 SEK)	1335.4	2607.7	2430.3	2587.9	1818.6	2214.8
	(1448.7)	(1709.8)	(1624.1)	(1636.6)	(1424.9)	(1420.7)
Disposable income (in 100 SEK)	1369.2	2350.1	2142.9	2276.7	1663.4	1876.3
	(1543.2)	(2183.9)	(1505.3)	(1467.6)	(1299.0)	(993.6)
Born outside of Sweden	0.17	0.25	0.23	0.21	0.28	0.30
	(0.38)	(0.43)	(0.42)	(0.41)	(0.45)	(0.46)
Health characteristics						
LIT (at any point)	0.04	0.68	0.75	0.71	0.06	1.00
	(0.19)	(0.47)	(0.43)	(0.46)	(0.24)	(0.00)
IVF (at any point)	0.04	0.69	0.67	0.69	0.07	0.51
	(0.20)	(0.46)	(0.47)	(0.46)	(0.25)	(0.50)
LIT or IVF (at any point)	0.06	1.00	1.00	1.00	0.09	1.00
	(0.23)	(0.04)	(0.03)	(0.02)	(0.29)	(0.00)
Having had a mental health Rx	0.23	0.26	0.19	0.21	0.17	0.16
	(0.42)	(0.44)	(0.39)	(0.41)	(0.37)	(0.37)
Number of individuals	1,807,328	85,110	40,818	43,165	182,746	6,300
Number of observations	13,370,388	85,110	40,818	43,165	5,911,461	72,733

Notes: The table shows summary statistics for our analytic samples. The "at-risk" sample in Column (1) includes all women in our data (see Section 2 for the description of data sources and Appendix A for the definition of key variables) who are age 16–45 and childless in years 2006–2019. The initiator sample in Column (2) includes all women who are in the at-risk sample (Column (1)) and also initiate ART (IVF or LIT) treatment between July 2006 and November 2019. The sample in Column (3) is a subset of Column (2) restricted to individuals initiating ART treatment by December 2012. The sample in Column (4) includes individuals who initiate LIT or IVF treatment between July 2006 and February 2016 and conceive within 3 years of treatment initiation. The regression discontinuity (RD) sample in Column (5) is identical to the at-risk sample in (1) but restricted to observations within 2 years of the ART insurance eligibility age cutoff. The RD sample in (6) is a version of (5) that selects women who have done LIT in the two years prior to the observation. Characteristics in Columns (2), (3), and (4) are measured in the year before initiation, and in Columns (5) and (6) at age 32. Income is measured in 2010 Swedish Krona.

Table 2: **Event Study Estimates**

	First stage	Women's outcomes				Partner outcomes		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Time period	Having had a birth	Having had a mental health Rx	Antibiotics Rx	Work income (in 100 SEK)	Having divorced	Having had a mental health Rx	Antibiotics Rx	Work income (in 100 SEK)
Q1–Q4	−0.24***	0.01***	−0.00	131.7***	0.00**	0.00***	0.00	7.8
(≈ Pregnancy)	(0.00)	(0.00)	(0.00)	(9.8)	(0.00)	(0.00)	(0.00)	(20.1)
Q5–Q8	−0.70***	0.03***	0.03***	1553.5***	0.01***	0.01***	−0.00	250.4***
(≈ First year of life)	(0.00)	(0.00)	(0.00)	(15.1)	(0.00)	(0.00)	(0.00)	(29.7)
Q9 and later	−0.33***	0.04***	−0.01***	69.1***	0.02***	0.01***	−0.01***	47.8
	(0.00)	(0.00)	(0.00)	(16.1)	(0.00)	(0.00)	(0.00)	(35.3)
Mean of dep. var. at $\tau = -1$ [a]	0.00	0.25	0.09	2772.0	0.04	0.17	0.05	3498.4
No. of observations[b]	1,301,451	1,183,258	1,183,258	612,995	251,944	697,294	697,294	361,602
No. of individuals	43,165	43,165	43,165	43,165	18,001	26,365	26,365	26,357

[a] Among women whose first conception does not result in a viable birth (treatment group).

[b] Observations are at the quarterly level for drug-based outcome and annual level for work income and divorce. For the "having divorced" outcome, we restrict the sample to women who are married in the year before conception ($\tau = -1$).

Notes: This table reports coefficients from estimating a version of the event study specified in Equation (1) with relative time (year or quarter) indicators replaced by indicators for three time periods: Q1–Q4 which capture the pregnancy quarters, Q5–Q8 which capture the first year of life if the conception was successful, and Q9 (two years) and beyond. The sample includes childless women who initiate LIT or IVF treatments between July 2006 and February 2016 (in columns 1–5) and their partners (in columns 6–8). Column (1) shows the first stage effect of an unsuccessful first conception on subsequent fertility. The outcome in Columns (2) and (6) is a stock measure for having a history of drug claims with ATC codes N05 (psycholeptics, commonly used to treat anxiety) or N06 (analeptics, commonly used to treat depression). The outcome in Columns (3) and (7) is the rate of antibiotics claims. The outcome in Columns (4) and (8) is annual individual work income in 100 SEK. Column (5) captures the effect on the history of a divorce. The control group includes women whose first conception after ART initiation resulted in a live birth. The treatment group are all women whose first conception after ART initiation was not successful. Standard errors (in parentheses) are clustered at the individual level. *, **, and *** indicate statistical significance at the 1, 5, and 10% level, respectively.

Table 3: **Regression Discontinuity Estimates**

	Full sample			By income	
				Below median	Above median
	(1)	(2)	(3)	(4)	(5)
$\hat{\beta}_1$	-11.9^{***}	-12.3^{***}	-10.1^{***}	-10.9^{***}	-12.8^{***}
	(0.7)	(0.7)	(1.1)	(0.9)	(1.4)
Mean dep. var					
overall	17.0	17.5	16.7	12.5	23.8
in $\tau = -12$ to $\tau = -1$	23.2	23.9	22.5	17.6	30.9
No. of observations	5,911,461	5,700,662	4,701,653	2,821,191	2,059,064
Controls	N	Y	N	Y	Y
Omit -5 to 5	N	N	Y	N	N

Notes: This table shows the results of estimating a linear parametric regression discontinuity specification in Equation (2), with IVF initiation as the outcome variable. In all regressions, we exclude the month when eligibility changes as we do not observe age at the daily level. Except as specified in Column (3) that reports the "donut" results and excludes 5 months around the threshold, we otherwise include 24 months before and 24 months after the cutoff in all specifications. Column (1) has no controls, while Columns (2), (4), and (5) include the full set of controls (fixed effects for the region of residence, calendar years, place of birth, and education categories (high school only, some college, college degree or more)). In Columns (1)–(3), the sample includes all childless women (of age -24 to 24 as specified on the x-axis of Figure 6) in July 2006 to November 2019 who have not yet initiated IVF. In Columns (4)–(5), we split up the sample by above and below median of women's disposable income rank at age 32. Standard errors reported in parentheses are clustered at the individual level.

Appendix

A Sample Construction and Definitions

Low-intensity treatment (LIT). The following describes the construction of the sample of women who initiate low-intensity treatment (LIT) between July 2006 and November 2019 at ages 16–45 using the Swedish prescription drug claims data. For each month m of drug claims, $m = 2006/07, \ldots, 2019/11$, we carry out the following five steps:

1. We find all women with an LIT drug claim (ATC codes G03GB, L02BA, L02BG, G03GA) in month m.
2. We keep person-year observations for women who are between 16 and 45 years old in the year of month m.
3. We drop all observations for women with a claim for an LIT or an IVF drug (ATC code G03GB, L02BA, L02BG, G03GA, L02AE, H01CA, H01CC) within 1 year (365 days) before the claim in (1).
4. We drop all observations for women with a breast cancer diagnosis (as the main diagnosis, ICD-10 code C50) in inpatient or outpatient records within 2 years (730 days) before the claim in (1). We drop women with chemotherapy or radiation treatment within 6 months (183 days) after the claim in (1)
5. We drop observations for women with a GnRH agonist/antagonist claim (ATC code L02AE, H01CA, H01CC) within 30 days after the claim in (1), including the date of the claim.

For women with more than one record satisfying the algorithm above, we keep the first record and define it as the initiation date.

IVF treatment. The following describes the construction of the sample of women who initiate IVF between July 2006 and November 2019 at ages 16–45. For each month m of drug claims, $m = 2006/07, \ldots, 2019/11$, we carry out the following five steps:

1. We find all women with a GnRH agonist/antagonist claim (ATC code L02AE, H01CA, H01CC) in month m
2. We keep person-year observations for women who are between 16 and 45 years old in the year of month m.
3. We drop all observations for women with a claim for a GnRH agonist/antagonist (ATC code L02AE, H01CA, H01CC) within 1 year (365 days) before the claim in (1).

4. We drop all observations for women with a breast cancer diagnosis (as the main diagnosis, ICD-10 code C50) in inpatient or outpatient records within 2 years (730 days) before the claim in (1). We drop women with chemotherapy or radiation treatment within 6 months (183 days) after the before the claim in (1).
5. We only keep women who used an ovulation drug (ATC code G03GB, L02BA, L02BG, G03GA) within 30 days before/after the claim in (1).

For women with more than one record satisfying the algorithm above, we keep the first record and define it as the initiation date.

Fertility outcomes. To impute conceptions, we combine data from the inpatient and outpatient registers, as well as the birth register. Conceptions can end in a miscarriage, stillbirth, or live birth.

To infer conceptions resulting in a miscarriage, we find all abortion diagnoses in the inpatient and outpatient registers (ICD-10 codes O00–O08). If an individual has not initiated IVF or LIT in the 22 weeks before the diagnosis date, we assume the diagnosis happens exactly 12 weeks after conception. If an individual has initiated IVF or LIT within the 22 weeks before the diagnosis, we assume the day of conception coincides with the day of the last LIT/IVF drug claim before the diagnosis.

To infer conceptions resulting in a live birth or stillbirth, we use birth entries from the birth register. Whenever available, we use the estimated date of delivery based on ultrasound examination during pregnancy and subtract 280 days (40 weeks). If the estimate based on ultrasound is not available, we use the estimated date of delivery based on the record of the first day of the last menstrual period and subtract 280 days. If this estimate is not available either, we use the year-month of delivery and subtract 280 days from the last day of that month. We use the birth register definition of stillbirths that includes births where the newborn dies within 27 days.

Drug-based health outcomes. We use the data on prescription drug purchases from July 2005 to December 2019 to define the drug-based outcomes in the event studies of Section 4. For each individual in the sample, we define two indicators at the calendar quarter level. The "mental health Rx" indicator is equal to one if an individual filled a prescription for a drug with an ATC-code starting with N05 (Psycholeptics) or N06 (Psychoanaleptics) *in or before* a calendar quarter and zero otherwise. I.e., the indicator is equal to one if an individual has filled a mental health prescription (which tend to be chronic) anytime since July 2005. The "antiobiotics Rx" indicator

is equal to one if an individual filled a prescription (which tend to be acute) for a drug with an ATC-code starting with J01 in a calendar quarter and zero otherwise.

Income measures. Throughout the paper, we use four measures of income:

1. Annual individual work-related income, which is defined as the sum of gross wages from employment (*loneink*) as well as positive net income from active self-employment (*fink* and *inkfnetto*).

2. Annual individual disposable income (*dispink*). Disposable income is defined by Statistics Sweden as salary net of social security contributions plus other incomes (such as business income, capital income, and rental incomes) and received taxable and tax-free transfers (e.g., student grants, care allowances, housing benefits, and sickness benefits) minus taxes and other paid transfers (such as student loan payments or pension savings).

3. Annual parental total (AGI-like) income. Total (AGI-like) income is defined as similar in spirit to the US adjusted gross income and includes earned labor market income (*forvers*), income from passive business activities (*pasnar*), study-related transfers (*stud*), care allowances (*vardbidr* and *komvardbidr*) as well as parental leave (*forpeng*), transfers due to unemployment and other labor market policies (*arblos* and *ampol*), capital income, and allowances related to military service (*vplers* and *gmuers*). To calculate parental income percentiles, we exclude negative and zero values and find an individual's father's and mother's income at ages 20 and 21 of the individual. We then pool the income over both parents and years and rank it within birth cohorts (of the child).

B Estimating Prevalence of Infertility

We consider a woman to be experiencing primary infertility if she is childless and uses an assisted reproductive technology (ART)—IVF or LIT—based on the prescription drug measure as defined in Appendix Section A. Using this definition of infertility, we construct two measures of how common infertility is—incidence at a point in time (Figure 1, panel (a)) and prevalence over fertile lifespan (panel (b) of Figure 1, panel(b)):

Incidence of infertility. For all childless women ages 16–45 residing in Sweden in 2018 (last full year of our data), where childlessness is measured in December 2017, we compute the share of

individuals at a given age, who initiate IVF, LIT, or any of the two (i.e., ART). Hence, our incidence of infertility measure is defined as:

$$P(\text{Infertile}|\text{Childless})_a = \frac{\text{N(Childless and ART)}_a}{\text{N(Childless)}_a}$$

for each age $a = 16, 17, \ldots, 45$. For example, taking $a = 30$, $\text{N(Childless and ART)}_{30}$ is the number of women who are 30 in 2018, do not have children as of December 2017, and initiate any ART in 2018. N(Childless)_{30} is the number of the number of women who are 30 in 2018 and do not have children as of December 2017. These resulting probabilities are displayed in panel (a) of Figure 1.

Lifetime prevalence of infertility. As we do not observe most cohorts of women over their full fertile lifespan, we use the cross-sectional estimates of incidence in 2018 to estimate lifetime prevalence. The approach is analogous to the aggregation of mortality rates to period life expectancy. We start with a sample of childless 16 year old woman. All these women are at risk of experiencing infertility at age $a+1$. In each subsequent age, the woman can *exit* the at risk sample by either giving birth (Figure A2, panel (a)) or by experiencing infertility (defined as doing ART). The probability of being at-risk at each age is then the probability of not having "exited" prior to that age:

$$\begin{aligned} P(\text{At risk})_a &= (1 - P(\text{Exit})_{16}) * (1 - P(\text{Exit})_{17}) * (1 - P(\text{Exit})_{18}) * \cdots * (1 - P(\text{Exit})_{a-1}) \\ &= \prod_{\alpha=16}^{a-1} (1 - P(\text{Exit})_\alpha) \end{aligned}$$

We let $P(\text{At-risk})_{16} = 1$. Next, we define the probability of being childless at a, not having experienced infertility until a, and experiencing infertility at a as $P(\text{At risk})_a * P(\text{ART})_a$. Figure A2, panel (b) shows this probability by age. Then, to compute the probability of experiencing infertility over an individuals fertile lifetime (16–45) we can sum up the age-specific probabilities:

$$P(\text{Experiencing infertility between 16 and 45} \mid \text{At risk at 16}) = \sum_{\alpha=16}^{45} (P(\text{At risk})_\alpha * P(\text{ART})_\alpha)$$

Similarly, we can compute the probability of experiencing infertility over an individual's *remaining* fertile lifetime, i.e., between a given age a and age 45, conditional on being childless and having not

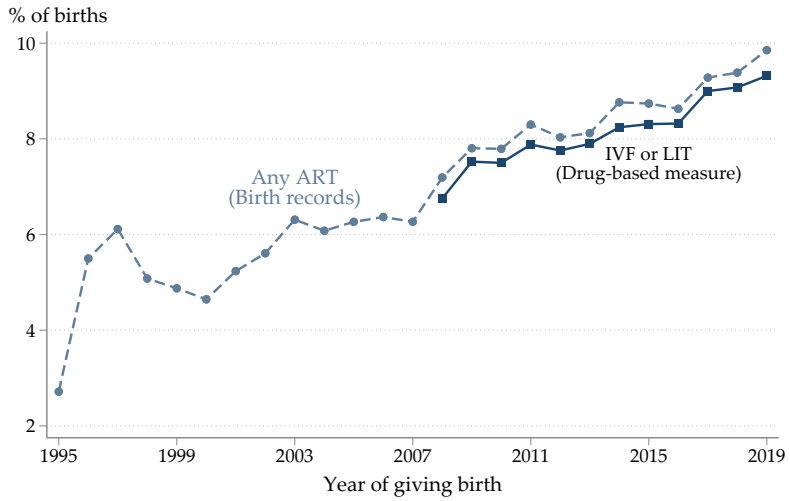
encountered infertility by age a:

$$P(\text{Experiencing infertility between } a \text{ and } 45 \mid \text{At risk at } a) = \frac{\sum_{\alpha=a}^{45}(P(\text{At risk})_\alpha * P(\text{ART})_\alpha)}{P(\text{At risk})_a}$$

These resulting probabilities are displayed in panel (b) of Figure 1.

C Additional Tables and Figures

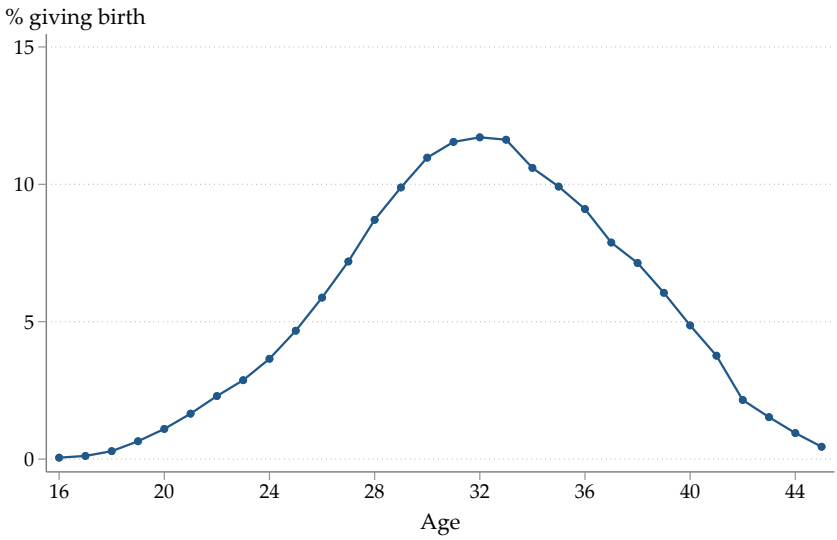
Figure A1: **The Share of First Births with Assisted Conception**



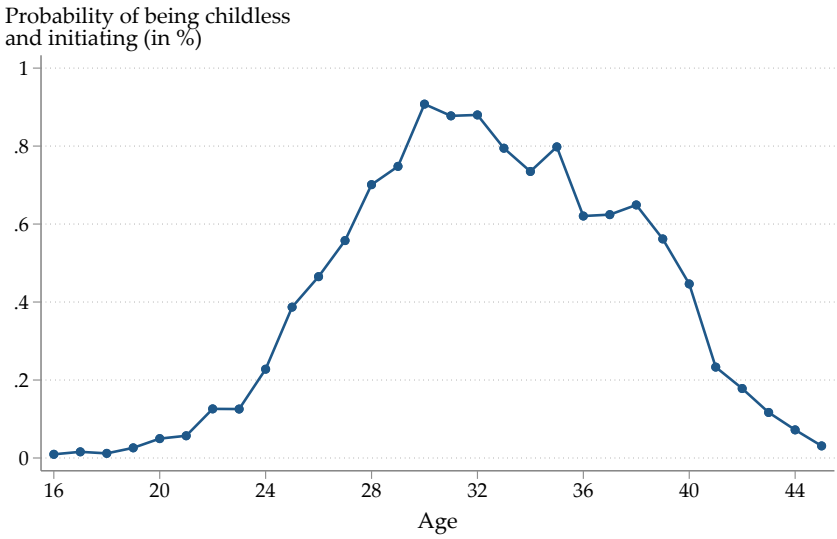
Notes: The figure shows the share of first births (in %) from pregnancies that used Assisted Reproductive Technologies (ART) among women aged 16 through 45 at the time of birth. The dashed line uses data from the Medical Birth Records, which includes flags for a birth being associated with an assisted conception starting in 1995. The solid line uses the prescription drug-based measures of high- and low-intensity infertility treatments (IVF and LIT) developed in this paper. See Appendix A for detailed definitions of LIT and IVF. The drug-based measure of LIT or IVF is available from July 2006 onwards. The solid blue line plots the share of births for which we observe the use of either LIT or IVF in the twelve months prior.

Figure A2: **Infertility Prevalence – Supplemental Figures**

(a) Share of childless women giving birth in 2018

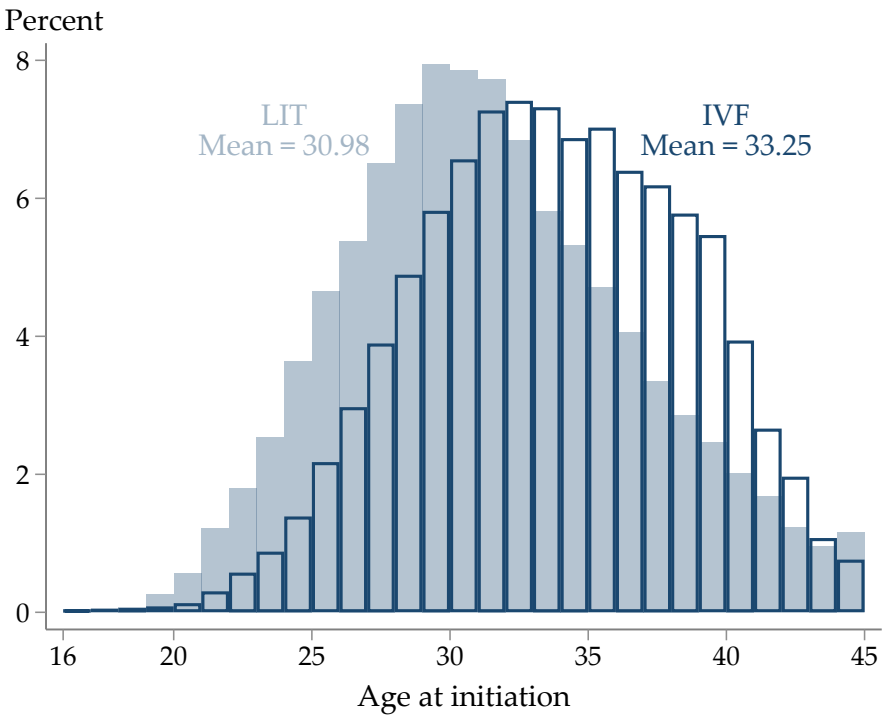


(b) Probability of being childless and initiating ART



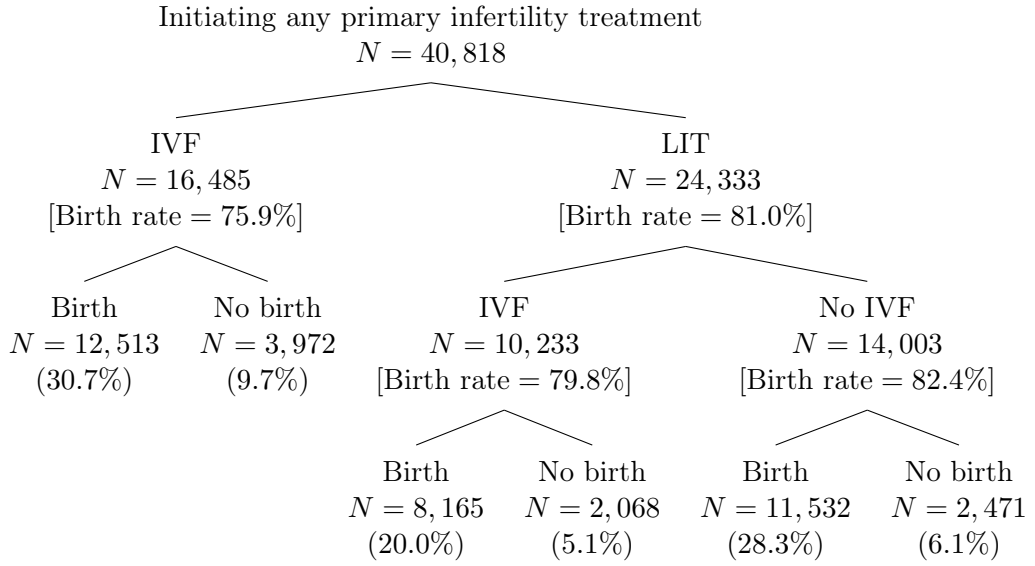
Notes: Panel (a) shows the share of women in year 2018 cross-section who give birth for the first time, separately for each age between 16 and 45. Panel (b) is also based on 2018 data and illustrates our estimates of the probability that a woman is "at-risk" (i.e., is childless and has not initiated ART yet) at age a (x-axis) and initiates any ART at that age a. Section B of the Appendix describes how these estimates enter our calculation of lifetime prevalence of infertility.

Figure A3: **Age at ART Initiation**



Notes: The figure shows the distribution of age at which women initiate either low-intensity (LIT) or IVF treatment as defined from the prescription drug records. See Appendix A for detailed definitions of LIT and IVF. The sample includes the universe of childless women aged 16–45 who initiate low-intensity treatment (LIT, gray filled bars) or IVF (bars with dark blue outline) in our primary analytic "initiators" sample. The sample of these women is characterized in Column (2) of Table 1.

Figure A4: **Infertility Treatment Pathways at 8 Year Follow-Up**

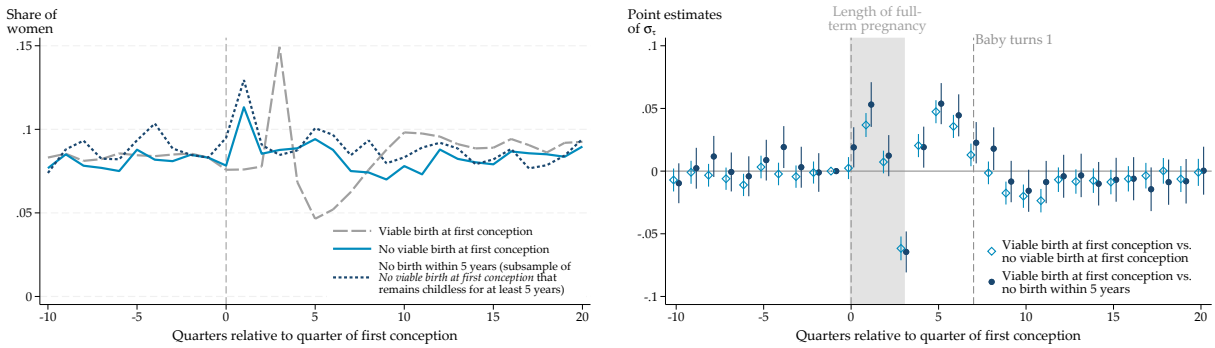


Notes: The figure tracks the progression of infertility treatments over an eight-year time horizon after the year of treatment initiation, for the universe of childless women in Sweden age 16–45 who initiate low-intensity treatment (LIT) or IVF between July 2006 and December 2012. The characteristics of this sample of women are reported in Column (3) of Table 1. Fertility outcomes are evaluated at eight years after the initiation of infertility treatment without requiring a birth outcome to directly follow a treatment cycle. The total birth rate within eight years after treatment initiation is 78.9%. IVF initiation after LIT is evaluated at 6 years and 11 months after LIT initiation, allowing for at least one year follow-up after the IVF treatment.

Figure A5: **First Conception after ART Initiation and Subsequent Antibiotics Use**

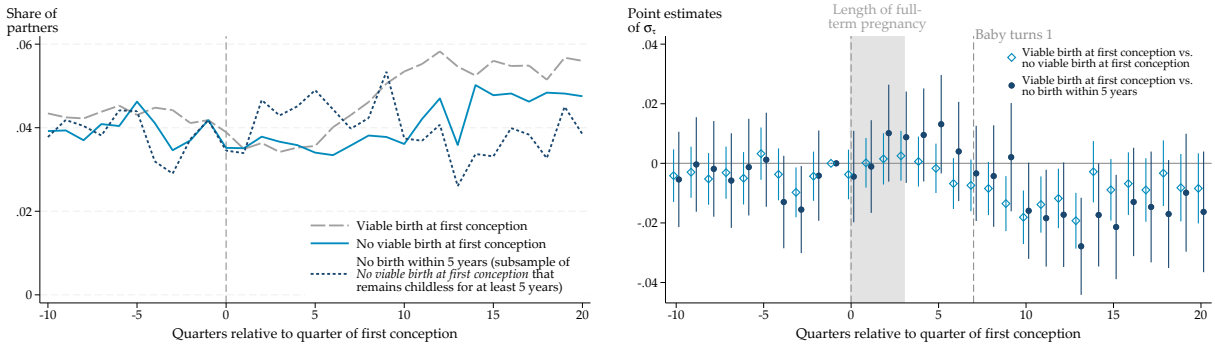
(a) Women, raw data

(b) Women, event study



(c) Partners, raw data

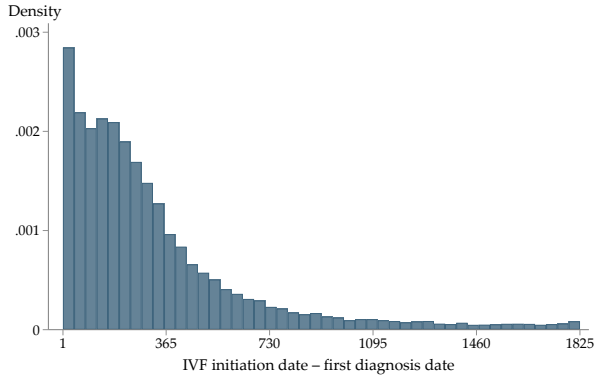
(d) Partners, event study



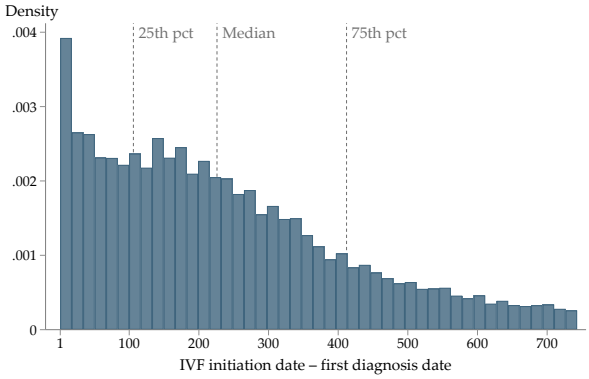
Notes: The figure shows raw data (panels on the left, normalized to the mean in time −1) and event study estimates as specified in Equation (1) (panels on the right) for the effect of having an unsuccessful first conception (which predicts persistent infertility, see Figure 3) on the consumption of antibiotics by women (panels a–b) and their partners (panels c–d). The sample includes childless women who initiate LIT or IVF treatments between July 2006 and February 2016 and conceive within 3 years of treatment initiation. The characteristics of this sample of women are reported in Column (4) of Table 1. The x-axis indexes time relative to the first conception in quarters. The outcome in all panels is an indicator for filling an antibiotics prescription (ATC code: J01). Event study panels show the results of two separate event studies. The control group in both are those women whose first conception after ART initiation resulted in a live birth (long dashed gray line in the raw data). We provide estimates for two treatment groups: (i) for all women whose first conception was not successful (solid line in the raw data; light blue estimates) and (ii) for a subset of (i) where no birth happens within 5 years of that conception, i.e., those couples who remain infertile in the longer-run (short-dashed line in the raw data; dark blue estimates). Vertical lines in the event studies denote 95% confidence intervals. Standard errors are clustered at the individual level.

Figure A6: **Time from an Infertility Diagnosis to IVF Initiation**

(a) Distribution of wait time



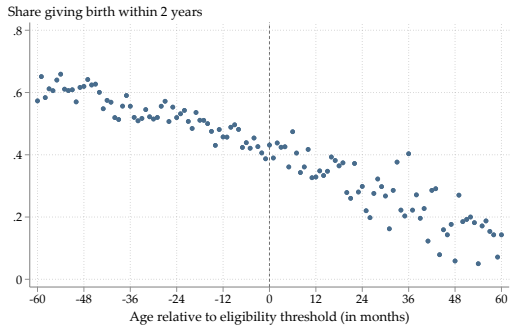
(b) Panel (a) cut at the 90th percentile



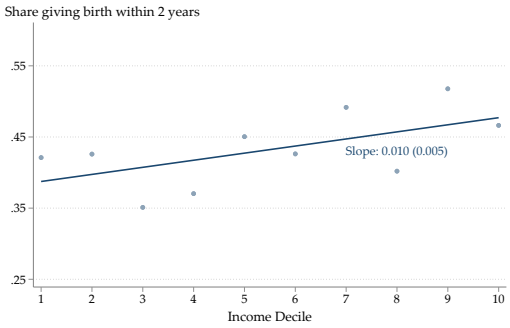
Notes: The figure illustrates the time between an individual's infertility (ICD-10 code N97) diagnosis and initiation of IVF in days. The sample includes a subset of the "initiator" sample (Column (2) of Table 1) that is restricted to childless women age 16–45 who initiate IVF between 2006 and 2019 and who had an infertility diagnosis in the 5 years before IVF initiation. "First diagnosis date" refers to the earliest diagnosis within 5 years before IVF initiation excluding the day of IVF initiation. Panel (b) drops the top 10% of wait times and displays a zoomed in version of panel (a).

Figure A7: **Impact of IVF Insurance Coverage: Supplemental Analyses**

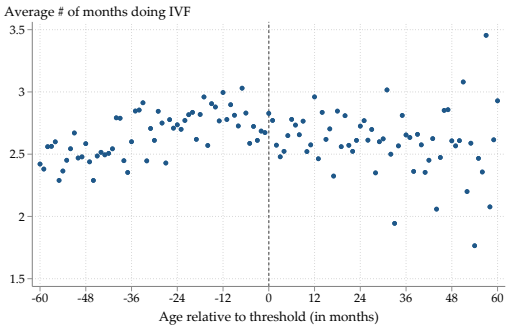
(a) Share of IVF initiators giving birth



(b) Share of IVF initiators giving birth, by income



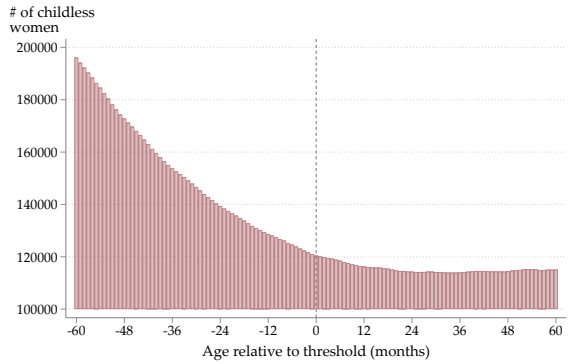
(c) Number of months with IVF treatment



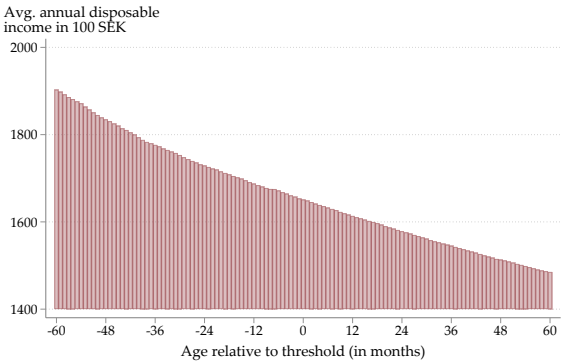
Notes: This figure presents additional analyses on the effects of public health insurance coverage of IVF. The x-axis in all panels indexes a woman's age in months relative to the insurance eligibility cutoff age, which is specific to the woman's region of residence and year. Panel (a) shows the share of women giving birth within two years of IVF initiation in the sample of childless women who are within five years of the insurance age eligibility cutoff and who have initiated an IVF treatment in a given relative month – this corresponds to the ratio of the gray bar and the dark blue bar in Figure 6b. Panel (b) shows the same quantity, i.e. the average share of women giving birth within two years of IVF initiation, for ten subsamples of women, split up by their income decile at age 32. Panel (c) shows the average number of months in which the women who initiate IVF in a given relative month (as indexed on the x-axis) claim IVF-related prescription drugs and remain childless within two years of initiation. Individuals' dates of birth are at the year-month level, thus in month 0 they can be either eligible or not, depending on their exact day of birth. Running an RD-style regression of the share shown in panel (a) on relative age $(a - A)$, an indicator for age being above the cutoff $(a > A)$, and an interaction between the two yields an estimate of 0.02 (p-value 0.25) on the indicator for being on the right side of the cutoff (using a bandwidth from $a = -24$ to 24). Running an analogous regression for panel (c), we get an estimate of -0.10 (p-value 0.24) on the indicator for being on the right side of the cutoff.

Figure A8: **Density and Characteristics of the Regression Discontinuity Sample**

(a) Density



(b) Individual disposable income

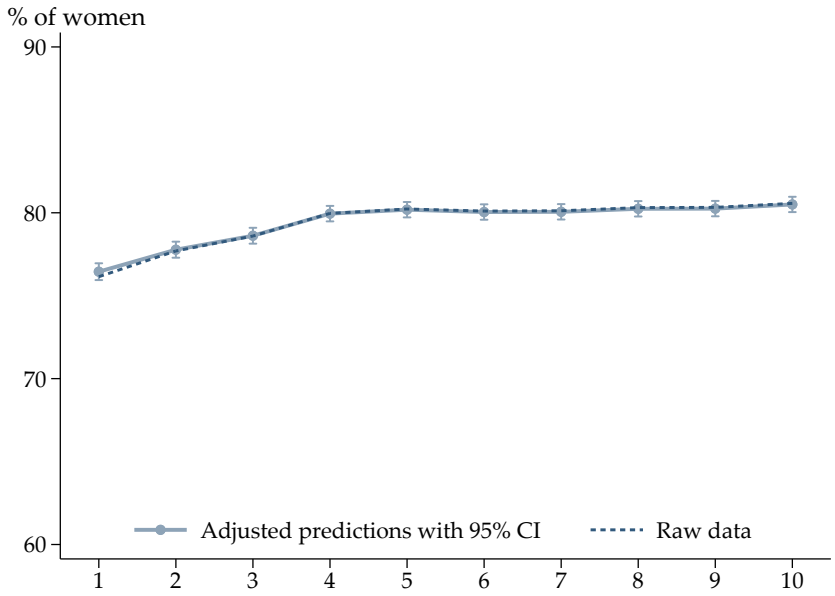


(c) College degree

(d) Foreign-born

Notes: Panel (a) of the figure shows the number of women in the regression discontinuity sample that underlies Figure 6, by a woman's age in months relative to the IVF insurance eligibility cutoff age (pooled over all months, this is the sample from Column (5) of Table 1). These women are the age-based subset of all age 16–45 childless women in 2006–2019 who have not initiated IVF treatment by a given relative month. A formal test that uses the number of individuals as the dependent variable as the outcome in an RD specification restricted to 24 months around the cutoff (as in Table 3), results in the following estimates (and standard errors): 196.08 (231.11) for a linear polynomial and 106.24 (152.69) for a quadratic polynomial. Panels (b)–(d) show different observable characteristics of women in the same sample at each age (in months) relative to the cutoff. All observable characteristics are measured at age 32. Panel (b) shows average annual individual disposable income (in 100 SEK, CPI-adjusted to 2010) measured at age 32. Panel (c) shows the share of women with a full-college degree by age 32. Panel (d) shows the share of women that were born outside of Sweden.

Figure A9: **Share of Women with a Non-ART Child**



Notes: The figure shows the share of women who have given birth without the use of ARTs between 1995 and 2019. We use the Medical Birth Register to flag non-ART births. The sample includes all women born between 1970–1975 who are observed in our data (i.e., these estimates are not restricted to being childless in any year). These women are 44-49 years old by 2019 and thus have mostly completed their fertility. Income deciles are based on total *parental* income measured at women's age 20 and 21, as described in Appendix A. The dashed line shows the raw data and the solid line show adjusted predictions after controlling for region and year of birth fixed effects. Predicted values plotted with 95% CIs are adjusted means of the outcome for each income decile at the mean of all other covariates.

Table A1: **Age Cutoffs for IVF Insurance Coverage**

Region	2008	2009	2010	2011	2012	2013	2014	2015-
Stockholm	39	39	39	39	39	39	39	39
Uppsala	39	39	39	39	39	39	39	39
Sörmland	37	37	37	37	37	37	37	39
Östergötland	37	37	37	37	37	37	37	39
Jönköping	37	37	37	37	37	37	37	39
Kronoberg	38	38	38	38	38	38	38	39
Kalmar	37	37	37	37	37	37	37	39
Gotland	39	39	39	39	39	39	39	39
Blekinge	38	38	38	38	38	38	38	39
Skåne	38	38	38	38	38	38	38	39
Halland	38	38	38	38	38	38	38	39
Västra Götaland	39	39	39	39	39	39	39	39
Värmland	39	39	39	39	39	39	39	39
Örebro	39	39	39	39	39	39	39	39
Västmanland	39	39	39	39	39	39	39	39
Dalarna	39	39	39	39	39	39	39	39
Gävleborg	38	38	38	38	38	38	38	39
Västernorrland	36	36	36	36	36	36	36	39
Jämtland	36	36	36	36	36	36	36	39
Västerbotten	36	36	36	36	36	36	39	39
Norrbotten	36	36	36	36	36	36	36	39

Notes: The table shows the maximum age at which women were eligible for public health insurance coverage of IVF treatments in Sweden from 2008 onwards for each of the 21 Swedish regions. The data is based on government reports, media coverage, and email correspondences with region representatives.

Sources:

Riksförbundet för sexuell upplysning (RFSU). "Sverigebarometern 2011." Accessed March 21, 2024. rfsu.se/globalassets/pdf/sverigebarometer-2011.pdf

Sveriges Kommuner och Landsting (SKL). "Assisterad befruktning Kartläggning och Landstingens Erbjudanded November 2012 – Version 2."

Sveriges Kommuner och Landsting (SKL). "Assisterad befruktning uppföljningsrapport med definitioner, rekommendationer och utvevcklingsområden – Maj 2014 Version 2." Accessed March 21, 2024. skr.se/download/18.1f376ad3177c89481f73e1af/1615204927279/SKL$_$assisterad$_$befruktning$_$uppfoljningsrapport$_$052014.pdf

Sveriges Kommuner och Regioner. "Rekommendation om enhetlighet i regionernas erbjudande av offentlig finansierad assisterad befruktning." Accessed March 21, 2024. skr.se/download/18.32563d7d1784aa279ecac6bb/1618304676440/Rekommendation-assisterad%20befruktning-ver2-nov2020.pdf

Sveriges Television (SVT). "Generösare regler för IVF-behandling på gång." Accessed March 21, 2024. svt.se/nyheter/lokalt/ost/hojda-aldersgranser-for-ivf-behandling-i-sormland-och-ostergotland

Sveriges Television (SVT). "'För gammal för barn' – lille Bore bevisar motsatsen." Accessed March 21, 2024. svt.se/nyheter/lokalt/ost/ansags-for-gammal-for-att-kunna-bli-gravid-lilla-bore-bevisar-motsatsen

Sveriges Television (SVT). "Vill se samma regler för assisterad befruktning." Accessed March 21, 2024. svt.se/nyheter/lokalt/vastmanland/enhetliga-regler-for-assisterad-befruktning

Sveriges Television (SVT). "Åldergränserna för IVF är orättvisa." Accessed March 21, 2024. svt.se/nyheter/lokalt/ost/aldergranserna-for-ivf-ar-orattvisa

Sundsvalls Tidning. "Hårda regler i länet för provrörsbefruktning." Accessed March 21, 2024. st.nu/artikel/harda-regler-i-lanet-for-provrorsbefruktning

Gefle Dagblad. "Åldersgränsen höjs för IVF."

Kvinnoklinikerna i Sydöstra sjukvårdsregionen. "Assisterad befruktning. Gemensamma regler och grundkrav för IVF i Sydöstra sjukvårdsregionen."

Södra regionsvårdsnämnden. "Rekommendation om enlighet i landstingens och regionernas erbjudande av offentligt finansierad assisterad befruktning."

Region Skåne. "Regionala riktlinjer för assisterad befruktning."

Table A2: **Event Study Sample Composition**

Initiating ART between 07/2006 and 02/2016 $N = 56,822$			
Conception within 3 years $N = 43,165$			No conception within 3 years $N = 13,657$
1st conception $\not\Rightarrow$ viable birth $N = 8,710$		1st conception $\Rightarrow$ viable birth $N = 34,455$	
Birth within 5 years $N = 6,443$	No birth within 5 years $N = 2,267$		

Notes: The table describes the composition of the sample used in the event study specifications (Equation (1)). This sample consists of all childless individuals initiating ART between 07/2007 and 02/2016; those who conceive within 3 years of initiation are included in the analysis in Table 2 ($N = 43,165$, see Column (4) of Table 1 for characteristics of this sample of women). Of these women, the first conception within 3 years of ART initiation leads to a viable birth for 34,455 individuals. This is the control group in Figure 4. For 8,710, the first conception does not lead to a viable birth. This is the first treatment group in Figure 4. Among these individuals, 2,267 individuals remain childless for at least five years. This is the second (i.e., the complier) treatment group in Figure 4. The event studies in Figure 4 involve comparing outcomes of the control group to the first treatment group (light blue estimates) as well as the second treatment group (dark blue estimates).

Table A3: **Regression Discontinuity Estimates: LIT Sample**

	(1)	(2)	(3)
$\hat{\beta}_1$	-167.1^{***}	-168.0^{***}	-136.8^{***}
	(26.5)	(26.7)	(43.0)
Mean dep. var			
overall	336.6	339.5	343.2
in $\tau = -12$ to $\tau = -1$	410.0	414.2	423.2
No. of observations	72,733	71,900	58,425
Controls	N	Y	N
Omit -5 to 5	N	N	Y

Notes: This table is similar to Table 3 and shows the results of estimating a linear parametric regression discontinuity specification in Equation (2), with IVF initiation as the outcome variable. Unlike in Table 3, the sample is restricted to all childless women (of age -24 to 24 as specified on the x-axis of Figure 6) in July 2006 to November 2019 who have not yet initiated IVF, and who also have a history of an LIT treatment in two years prior. In all regressions, we exclude the month when eligibility changes as we do not observe age at the daily level. Except as specified in Column (3) that reports the "donut" results and excludes 5 months around the threshold, we otherwise include 24 months before and 24 months after the cutoff. Column (1) has no controls, while Column (2) includes the full set of controls (fixed effects for the region of residence, calendar years, whether the individual is foreign-born, and education categories (high school only, some college, college degree or more)). Standard errors reported in parentheses are clustered at the individual level.

D Alternative Event Study Specifications

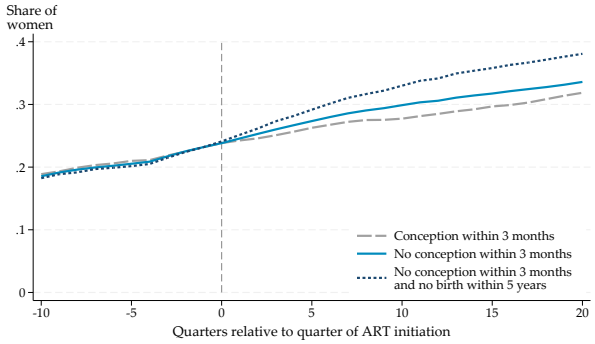
In this section, we estimate a version of our event study specification on the sample of all women who initiate LIT or IVF between July 2006 and February 2016 irrespective of whether they conceive or not. Specifically, we estimate

$$Y_{it} = \alpha_i + \sum_{\tau} \kappa_\tau D_{\tau,it} + \sum_{\tau} \sigma_\tau D_{\tau,it} * \text{No conception}_i + \gamma_t + \beta * \mathbf{X}_{it} + \epsilon_{it} \tag{3}$$

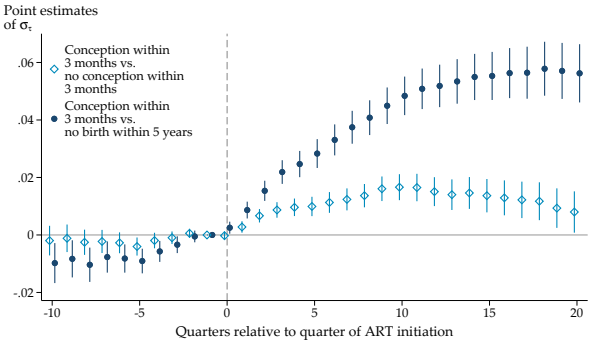
where No conception$_i$ is equal to 1 if individual i does not conceive within *three months* of treatment initiation and 0 otherwise. In Figure A10, we show the results of estimating Equation (3) (light blue estimates). Figure A10 also shows the results of estimating (3) when the treatment group consists of individuals who do no conceive within three months *and* who do not give birth within five years of treatment initiation (dark blue estimates). Figure A11 shows results of these specifications for partners (who are married to or co-habitating with an ART initiator in the year before initiation).

Figure A10: **Event Study Estimates (No Conception vs Conception)**

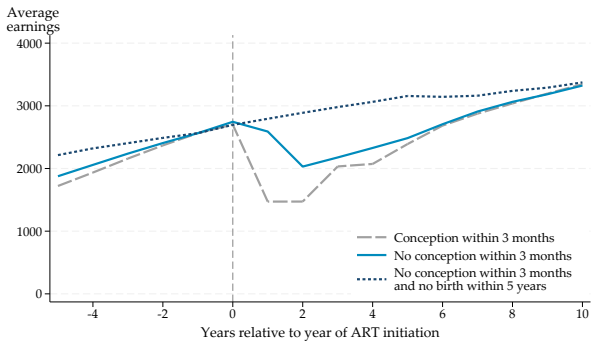
(a) Mental health R_x, raw data



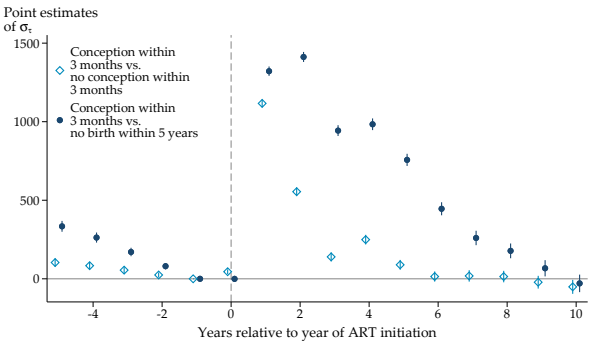
(b) Mental health R_x, event study



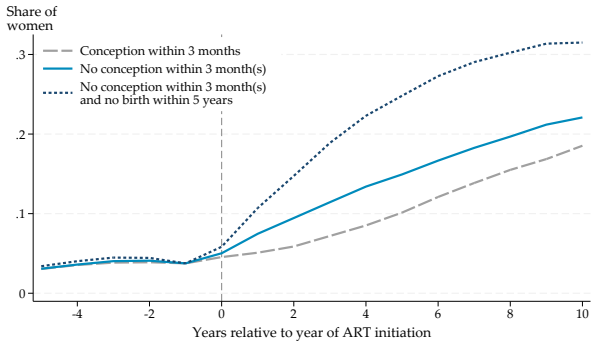
(c) Work income, raw data



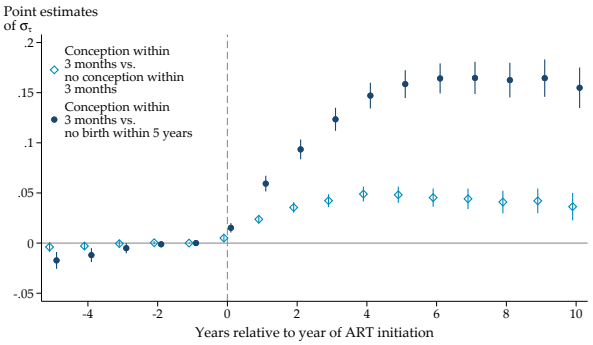
(d) Work income, event study



(e) Divorce, raw data



(f) Divorce, event study

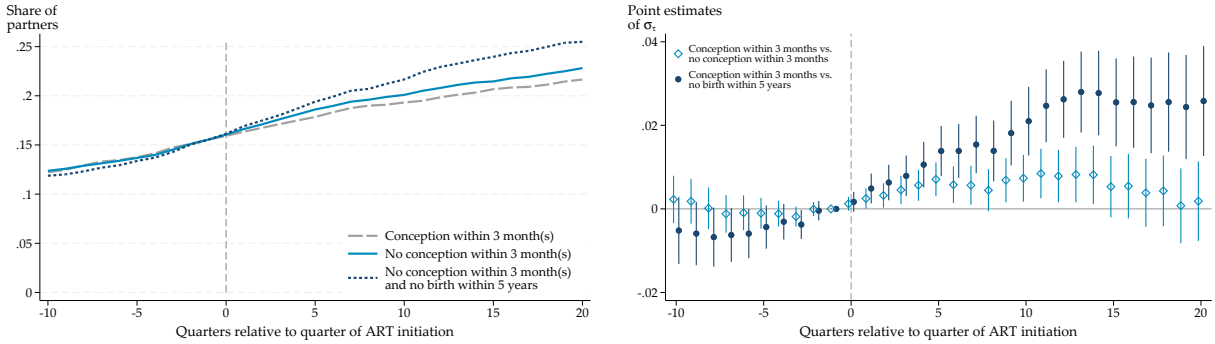


Notes: The figure shows raw data (panels on the left, normalized to the mean in time −1) and event study estimates as specified in Equation (3) (panels on the right) for the effect of having no conception within 3 months of an ART treatment initiation on outcomes of women. The sample includes childless women who initiate LIT or IVF treatments between July 2006 and February 2016. The x-axis indexes time relative to treatment initiation in either quarters (panels a–b) or years (panels c–f). The outcome in panels (a) and (b) is a stock measure for having a history of drug claims with ATC codes N05 (psycholeptics, commonly used to treat anxiety) or N06 (analeptics, commonly used to treat depression). In panels (c) and (d) the outcome is annual individual work income in 100 SEK. In panels (e) and (f) the outcome is the history of a divorce. Event study panels show the results of two separate event studies. The control group in both are those women whose initiation resulted in a conception within 3 months (long dashed gray line in the raw data). We provide estimates for two treatment groups: (i) for all women whose initiation did not lead to conception within 3 months (solid line in the raw data; light blue estimates) and (ii) for a subset of (i) where no birth happens within 5 years of that initiation, i.e., those couples who remain infertile in the longer-run (short-dashed line in the raw data; dark blue estimates). Vertical lines in the event studies denote 95% confidence intervals. Standard errors are clustered at the individual level.

Figure A11: **Event Study Estimates (No Conception vs Conception, Partners)**

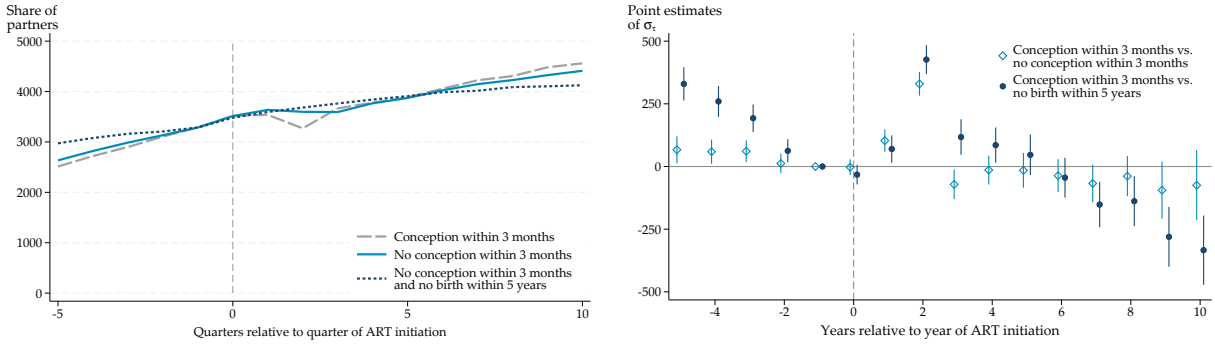
(a) Mental health R_x, raw data

(b) Mental health R_x, event study



(c) Work income, raw data

(d) Work income, event study



Notes: The figure shows raw data (panels on the left, normalized to the mean in time −1) and event study estimates as specified in Equation (3) (panels on the right) for the effect of having no conception within 3 months of an ART treatment initiation on outcomes of partners. The sample includes married or co-habitating partners of women who initiate LIT or IVF treatments between July 2006 and February 2016. The x-axis indexes time relative to treatment initiation in either quarters (panels a–b) or years (panels c–d). The outcome in panels (a) and (b) is a stock measure for having a history of drug claims with ATC codes N05 (psycholeptics, commonly used to treat anxiety) or N06 (analeptics, commonly used to treat depression). In panels (c) and (d) the outcome is annual individual work income in 100 SEK. Event study panels show the results of two separate event studies. The control group in both are those partners whose female partners' initiation resulted in a conception within 3 months (long dashed gray line in the raw data). We provide estimates for two treatment groups: (i) for all partners of women whose initiation did not lead to conception within 3 months (solid line in the raw data; light blue estimates) and (ii) for a subset of (i) where no birth happens within 5 years of that initiation, i.e., those couples who remain infertile in the longer-run (short-dashed line in the raw data; dark blue estimates). Vertical lines in the event studies denote 95% confidence intervals. Standard errors are clustered at the individual level.