The Impact of Extended Reproductive Time Horizons: Evidence from Israel's Expansion of Access to IVF

Naomi Gershoni* and Corinne Low†

March 21, 2018

Abstract

Women's time-limited fertility window, compared to men's longer period of fecundity, could be a key constraint shaping the gender gap in career choices and hence outcomes. Israel's 1994 policy change to make *in vitro* fertilization free provides a natural experiment for how fertility time horizons impact women's investment choices. We find that following the policy change women marry later, complete more college education, achieve more post-college education, and have better labor market outcomes. Additionally, we find the "penalty" in spousal quality for women who get married after thirty substantially dissipates. This suggests that persistent labor market inequality may be partly rooted in biological asymmetries, and policies that protect against age-related infertility or relieve the career-family tradeoff could have far-reaching impacts.

JEL Codes: D10, J12, J13, J16, J24.

*Ben-Gurion University, naomige@bgu.ac.il.

†University of Pennsylvania, The Wharton School, corlow@wharton.upenn.edu. We are grateful for the advice of Analia Schlosser, Pierre-André Chiappori, Cristian Pop-Eleches, and Yoram Weiss. We thank Tal Gross, David Weiss, Moshe Hazan, Alma Cohen, Neale Mahoney, Maya Rossin-Slater, and the participants at the NBER Summer Institute for helpful comments, and Jennie Huang and Ting Wang for excellent research assistance. We thank the staff of Tel Aviv University and the Israeli Central Bureau of Statistics for facilitating data access. The authors gratefully acknowledge research funding from the Wharton Dean's Research fund, the Wharton Pension Research Council/Boettner Center for Pensions and Retirement Research, and the Population Studies Center of the University of Pennsylvania. Opinions and views are solely the responsibility of the authors.

1 Introduction

The persistence of the gender wage gap has been the subject of much exploration in the economics literature. It is now clear that a significant part of this gap is a direct consequence of women's choices, including over college major, career track, working hours, and workplace amenities. However, it is less well established whether these divergent choices result from inherent differences in women's preferences or abilities,[1] or if they are the product of differing constraints. In this paper, we test the hypothesis that women's time-limited fertility window is a key constraint shaping women's career investment choices, using a large expansion of access to *in vitro* fertilization (IVF).

Women's fertility begins to sharply decline in their mid-thirties, whereas men can successfully reproduce for many years after.[2] Anticipating this decline, women may cut short career investments in order to search for a partner, marry, and have their desired number of children before the fertility window closes. To test whether this reproductive time horizon is materially important in women's planning and decision-making requires variation in women's beliefs about later-life fertility. We know that control over fertility is important to women's ability to invest in their careers from literature showing that the introduction of modern contraception increased women's education and career outcomes (Goldin and Katz, 2002; Bailey, 2006; Bailey, Hershbein and Miller, 2012). However, unlike the pill, which was used concurrently by the women making career investments, if IVF were to affect such investments, it would be through the *anticipation* of a longer time horizon for fertility while making *early* life educational and career decisions. Thus, testing its impacts requires a large enough change in access and knowledge of the technology to influence women who may only possibly need it in another 10 to 15 years. In light of this, Israel's unprecedented decision to make IVF completely free to all citizens in 1994 provides an ideal natural experiment.

The 1994 Israeli IVF policy change offers a unique advantage over previously studied IVF policies, by providing an opportunity to study the impact of an exogenous, generalized shift in beliefs about fertility horizons (rather than local changes from state-level insurance mandates as in the United States). Key features of the Israeli policy change that allow us to identify its impact are that it came at a time when IVF was relatively new and unknown, it was widely covered in the Israeli press, and it can be linked to a discontinuous increase in usage. In the three years following the policy change, live deliveries using IVF more than tripled; by 2002, 8 years after the policy change, 1,657 IVF treatment cycles per million people were performed in Israel, compared to 126

[1] E.g., taste for competitiveness (Niederle and Vesterlund, 2007) or even distribution of intelligence (Wai et al., 2010).

[2] Women lose 97% of eggs by 40 (Kelsey and Wallace, 2010), while remaining egg quality declines (Toner, 2003). Pregnancies are rarer (Menken, Trussell and Larsen, 1986), more likely to end in miscarriage (Andersen et al., 2000), and more likely to result in fetal abnormalities (Hook, Cross and Schreinemachers, 1983) later in life, before the complete cessation of fecundity in menopause. While it is difficult to separate fecundity from fertility choices, even prospective studies of women trying to conceive show an accelerating decline in fecundity by age 40 for women, whereas men's fertility is relatively stable. For example, Rothman et al. (2013) finds that women 35-40 years old will become pregnant 77% as frequently as women age 20-24, whereas for men this ratio is 95%.

in the United States (Collins, 2002). Thus, even without direct knowledge of the policy change, the large amount of media attention to older women having children, and first-hand experience of observing motherhood at older ages, could have facilitated a general shift in beliefs regarding the time horizon of fertility.

The challenge inherent in our setting is that all Israelis were exposed to the policy change at the same time, and thus there is no spatial or other statutory variation induced by the policy. However, not everyone within Israel would have received the same update about their *own* fertility time horizons. Naturally, men's later life fertility and hence planning horizons were not directly affected by IVF. In addition, Arab-Israeli women, who marry and start families at a very young age and who had limited access to IVF due to religious restrictions, are expected to be much less affected. These groups thus allow us to separate the specific effect of IVF on Jewish women's time path from general trends for others within Israel. In addition, we then present multiple robustness checks designed to account for potential confounding factors and macroeconomic trends. Despite its limitations, this approach to identification is called for by our research question, because some of the most important impacts of extending reproductive time horizons may only occur when a policy affects widespread perceptions. In other words, a policy that allowed for comparison between areas or small groups would necessarily also have a limited impact on beliefs, and thus only be able to address the impact of actual access to fertility extending technology, but not of a changed planning horizon.

We hypothesize that the expectation of the option to use IVF in the future changed the perceived cost of career investments to younger women, causing them to delay marriage and pursue greater educational investments. In addition, we expect an improvement in older women's marriage outcomes, via a change in potential partners' estimation of women's fecundity horizons. Importantly, our theory of impact does not rely on the affected women actually using the technology themselves. Rather, the future availability provides a form of insurance against age-related infertility. In this way, our work is related to literature on how life expectancy impacts financial planning (Skinner, 1985; Jayachandran and Lleras-Muney, 2009) and how health assets affect other decisions and investments (Delavande and Kohler, 2015; Hugonnier, Pelgrin and St-Amour, 2012). A longer time period before the loss of fertility can be seen as slowing the depreciation of "reproductive capital" (Low, 2017). We thus expect to see later age at first marriage, increased rates of education, improvements in labor market outcomes, and improved marriage outcomes for those who marry late.

To identify the impact on these outcomes, we use a difference-in-differences (DID) approach with the 2008 Israeli population census and administrative tax data. The "treated" group is Jewish-Israeli women born in Israel, who would have experienced a sharp change in fecundity horizons and perceptions due to the 1994 policy change. A valid control group for the purpose of our analysis, has to be "un-treated" by the IVF policy change while otherwise responding to similar forces that could potentially confound our analysis. Thus, for the analysis of marriage timing and educational investment, we first compare Jewish-Israeli women to Jewish-Israeli men, who face similar military requirements, a similar labor market, and the same marriage market, but whose fertility expectations would not be impacted by the policy change.[3] Supporting the use of men as a comparison group, we show that men and women in Israel exhibited remarkably parallel trends in age at first marriage prior to the policy change, with this outcome moving in lockstep for men and women in the "pre" period. Therefore, although men's decisions may be impacted in equilibrium by women's decisions, thus limiting our ability to interpret the magnitudes of our effect at face value, a differential increase in women's marriage age or education precisely in 1994 would still provide evidence of a causal impact of fertility horizons on women's career investments.

In addition to shocks or trends affecting Israelis as a whole, there might be some potential confounding factors that impact particularly *women*. For this reason, we use as a second control group Arab-Israeli women, who were much less likely to use IVF at the time of the policy change due to religious restrictions, as well as less likely to be on the margin for large career investments, but *were* impacted by policies and shocks affecting specifically female Israelis. For example, we know that over the last decades there has been a broad trend, both globally and within Israel, towards increased female education. This general trend, as well as the substantial expansion of higher education in Israel, impacted Arab-Israeli women just as much, if not more than, Jewish-Israeli women.

For marriage market outcomes, we can compare within Jewish women, differentiating based on age at marriage. We compare spouses of women who got married over thirty to those who married when they were younger, thus estimating the change in the "penalty" for older marriage in terms of spousal quality. The treatment and comparison groups in this analysis should be impacted by the same changes to marriage patterns over time, but younger brides do not experience the same change in their fecundity following the change in IVF policy. Here also, the magnitude of our estimates must be interpreted with caution, since an improvement in the marriage market for older women would impact younger women in equilibrium. This concern is accentuated by our findings that the policy affected marriage age and hence selection into marriage "over 30". However, results remain stable when we control for observable characteristics of the women.

We find striking evidence of the policy's impact, starting with a discontinuous jump in women's age at first marriage. While women delayed marriage, they did not forego it entirely, but rather shifted from marrying in their late twenties to their early thirties.[4] College and graduate education completion increased substantially for the cohorts entering these educational levels at the time of

[3] Although the policy also impacted treatment of male infertility via access to sperm donation, this is "primary" rather than "secondary," meaning age-related, infertility, and thus does not affect time horizons.

[4] In contrast to other developed countries, in Israel marriage rates have not been decreasing. According to the Israeli Central Bureau of Statistics report, in 2017, only 5% of cohabiting couples are un-married. A significant share of these couples are still in their 20s, and are expected to eventually marry.

the change. The impact on graduate education are larger than those for college, supporting our hypothesis that the fertility constraint played an active role in these changes. These investments paid off in terms of career outcomes, with this more educated cohort going on to have more full-time employment, higher income, and greater participation in prestigious (high-investment) occupations. Moreover, we find this reflected in the marriage market equilibrium, where the stark penalty associated with marrying older dissipated almost entirely following the policy change. This in turn could have further encouraged investment.

To rule out potential alternative explanations for our findings, we employ a number of robustness checks and placebo tests. Two particular concerns are the substantial expansion of access to higher education in Israel over the 90s and the global trend toward greater education for women. We first use a Quandt-Likelhood Ratio breakpoint test and event-study analysis to pin down the precise timing of the change, showing that any alternative event would need to have an impact precisely at 1994. Then, we use international census data to show that the change in education rates for women in other, similar countries, including the United States, have been smooth over time, unlike the trend break Israel exhibits. We next directly address educational expansion in Israel by controlling for its rate, using data on the precise years of new academic colleges approval, also allowing this expansion to have a gender-differential impact. This is in addition to the use of the Arab control group, which should control for any factors that affected women separately from men. We also show there was no similar gender-differential increase in high school completion or the taking of matriculation exams, "Bagrut," which is a decision less likely to be affected by fertility planning. Finally, we use the 1995 Israeli Census to perform a placebo test, demonstrating that the structural breaks we show in the data are unprecedented historically, and use a permutation analysis to demonstrate that our effects are uniquely empirically large.

Together, our findings indicate that mitigating women's concerns for age-related infertility alters women's educational and marriage decisions, leading to better career outcomes. This bolsters the theory that fertility time horizons are an important factor in women's family and career decisions. More broadly, this research stresses the role of biological differences in divergence of economic outcomes, and thus the potential role of policies in blunting this effect. Such policies would not necessarily need to focus on fertility extending technology alone, but could rather aim to relieve the stark career-childbearing tradeoff, thus making the fertility time horizon less salient to investments.

The remainder of the paper proceeds as follows: Section 2 discusses prior literature; Section 3 describes the empirical setting for our project and the data we use; Section 4 presents results and tests their robustness, and Section 5 concludes.

2 Related Literature

Numerous papers document and aim to explain the evolution of the gender pay gap and its persistence over the last two decades in the US and in other developed economies.[5] A significant part of the lack of convergence between genders is attributed to human capital factors, including choices of college majors (Zafar, 2013; Bronson, 2015), entry into different professions and industries (Bertrand, Goldin and Katz, 2010; Goldin, 2014; Buser, Niederle and Oosterbeek, 2014), and cumulative labor force experience, driven by part time work and time out of the labor force (O'Neill and Polachek, 1993).

Previous literature establishes that control over fertility plays an important role in women's career choices. Generally, this control can be divided into two types: for young women, control over fertility means being able to avoid unwanted pregnancies; however for older women, in their late thirties and forties, control over fertility means actually being able to conceive and give birth *when* they want to. While there is a vast body of empirical evidence on how "too much" fertility affects women's educational, career and marital prospects, the impact of "too little" fertility has not been sufficiently explored.

Goldin and Katz (2002) (and later Bailey (2006); Bailey, Hershbein and Miller (2012)) use the expansion of access to oral contraception to demonstrate that the ability to delay motherhood enabled women to make greater educational and labor market investments. Numerous additional studies support these findings, and use various methods to establish and quantify the tradeoff between family and career for women (Loughran and Zissimopoulos, 2009; Buckles, 2008; Blackburn, Bloom and Neumark, 1993; Taniguchi, 1999; Gustafsson, 2003; Miller, 2011; Avellar and Smock, 2003; Wilde, Batchelder and Ellwood, 2010). Moreover, recent work directly connects raising children to substantial wage declines for women (Adda, Dustmann and Stevens, 2017; Kleven, Landais and Sogaard, 2016; Angelov, Johansson and Lindahl, 2016). Lundborg, Plug and Rasmussen (2017) also shows a substantial "motherhood penalty," using quasi-random successful IVF treatments.

On the other hand, women who choose to delay fertility in favor of career investments risk not achieving their desired family size. In addition to individual utility consequences of potentially lower fertility, women who marry when older may experience difficulties on the marriage market, since fecundity may be a trait that potential spouses value (Siow, 1998; Dessy and Djebbari, 2010; Bronson and Mazzocco, 2015). Díaz-Giménez and Giolito (2013) demonstrate that the gender difference in fecundity horizons can account for patterns of the spousal age gap, and hence that fecundity is fundamentally important in explaining the timing of marriage, even for young people. Low (2017) presents direct evidence on the impact of fertility on marriage markets, using an online experiment to demonstrate that when age is randomly assigned to dating profiles, men, but not

[5]See Blau and Kahn (2017) for updated evidence on the US and an extensive discussion and review on the possible sources of the phenomenon.

women, prefer younger partners. This effect is especially driven by men who have no children of their own and have accurate knowledge of the age-fertility relationship for women. An accompanying theoretical model shows that increasing women's expected fertility will increase career investment and improve marriage matches for women who choose to invest.

This literature suggests that technology alleviating the problem of age-related infertility could increase women's educational investment and improve marriage outcomes for women who do invest. However, the research on assisted reproduction technology (ART) has, to date, mainly focused on outcomes of women who actually use the technology, rather than younger women who perceive it as offering insurance for infertility later in life. A series of papers uses the variation in mandated insurance coverage of ART (including IVF) across US states and over time to determine how more coverage affects IVF usage and fertility outcomes (Velez et al., 2014; Hamilton and McManus, 2012; Bitler and Schmidt, 2012, 2006; Bundorf, Henne and Baker, 2007; Buckles, 2013; Schmidt, 2007, 2005), providing evidence that when coverage goes up, more women use IVF, fertility rates for older mothers go up, and multiple births rise.

A much more limited literature uses the same state-year variation to explore the impact on the timing of marriage and childbearing, supporting the hypothesis that access to infertility treatments (that will primarily be helpful later in life) may influence the decisions of younger women. The approach of using state-year variation in IVF coverage mandates has limitations, especially when discussing general equilibrium shifts in perceptions of both men and women. Since these are small and localized policy changes, awareness may not be widespread, particularly with young women who may not even be managing their own insurance yet. Due to the gradual and moderate increase in IVF usage, it is not clear if knowledge about the change could reach the young population by observation and discontinuously shift their perceptions. Moreover, insurance status may be the result of marital and career choices. This limits the ability to study the potential impact on *young* women's educational and career choices. In addition, it is impossible to establish that the timing of mandates was exogenous and that pre-trends are parallel. More generally, there is mixed evidence on how state health insurance mandates influence the insurance and labor market equilibrium: mandates may increase insurance premiums more significantly for the most affected workers and therefore negatively affect their wages and employment (Lahey, 2012).[6]

Hence, the only evidence on career outcomes comes from an unpublished work by Buckles (2007), who finds suggestive evidence that infertility insurance mandates led to increased labor force participation for women under the age of 35 (but a decrease for older women), and from Kroeger and La Mattina (2017), who present evidence on changes in occupational choices towards "more professional" careers for college educated women who were younger than 35 at the time of

[6]With any employer-provided insurance benefit, wages may fall to reflect the presence of the benefit to employees, and cost to employers. Lahey (2012) presents evidence on infertility mandates suggesting that in addition to these falling wages, because wage changes will not fully offset the increased premium costs for women in affected age groups, employment opportunities (and thus labor force participation) for this group decreases.

the change. No evidence is presented on educational completion, or specifically for young women who were making their choice of college majors concurrently with the policy change. Regarding marriage and birth timing, unpublished work by Ohinata (2011) finds that ART mandates resulted in 1-2 year delays in first birth among highly educated white women, and Abramowitz (2017, 2014) shows that increased access is associated with marriage and childbearing delays for white women. All of these studies are based on a DID strategy over a long time period, not accounting for possible differential time trends between states that did and did not implement coverage mandates. Nonetheless, these papers find effects even with this more limited variation, and thus suggest an important potential contribution in testing the hypothesis that access to IVF may affect women's early-life decisions using a more discrete policy event.

The Israeli policy change thus provides a unique opportunity, in applying equally to all, and being widely discussed publicly. Moreover, given that the coverage is publicly funded, there are no concerns that the observed changes in women's career investment are driven by a shift in employers' costs and preferences for employing older women. This paper is the first to study a large-scale policy change that affected not just the actual chance of getting pregnant when older, but, crucially, the beliefs about this chance by both young women considering career investments and their potential marriage partners. Moreover, this is the first paper to empirically study the impacts of a shift in later-life fertility potential on outcomes resulting from the decision to delay childbearing, including educational investments and marriage match quality.

3 Setting and Empirical Approach

3.1 IVF in Israel

Since the emergence of IVF technology in the early 80s, Israel has been on the forefront of IVF research: the first Israeli "test tube baby," born in 1982, was only the fifth IVF birth worldwide. However, until the early nineties, usage of the technology was still relatively low, and technological advances were slow in coming. IVF treatments were covered at least to some extent by Israel's four main health plans,[7] but in practice, most couples had to pay substantial fees to access services.[8] The extent of coverage and terms of eligibility varied between health plans and, in many cases, were vague or a priori undetermined.[9]

[7]The four health plans were partially subsidized by the government, but mostly relied on the membership fees they collected. Approximately 5% of the population, had no health insurance at all.

[8]See, for example "We will have to forgo having a child since we cannot afford fertility treatments," *Yedioth Ahronoth*, June 14, 1992, which tells the story of a couple who could not afford treatment because, although ostensibly covered by their health plan, the hospital required they pay 2,000 Israeli Shekels $approximately 1,300 in current US dollars$ as a "donation."

[9]The most generous coverage was offered by the largest health plan ("Clalit"), which placed almost no limitations on usage, but due to difficulty tracking treatments, rather than official policy (Birenbaum-Carmeli, 2004). The other health plans offered a limited number of treatment cycles and placed age restrictions on use, as well as requiring long qualification periods. For example in the "Leumit" health plan the number of treatment cycles was limited to

Following a widely covered, public debate, the Knesset enacted the 1994 National Health Insurance Law (NHI), which included IVF tests and treatments in a "basket" of free health services that all health plans must provide. The law provides all Israeli citizens with guaranteed access to:

> "IVF treatments for the purpose of the birth of two children for couples who do not have children from their present marriage, as well as for childless women who wish to establish a single-parent family."

The law, as originally written, did not place any restrictions on the age of women, or the number of attempts that could be made, and provided coverage for up to two "take-home babies." This is in stark contrast to most IVF coverage policies, which usually entitle beneficiaries to a certain number of *treatments*, rather than a certain outcome. The 1994 law thus provided access to IVF that is unmatched anywhere in the world, ushering in an era of expanded usage and technological improvement.

Importantly, the passing of the law was driven by a pro-natalist agenda, rooted in the Jewish tradition of familism, rather than pro-woman or "feminist" impulses which may have carried other effects.[10] As an example of the policy objectives behind the law, Israel's supreme court ruled (during numerous debates over the implications of the policy) that becoming a parent is a fundamental human right.[11]

The new and unique Israeli funding policy facilitated fast adoption and increased usage of fertility-enhancing technologies. Figure 1 shows that the number of IVF treatment cycles (both total and relative to the population of fertile women) more than doubled in the 6 years following the approval of the new policy.[12] Although the benefits of the law came into effect in 1995, the increase in the number of IVF treatment cycles began already in 1994, with the large amount of press coverage and increased knowledge of IVF availability. The figure on the right hand side shows that in the year after usage increased, there was a sharp increase in live deliveries using IVF.[13]

Figure 2 shows the direct impact on older women's fertility, by measuring the increase in women over 40 (a significant portion of whom would require fertility technology of some kind to be able to conceive) with children under one year old. The data for this figure is from the Israeli Annual Labor

six and the maximal age was 40. Membership in health plans was mostly based on political affiliation and parents' membership, and switching between them was very difficult.

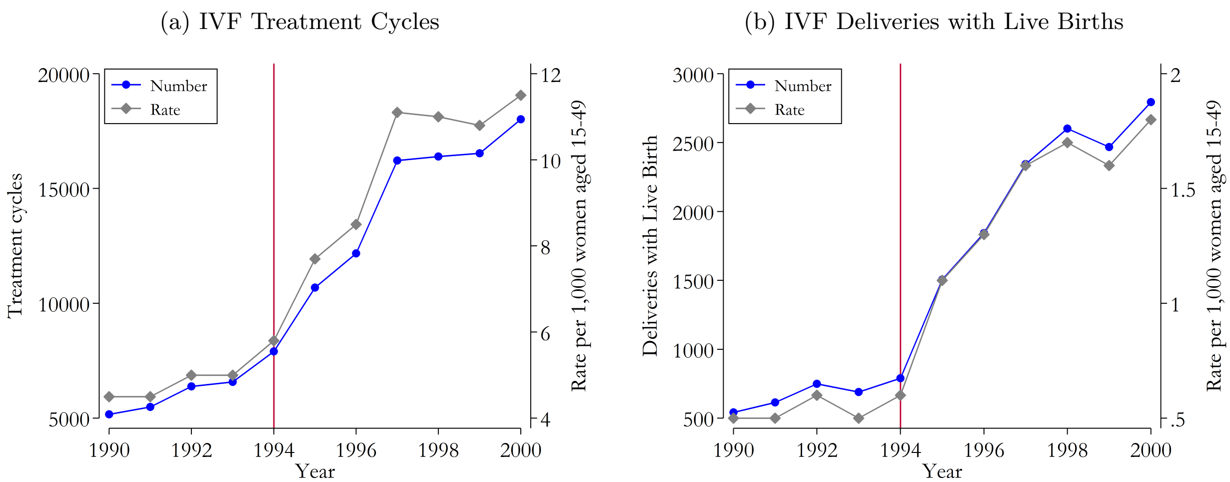
[10]Other examples of such policies are governmental child allowances and maternity grants, broad legal protection of working mothers' rights, extended funding of prenatal care and various tax benefits for parents. For a thorough discussion of those policies and their evolvement over time, see Birenbaum-Carmeli (2003).

[11]See High Court 7052/03 Adalla vs. Ministry of Interior.

[12]The Israeli parliament "Kneset" issued a report in 2012 that attributes this dramatic change to the regularization and expansion of IVF funding under the NHI law.

[13]The common measure of usage is the number of IVF treatment cycles relative to the size of fertile women population. Since there is no documentation of the number of women treated each year, it is impossible to assess whether the sharp increase in usage stems from an increase in the number of women undergoing IVF treatments, or from an increase in the number of attempts each IVF patient makes. It is reasonable to assume that it is a result of a combination of these two, especially given the large increase in IVF-assisted births.

Figure 1: Direct Impacts of IVF Access



Notes: Administrative data from Israeli Ministry of Health, covering all women in Israel.

Force Survey (LFS). This graph shows a large increase in older motherhood in 1995, immediately after women dramatically increased usage of IVF technology. This immediate change, amounting to more than 30% of the initial level of older motherhood,[14] persisted and kept expanding rapidly in the years that followed.

As a result of this shift, the media was flooded with IVF success stories, such as extreme cases of women having children at advanced ages, further raising awareness of the new technology.[15] In the Israeli press, local success stories were celebrated as "national accomplishments and symbols of local scientific excellence" (Birenbaum-Carmeli, 2004). The IVF law itself was also heavily covered in the press, and continued to be covered as debates ensued on whether to limit coverage.[16] Both 1994 and 1995 saw a spike in newspaper coverage of IVF and funding issues surrounding it.[17] Advocates of the health reform, including the minister of health, publicly touted the benefits of IVF funding in interviews as a reason to support the policy change.[18]

In the years following the policy change, there was expanding access to IVF services, and a standardization of practices surrounding IVF and its funding.[19] Nowadays, there are 26 IVF clinics

[14] We refer to 1993 as the pre-change reference since 1994 may have already been partially impacted

[15] For example, "World record: woman aged 60 gave birth to girl," *Yedioth Ahronoth*, February 22, 1994. (After 44 failed test-tube fertilizations, a 60-year-old woman gave birth to a baby girl in 1994.)

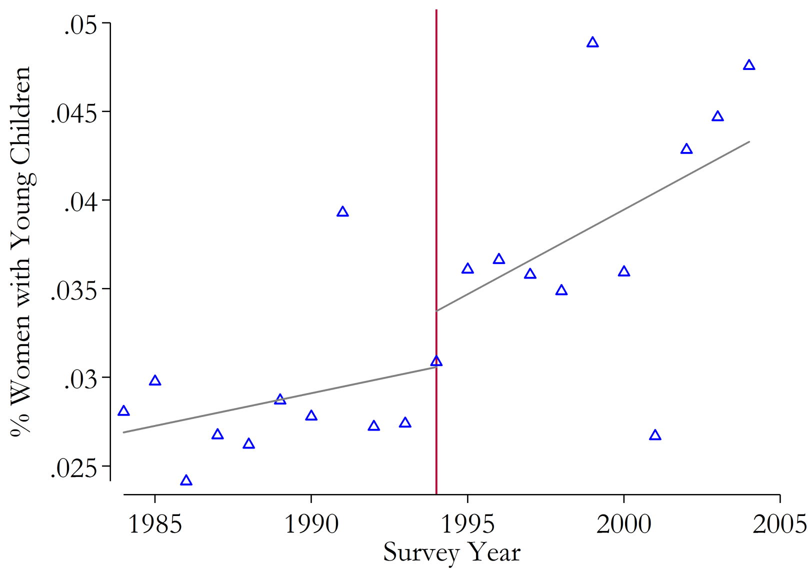
[16] The Ministry of Health expressed its intent to limit coverage to seven treatment cycles and provoked public protest. The press covered this conflict using personal stories of women over 40 that had children only following dozens of IVF treatment cycles and others who are still trying after a number of failures (Birenbaum-Carmeli, 2004).

[17] Based on a news search of Israel's leading Hebrew newspaper, 1994-1995 saw the largest number of articles on fertility, including success stories, reports about the law and the arrangements surrounding it (e.g. how to fund treatments for couples with different HMOs)

[18] See for example an interview with the minister of health Dr. Ephraim Sneh, *Yedioth Ahronoth*, December 15, 1994.

[19] The most distinct example is the 1996 Embryonic Carrying Agreement Law, officially legalizing and regulating surrogacy for the first time in the world (Simonstein, 2010).

Figure 2: Percentage of Women over 40 with Children $\leq$1 year, Labor Force Survey



Notes: The figure presents the percentage of women above age 40 (>40 and <=47, since in practice very few women above 47 have young children) with children of age 1 or below. Data from the Israeli Annual Labor Force Survey 1984-2004, restricted to Israeli-born Jews.

spread throughout Israel, making treatment very easily accessible for most residents of Israel. Israel has become the world leader in the rate of IVF treatment cycles and in the percentage of babies born following IVF treatments: approximately 4% of all babies born in Israel are conceived using IVF (Hashiloni-Dolev, 2013).[20] Meanwhile, IVF technology was steadily improving during this period, increasing success rates and the range of fertility problems that could be treated. The three forces of improved access, technological improvement, and publicity reinforced each other, driving a rapid and ongoing change in Israelis' attitudes and perceptions regarding IVF success rates, and thus the fertility time horizons for women.

Hashiloni-Dolev, Kaplan and Shkedi-Rafid (2011) surveyed Israeli undergraduates, both male and female, in 2009 about their perceptions of natural fertility and IVF success. Table 1 presents an estimated comparison between perceived natural fertility versus IVF-enabled fertility for women of different ages, constructed using these survey responses.[21] The survey responses show, first, that college-aged men and women were aware of the natural decline in fertility with age, although they somewhat over-estimated older-age success rates.[22] Next, it shows that not only were college-age students aware of IVF technology, they, if anything, believed it to be much more effective than true medical rates.[23] Students believed that the addition of IVF technology made 36-39 year-old mothers as fertile as those age 20-35, demonstrating the belief that IVF allowed a substantial delay in fertility commencement. Moreover, the students estimated that IVF created a substantial improvement in fertility in each age range. This data allows us to approximate the extent to which young women's expected fertility increased following the introduction of IVF.

3.2 Empirical Approach and Data

We examine the impact of knowledge of IVF access on timing of marriage, higher education attainment (both college and graduate), career outcomes, and spousal quality for women who marry when older. The impact of increased access to fertility extending technologies could be expected to be particularly salient to Israelis, who tend to marry 3 years younger and have 1.5 more children than in other OECD countries. Moreover, as Israelis tend to complete education 2.5 years later than in these countries, due to mandatory military service, even early educational and career investments,

[20] Compared to approximately 1-2% of the children born in other countries where IVF use is prevalent. The annual number of IVF cycles per million persons in Israel is the highest in the world and amounts to almost 3,500, compared to 2,000 in Denmark, which is second (Birenbaum-Carmeli, 2010).

[21] The survey did not ask about the same age ranges for natural fertility versus IVF success. Thus, we fit a flexible polynomial to the responses for natural fertility to construct estimates of the perceived natural fertility for the same age ranges for which students were asked about IVF success. We then assumed that IVF was to be used in the case natural conception was unsuccessful, so added the potential additional IVF success (IVF success rate multiplied by the chance of natural conception failure) to the natural conception rate. There is no similar study conducted in Israel prior to 1994.

[22] Note that the over-estimation of natural fertility rates could also be attributed to increased availability of ARTs, because while older-age pregnancies are observable, the usage of fertility extending technology is not.

[23] For example, students believed that IVF would be 32% effective for women 40-43, whereas the true rate is around 20% (Hashiloni-Dolev, Kaplan and Shkedi-Rafid, 2011).

Table 1: Beliefs of Israeli Students about IVF Success Rates, 2009

Woman's age	Natural fertility success rate %	Success rate with IVF %	Improvement from IVF %
20-35	74.6	90.8	21.7
36-39	58.1	75.9	30.5
40-43	46.9	63.9	36.3
44-47	36.8	52.8	43.2
48-52	28.4	41.5	45.8
53-58	17.6	29.5	67.4

Notes: Imputed estimates from Hashiloni-Dolev, Kaplan and Shkedi-Rafid (2011) survey of Israeli male and female college students. Estimates for natural fertility success rates for given age ranges created by fitting a fifth-order polynomial to survey responses, which were for different age ranges. Total success rates computed by multiplying IVF success rates from the survey by the natural fertility failure rate, then adding to the natural fertility success rate. Percent improvement is the success attributed to IVF divided by baseline success.

such as completing college, may infringe on a woman's planned reproductive years, and potentially limit family size.[24] This makes Israel an ideal setting for the study of the impact of extended reproductive time horizons offered by IVF technology on women's decisions and outcomes.

As previously mentioned, the 1994 Israeli IVF policy applied to all Israeli citizens. This provides the advantage of potentially shifting widespread beliefs about reproductive time horizons, but the disadvantage of not providing statutory variation for identification. We thus employ a difference-in-differences strategy, comparing groups which are expected to be more versus less affected by the policy, within the country. We use multiple control groups and a number of robustness and placebo checks to present a collage of evidence that the policy indeed causally impacted the outcomes we examine.

Our data comes from the 2008 Israeli population census, covering approximately 20% of Israeli households.[25] This data is combined with administrative tax data, to accurately measure income. Our analysis is restricted to native born Israelis, to avoid potential bias from substantial immigrant inflows over time, including the mass migration from the former USSR during the late 80s and early 90s. Our "treated" group is Jewish-Israeli women, who are most likely to be responsive to a change in IVF access. When we choose the appropriate comparison groups, we rely on the fact that different population groups responded differently to the introduction of IVF (Remennick, 2010), in addition to having significantly different gender and family norms (Danziger and Neuman, 1999).

[24]These estimated differences are based on UN, OECD and Israeli CBS data and reports.

[25]The survey began at the end of 2008 and was concluded in July 2009.

Religion plays a significant role in these differences, and while Judaism allows essentially all ART practices, the Roman Catholic church completely bans them and Islam places significant restrictions on use (especially on surrogacy and egg and sperm donations (Birenbaum-Carmeli, 2003)). We use Jewish men, Arab women, and young Jewish women as comparison groups for various parts of our analysis.

We use Jewish men to absorb changes that could have affected the overall marriage and educational market for Jewish Israelis, such as changing life expectancies, economic shifts or changes in labor demand, and demographic change. Given that men do not experience the same drop in fertility with age as women, IVF funding is unlikely to affect their expectations for age-related levels of fecundity.[26] They may, however, be affected in equilibrium by shifts in women's choices (e.g., for marriage age, having to search for a new partner if women choose to delay), but these effects would not exceed the initial impact on women. This does mean, however, that our effect should be considered a lower bound on the true impact.

We use Arab women to absorb changes that could have affected specifically *women* in Israel, distinctly from men. One important example is an ongoing reform in Israel that increased access to higher education, especially for disadvantaged populations in peripheral areas. This reform might impact women differentially, for example, if the cost of moving away from home to acquire education is more costly for women than for men, or if it helped remove cultural barriers to women's education. If so, though, Arab women would have certainly been impacted by this change. On the other hand, Arab-Israeli women make a suitable control group first and foremost because they are much less likely to use IVF, due to stronger religious restrictions on its use. Secondly, Arab women were at the time much less likely to be on the margin for time-costly career investments, due to lower baseline education levels and younger age at first marriage (although trends in marriage age were parallel).[27] Additionally, Arab women begin college on average 3-4 years younger than Jewish women, since they do not serve in the military, creating a lower concern for the impact of college education on fertility.[28] While this strategy may have its own potential confounding factors, they should be orthogonal to any issues presented by the male control group. Thus, if we estimate similar effects using the two strategies, it is unlikely that they are both caused by a single omitted factor.

[26] Similarly to how in Jayachandran and Lleras-Muney (2009), women's life expectancy was updated by declines in maternal mortality, while men's was unaffected.

[27] According to annual data published by the Israeli central bureau of statistics, at 1993 (just before the policy change) the median age at first marriage for Arab-Muslim Israeli women was 20 compared to 23.3 for Jewish Israeli women (average age was 21.1 compared to 24). Total fertility rates for Arab Israelis were also higher, at 4.5-4.7 births per women, compared to slightly below 3 for Jewish-Israeli women (as reported in CBS working paper no. 60, Fertility among Jewish and Muslim Women in Israel, by Ahmad Hleihel).

[28] See for example CBS report "Arabs in Higher Education in Israel - First Year Students for First Degree in 2011/12" issued October 21st 2014. It should also be noted that the variance of the age of college applicants is much larger for the Jewish population (based on CBS data processed and presented by Mr. Aviel Kranzler, Higher Education and Science department at the CBS.)

In addition to these main control groups, we also employ other control groups and tactics to help identify the impact of the policy. For the analysis on how older women's spousal quality changes following the policy change, we are able to use younger women as a control group, since women who marry younger would have their anticipated fertility largely unchanged by the new technology. To account for the potential that broader international trends, such as the global trend in women seeking more education, are responsible for our results, we do placebo analyses in several other countries. To verify that censoring due to the retrospective nature of the data does not bias the results, we compare our main findings to a placebo analysis using the exact same techniques in the 1995 Census, looking at the period prior to the true policy change. All of these exercises confirm that our effect is unique and empirically large.

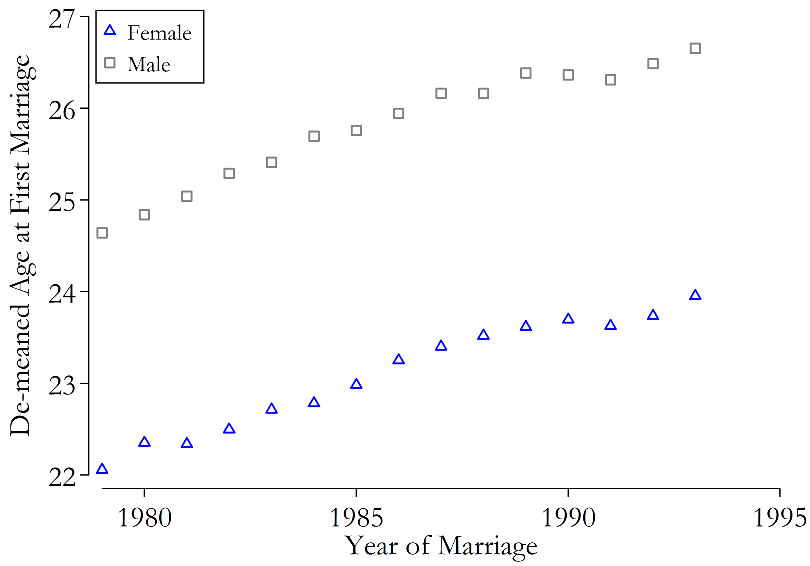
We begin our analysis by looking at age at first marriage, by year of marriage, over a 30-year study period, from 1979 to 2008. We consider 1994 to be the first year of the treatment period, as our treatment is knowledge of IVF availability in the future and the resulting change of expectations, rather than the actual funding change. We first use men as the "less affected" comparison group. Using this group to identify the causal impact of interest requires men and women's average age at first marriage to respond similarly to any changes in the environment except for the introduction of IVF. To assess the plausibility of this assumption, we plot the pre-period trends of age at first marriage for men and women separately. Figure 3 shows that men and women had strikingly parallel trends in age at first marriage during the pre-period, moving in lock-step and responding to common shocks. This may be partially attributable to the largely "closed" marriage market of Israeli-born Jews—immigrants and other ethnic groups tend to marry within their own groups. One limitation of the tight relationship between male and female age at first marriage, is that if women suddenly decided to delay marriage, it may create an "echo" effect on men's outcomes.This would make the estimated effect using this control strategy a lower bound on the true impact; however, this effect could not cause us to find an effect where none exists.

We first estimate a basic difference-in-differences specification, that measures the average change in the female-male difference in age at first marriage (AFM), between the "pre" and "post" periods, according to the following equation:

$$AFM_i = \beta_0 + \beta_1 fem_i + \beta_2 post_i + \beta_3 fem_i \times post_i + \beta_4 time_i + X_i'\gamma + u_i$$

where X is a vector of individual level controls, which includes indicators for religiosity (ultra-orthodox or not) and parents' origin (Europe and America, Asia and Africa, or Israeli born), to account for demographic shifts over time, fem is a dummy for female, $post$ is a dummy for marriage years 1994 and onwards, and $time_i$ is a linear time trend. We then test for both a change in levels at the time of the policy change and a change in the time-trend of the outcome variable, allowing

Figure 3: Female vs. Male Age at First Marriage – Pre-period



Notes: Figure shows average age at first marriage for women and men, by year of marriage, for the years prior to the policy change. Data from the 2008 Israeli population census, restricted to Israeli-born Jews.

us to examine the evolution of the effect over time:

$$AFM_i = \beta_0 + \beta_1 fem_i + \beta_2 post_i + \beta_3 fem_i \times post_i + \beta_4 time_i \\ + \beta_5 post_i \times time_i + \beta_6 fem_i \times time_i + \beta_7 fem_i \times post_i \times time_i + X_i'\gamma + u_i$$

We repeat the estimation of both equations adding year of marriage fixed-effects, to more flexibly control for time trends and account for transitory shocks that may affect the marriage market. We present the same estimation with the Arab women control in the appendix.

To study how the distribution of age at first marriage changed, we run a series of regressions using indicators for being married by a given age as the outcome variable, with the same specification. We expect to find that the shift in average age is driven by women in their mid to late twenties postponing marriage, rather than women who marry very young, who should not be affected by fecundity concerns. This also allows us to examine the extensive margin of marriage, by looking at those who "ever married" by their forties.

We then turn to examining the impact of the policy on education, first comparing women's outcomes to those of men.[29] Given that our data is from a single 2008 cross-section, we identify as treated in this case those from cohorts that were still at the relevant age for educational decisions

[29]It should be noted that just as with age at first marriage, men's education levels may be affected in general equilibrium (e.g., if partner education is complementary in marital surplus, and so more educated women leads to more educated men), but should be affected much less than women.

at the time they learned about the increased access to IVF. As our main specification, we use the median age of applicants (as reported in macro data) to determine the first treated cohort. We use a data range from the 1951 cohort to the 1977 cohort, for college education, and the 1974 cohort for graduate education, to avoid censoring.[30]

Since there is a slight upward trend in women's college education relative to men's pre-treatment, we control for a gender-specific linear time trends in the standard DID specification, and estimate the following equation:

$$Education_i = \beta_0 + \beta_1 fem_i + \beta_2 post_i + \beta_3 fem_i \times post_i + \beta_4 time_i + \beta_5 fem_i \times time_i + X_i' \gamma + u_i$$

where *Education* is an indicator for having a college degree (Bachelor's degree) or a graduate degree, *post* indicates cohorts young enough to enter college or graduate school after the policy change and *time* allows for a linear trend over cohorts (also interacted with fem to account for gender specific pre-trends) . In addition, we estimate a specification which allows for changes both in level and trend (as with age at first marriage) and add year of birth fixed effects to both specifications.

We then re-estimate this section using Arab women as the control group, to account for any possible changes in women's access to or demand for education over the same time period. We also examine this outcome in repeated cross-section data, the Israeli Labor Force Survey, and conduct a placebo test on high school completion.

Finally, we examine the impacts of the policy on marriage market outcomes for older women, measured by spousal quality. We compare women who marry between 30 and 34 to those who marry between 25 and 29. We place the cutoff between the two groups at 30 to exploit the perceived discontinuity in expected fertility exhibited at this age. Following the IVF policy change, the group of "older" brides (and their potential spouses) may expect greater fecundity following the policy change, while the marital fecundity of the "younger" brides remains essentially unchanged. We measure the impact of this difference over year of marriage, expecting to observe a lag in this outcome's response to the policy change, since reaching a new equilibrium in the marriage market takes time. Nevertheless, we use 1994, the year of the new policy introduction, as the first year in the post-treatment period.

The main measure we use for spousal quality is income. Since our data is cross-sectional, we use several methods to alleviate concerns for bias due to comparing spouses of different ages. First, the restriction of brides' age range (25-34) also places some bounds on grooms' age range. Second, in all the specifications for this outcome, we control for the age of the spouse, using a flexible

[30] 1951 is chosen as the beginning of the range because it is the first cohort to have a reasonable number of observations for Israeli-born Jews (the state of Israel was founded in 1948). The end of our data range is chosen to avoid censoring in educational outcomes among individuals who may not have completed their education by the time of the 2008 Census, thus we analyze individuals no younger than 31 years old for college education and individuals no younger than 34 years old for graduate education.

polynomial in age (in the appendix, we show that results hold when we alternatively control for spousal year-of-birth fixed effects or use income-for-age percentiles as the outcome). Third, we use additional measures of spousal quality which are less sensitive to age, e.g., level of education and occupation. Last, we restrict our sample period to 18 years between 1986-2003, in order to avoid measuring income of students or retirees. This range of years was determined based on men's full-time employment patterns over age together with the distribution of their age at marriage. Due to this restriction and to the expected lag in the development of the new marriage market equilibrium, we have too few post-treatment periods to identify a change in trend and thus focus on a difference-in-differences analysis for levels only:

$$Spouse\_inc_i = \beta_0 + \beta_1 older_i + \beta_2 post_i + \beta_3 older_i \times post_i + \beta_4 time_i + X_i'\gamma + u_i$$

Here, *post* indicates marriage years 1994 and onwards and the demographic controls contained in X address characteristics of the *bride* (although the outcome measure is for the groom), in addition to controlling for a fifth degree polynomial function of spousal age, to account for the cross-sectional nature of the data. In the appendix we present results for specifications with additional controls.

For this outcome, we can also use men as a second control group in a triple-differences specification:

$$\begin{aligned} Spouse\_inc_i = {} & \beta_0 + \beta_1 older_i + \beta_2 post_i + \beta_3 older_i \times post_i \\ & + \beta_4 fem_i + \beta_5 older_i \times fem_i + \beta_6 post_i \times fem_i + \beta_7 older_i \times post_i \times fem_i + X_i'\gamma + u_i \end{aligned}$$

The coefficient of interest here is β_7 which can be interpreted as the change in spousal quality for "older" brides relative to younger brides, compared to those same differences in spousal quality for *grooms* marrying over this time period. This helps determine that the evolution we observe is driven by a female specific change, and not by some general shift for older marriages. However, the quantitative results of this estimation should be interpreted cautiously, since we are observing a shift in equilibrium which will also impact men's outcomes.

For all of the outcomes and specifications described above, we use two alternative methods to calculate standard errors. The first method, has the advantage of accounting for cross-sectional correlation in outcomes as we cluster at the year $\times$ group level. Our second set of standard errors aims to deal with potential serial correlation in the outcome variables, which could lead to over-rejection of the null hypothesis in a DID framework (Bertrand, Duflo and Mullainathan, 2004). We divide our sample into sub-groups, based on the standard classification of Israel into 51 "natural regions" and cluster at the geography $\times$ group level (examples of similar sub-group clustering can be found in Agarwal et al. (2014), clustering at the product level, and Hanlon (2015), clustering at the patent level). Since the regions are defined to be as homogenous as possible, both in terms of geographical characteristics and in terms of demographic and economic traits of their population, we expect any shocks related to women's access to education or health services to occur within these regions. Moreover, structural shifts in labor and marriage markets, which directly affect our outcomes of interest, are likely to have significant regional components.

To further address the concern for serial correlation, we then re-estimate our main specification for each outcome with an explicit AR(1) error structure, by collapsing our data into year-group cells and estimating via Generalized Least Squares (GLS), allowing for correlation both *across* and *within* panels (as in Chandra, Gruber and McKnight (2010)). The within-panel correlation factor accounts for serial correlation, assuming an AR(1) process with a unique autocorrelation parameter for each panel (i.e., gender). Finally, in the appendix, we show permutation tests for each of our main results, demonstrating that our effects are "large" relative to the actual variation present in the data.

Because there may have been other long-term societal trends that could have divergent effects for men and women, we perform several analyses to provide further evidence that the 1994 IVF policy change drives our results. First, we use a Quandt Likelihood Ratio (QLR) test[31] to search over all possible break dates, and show that our "treatment year" is indeed identified as the break among candidate dates. Second, we use event study graphs, charting the impact over time around the time of the policy change, to show that a pre-trend is not driving our results, but rather that the observed effects only become significant after the policy change. We additionally present a variety of placebo tests, discussed above, as well as a number of robustness checks to control for possible omitted factors. Finally, we rely on the specific combination of outcome variables to bolster the evidence that IVF access is the driver behind the changes. While there are a few other mechanisms that may have an effect on one of the outcomes we study, none of those can be expected to impact *both* women's educational and marriage decisions *and* marriage outcomes for older women. We review candidate alternative explanations in more detail in section 4.4.

Table 2 shows summary statistics for our sample, comparing Jewish women to Jewish men (our main comparison) as well as Arab women, and then women who marry while young (25-29) to those who marry older (30-34). In addition to showing means for our key outcomes and controls, Table 2 also compares pre-trends in outcomes for the different groups. As mentioned above, our specifications control for group-specific trends when needed.

[31]See Andrews (1993).

Table 2: Summary Statistics

Marrying pre-1994:	Jewish women N=38,370		Jewish men N=33,949		Arab women N=14,901	
	Mean	SD	Mean	SD	Mean	SD
Ultra-Orthodox	0.09	0.28	0.09	0.29	N/A	N/A
European-born mother	0.24	0.43	0.28	0.45	N/A	N/A
Asian/African-born mother	0.54	0.50	0.52	0.50	N/A	N/A
Income (Shekels)	95,629	92,393	186,757	173,543	53,299	49,999
Age	44.81	5.39	47.46	5.34	42.24	5.66
Age at first marriage	23.15	3.91	25.86	3.86	21.03	4.11
AFM pre-trend (SE)	0.13	(0.00)	0.14	(0.00)	0.08	(0.01)
College age pre-1994:	N=61,000		N=58,704		N=22,278	
	Mean	SD	Mean	SD	Mean	SD
College Educated	0.31	0.46	0.28	0.45	0.08	0.27
College pre-trend (SE)	0.0057	(0.00)	0.0041	(0.00)	0.0035	(0.00)
Grad-school age pre-1994:	N=46,428		N=44,355		N=16,449	
	Mean	SD	Mean	SD	Mean	SD
Highly Educated	0.11	0.31	0.11	0.31	0.01	0.10
Highly Ed. pre-trend (SE)	0.0009	(0.00)	0.0008	(0.00)	0.0033	(0.00)

Marrying pre-1994:	Married 30-34 N=3,549		Married 25-29 N=11,227	
	Mean	SD	Mean	SD
Ultra-Orthodox	0.03	0.18	0.03	0.16
European-born mother	0.31	0.46	0.24	0.43
Asia/African-born mother	0.25	0.43	0.30	0.46
Income (Shekels)	92,324	89,096	99,632	96,644
Spousal Income	166,626	170,094	192,747	181,682
Sp. Income pre-trend (SE)	-1494	(1,734)	504	(954)

Notes: 2008 Israeli population census (20% sample). Restricted to Israeli-born. Sample "marrying pre-1994" is those married 1979 - 1993, inclusive. "College age pre-1993" is those born 1951 - 1970 for Jewish population, and 1954 - 1973 for Arab population. "Grad-school age pre-1994" is those born 1951 - 1966 for Jewish population, 1954 - 1969 for Arab population. Table by marriage age uses sample of women married 1986 - 1993 (shorter range due to income censoring for older individuals) with spousal matches in 2008 census.

4 Results

4.1 Marriage Timing

The first outcome we examine is women's marriage timing. We hypothesize that when women feel more certain about their reproductive prospects later in life, they are more willing to delay marriage, which could in turn allow greater career investments. In contrast to the other outcomes we consider, the decision to delay marriage is completely controlled by the individual and does not require meeting certain standards or going through some process, such as being accepted to a university. It also does not require the consent of a potential partner or any update to men's beliefs, as would be the case for the choice of spouse, for example.[32] This should enable us to identify a clean and immediate effect precisely at the year when the policy was introduced. Note that even women unaware of the policy change itself may have altered their marriage decisions due to updated beliefs about later life fertility, as knowledge of IVF swiftly spread through media coverage of the policy and related success stories. This increased sense of control over later life fertility both increased expected utility from delayed childbearing while potentially alleviating concerns about being penalized on the marriage market for delaying marriage, as potential spouses also value higher fecundity. Both would increase the willingness to delay marriage.

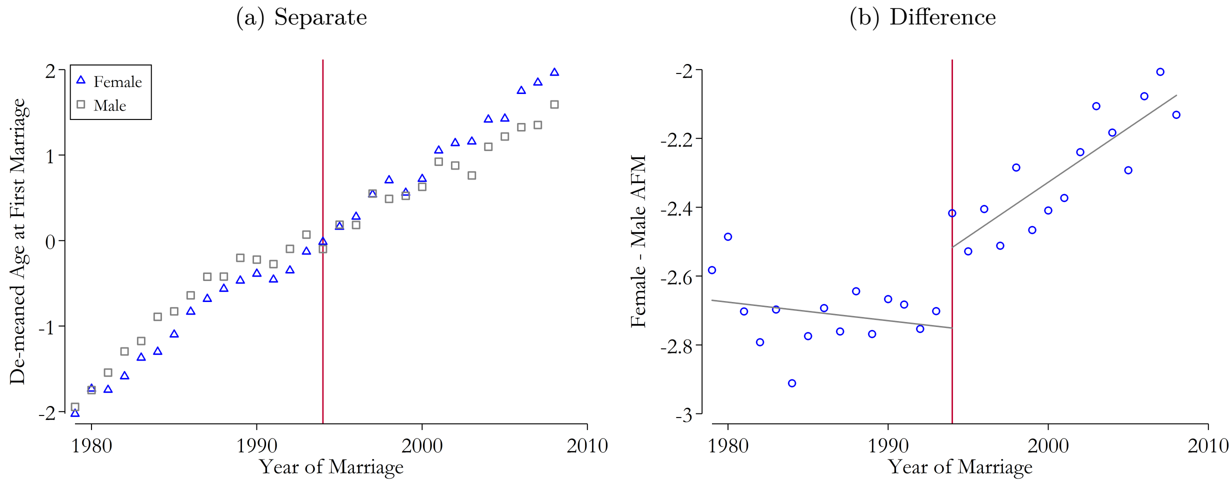
For this part of the analysis, time represents the year the marriages are taking place, and thus, the treatment year is 1994, the year of the policy change. Figure 4 clearly shows that pre-trends for men and women were parallel, in the lefthand figure that graphs outcomes separately (with age at first marriage de-meaned), implying that women's marriage age was practically constant relative to men's until 1994. Men and women's age at first marriage also appears to respond to common shocks in the pre-period. Starting in 1994, women's marriage age increases discontinuosly relative to men.[33] The graph of the difference in outcomes shows a sharp increase in level immediately at 1994, followed by a substantial positive change in trend. This striking discontinuity in age at first marriage is unprecedented, is unique amongst other countries, and has no clear explanation other than the introduction of free IVF.

Table 3 analyzes this change using a regression, in both a simple DID framework (columns 1 and 2), and an analysis demonstrating the change both in level and trend (columns 3 and 4). The latter indicates an increasing change in the outcome over time, which correlates with the gradual change in perceptions, rather than just a one-time jump. We find that women's age at marriage jumps by three months, relative to men, immediately following the policy change, and continues steadily increasing thereafter. Both the discontinuity as well as the slope change are significant, although at a lower level under the geographically clustered standard errors. Columns 5 and 6

[32]It should also be noted that in Israel, couples tend to marry very soon after becoming engaged, and so there is not an extensive "lag" between the decision to get married and marriage itself.

[33]Although men's age at first marriage appears to decline in this year, in all previous shocks, men's and women's outcomes have responded in tandem. Thus, relative to the common shock, women's age increases.

Figure 4: Female vs. Male Age at First Marriage



Notes: Figure (a) shows average age at first marriage for women and men, by year of marriage, de-meaned so that the relative changes can be seen more clearly. Figure (b) presents the difference in average age at first marriage between women and men, as well as fitted lines for the pre (1979-1993) and post (1994-2008) periods. Data from the 2008 Israeli population census, restricted to Israeli-born Jews.

show the Generalized Least Squares (GLS) specification, where data is collapsed to year-group cells, for the DID with slopes. As mentioned above, the standard errors are adjusted to account for cross-sectional correlation and serial correlation, assuming a panel-specific AR(1) process. Once again, both the intercept change and slope change at 1994 are significant, and of similar magnitude to the OLS estimates. Note, these large effects were most likely caused by some women delaying substantially (perhaps to pursue education), and others being unaffected.

To understand how the distribution of marriage age was affected, rather than just the average, we run a series of regressions using the column 3 specification, but replacing the outcome variable with an indicator for being married by a certain age. Figure 5 shows the point estimates and confidence intervals for the two coefficients of interest on the interaction terms $fem \times post$ and $fem \times post \times time$, for each age cutoff. Figure 5(a) presents estimates for the immediate change in level (i.e. change in the percentage of women married by the specified age) and figure 5(b) shows the estimated change in trend. The two graphs show no decrease in marrying by age 22, which provides a useful falsification test, since we would not expect women inclined to marry and begin childbearing by age 22 to be concerned about fertility in their late thirties, and hence to be affected by access to IVF. We see the largest reduction in marriage by age 26, and from there a steadily decreasing impact, until the total effect reaches zero at age 38. Importantly, the lack of reduction in marriage by age 38 suggests that women are delaying marriage, but not forgoing it entirely. Overall, this analysis suggests that the decrease in average marriage age after the policy change is mostly driven by women delaying marriages from their mid- and late-twenties into their thirties

Table 3: Age at First Marriage

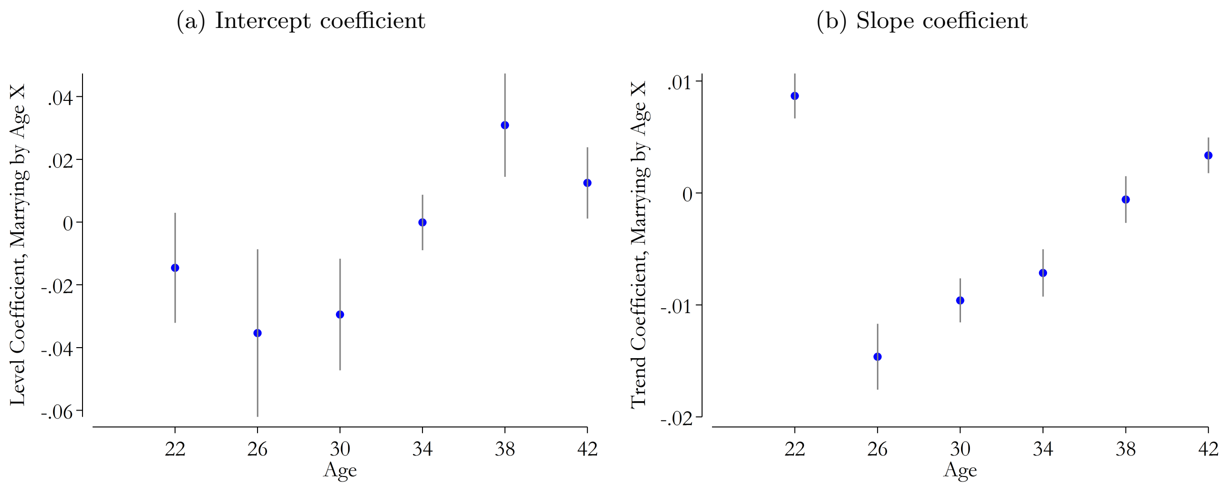
	Dependent variable: Age at First Marriage					
	DiD		Slope-Change DiD		GLS Slope-Change DiD	
	(1)	(2)	(3)	(4)	(5)	(6)
fem × post	0.412	0.415	0.241	0.246	0.247	0.195
	(0.073)***	(0.036)***	(0.127)*	(0.044)***	(0.082)***	(0.058)***
	[0.221]*	[0.220]*	[0.135]*	[0.133]*		
fem × post × time			0.039	0.040	0.035	0.036
			(0.013)***	(0.005)***	(0.010)***	(0.007)***
			[0.021]*	[0.022]*		
fem × time			-0.008	-0.008	-0.009	-0.007
			(0.012)	(0.003)**	(0.008)	(0.005)
			[0.015]	[0.015]		
post × time			0.001		-0.002	
			(0.010)		(0.013)	
			[0.014]			
post	-0.440		-0.322		-0.395	
	(0.084)***		(0.091)***		(0.108)***	
	[0.141]***		[0.079]***			
female	-2.649	-2.651	-2.710	-2.715	-2.783	-2.757
	(0.059)***	(0.016)***	(0.114)***	(0.031)***	(0.067)***	(0.045)***
	[0.236]***	[0.235]***	[0.253]***	[0.250]***		
time	0.176		0.168		0.138	
	(0.004)***		(0.008)***		(0.010)***	
	[0.011]***		[0.011]***			
Constant	26.319	23.947	26.251	23.521	26.856	24.628
	(0.070)***	(0.078)***	(0.096)***	(0.063)***	(0.088)***	(0.022)***
	[0.198]***	[0.187]***	[0.188]***	[0.185]***		
YOM FEs		YES		YES		YES
Observations	167416	167416	167416	167416	60	60
R-Squared	0.246	0.246	0.246	0.247		

Notes: Columns 1–4: Ordinary least-squares difference-in-differences regression using micro data, including controls for religiosity and parents' origin. Robust standard errors clustered at the gender × year level in parentheses; robust standard errors clustered at the gender × geography level in square brackets. Columns 5–6: Generalized least squares regression with data collapsed to the gender-year-of-marriage level. Robust standard errors that allow for cross-sectional correlation and for panel-specific serial correlation (i.e. estimate a unique autocorrelation parameter for each group) , in parentheses. Data from the 2008 Israeli population census, restricted to Israeli-born Jews.

*** p<0.01, ** p<0.05, * p<0.1

and even late thirties.

Figure 5: Regression Coefficients for Effect of IVF Law on Marrying by a Given Age



Notes: The figures presents the point estimates and confidence intervals of the coefficients on (a) the interaction term fem×post and (b) the interaction term fem×post×time, for regressions where the outcome is a binary variable indicating whether or not the individual got married at or before a certain age, and the specification is as in column (3) in table 3. Data from the 2008 Israeli population census, restricted to Israeli-born Jews.

Quandt Likelihood Ratio breakpoint test To confirm that what we are picking up is truly a discontinuous shift in age at first marriage—a break in the time series—rather than more gradual time trends, we perform a Quandt Likelihood Test to "search" for the most likely break year in the data, over our entire sample period except for the standard 15% "trimming" on either end, to account for limited data at the beginning and end of the sample period. We perform this test for age at first marriage, because it is the outcome measure that should have the cleanest "break" at 1994, since the education outcomes rely on cohorts entering in 1994, which may be imprecise, and spousal income relies on shifts in the marriage market, which may take more time.

To implement the test, we run a loop of regressions identical in specification to our columns 3 and 4 regressions, except the "break" year changes in each regression. We then perform an F-test for whether the two "break" parameters—slope and intercept—are different from zero. Finally, we search for the maximal F-stat among these tests.

As shown in Table 4, the test returns the highest F-statistic for 1994, which indicates that the year of the policy change and hence our treatment year is the most probable break year. Moreover, this break is significant even when accounting for the multiple tests used to identify it. The procedure for the QLR specifies comparing this "sup(F-stat)" to a table of critical values adjusted for the number of tests: the critical value for two restrictions and 15% trimming is 5.86, whereas the QLR statistic for age at first marriage for the "break" year is 10.38 or 10.78, depending on whether fixed effects are used or not. This shows strong evidence of a break specifically at 1994.

Table 4: Quandt Likelihood Ratio test for break point

	F-Statistic	
Year of Marriage	No FEs	With YoM FEs
1983	7.05	7.40
1984	7.28	7.67
1985	7.29	7.69
1986	7.40	7.85
1987	7.83	8.27
1988	8.06	8.35
1989	8.12	8.36
1990	8.09	8.27
1991	8.03	8.16
1992	7.83	8.11
1993	8.57	8.91
1994	**10.38**	**10.78**
1995	7.19	7.49
1996	6.49	6.75
1997	5.61	5.86
1998	5.97	6.14
1999	3.95	4.07
2000	4.52	4.69
2001	4.02	4.16
2002	3.95	4.03
2003	2.60	2.64
2004	1.03	1.04

Notes: Table reports F stats from a regression according to the specification in columns 3 and 4 of Table 3, where the hypothesis is that the coefficients on $post \times fem$ and $post \times fem \times time$ equal 0, and "post" is defined as being greater than or equal to the indicated year. Standard errors are not clustered in this case, as clustering is not conventional in QLR models, but similar results are obtained with clustering.

4.2 Education and Career Choices

College Education We now turn to women's educational investments, and subsequent career outcomes. First, we examine the impact of the policy on college completion. At the time of the policy change, Israeli women's median age for college entry was 22.5.[34] This, combined with high average fertility rates, means that changes in reproductive time horizons could be expected to affect

[34] As reported by the Israel Central Bureau of Statistics. This later college entry is due to the two to three years period of mandatory military service following high school, and the typical one year period to take entry exams following that. In addition, military service may start and end "off cycle" with the academic year, further delaying college entry.

even college educational decisions (as opposed to in the US, where decisions about college are made earlier).

Because our data is cross-sectional, we identify as "treated" birth cohorts that were young enough to still be making the relevant educational decisions after the policy was announced (and beliefs were updated). Figure 6 shows the raw data used in this analysis, charting women's college completion compared to men, by year of birth. In our main specifications, we use the median age of college entry to define the post period as starting in 1971.[35] We assume cohorts born after this point are able to be influenced in their college-going decisions by new information about reproductive time horizons, whereas older cohorts would have already committed to an educational path.[36] However, since there is some variation in college-entry age, the treatment may be somewhat "fuzzy." To address this, we additionally construct a continuous treatment variable that measures the level of each cohort's exposure to treatment, i.e. the percent of the cohort that has not yet entered college.[37] Using this measure, in figure A1, we draw a "treatment interval" starting at the 1969 cohort, where approximately 25% are treated and ending at the 1973 cohort with about 75% treated. Our choice for the "sharp" cutoff, namely 1971, rests precisely in the middle of this term, during which the treatment was very rapidly deployed.

Both figures show that men and women's college education moves roughly in parallel prior to the 1994 time change, seemingly responding to common shocks, with the exception of the earliest three cohorts, in which a different pattern is present. This creates a slight upward pre-trend in women's education relative to men's. We therefore include group-specific time trends in the DID regression specification. Following the policy change, there is a sharp change in the slope of women's college completion relative to men. The "fuzziness" of the timing, in addition to the continued increase in knowledge over time, may be partially responsible for the pattern being more of a trend break than a discrete before-and-after jump. We interpret this effect as causal in light of the strong evidence for a trend break in age at first marriage, which suggests a sharp disruption to outcomes in 1994. This break is likely merely distributed over multiple cohorts in the educational results.

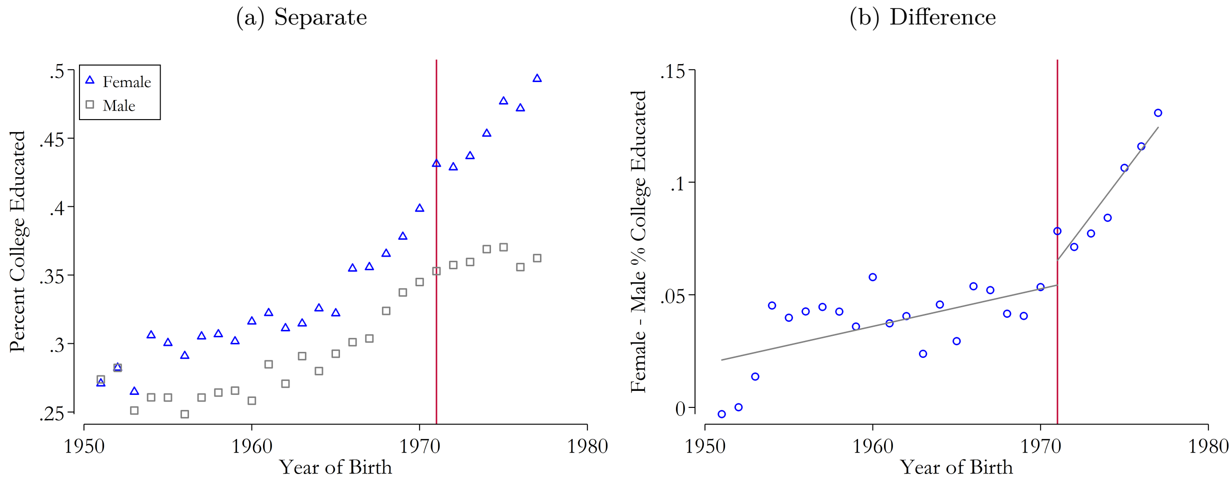
These results are presented formally in a regression in Table 5. Columns 1 and 2 show estimates for a simple difference-in-differences specification, with men as the control group and gender-specific time trends. The interaction between being female and of college-entering-age after the year of the policy change is positive and significant, showing a 2.5 percentage point increase in rates of college completion for women, relative to men, between the pre and post period. This represents an 8% increase in rates of college education from the average 31% level for women in the pre period. This

[35]Men's median age for college entry is 24. Because men enter college slightly later, we experiment with shifting the treatment year for men as one of our robustness checks, shown in the appendix. See table A4.

[36]We find additional support for this choice of cutoff based on our findings regarding marriage age, where 23 (the age of the 1971 cohort in 1994) is in the age range with the most substantial tendency to delay marriage following the policy change.

[37]The calculations of the continuous treatment variable are based on the age distribution of college entrants in 1993, as reported in the 1995 Census data.

Figure 6: Female vs. Male College Education



Notes: Figure (a) shows average college attainment for women and men by birth cohort. Figure (b) presents the difference in college attainment between women and men, as well as fitted lines for the pre (1951-1970 birth cohorts) and post (1971-1977 birth cohorts) periods. Data from the 2008 Israeli population census, restricted to Israeli-born Jews.

effect remains stable when year of birth fixed effects are introduced. Then, in columns 3 and 4, differential post trends are introduced, revealing that the effect is driven primarily by the change in slope. This is expected since, as discussed above, the exposure to treatment is in fact gradual, and educational decisions may be path dependent based on earlier choices. This specification is repeated using GLS estimation in columns 5 and 6. For this purpose our unit of observation is the group of same gender individuals born in a specific year (e.g. men born in 1968), and we collapse the data to means accordingly. As mentioned above, we correct standard errors for potential cross-sectional correlation and assume a gender-unique AR(1) process to account for potential serial correlation.

To account for the "fuzziness" in treated status based on cohort, in Table A1, we present a regression that classifies a portion of each cohort as treated based on which percentage of individuals would have not yet entered college, according to data from the 1995 Census on college entry ages. These percentages are allowed to be different for men and women, which accounts for the fact that men on average enter college later. These results again show a significant impact of being in the "treated" cohorts.[38]

To rule out a general change in education-seeking, either by the 1971 cohort or based on other things happening in 1994, we perform a placebo test using high school completion. Those on the margin of completing high school are unlikely to make this choice based on older-age fertility prospects. Figure A2 demonstrates that, as expected, there is no impact on this outcome either

[38]The differential and gradual exposure to treatment specification captures a separate non-linear time trend for men and women, hence we estimate a standard DID only in this case.

Table 5: College Graduation Rates

	Dependent variable: College Education					
	DiD with GSTT		Slope-Change DiD		GLS Slope-Change DiD	
	(1)	(2)	(3)	(4)	(5)	(6)
fem × post	0.025	0.025	0.009	0.009	0.020	0.013
	(0.011)**	(0.006)***	(0.009)	(0.008)	(0.012)*	(0.009)
	[0.016]	[0.016]	[0.015]	[0.015]		
fem × post × time			0.007	0.007	0.006	0.007
			(0.002)***	(0.002)***	(0.003)**	(0.002)***
			[0.004]*	[0.004]*		
fem × time	0.003	0.003	0.002	0.002	0.002	0.002
	(0.001)***	(0.000)***	(0.001)***	(0.000)***	(0.001)**	(0.000)***
	[0.001]***	[0.001]***	[0.001]**	[0.001]**		
post × time			-0.009		-0.001	
			(0.001)***		(0.003)	
			[0.003]***			
post	-0.015		0.007		0.021	
	(0.009)		(0.006)		(0.011)*	
	[0.012]		[0.010]			
female	0.059	0.059	0.055	0.055	0.051	0.053
	(0.006)***	(0.004)***	(0.005)***	(0.004)***	(0.009)***	(0.005)***
	[0.025]**	[0.025]**	[0.026]**	[0.026]**		
time	0.005		0.006		0.004	
	(0.000)***		(0.000)***		(0.001)***	
	[0.001]***		[0.001]***			
Constant	0.505	0.394	0.512	0.392	0.326	0.257
	(0.007)***	(0.007)***	(0.005)***	(0.009)***	(0.009)***	(0.003)***
	[0.026]***	[0.024]***	[0.026]***	[0.024]***		
YOB FEs		YES		YES		YES
Observations	173790	173790	173790	173790	54	54
R-Squared	0.108	0.109	0.109	0.109		

Notes: Columns 1–4: Ordinary least-squares difference-in-differences regression using micro data, including controls for religiosity and parents' origin. Robust standard errors clustered at the gender × year level in parentheses; robust standard errors clustered at the gender × geography level in square brackets. Columns 5–6: Generalized least squares regression with data collapsed to the gender-year-of-birth level. Robust standard errors that allow for cross-sectional correlation and for panel-specific serial correlation (i.e. estimate a unique autocorrelation parameter for each group), in parentheses. Data from the 2008 Israeli population census, restricted to Israeli-born Jews.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

if we are timing the break at the same cohort as for college graduation (1971—in case there was a shock to this cohort's educational outcomes for non-IVF reasons) or if we use those who would be entering their junior year (and thus still able to change their high school completion decision) in 1994 (1978 cohort—in case there was a shock to all educational outcomes in 1994 not driven by the IVF policy change). Figure A3 confirms this result using administrative data on the number of high school students passing the matriculation exams ("Bagrut") by gender and birth year. While the number of students passing the test increases for both genders, the ratio of women to men remains stable for the Jewish population of Israel (while there is an increase in this ratio for the Arab population of Israel).

Graduate Education We next examine whether more women completed graduate education following the policy change. For this outcome measure, we again use the median age of students entering that educational level to guide us, therefore treating the 1966 cohort as the first treated year.[39] The raw data is shown in Figure 7, showing again a clear increase in women's completion relative to men starting at the cohorts who have not completed their educational decisions before they learn of expanded access to IVF. While women's educational outcomes remain on a moderately increasing relatively stable trend in the pre-period, for men we find inconsistently low rates of graduate education for the 1954 to 1958 cohorts.[40] Nevertheless, women demonstrate an upward shift at the 1966 cohort, while men seem to follow approximately the same pre-trend (if we disregard the aforementioned irregularity).

This is confirmed by the findings presented in Table 6. Columns 1–2 show estimates for the simple DID specification with gender-specific time trends. Women in the "treated" cohorts are significantly more likely to complete graduate degrees than before. The approximately 1.8 percentage point increase is a much larger effect than that for college education, as the baseline level of graduate education among women in the pre period is 11%. This effect thus represents a 16% increase in completion of graduate education. When allowing for a discontinuous slope change, the main effect is very similar and the slope change, while positive, is not significant. This is likely due to the fact that we have fewer post period cohorts (to avoid censoring, since graduate education may be completed later in life), in addition to the unstable trend exhibited for men in the pre period (as discussed above). Nevertheless, the positive effect remains stable when year-of-birth fixed effects are included, and the GLS specification confirms the magnitude of the results, although due to having very few observations (46), it is only marginally significant in column 5, and not significant in column 6.[41]

[39] The median age for second degree applicants in Israel is 28.2 for women and 29.7 for men.

[40] We can speculate that these cohorts may have been entering military service during the 1973 Yom Kippur war and the period of hostility that followed, which may have impacted their long-term educational attainment, but have not found any literature indicating a reason for this decline.

[41] Using the GLS model with a standard DID specification, as in columns (1) and (2), renders effects that are significant at the 5% level.

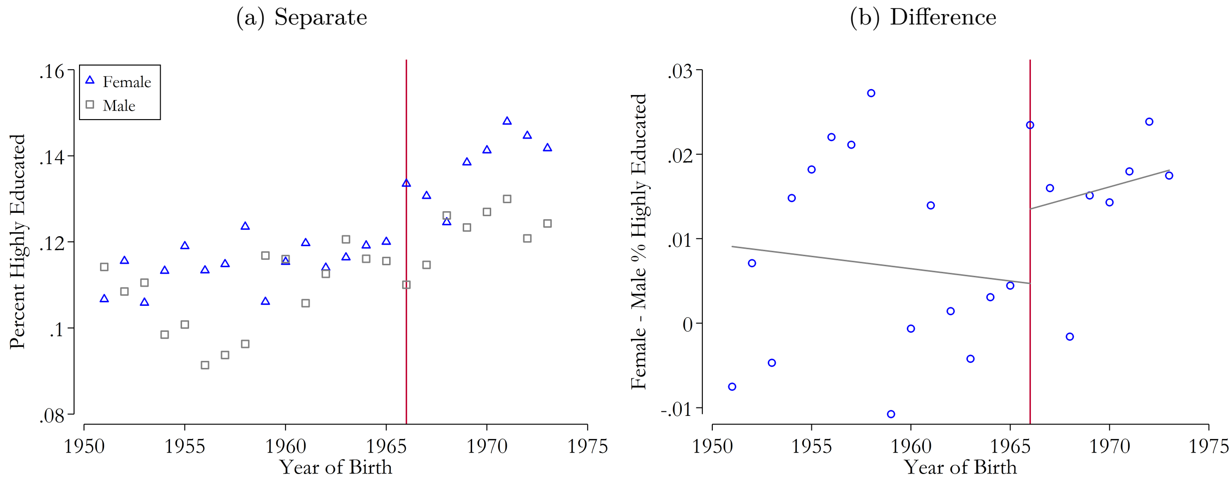
Table 6: Rates of Graduate Education

	Dependent variable: Graduate Education					
	DiD with GSTT		Slope-Change DiD		GLS Slope-Change DiD	
	(1)	(2)	(3)	(4)	(5)	(6)
fem × post	0.018	0.018	0.016	0.016	0.016	0.014
	(0.007)**	(0.005)***	(0.006)**	(0.005)***	(0.009)*	(0.009)
	[0.009]**	[0.009]**	[0.008]*	[0.008]*		
fem × post × time			0.001	0.001	0.001	0.001
			(0.001)	(0.001)	(0.002)	(0.002)
			[0.002]	[0.002]		
fem × time	-0.000	-0.000	-0.001	-0.001	-0.001	-0.001
	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
	[0.001]	[0.001]	[0.001]	[0.001]		
post × time			-0.003		0.001	
			(0.001)***		(0.001)	
			[0.002]**			
post	-0.014		-0.008		-0.003	
	(0.005)***		(0.005)*		(0.007)	
	[0.006]**		[0.006]			
female	0.000	0.000	-0.001	-0.001	-0.001	-0.000
	(0.005)	(0.003)	(0.004)	(0.003)	(0.007)	(0.006)
	[0.012]	[0.012]	[0.011]	[0.011]		
time	0.002		0.003		0.001	
	(0.000)***		(0.000)***		(0.001)*	
	[0.001]***		[0.001]***			
Constant	0.218	0.190	0.223	0.212	0.117	0.104
	(0.005)***	(0.004)***	(0.004)***	(0.004)***	(0.005)***	(0.003)***
	[0.015]***	[0.013]***	[0.014]***	[0.016]***		
YOB FEs		YES		YES		YES
Observations	138953	138953	138953	138953	46	46
R-Squared	0.0463	0.0465	0.0464	0.0465		

Notes: Columns 1–4: Ordinary least-squares difference-in-differences regression using micro data, including controls for religiosity and parents' origin. Robust standard errors clustered at the gender × year level in parentheses; robust standard errors clustered at the gender × geography level in square brackets. Columns 5–6: Generalized least squares regression with data collapsed to the gender-year-of-birth level. Robust standard errors that allow for cross-sectional correlation and for panel-specific serial correlation (i.e. estimate a unique autocorrelation parameter for each group), in parentheses. Data from the 2008 Israeli population census, restricted to Israeli-born Jews.

*** p<0.01, ** p<0.05, * p<0.1

Figure 7: Female vs. Male Graduate Education



Notes: Figure (a) shows average graduate school attainment for women and men by birth cohort. Figure (b) presents the difference in graduate school attainment between women and men, as well as fitted lines for the pre (1951-1965 birth cohorts) and post (1966-1973 birth cohorts) periods. Data from the 2008 Israeli population census, restricted to Israeli-born Jews.

We also look at whether rates of women gaining graduate education *conditional* on obtaining college education have increased, in table 7. In doing this, we seek to understand whether graduate education has increased as a natural consequence of the increase in college education, or whether there has been an increase in graduate education over and above the mechanical impact of increasing the pool of college graduates. Because we are conditioning on an endogenous variable, as college education is also impacted by the treatment, the magnitude of this effect should be interpreted with caution. Nonetheless, the fact that we see a positive and significant increase indicates the change in graduate education is not merely driven by a greater pipeline of college graduates. The larger results for graduate education, and the further increase conditional on college education, supports our main hypothesis that extended later-life fertility for women drives the observed shifts, since decisions on graduate education are made at an older age when expected fertility plays a much more important role.

Repeated-cross-section data One may be concerned about data censoring in the educational outcomes, since we use data collected in a single year and therefore compare individuals of different ages over time. To minimize this problem, the youngest cohort we use in our estimation is 31 at the census year, as discussed in section 3.2. In this section, we use a different data set comprised of annually repeated cross-sections, the Israeli Labor Force Survey (LFS), to verify that censoring is not what drives the result. This data allows us to compare individuals of the same age before and after the policy change. Because of the limited years of availability, and small sample size, we

Table 7: Conditional Rates of Graduate Education

	Dependent variable: Graduate Education | College					
	DiD with GSTT		Slope-Change DiD		GLS Slope-Change DiD	
	(1)	(2)	(3)	(4)	(5)	(6)
fem × post	0.062	0.062	0.056	0.056	0.050	0.049
	(0.021)***	(0.014)***	(0.018)***	(0.015)***	(0.026)*	(0.027)*
	[0.022]***	[0.021]***	[0.022]**	[0.021]***		
fem × post × time			0.003	0.003	0.003	0.003
			(0.003)	(0.002)	(0.005)	(0.005)
			[0.004]	[0.004]		
fem × time	-0.005	-0.004	-0.005	-0.005	-0.005	-0.005
	(0.001)***	(0.001)***	(0.001)***	(0.001)***	(0.002)**	(0.002)**
	[0.002]***	[0.002]***	[0.002]***	[0.002]***		
post × time			-0.009		-0.006	
			(0.002)***		(0.004)	
			[0.004]**			
post	-0.054		-0.039		-0.031	
	(0.017)***		(0.015)**		(0.022)	
	[0.017]***		[0.016]**			
female	-0.055	-0.055	-0.059	-0.060	-0.060	-0.059
	(0.015)***	(0.010)***	(0.012)***	(0.010)***	(0.018)***	(0.018)***
	[0.014]***	[0.014]***	[0.014]***	[0.014]***		
time	0.001		0.003		0.002	
	(0.001)		(0.001)**		(0.002)	
	[0.001]		[0.002]**			
Constant	0.463	0.431	0.476	0.469	0.417	0.405
	(0.014)***	(0.008)***	(0.012)***	(0.006)***	(0.015)***	(0.009)***
	[0.012]***	[0.023]***	[0.012]***	[0.013]***		
YOB FEs		YES		YES		YES
Observations	45609	45609	45609	45609	46	46
R-Squared	0.0128	0.0136	0.0132	0.0136		

Notes: Columns 1–4: Ordinary least-squares difference-in-differences regression using micro data, including controls for religiosity and parents' origin. Robust standard errors clustered at the gender × year level in parentheses; robust standard errors clustered at the gender × geography level in square brackets. Columns 5–6: Generalized least squares regression with data collapsed to the gender-year-of-birth level. Robust standard errors that allow for cross-sectional correlation and for panel-specific serial correlation (i.e. estimate a unique autocorrelation parameter for each group), in parentheses. Data from the 2008 Israeli population census, restricted to Israeli-born Jews.

*** p<0.01, ** p<0.05, * p<0.1

do not use this data as our main source.[42]

We use three-year age groups to increase the number of observations per year and decrease variation (although the sample remains quite small). We choose ages to be high enough so we can be confident that there is minimal censoring due to ongoing education, but also young enough that we can follow what happens to this age group for several years after 1994.[43]

Figure 8 presents levels of education, for women and men of the same age range, by their year of birth. Figure (a) shows percentage of college graduates in each cohort, separately for men and for women. We clearly see that while men stay on the same moderately increasing time trend, women's rate of college completion sharply increases for the cohorts most affected by the policy change, although the largest change here appears for the 1970 cohort.[44] The same analysis is presented in figure (b) for graduate level education. Interestingly, we observe two "jumps" for women, the first for the 1966 cohort and the second for the 1971 cohort. It seems reasonable that the first increase is driven by women who had already completed college when the policy was introduced and due to the policy faced a decreased cost of attending graduate education. The second increase correlates with the increase in college attainment and is at least partially driven by the higher rates of women who are college graduates and can actually consider post college education. These findings refute the possibility that the results we presented above are the result of data censoring differentially impacting men's outcomes.

Moreover, this data set is sampled and assembled completely differently from our principal data, and yet shows remarkably similar results. This provides additional evidence that there was a differential increase in women's investment in higher education, starting with the cohorts who had the opportunity to enter either college or graduate school in 1994.

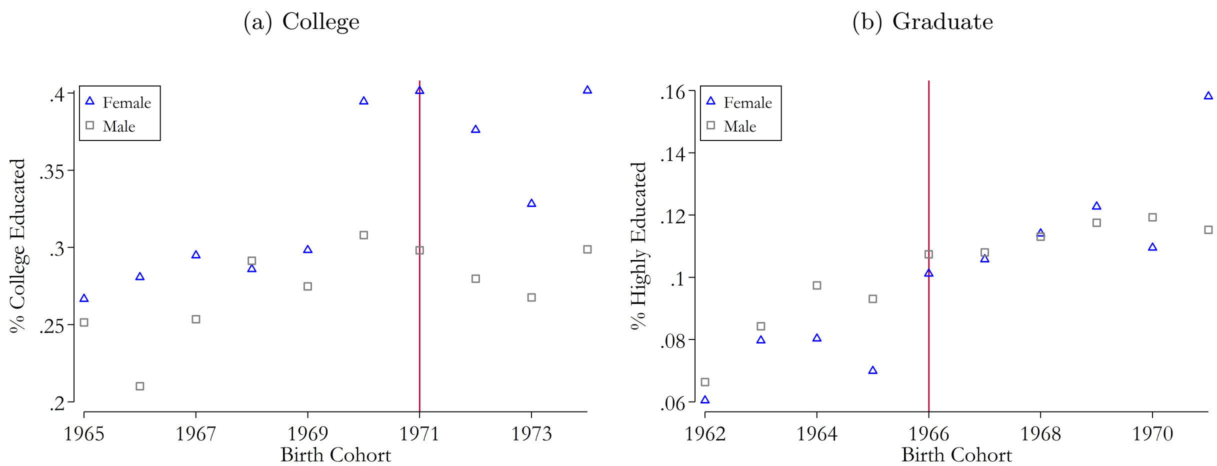
Alternative control: Arab-Israeli women Our identification strategy relies on the post-1994 time-path of men's outcomes being similar to women's in the absence of the IVF policy change (once pre-trends and level effects have been controlled for). A threat to this identification would be a policy, or any other exogenous shock, that affected Israeli women, but not men, commencing at or around the time of the 1994 IVF policy change. The QLR analysis above, as well as the event studies presented in Section 4.4, show that such a change would need to be very precisely timed to coincide with the IVF policy change in order to produce similar results. One specific concern is a reform in higher-education, that rolled out during the eighties and ninties. The reform, which obviously had an impact on Israelis' college graduation rates, could potentially have a differential

[42] We utilize The Israeli Annual Labor Force Survey for the years 2001 to 2011. The sample is representative of the population and follows the evolution of the labor force in Israel at the household level.

[43] It should be noted that our results are not sensitive to this choice and the same pattern appears for a variety of age ranges.

[44] The largest effect in the Census data appears at the 1971 cohort, however, these data are collected using a different sample and methods than the Census, and we expect a "fuzzy" cutoff for treatment status for educational outcomes.

Figure 8: Female vs. Male College Education, Labor Force Survey



Notes: Figure (a) presents the fraction of men and women with college education by birth cohort. Figure (b) presents the fraction of men and women with graduate education by birth cohort. For college, all individuals are between age 35 and 37. For graduate education, all individuals are between age 38 and 40. The age ranges were chosen to avoid censoring while focusing on the cohorts which are most relevant to the policy change. Data from the Labor Force Survey 2001-2011, restricted to Israeli-born Jews.

effect on women.

We explore the possibility that the higher education reform differentially affected women using the Arab population of Israel, which would have been unlikely to respond substantially to the IVF reform, but *was* targeted by the education expansion. Arab-Israelis were much more likely than Jewish-Israelis to be affected by the higher education reform due to lower high-school achievements (on average) and higher concentration in peripheral areas.[45] On the other hand, as we mention in section 3.2, Arab-Israelis were less likely (if at all) to respond to the change in IVF funding for three main reasons. First, most Arab-Israelis are Muslim and Islam places more stringent restrictions on the use of IVF than does Judaism.[46] Second, Arab-Israeli women were much less likely to be on the margin of large career investments in the 1990s, as average educational levels were substantially lower than in the Jewish population. At our baseline year 1993, there is a 25 percentage point difference in the rate of women's college attainment between Jews and Arabs, and a 10 percentage point difference in the same figures for graduate education (Arab women's attainment in graduate

[45] This effect is described in Volanski (2005) and also in various reports issued by the Israeli council for higher education (e.g. Higher education in Israel 2014, pp. 29-31).

[46] For example, Islam prohibits the use of donor eggs or sperm, the former being extremely important and even crucial for women in their forties, and requires significant precautions to guarantee the egg and sperm come from the correct mother and father and cannot have been contaminated. In addition, the Israeli Jewish religious leadership very quickly addressed the innovative IVF technology and approved usage with practically no limitations, whereas other religions took longer to respond.

education was actually very close to zero at that time). Third, since the Arab population is not subject to obligatory military service, they tend to make decisions about career investment when they are 2-3 years younger, making fertility considerations less relevant for educational decisions.[47]

We thus re-estimate the impact of the policy change on Jewish women's education using Arab women, rather than Jewish men, as the comparison group. This complementary strategy allows us to difference out the impact of being female following the policy change. Since Arab-Israeli women would have been targeted by, and affected by, any government program that has the potential to increase women's education, this helps rule out that other policies, or any other unspecified gender-specific social trends (e.g., "women's liberation"), may have been responsible for the observed changes. At the same time, the Arab population may also be affected by events that are specific to their community and irrelevant to the Jewish population in Israel, and therefore this strategy would certainly be imperfect on its own. However, any bias caused by factors specific to the Arab population would be uncorrelated with any bias caused by features that relate to gender. Thus, finding a consistent differential effect on Jewish women using both control groups provides a compelling piece of evidence that this effect is caused by the increased access to IVF.

To capture any effects that could have possibly impacted women entering college in 1994, we need to use the birth cohort of Arab women that would be entering college at the same time as Jewish women. Thus, in our figures and the following regression analysis we compare Jewish women to three-years-younger Arab women and align the affected cohorts for the two groups to match college entry at 1994.[48] All of the results hold and are qualitatively similar when we do not adjust for this difference and conduct the analysis using year of birth.

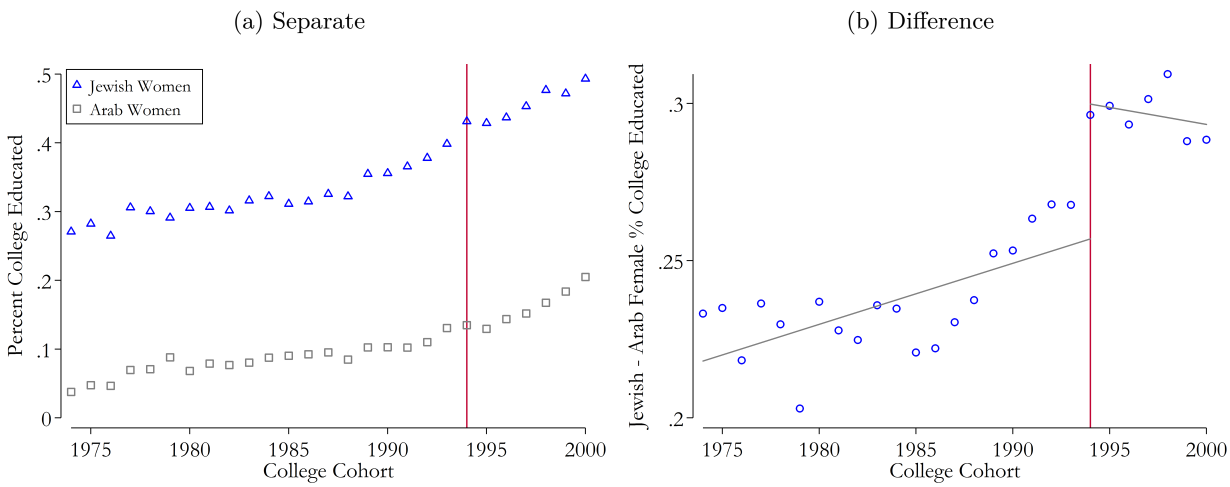
Figures 9 and 10 show the difference between Jewish and Arab women using raw data on college and graduate education completion. These figures also show these outcomes separately by population group. Although the pre-trends are not parallel, as with the male control group, a similar increase in both types of education is clearly observed at 1994 *only* for Jewish women. Tables 8 and 9 confirm that the main findings for education hold using the alternate Arab-female control group. The magnitude of the coefficients is quite similar to the results for the male control group (3.1 percentage points in the college DID specification here, versus 2.5 in the male control; 2.2 percentage points in the graduate education DID, versus 1.8 in the male control.). The slight negative trend seen in columns 3 and 4 is most likely because the Jewish-Israeli population take longer to complete their degrees and have a higher variation in age at college entry, but could also

[47]In addition, the Arab population has a much lower average age at first marriage for women, a higher average birth rate, and a much lower labor force participation rate for women. In our baseline year 1993, for example, Arab women marry 2.5 years earlier than Jewish women. Labor force participation rates in the early-mid 90s was approximately 13% for Arab-Israeli Muslim women compared to around 55% for Jewish-Israeli women (based on data from the Labor Force Surveys and macro data reported by the Bank of Israel.

[48]Military service for Jewish women is two years long, but macro data shows a three year difference in the median age of college applicants between the two populations, most likely due to waiting periods before and after entering military service.

be due to more recent increases in education among the Arab population, from policies designed to encourage education in under-served areas. However, the *Jewish* × *post* coefficient is positive and significant in all specifications and the overall effect remains positive throughout the 'post' period".

Figure 9: College Education for Arab vs. Jewish Women



Notes: The figure compares college completion rates between Jewish and Arab women over time. The cohorts are aligned based on anticipated year of college entry, since Arab women do not serve in the military, and tend to enter college three years younger than Jewish women (as calculated in 1995 census data, and reported by the Israeli Central Bureau of Statistics). Therefore, the first "treated" cohort is the 1971 cohort for Jewish women, and the 1974 cohort for Arab women. Figure (a) shows average college attainment by population group. Figure (b) presents the difference in college attainment between Jewish and Arab women, as well as fitted lines for the pre and post periods. Data from the 2008 Israeli population census.

Although we use the Arab control principally to address educational results, Table A2 shows the results for age at first marriage, where we see that Jewish-Israeli women experience a differential increase in age at first marriage, beginning in 1994, compared to Arab-Israeli women.[49] Note, that the estimated effect is considerably larger then the one reported in Table 3, which may be because Arab women are not affected in equilibrium by Jewish women postponing marriages, while Jewish men, the marriage partners of Jewish women, are likely to be.

Combining the results in this section with the ones in our main specifications establishes that the most likely cause for the observed change in Jewish women's educational and marriage choices is the increased access to IVF. Any other explanation would have to induce both a gender and religion-divergent impact precisely in 1994, in addition to affecting all of the outcomes we consider.

Women's Labor Market Outcomes To supplement our findings regarding educational outcomes, we examine the impact of the policy change on treated cohorts' full-time employment,

[49]For spousal income, we did not use men as the principal control group, but rather women who married younger, and thus there is no scope for the alternate Arab control.

Table 8: College Graduation Rates (Arab Control)

	Dependent variable: College Education					
	DiD with GSTT		Slope-Change DiD		GLS Slope-Change DiD	
	(1)	(2)	(3)	(4)	(5)	(6)
jewish × post	0.031	0.030	0.043	0.042	0.038	0.042
	(0.013)**	(0.007)***	(0.012)***	(0.005)***	(0.011)***	(0.010)***
	[0.013]**	[0.013]**	[0.015]***	[0.015]***		
jewish × post × time			-0.005	-0.005	-0.005	-0.005
			(0.002)**	(0.001)***	(0.003)*	(0.002)**
			[0.004]	[0.004]		
jewish × time	0.002	0.002	0.002	0.003	0.003	0.002
	(0.001)***	(0.000)***	(0.001)***	(0.000)***	(0.001)***	(0.001)***
	[0.001]**	[0.001]**	[0.001]**	[0.001]**		
post × time			0.009		0.010	
			(0.001)***		(0.002)***	
			[0.002]***			
post	0.026		0.004		0.003	
	(0.009)***		(0.006)		(0.007)	
	[0.007]***		[0.010]			
jewish	0.243	0.243	0.246	0.246	0.248	0.246
	(0.009)***	(0.005)***	(0.009)***	(0.004)***	(0.008)***	(0.006)***
	[0.027]***	[0.027]***	[0.028]***	[0.028]***		
time	0.004		0.004		0.004	
	(0.000)***		(0.000)***		(0.000)***	
	[0.000]***		[0.001]***			
Constant	0.134	0.058	0.129	0.142	0.129	0.060
	(0.005)***	(0.005)***	(0.005)***	(0.003)***	(0.004)***	(0.006)***
	[0.013]***	[0.011]***	[0.013]***	[0.020]***		
YOB FEs		YES		YES		YES
Observations	125229	125229	125229	125229	54	54
R-Squared	0.0742	0.0748	0.0744	0.0748		

Notes: Columns 1–4: Ordinary least-squares difference-in-differences regression using micro data (no controls included since religiosity and parents' origin controls used only apply to Jewish population). Robust standard errors clustered at the group × cohort level in parentheses; robust standard errors clustered at the group × geography level in square brackets. Columns 5–6: Generalized least squares regression with data collapsed to the population group-college cohort level. Robust standard errors that allow for cross-sectional correlation and for panel-specific serial correlation (i.e. estimate a unique autocorrelation parameter for each group), in parentheses. Data from the 2008 Israeli population census, restricted to Israeli-born.

*** p<0.01, ** p<0.05, * p<0.1

Table 9: Graduate Education Rates (Arab Control Group)

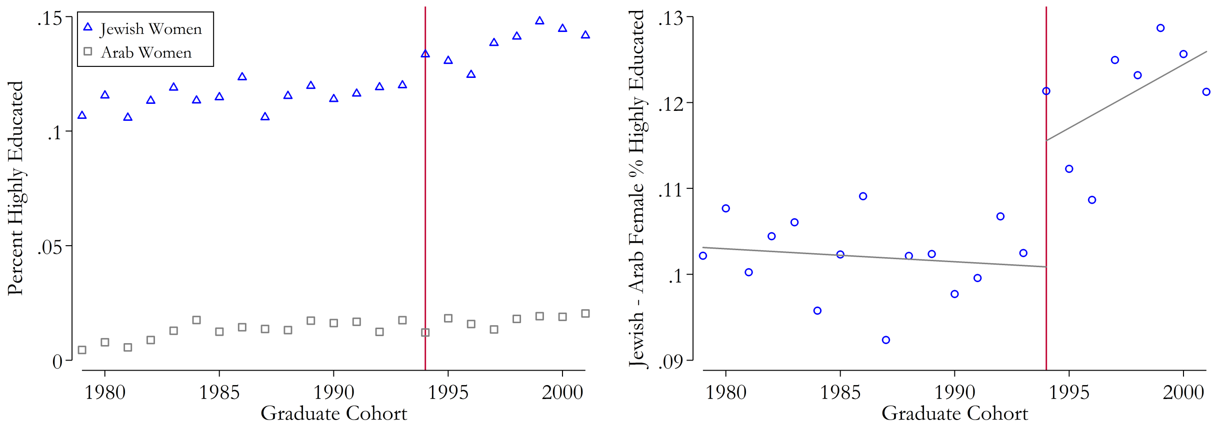
	Dependent variable: Graduate Education					
	DiD with GSTT		Slope-Change DiD		GLS Slope-Change DiD	
	(1)	(2)	(3)	(4)	(5)	(6)
jewish × post	0.022	0.021	0.018	0.018	0.018	0.019
	(0.005)***	(0.003)***	(0.005)***	(0.003)***	(0.004)***	(0.004)***
	[0.007]***	[0.007]***	[0.007]**	[0.007]**		
jewish × post × time			0.002	0.002	0.002	0.002
			(0.001)*	(0.001)**	(0.001)***	(0.001)**
			[0.002]	[0.002]		
jewish × time	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)**	(0.000)	(0.000)
	[0.001]	[0.001]	[0.001]	[0.001]		
post × time			0.000		0.000	
			(0.001)		(0.001)	
			[0.001]			
post	-0.005		-0.006		-0.006	
	(0.003)*		(0.003)*		(0.003)*	
	[0.004]		[0.004]			
jewish	0.096	0.096	0.094	0.094	0.094	0.094
	(0.003)***	(0.002)***	(0.003)***	(0.002)***	(0.003)***	(0.003)***
	[0.010]***	[0.010]***	[0.010]***	[0.010]***		
time	0.001		0.001		0.001	
	(0.000)***		(0.000)***		(0.000)***	
	[0.000]***		[0.000]***			
Constant	0.022	0.020	0.022	0.020	0.022	0.004
	(0.002)***	(0.001)***	(0.002)***	(0.003)***	(0.002)***	(0.002)
	[0.003]***	[0.005]***	[0.003]***	[0.005]***		
YOB FEs		YES		YES		YES
Observations	100724	100724	100724	100724	46	46
R-Squared	0.0267	0.0269	0.0267	0.0269		

Notes: Columns 1–4: Ordinary least-squares difference-in-differences regression using micro data (no controls included since religiosity and parents' origin controls used only apply to Jewish population). Robust standard errors clustered at the group × cohort level in parentheses; robust standard errors clustered at the group × geography level in square brackets. Columns 5–6: Generalized least squares regression with data collapsed to the population group-graduate cohort level. Robust standard errors that allow for cross-sectional correlation and for panel-specific serial correlation (i.e. estimate a unique autocorrelation parameter for each group), in parentheses. Data from the 2008 Israeli population census, restricted to Israeli-born.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Figure 10: Graduate Education for Arab vs. Jewish Women

(a) Separate (b) Difference

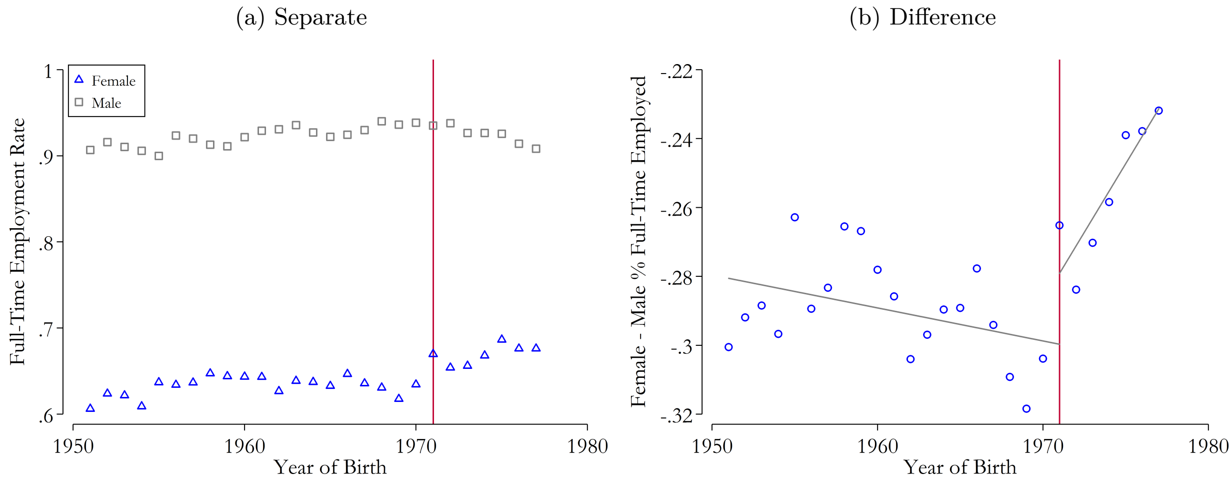


Notes: The figure compares graduate education completion rates between Jewish and Arab women over time. The cohorts are aligned based on anticipated year of graduate school entry, since Arab women do not serve in the military, and tend to enter college three years younger than Jewish women (as calculated in 1995 census data, and reported by the Israeli Central Bureau of Statistics). Therefore, the first "treated" cohort is the 1966 cohort for Jewish women, and the 1969 cohort for Arab women. Figure (a) shows average graduate school attainment by population group. Figure (b) presents the difference in graduate school attainment between Jewish and Arab women, as well as fitted lines for the pre and post periods. Data from the 2008 Israeli population census.

income, and participation in "prestigious" (high-investment) occupations, using the same cohorts and treatment definition as for college graduation. Figure 11 plots women's rate of full-time employment, relative to men's, and shows a clear upwards shift for women of the same cohorts whose higher education levels increased following the policy change. This supports our hypothesis that the extended horizons offered to women by IVF expansion enabled women to enter more demanding career paths and shift from part-time employment (characteristic of women with family constraints) to full-time employment. These findings are corroborated by an increase in female earnings relative to male earnings, presented in figure A4. Moreover, there is a sharp differential increase in the percentage of women in prestigious occupations requiring high human capital investments, such as management, engineering, law, medicine, and academia, presented in figure A5. These occupations are notable for all requiring significant on-the-job training and long work hours, in addition to educational preparation. If we break down these occupations, the positive shift appears to be driven especially by lawyers and managers, which were found by Cortes and Pan (2017) to be the occupations with the highest returns to working long hours.

Table 10 presents regression results for these three outcomes, using a DID specification that allows for a differential slope change in the post period, with and without year of birth fixed effects. Columns 1 and 2 show a significant increase of 3 percentage points in women's full-time employment,

Figure 11: Female vs. Male Full-Time Employment



Notes: Figure (a) shows rates of full-time employment out of the employed population for women and men by birth cohort. Figure (b) presents the difference in these rates between women and men, as well as fitted lines for the pre (1951-1970 birth cohorts) and post (1971-1977 birth cohorts) periods (cohorts correspond to our analysis of college attainment). Data from the 2008 Israeli population census, restricted to Israeli-born Jews.

on an initial level of approximately 60%, almost 30 percentage points less than men. In addition, full-time employment kept increasing for younger cohorts, by roughly 1 percentage point per year amounting to a total of 10 percentage points for the last cohort in our analysis. For income and participation in prestigious occupations, we observe a positive and significant differential change in trend (for prestigious occupations it is only significant when fixed effects are added in column 6). The increase in income, listed in Shekels, represents a gain of more than $1,000 per year for women's income relative to men. Based on the coefficient on "female," the initial gender wage gap is around $20,000. Thus, in the seven years following the policy change, almost a third of this gap was eliminated. This reduction in the pay gap is much larger (in proportion) than the one we observe for full-time employment, suggesting that women's wages increased, regardless of working hours.[50] One possible channel for this is the concordant increased participation in prestigious, high-paying occupations.

These results show that the cohorts affected by the IVF policy change not only got more education, but they made use of that education to participate more in the labor force at a higher level, thus bolstering our hypothesis that the overall treatment effects are driven by a desire for greater career investment. The educational effects of the policy translated into better career outcomes, rather than women achieving high-powered education, but then dropping out of professional occupations, as in Bertrand, Goldin and Katz (2010). This suggests that differential fertility time

[50]Estimating the same specifications for women's income conditional on full-time employment yields positive and significant results that support this conclusion.

horizons can meaningfully contribute to disparities in labor market outcomes.

Table 10: Women's Labor Market Outcomes

Dependent Variable:	Full-Time Employment		Annual Income		Prestigious Occupation	
	(1)	(2)	(3)	(4)	(5)	(6)
fem × post	0.032	0.032	804.5	868.9	0.001	0.001
	(0.012)***	(0.009)***	(3452.5)	(2017.8)	(0.008)	(0.004)
	[0.008]***	[0.008]***	[3615.2]	[3585.4]	[0.012]	[0.012]
fem × post × time	0.010	0.010	5443.2	5432.9	0.002	0.002
	(0.002)***	(0.002)***	(450.0)***	(314.2)***	(0.001)	(0.001)***
	[0.002]***	[0.002]***	[927.4]***	[927.0]***	[0.003]	[0.003]
fem × time	-0.001	-0.001	560.459	563.438	0.001	0.001
	(0.001)*	(0.000)**	(257.840)**	(130.394)***	(0.001)	(0.000)**
	[0.001]	[0.001]	[303.143]*	[301.043]*	[0.001]	[0.001]
post × time	-0.006		-8712.434		-0.010	
	(0.001)***		(368.996)***		(0.001)***	
	[0.001]***		[930.193]***		[0.002]***	
post	-0.002		-4934.203		0.002	
	(0.003)		(2756.126)*		(0.004)	
	[0.004]		[3042.923]		[0.009]	
female	-0.297	-0.297	-79535.047	-79557.940	-0.068	-0.068
	(0.006)***	(0.005)***	(3012.548)***	(1468.201)***	(0.006)***	(0.004)***
	[0.011]***	[0.011]***	[6956.278]***	[6952.937]***	[0.017]***	[0.017]***
time	0.002		-1081.671		0.002	
	(0.000)***		(192.965)***		(0.000)***	
	[0.000]***		[262.103]***		[0.001]***	
Constant	0.948	0.901	202462.2	221179.2	0.427	0.384
	(0.003)***	(0.009)***	(3388.6)***	(3681.9)***	(0.005)***	(0.005)***
	[0.006]***	[0.007]***	[8658.3]***	[11145.1]***	[0.016]***	[0.022]***
YOB FEs		YES		YES		YES
Age Poly.			YES			
Observations	133206	133206	153522	153522	147231	147231
R-Squared	0.135	0.135	0.123	0.124	0.0480	0.0482

Notes: All columns: Ordinary least-squares difference-in-differences regression using micro data, including controls for religiosity and parents' origin. Robust standard errors clustered at the age group × year level in parentheses; robust standard errors clustered at the age group × geography level in square brackets. Columns 3 and 4 control for a flexible polynomial in age. Data from the 2008 Israeli population census, restricted to Israeli-born Jews.

*** p<0.01, ** p<0.05, * p<0.1

4.3 Marriage Market Equilibrium

The additional reproductive years afforded by access to assisted reproduction technologies may have impacted not only women's decisions, but also men's marriage choices. Low (2017) shows that men respond to prospective mates' expected fertility when choosing a partner, trading off between so-called "reproductive capital" and more traditional human capital traits like income and education.

As a result, women who are high-earning, but older, may marry poorer men than lower-earning, but younger, women. These marriage market consequences to investment are crucially important because they may in turn deter women's investments, by adding an additional cost to women's own utility impact of foregone fertility.

Fertility extending technology thus has the potential to blunt both the individual impacts and the marriage market effects of increased investment. Because the increase in access to IVF technology lessens the perceived fertility cost of waiting to marry, "high-quality" men may have been more willing to marry older women following the policy change. If this is the case, we can expect equilibrium matching to adjust so that these women will match with higher quality partners. This equilibrium shift would be accelerated by women also changing their marriage timing, creating a thicker marriage market for older women.

We test this by examining the spousal quality of women who marry older versus younger before and after the policy change. If women's reproductive fitness is taken into account by men, we would expect the "spousal quality penalty" to older women to lessen once access to IVF expands.[51] We use administrative data on income, provided by the tax authority and paired with the Census's survey data. Women are matched to their spouses in the data based on relationship to household head. We only have spousal income data for the current spouse, so this analysis excludes women who are divorced or widowed before the census year.

We compare spousal quality, measured in a variety of ways, for married women who were between 25 and 29 at the time of marriage versus women who were between 30 and 34 at the time of marriage. Restricting to women who were between 25 and 34 at marriage ensures that we compare relatively similar groups of women. At the same time, placing the cutoff at the age of 30 helps us identify the response to the change in expected fecundity for older women, which becomes significantly more relevant when a woman hits her thirties.

Our main proxy for spousal quality is husband's income, as it is well established that income is an important quality that male spouses bring to the relationship (see, for example, Fisman et al. (2006)). To control for age effects, since we use cross-sectional data, and women who marry at different ages may have spouses of different ages as well, we include a flexible polynomial in age as a control. Because for very young or very old spouses income may provide a distorted measure of quality, even with age controls, as they may still be completing school and career training, or already retired, we restrict our analysis to the marriage years 1986 to 2003, which in turn restricts the age range of spouses. This year range was chosen based on an analysis of men's rate of full-time employment over age and the distribution of average husband's age by wife's age. Our qualitative results are not sensitive to the exact year-range chosen. In the regression results, we present results

[51] Note that although the selection into who marries older versus younger is changing, Low (2017) shows it would be lower human capital women who are now willing to delay marriage. Despite that, in equilibrium we would expect on average an improvement in older women's marriage outcomes as reproductive time horizons are extended. We account for selection by directly controlling for wives' characteristics, and showing our results are unchanged.

with additional controls as well as alternate quality measures, such as education and occupational choice.

Figure 12 shows how our the residual of spousal income regressed on a flexible age polynomial evolves over time, by year of marriage. Starting in 1996, two years after the policy change was announced, we observe a distinct shift upward in spousal income for women who marry over 30, relative to the same measure for women who marry younger, which remains very stable over time. The fact that this break appears with a lag may be attributed to men's perceptions taking longer to update or simply to the fact that reaching this new matching-equilibrium takes time.

Figure 12: Spousal Income for Older vs. Younger Women, 1986 - 2003

Notes: This figure compares current spousal income for women who married at age 30-34 to women who married at age 25-29. In order to control for age effects, the measure of spousal income used is the residual from a regression of income on a flexible polynomial in age. We use a narrower time range in these graphs to prevent censoring from individuals either still being students, for younger cohorts, or entering retirement, for older cohorts. Figure (a) shows average income-age-residuals separately by age group, by year of marriage. Figure (b) presents the differences in spousal income-age-residual, as well as fitted lines for the pre (1986-1993) and post (1994-2003) periods. Data from the 2008 Israeli population census, restricted to Israeli-born Jews.

Table 11 presents the results from a regression of spousal income on marrying older, before and after the policy change, controlling for spousal age with a flexible polynomial (for the GLS specification, we first regress income on the flexible polynomial in age, then take the residuals, then collapse to cells). The results confirm that there is a significant "penalty" in terms of spousal income for women who choose to marry over thirty, and that this penalty significantly decreases in the post period. Depending on the specification, we observe a reduction of between 60 and 100 percent in the "older marriage penalty" for women. This change is significant with both types of clustering, and in the GLS specification. Note that the magnitude of the increase in spousal income is very similar throughout the specifications, however the estimated initial penalty for marrying older is larger when we use GLS. As previously mentioned, we do not include a specification that allows for a change in slope for this outcome measure, due to the narrow band of years and the lagged response.

Table 11: Spousal Income

	Dependent variable: Spousal Income			
	DiD		GLS DiD	
	(1)	(2)	(3)	(4)
older × post	19316.144	20697.644	17540.641	16639.053
	(7616.659)**	(6392.176)***	(4907.371)***	(5191.417)***
	[7862.496]**	[8196.525]**		
post	1564.801		6849.849	
	(6146.128)		(6418.767)	
	[6367.385]			
married older	-19882.047	-20813.945	-29654.893	-29069.648
	(6133.012)***	(5872.047)***	(4131.902)***	(4377.880)***
	[13070.829]	[12923.828]		
time	-2786.126		-710.612	
	(766.338)***		(563.463)	
	[710.951]***			
Constant	-393935.497	-589998.553	48803.834	46272.107
	(939517.549)	(930193.480)	(3900.803)***	(991.128)***
	[835939.612]	[789565.836]		
YOM FEs		YES		YES
Observations	18543	18543	36	36
R-Squared	0.0518	0.0530		

Notes: All columns control for a flexible fifth-order polynomial in spousal age. Columns 1–2: Ordinary least-squares difference-in-differences regression using micro data, including controls for religiosity and parents' origin. Robust standard errors clustered at the age group × year level in parentheses; robust standard errors clustered at the age group × geography level in square brackets. Columns 3–4: Generalized least squares regression with data collapsed to the age group-year-of-marriage level. Robust standard errors that allow for cross-sectional correlation and for panel-specific serial correlation (i.e. estimate a unique autocorrelation parameter for each group), in parentheses. Data from the 2008 Israeli population census, restricted to Israeli-born Jews.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

In the appendix we use a triple-differences specification to compare the older-versus-younger results to those of *men's* spouses based on their age at marriage. In this specification, the spousal income for women who marry older versus younger, before and after the policy change are compared to the same metrics for older versus younger men before and after the policy change. The results in Table A3, show that not only do older women's marriage outcomes (in terms of spousal quality) improve relative to younger women's following the policy change, but they also improve relative to the change in older *men*'s spousal quality. However, because these effects will also reflect women's increased education and general equilibrium changes in matching patterns, the magnitude should be interpreted with caution.

To test the robustness of these findings, we include additional controls and use alternative

Table 12: Alternative Spousal Quality Specifications

Dep. Var.:	Income Wife Controls		Income Sp. YOB FEs		Income Pctile	
	(1)	(2)	(3)	(4)	(5)	(6)
older × post	18407.16	17905.36	20021.48	21651.21	0.026	0.026
	(9397.97)*	(8160.31)**	(8171.83)**	(6943.97)***	(0.011)**	(0.007)***
	[8359.51]**	[8408.73]**	[7675.44]**	[8034.79]***	[0.015]*	[0.015]*
married older	-24604.60	-26872.22	-20052.75	-21020.23	-0.051	-0.050
	(11038.23)**	(10227.43)**	(6380.26)***	(5965.34)***	(0.006)***	(0.005)***
	[14415.39]*	[14811.97]*	[13088.46]	[12978.08]	[0.016]***	[0.015]***
Constant	-167890.13	-520406.54	81715.40	94161.35	0.619	0.622
	(721687.21)	(756961.82)	(10384.74)***	(9848.97)***	(0.011)***	(0.010)***
	[744709.09]	[749225.14]	[15645.78]***	[18413.96]***	[0.012]***	[0.018]***
Age Poly.	YES	YES				
Wife Controls	YES	YES				
Sp. YOB FEs			YES	YES		
YOM FEs		YES		YES		YES
Observations	16693	16693	18543	18543	18543	18543
R-Squared	0.104	0.105	0.0547	0.0557	0.0496	0.0510

Dep. Var.:	College Education		Full-time Employment		Prestigious Occupation	
	(1)	(2)	(3)	(4)	(5)	(6)
older × post	0.039	0.039	0.045	0.043	0.053	0.053
	(0.021)*	(0.019)**	(0.018)**	(0.012)***	(0.018)***	(0.015)***
	[0.023]*	[0.022]*	[0.014]***	[0.014]***	[0.017]***	[0.017]***
married older	-0.051	-0.051	-0.052	-0.050	-0.042	-0.042
	(0.015)***	(0.015)***	(0.017)***	(0.010)***	(0.014)***	(0.012)***
	[0.027]*	[0.027]*	[0.013]***	[0.013]***	[0.027]	[0.027]
Constant	0.555	0.549	0.946	0.946	0.473	0.484
	(0.017)***	(0.014)***	(0.008)***	(0.006)***	(0.014)***	(0.013)***
	[0.024]***	[0.039]***	[0.008]***	[0.016]***	[0.022]***	[0.026]***
YOM FEs		YES		YES		YES
Observations	19511	19511	16730	16730	18178	18178
R-Squared	0.0782	0.0791	0.114	0.116	0.0389	0.0400

Notes: Ordinary least-squares difference-in-differences regression using micro data, including controls for religiosity and parents' origin. Robust standard errors clustered at the age group × year level in parentheses; robust standard errors clustered at the age group × geography level in square brackets. Top panel: Columns 1 and 2 control for wife's age, income and education. Columns 3–4 control for spouse's age through year-of-birth fixed effects. Columns 5–6 use spouse's income-for-age percentile as the outcome. Bottom panel: Columns 1–2 use spouse's college education as the outcome variable. Columns 3–4 use spouse's full-time employment, defined on entire male population. Columns 5–6 use spouse's participation in prestigious occupations such as management, law, medicine, engineering, and academia. All specifications include coefficients for post and time. Data from the 2008 Israeli population census, restricted to Israeli-born Jews.

*** p<0.01, ** p<0.05, * p<0.1

measures of spousal quality. First, in columns 1 and 2 of the top part of table 12, we repeat the spousal income regressions controlling for additional wife's characteristics, namely education, income and age. The purpose of this exercise is to ensure that the increase in spousal income is not a result of the improved quality of older brides themselves (given our results regarding women's increased education levels following the policy change). We find a slightly smaller reduction of the "older marriage penalty", mostly because the estimate for the initial penalty is about 20% larger than the one reported in columns 1 and 2 of table 11. Second, in columns 3 and 4 we control non-parametrically for spousal age by adding spouse's year of birth fixed effects (instead of the spousal age polynomial in our main specification). Our results remain significant and of the same magnitude. Then, in columns 5 and 6, we use spouse's income-for-age percentile, essentially a "ranking" of this spouse amongst his age cohort, which avoids income level comparisons across cohorts. The effect remains positive and significant using this measure.

In the bottom half of Table 12, we use three alternative measures of spousal quality that should be less sensitive to age effects: spouse's college education, full-time employment, and occupation (measured by participation in a "prestigious," high-paying field such as management, medicine, and law). All three of these alternative approaches to measuring spousal quality provide positive, significant results for the change in spousal quality for brides over thirty. In terms of magnitude, these estimates are on the high end of our main results, indicating that 80-100 percent of the "older marriage penalty" dissipated in the post period.

Together, these results provide strong evidence that spousal quality improved for older brides following the policy change. The stark change in the marriage market equilibrium following the significant change we observe in marriage timing and educational investment are consistent with women and men changing their beliefs regarding older women's fertility prospects. The changed beliefs could have stemmed from both the direct effects of the policy, greater access to IVF, and the broader awareness of fertility-extending technology brought on by the policy, as well as direct observation of the increase in older motherhood that occurred. The combination of results across different outcomes, women's marriage timing, education, career outcomes, and marital outcomes, strengthen our conclusion that the expansion of IVF was responsible for each individual change.

4.4 Robustness Checks and Alternative Explanations

In this section, we perform robustness checks of our results and examine some potential alternative explanations for our findings.

Permutation approach to significance levels In micro-data difference-in-differences studies, one may be concerned that a high degree of intra-group correlation or correlation across time periods is driving the significance of the results. We have addressed this issue by providing two alternative ways to cluster the standard errors, as well as GLS estimates that are collapsed to the year-group level and allow for correlation within groups and across years. As an additional check that our estimates are large relative to the true variation in the data, we also perform two types of permutation analyses on our coefficients for the DID specifications (since these have only a single coefficient of interest, unlike the regressions that allow for a slope change), presented in Figures A6 – A9.

For each outcome, we first perform a permutation test that respects the potential serial correlation in the data, by implementing a standard DID specification (like in column (1) of the main results tables) for only ten years of data, five years pre and five years post, with 1994 (or the corresponding school entry cohort) as the treatment year. We then compare this coefficient to the coefficient obtained from taking each possible sequential interval of ten years within our study period, and each corresponding false "treatment" year in the middle of the interval. This test does not yield a normal distribution of coefficients, as there are a limited number of ten-year intervals in the study period. Our true effect is larger than the effect of any other treatment year for age at first marriage and graduate education. For college education only one of the coefficients is larger (which amounts to approximately 6% of the values). For spousal income, we see higher estimated changes for two years following our "true" treatment year (1996 and 1997), which corresponds to the "delayed response" of the marriage market equilibrium (as discussed above and apparent in figure 12).

We then perform a more standard permutation test, where we randomly draw a number of years equal to our true number of treated years from the entire study period, and run our baseline regression with these randomly selected years as the "treatment" period (for an example of this approach, see Agarwal et al. (2014)). This approach does not respect the underlying serial correlation in the data, since the years are drawn randomly, but does account for intra-group correlation or other non-standard error structures. We perform 1,000 such random draws, and compare our true treatment coefficient to the resulting normal distributions. Our true effect is outside of the curve or in the far right tail, with less than 5% of the values above it, for every outcome measure.

Event Study analysis The next potential confounding factor we explore is that long term time trends may be responsible for the effects we see. This is already partially addressed by the inclusion of group-specific time trends in our regressions. However, to further address this possibility, we perform an event study analysis (also known as dynamic lag analysis), to pinpoint the timing of the changes we observe. We do this for our main outcome measures: college education, graduate education, age at first marriage, and spousal income for women who marry when older.

The event study graphs depicted in Figure 13 are created by regressing our key outcome variable on a series of dummies for each year, interacted with gender or age respectively (gender for educational outcomes, age-at-marriage for spousal income), and including the same controls as our regression specifications. Because *each* year of the entire sample period is allowed to have a group-specific effect, this is a demanding test. The coefficients graphed represent the effect of each time period, controlling for all other time periods. The coefficient on the lag just before the policy change is normalized to zero, so that subsequent effects show the relative difference in the affected group's outcomes compared to the period just before the policy change. Note that the time periods we are looking at are quite small relative to our other graphs, as we are zeroing in on four years before and six years after the policy change only.

Appendix Figure A10 shows an alternate format of event study, a "distributed lag" analysis, where coefficients for being born at or above a certain year (rather than *in* a certain year) are graphed, controlling for all other years. This, essentially, measures the permanent shock that occurs in each year to the outcome variables. This alternate analysis confirms the results presented here.

All event studies show that there are no significant pre trends driving our results, and that outcomes were relatively flat in the years immediately preceding the policy change. The event studies for age at first marriage, college education, and graduate education show that results become significant after the policy change, either immediately or over time. Figure 13, panel (a), shows that prior to the policy change the difference between men and women's age at first marriage was relatively constant, showing no apparent pre-trend. Then, in the year of the policy change, there is a large and permanent change to subsequent outcomes.

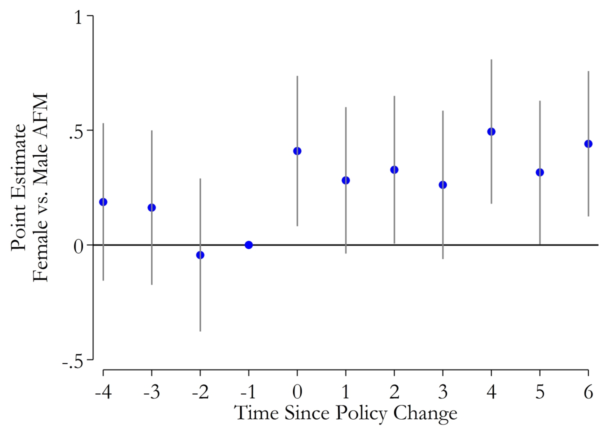
For the education event studies, we cannot expect that the impact will necessarily be isolated to the "first treated" cohort, since preceding cohorts may have been exposed (at least partially). Indeed, figures 13, panel (b) shows that the change appears to happen more gradually, only becoming significant at the end of the period. However, part of the reason for this less significant outcome is that the 1970 cohort, immediately preceding the first "treated" cohort of 1971, was also partially treated, as some of these individuals were still young enough to alter their college-going plans. This is confirmed by the distributed lags analysis in Figure A10, panel (b) where *both* the 1970 and the 1971 cohorts (years -1 and 0) appear to experience a permanent positive shock of approximately equal magnitude.

Figure 13, panel (c), shows that, in the case of graduate education, the effect is more clearly isolated in the "first treated" cohort, the 1966 birth cohort, remaining permanently high after that point.

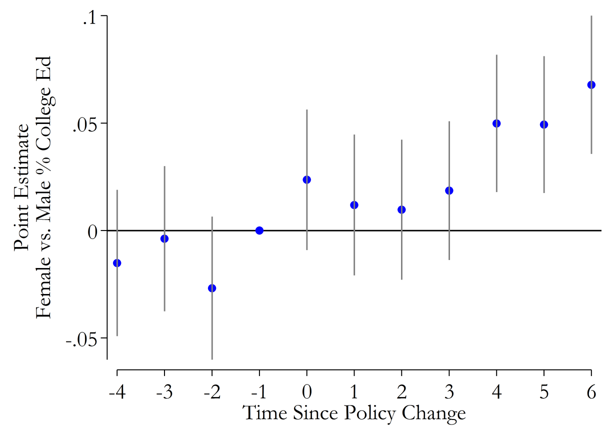
Figure 13, panel (d), shows an event study analysis of the change in spousal income for women who marry older versus younger. This graph shows no significant effect in the period following the policy change, although it does also affirm there do not appear to be significant pre-trends prior to 1994. The weaker results in this event study may be due to men's beliefs about the fertility of older partners taking longer to update, as well as the shifting of the marriage market to a new equilibrium taking time.

Figure 13: Event Study Analysis

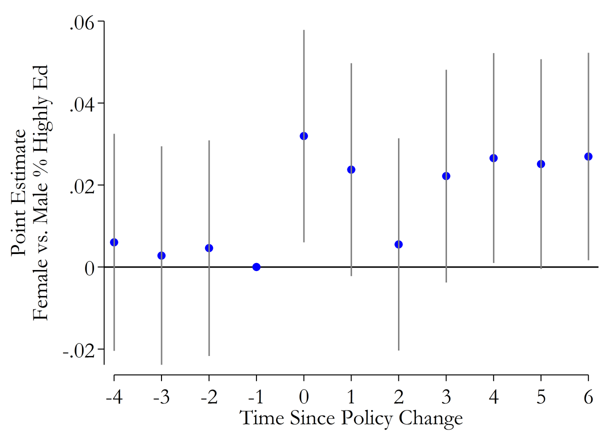
(a) Age at First Marriage



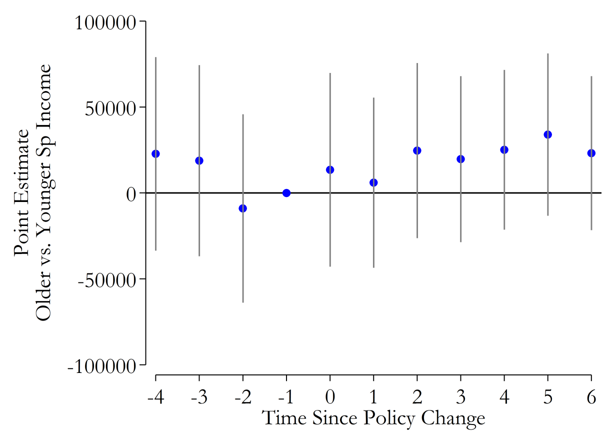
(b) College Education



(c) Graduate Education



(d) Spousal Income



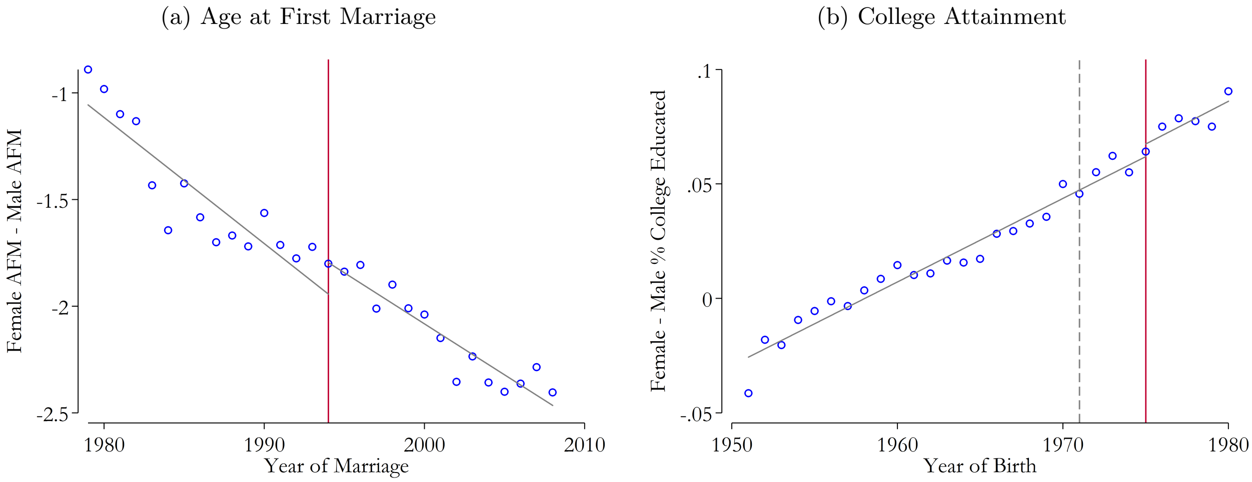
Notes: The figures present dynamic lag event study analyses for the principal regression specifications for our main outcome variables. Point estimates and confidence intervals are for the coefficients on yearly dummy variables interacted with dummy variables for the treated group. For age at first marriage, college education, and graduate education, the treated group is women, compared to men, while for spousal income it is women over 30 at first marriage, compared to those under 30. For age at first marriage and spousal income, time is year of marriage, with 1994 as time "zero," while it is birth cohort for college and graduate education. The regression equation includes all year-group interactions for the entire sample period, as well as a linear time trend and demographic controls. All regressions restricted to Israeli-born Jews.

Censoring or other data issues A different possibility is that the effect we observe is an artifact created by looking at outcomes retrospectively in cross-sectional data. We already minimize this concern as we carefully choose the years and cohorts which constitute the sample for each outcome, and re-examined our educational outcomes in repeated data, finding similar effects. However, to further verify that the retrospective nature of the analysis could not create similar breaks in the data *without* a real policy effect, we use the 1995 Israeli Population Census to conduct a placebo test. We replicate our analysis for a fake "policy change" in 1981 (14 years prior to the Census year, as the real 1994 policy change is 14 years before the 2008 Census) and find no evidence of a break in age at first marriage, college education, graduate education, or spousal income (age at first marriage, if anything, shifts in the opposite direction), as shown in Appendix Figure A11. There are no similar shifts in any of the outcomes in the 1995 Census as in our results using the true policy change.

A more specific concern in this context is the possibility that men's educational outcomes were more censored than women's outcomes as we look at years closer to the present day, since men are entering and completing college education later than women. It should be noted that we already presented one method which satisfyingly deals with this obstacle; in Appendix Table A1, we show that our results for college education hold when we define gradual exposure to treatment and allow men and women to differ in their level of exposure, according to the actual distribution of each group's age at college entry. Next, we further explore the robustness of our findings to this difference, by shifting data for men one year, to account for the average one year lag in male applicants' age compared to female applicants (this lag likely results from the extra year of mandated military service for men). As a result of this shift, men and women are aligned by college entry cohort, rather than by birth year. Appendix Table A4 shows these results, for both college and graduate education. The effects are slightly smaller in magnitude when accounting for this lag, yet there is still a significant increase in women's education after the policy change for all specifications.

Global shock After establishing that the break we observe is genuine, significant, and timed at the year of the policy change, we turn to explore the possibility that we have misattributed the source of this dramatic change. To verify that broader international trends during the nineties are not responsible for our effects, we conduct placebo tests in the United States as well as four other OECD countries with GDP per capita similar to Israel, and Census data availability: Greece, Spain, France, and Portugal. Results for college education in the four "comparable" countries are shown in Appendix Figure A12. The United States American Community Survey also contains information on marriage age, which allows us to look at both educational and marriage outcomes, shown in Figure 14. None of these placebo tests produce positive, significant results, indicating there was no broader global shock affecting those entering college in 1994.

Figure 14: United States College Attainment and Age at Marriage



Notes: Placebo test for global 1994 shift. Figure (a) shows the difference between age at first marriage for women and men, with fitted lines for the "pre" and "post" periods. Figure (b) shows the difference in college attainment between women and men, with the "treatment" year adjusted to reflect the cohort entering college in 1994 in the US, as well as a dotted line at the 1971 cohort (to visually check for a shock based on both birth year and college-entry year). Data from the 2010 American Community Survey.

Other health expansions Another possible explanation to the observed changes in women's outcomes could be the general change in health services provided by the NHI law. Although improved health services could have some effect on education (due to an increase in life expectancy for instance) or on marital outcomes (if we believe that age serves as a proxy for spousal health in general, rather than just fertility), it seems less likely to explain the gender differential change that we identify. Moreover, only a very dramatic change in access to health services would have an impact on either life expectancy or age related health perceptions. In practice, only 5% of the population had no health insurance prior to the reform. The remaining 95% were covered by one of the four HMOs, and experienced a very moderate change in coverage, except for very specific services such as IVF.[52] It should also be noted that during 1994, before the reform was actually executed, many people thought that it was bound to increase the cost of health care, and decrease the quality of the services.[53] Due to the complexity of the reform, mainly aimed at changing the institutional structure of the health system in Israel, the uncertainty over the actual results of the reform continued for decades, and even nowadays these consequences are being debated.[54]

Nonetheless, to check that our results are not driven by previously uninsured women accessing

[52]Additional significant changes were made in coverage of learning disabilities treatment and psychotherapy, as well as in tests for early detection of cancer.

[53]The slogan of the anti-reform campaign was "pay more, get less", see *Yedioth Ahronoth*, December 15, 1994.

[54]See for example reports issued by the Ministry of Health: "20 Years to the NHI Law" (May 2015) "Light and Shadow in the evolvement of the NHI Law" (February 2010).

additional health services, we repeat our main analysis excluding the groups who were most likely to have no insurance before the NHI law. The previously uninsured population consisted mostly of very poor households, immigrants, elderly individuals, and Arab-Israelis living in remote areas.[55] Since the three latter groups are already excluded from our main sample, we proceed by excluding the bottom quartile socio-economic groups, based on the geographic area of residence.[56] Appendix Table A5 reports the results of this estimation for our four main outcomes, using the main specification including fixed effects. The estimates for the main coefficients of interest remain essentially unchanged, suggesting that the effect we document was not unique nor especially dominant for the previously uninsured population.

Higher-education reform Finally, we consider the higher education reform in Israel, which materialized throughout the eighties and nineties, as an alternative explanation. Prior to the reform only universities could grant academic degrees. Starting in the seventies, colleges gradually began to receive permission to grant academic degrees equivalent to the ones given by universities. It is important to note that this reform was gradual, and thus is already partially ruled out as an alternative explanation by the discontinuous nature of the change in age-at-first marriage, as well as the event study analysis. This process accelerated during the eighties and early nineties, culminating in an official and comprehensive plan for the development of academic colleges. In the decade between 1992 and 2002 the number of students in academic programs approximately doubled (the effect of the reform was already apparent in the early nineties, but really started to build up in 1997-1998 (Volanski, 2005; Bernstein, 2002)). We already discussed the potential impact of the reform in Section 4.2, where we repeated our main analysis using Arab women as a control group to address the concern that this reform, or any other policy event, drives the gender differential change in higher-education completion. Arab-Israelis were much more likely than Jewish-Israelis to be affected by the higher-education reform, because it increased access to higher education for individuals with relatively low high-school achievements (Volanski, 2005), and for populations in peripheral areas. On the other hand, Arab women are not expected to react to the introduction of IVF, mostly due to significant religious constraints on usage (this is discussed further in Section 4.2). Hence, the fact that Jewish women, starting at 1994, divert to a higher and increasing educational trend in comparison to both Jewish-men and Arab-women, cannot be explained by the higher-education reform.

In addition, the two different reforms were likely to have affected different socioeconomic groups: IVF access is expected to have the largest impact for women interested in pursuing graduate

[55] See for example: the Netanyahu Commission report, "State Commission for Investigation of Functioning and Efficiency of the Health Care System in Israel," 1990; NHI law proposal, June 1993; "NHI Law: Equality, Efficiency, and Cost," report issued by Adva center, 1996.

[56] We use the smallest geographic unit available, "statistical area" which usually has 3-5 thousand residents. There is a standard classification of these areas to 20 socio-economic groups. We drop the bottom five.

education and demanding careers, while the education reform was targeted at those who did not currently have access to higher education. At the time of the education reform, women already constituted more than 50% of undergraduate students. The main purpose of the reform was to make higher education institutions more accessible to a lower socioeconomic status population, mostly concentrated in peripheral regions (Volanski, 2005; Shavit, Arum and Gamoran, 2007), and increase higher education supply to match the rapidly increasing demand.[57] Numerous studies were conducted to document the reform's consequences, none of which report a distinctive effect on women's participation in higher education (see for example Volanski (2005)). In fact, the percentage of female students in colleges (the previously "non-academic" institutions) was actually lower than in universities.[58]

Moreover, similar reforms in other countries were not found to affect women differently than men. One example is the higher education reform in Spain, which was enacted at approximately the same years as in Israel, and did not change the trend of women's education or of women's marriage decisions (Mora, 1996). It should also be noted that there is no reason to expect the reform to affect the way women's marriage outcomes depend on their age. We find that the previously existing penalty for older marriage practically disappears, even if we control for women's level of education. Combined together, it is hard to imagine that the driving force for all of the above described impacts is the increased supply of higher education rather than the increased availability of IVF technology, as we suggest.

To confirm this empirically, we conduct two additional tests to rule out this alternative explanation to our findings. To verify that our results are not due to an "in name only" change in the degree individuals received, we graph the percentage of any post secondary education graduates, which will include those whose degree status would have been switched into the academic "college" category after the reform. Appendix Figure A13 shows that even if we add non-academic degrees to our analysis, we get the same trends and the same change in trend only for women. This, together with our strong results for graduate education, eases our concerns for higher education reform driving the results by re-labeling once non-academic degrees.

Finally, to directly rule out this alternative explanation, in Appendix Table A6 we implement regressions controlling for a gender-specific impact of expanded educational access. We include controls for the number of academic colleges and teacher-training colleges by year,[59] interacted with gender, allowing the additional educational access to have a differential effect for men versus women. Even with these additional controls allowing for a direct impact of the educational reform,

[57] The demand increase stems from the growing rate of high-school graduates that received certificates in matriculation exams (which are needed when applying for college) (Shavit, Arum and Gamoran, 2007).

[58] The only exception is teacher's training colleges, where there is a vast majority of female students, however the academization process for those colleges took place in the early eighties. In addition, the students in these institutions constitute only a small share of the number of college students overall.

[59] Although the number of student slots varies by college, controlling for the number of students would be endogenous, and thus the number of institutions provides a better measure of the exogenous change.

the break at 1994 is still significant for all outcomes.

5 Conclusion

Increased access to *in vitro* fertilization offers women the security of a second-line option in case they do not naturally achieve their desired level of fertility. Like any insurance, this guaranteed access to IVF may influence individual behavior: In this case, women delay starting families, using the time to pursue additional education and potentially other career opportunities. The delay in starting families is shown by the stark increase in age at first marriage for women following the policy change. The productive use of this time is demonstrated by the rise in completion of college and graduate education, and subsequently improved career outcomes.

The policy change may also have impacted men's beliefs about older women's value as partners. We show evidence that older women marry higher quality partners after the policy. This shift in the marriage equilibrium may further reflect in women's decisions—knowing they will not lose as much "reproductive capital" by delaying marriage, and that their later-life marriage opportunities will be more favorable as a result, they will have fewer impediments to pursuing desired educational or career investments.

By testing what happens when the threat of later life infertility is attenuated, this research suggests time-limited fecundity as a key source of asymmetry between men and women. When better insured against later life infertility, women delay marriage, invest in more education, have better career outcomes, and marry higher quality partners despite the delay. In the absence of such insurance, this female-specific sharp decline in fertility may contribute to lower human capital investments by women during their reproductive years, and consequently the gender wage gap. In Israel, changing fertility horizons appears to substantially impact college and graduate education, both because women start families quite young and have relatively high desired fertility rates, and because obligatory military service already delays any decision women make by at least two years. In other OECD countries, however, this investment tradeoff may take place after women have completed their education, when further on-the-job investments are required in order to climb the corporate ladder: late nights at the law firm, medical residencies, or the tenure sprint. Thus, depreciating reproductive capital may help to explain the lack of women in higher-level management positions as well as the high-skill gender wage gap. A wide range of policies, such as increased support for child-rearing in two-career households and access to maternity leave and career re-entry, in addition to access to assisted reproduction technologies, could help alleviate this tradeoff.

In regard to the specific Israeli policy we evaluate, our findings demonstrate that the beneficiary population extends far beyond the women who actually use IVF or other assisted reproduction technologies. Rather, because the guaranteed access acts as insurance in case natural conception fails, all women considering further educational investments or delayed marriage may benefit. This is of critical importance because the cost per user of free IVF with Israel's generous coverage is enormous, and Israel is currently considering measures to limit the policy, having already placed age limits on use, and restricted the number of cycles for certain women. When taking into account the "insurance effect" of the policy, the potential benefits to be weighed against those costs expand considerably.

One slight caution in regards to this cost-benefit calculation is that the type of benefits we describe may not be what the Israeli government had in mind when they enacted the policy. The objectives of the policy were not to increase women's education and career outcomes, but were rather explicitly pro-natalist, aimed at increasing the birth rate of Israeli citizens.[60] Thus, policymakers should note that the behavioral response to IVF access may cause fertility effects to be attenuated, or even go in the opposite direction. If women do delay starting families, assured against the outcome of having zero children, they may nonetheless end up with a smaller overall family size, due to the late start. Moreover, since some evidence suggests individuals are *overly* optimistic about IVF's success rates, some women may delay and go on to use the technology, only to be unsuccessful conceiving.

The question of the tradeoff between further human capital investments and labor market productivity versus satisfaction derived from family and home life extend beyond Israeli policymaking. As more and more US companies consider measures such as paying for employees to freeze eggs, which similarly creates insurance against later life infertility, some women who are already planning to delay childbearing may be relieved by the benefit, while others could see a constantly moving finish line for how long they are expected to delay, and feel pressured to submit themselves to intrusive medical procedures and late parenthood. Thus, it is unclear if expanded access to IVF is the *best* policy to alleviate the one-sided burden of depreciating reproductive capital. What is clear, however, is that this burden plays a crucial role in women's decisions and outcomes.

[60]The policy was defended in courts and described as a part of the fundamental human right to give birth and build a biological family.

References

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

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

Adda, Jerome, Christian Dustmann, and Katrien Stevens. 2017. "The Career Costs of Children." *Journal of Political Economy, forthcoming.*

Agarwal, Sumit, Souphala Chomsisengphet, Neale Mahoney, and Johannes Stroebel. 2014. "Regulating consumer financial products: Evidence from credit cards." *The Quarterly Journal of Economics*, 130(1): 111–164.

Andersen, Anne-Marie Nybo, Jan Wohlfahrt, Peter Christens, Jorn Olsen, and Mads Melbye. 2000. "Maternal age and fetal loss: population based register linkage study." *BMJ*, 320.

Andrews, D.W.K. 1993. "Test for Parameter Instability and Structural Change with Unknown Change Point." *Econometrica*, 59: 817–858.

Angelov, Nikolay, Per Johansson, and Erica Lindahl. 2016. "Parenthood and the gender gap in pay." *Journal of Labor Economics*, 34(3): 545–579.

Avellar, Sarah, and Pamela J Smock. 2003. "Has the Price of Motherhood Declined Over Time? A Cross-Cohort Comparison of the Motherhood Wage Penalty." *Journal of Marriage and Family*, 65(3): 597–607.

Bailey, Martha J. 2006. "More Power to the Pill: The Impact of Contraceptive Freedom on Women's Life Cycle Labor Supply." *Quarterly Journal of Economics*, 121(1): 289–320.

Bailey, Martha J, Brad Hershbein, and Amalia R Miller. 2012. "The opt-in revolution? Contraception and the gender gap in wages." *American Economic Journal: Applied Economics*, 4(3): 225–54.

Bernstein, Sara. 2002. "The Rise and Fall of Private Higher Education in Israel." *Policy Studies*, 52.

Bertrand, Marianne, Claudia Goldin, and Lawrence F Katz. 2010. "Dynamics of the gender gap for young professionals in the financial and corporate sectors." *American Economic Journal: Applied Economics*, 2(3): 228–255.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-In-Differences Estimates?" *The Quarterly Journal of Economics*, 119(1): 249–275.

Birenbaum-Carmeli, Daphna. 2003. "Reproductive Policy in Context: Implications on Women's Rights in Israel, 1945–2000." *Policy Studies*, 24(2-3): 101–113.

Birenbaum-Carmeli, Daphna. 2004. "'Cheaper than a Newcomer': on the Social Production of IVF Policy in Israel." *Sociology of Health & Illness*, 26(7): 897–924.

Birenbaum-Carmeli, Daphna. 2010. "Genetic Relatedness and Family Formation in Israel: Lay Perceptions in the Light of State Policy." *New Genetics and Society*, 29(1): 73–85.

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

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

Blackburn, McKinley L, David E Bloom, and David Neumark. 1993. "Fertility Timing, Wages, and Human Capital." *Journal of Population Economics*, 6(1): 1–30.

Blau, Francine D., and Lawrence M. Kahn. 2017. "The Gender-Wage Gap: Extent, Trends, and Explanations." *Journal of Economic Literature, Forthcoming.*

Bronson, M. A., and M. Mazzocco. 2015. "Cohort Size and the Marriage Market: Explaining Nearly a Century of Changes in U.S. Marriage Rates." *Working Paper.*

Bronson, Mary Ann. 2015. "Degrees are forever: Marriage, educational investment, and lifecycle labor decisions of men and women." *Working Paper.*

Buckles, Kasey. 2007. "Stopping the Biological Clock: Infertility Treatments and the Career-family Tradeoff." *Boston University, Unpublished Manuscript.*

Buckles, Kasey. 2008. "Understanding the Returns to Delayed Childbearing for Working Women." *The American Economic Review*, 98(2): 403–407.

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

Bundorf, M Kate, Melinda Henne, and Laurence Baker. 2007. "Mandated Health Insurance Benefits and the Utilization and Outcomes of Infertility Treatments." *Working Paper*, , (12820).

Buser, Thomas, Muriel Niederle, and Hessel Oosterbeek. 2014. "Gender, competitiveness, and career choices." *The Quarterly Journal of Economics*, 129(3): 1409–1447.

Chandra, Amitabh, Jonathan Gruber, and Robin McKnight. 2010. "Patient cost-sharing and hospitalization offsets in the elderly." *The American Economic Review*, 100(1): 193–213.

Collins, John A. 2002. "An International Survey of the Health Economics of IVF and ICSI." *Human Reproduction Update*, 8(3): 265–277.

Cortes, Patricia, and Jessica Pan. 2017. "When Time Binds: Returns to Working Long Hours and the Gender Wage Gap among the Highly Skilled." *Working Paper*.

Danziger, Leif, and Shoshana Neuman. 1999. "On the Age at Marriage: Theory and Evidence from Jews and Moslems in Israel." *Journal of Economic Behavior & Organization*, 40(2): 179–193.

Delavande, Adeline, and Hans-Peter Kohler. 2015. "HIV/AIDS-related expectations and risky sexual behaviour in Malawi." *The Review of Economic Studies*, 83(1): 118–164.

Dessy, Sylvain, and Habiba Djebbari. 2010. "High-Powered Careers and Marriage: Can Women Have It All?" *The BE Journal of Economic Analysis & Policy*, 10(1).

Díaz-Giménez, Javier, and Eugenio Giolito. 2013. "Accounting for the Timing of First Marriage." *International Economic Review*, 54(1): 135–158.

Fisman, Raymond, Sheena S. Iyengar, Emir Kamenica, and Itamar Simonson. 2006. "Gender Differences in Mate Selection: Evidence From a Speed Dating Experiment." *Quarterly Journal of Economics*, 121(2): 637–697.

Goldin, Claudia. 2014. "A grand gender convergence: Its last chapter." *The American Economic Review*, 104(4): 1091–1119.

Goldin, Claudia, and Lawrence F. Katz. 2002. "The Power of the Pill: Oral Contraceptives and Women's Career and Marriage Decisions." *Journal of Political Economy*, 110(4): 730–770.

Gustafsson, Siv. 2003. "Optimal Age at Motherhood. Theoretical and Empirical Considerations on Postponement of Maternity in Europe." In *Family, Household and Work*. 345–367. Springer.

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

Hanlon, W Walker. 2015. "Necessity is the mother of invention: Input supplies and Directed Technical Change." *Econometrica*, 83(1): 67–100.

Hashiloni-Dolev, Yael. 2013. *The Fertility Revolution.* Ben Shemen: Modan Publishing House, in Hebrew.

Hashiloni-Dolev, Yael, Amit Kaplan, and Shiri Shkedi-Rafid. 2011. "The Fertility Myth: Israeli Students' Knowledge Regarding Age-related Fertility Decline and Late Pregnancies in an Era of Assisted Reproduction Technology." *Human Reproduction*, 0(0): 1–9.

Hook, Ernest B, Philip K Cross, and Dina M Schreinemachers. 1983. "Chromosomal abnormality rates at amniocentesis and in live-born infants." *JAMA: The Journal of the American Medical Association*, 249(15): 2034–2038.

Hugonnier, Julien, Florian Pelgrin, and Pascal St-Amour. 2012. "Health and (other) asset holdings." *Review of Economic Studies*, 80(2): 663–710.

Jayachandran, Seema, and Adriana Lleras-Muney. 2009. "Life expectancy and human capital investments: Evidence from maternal mortality declines." *The Quarterly Journal of Economics*, 124(1): 349–397.

Kelsey, Thomas W, and W. Hamish B. Wallace. 2010. "Human ovarian reserve from conception to the menopause." *PLoS One*, 5.

Kleven, Henrik Jacobsen, Camille Landais, and Jakob Egholt Sogaard. 2016. "Children and Gender Inequality: Evidence from Denmark." *Working Paper.*

Kroeger, Sarah, and Giulia La Mattina. 2017. "Assisted reproductive technology and women's choice to pursue professional careers." *Journal of Population Economics*, 30(3): 723–769.

Lahey, Joanna N. 2012. "The Efficiency of a Group-Specific Mandated Benefit Revisited: The Effect of Infertility Mandates." *Journal of Policy Analysis and Management*, 31(1): 63–92.

Loughran, David S, and Julie M Zissimopoulos. 2009. "Why Wait? The Effect of Marriage and Childbearing on the Wages of Men and Women." *Journal of Human Resources*, 44(2): 326–349.

Low, Corinne. 2017. "Pricing the Biological Clock: The Reproductive Capital – Human Capital Tradeoff on the Marriage Market." *Working Paper.*

Lundborg, Petter, Erik Plug, and Astrid Würtz Rasmussen. 2017. "Can women have children and a career? IV evidence from IVF treatments." *American Economic Review*, 107(6): 1611–37.

Menken, J, J Trussell, and U Larsen. 1986. "Age and infertility." *Science*, 233(4771): 1389–1394.

Miller, Amalia R. 2011. "The Effects of Motherhood Timing on Career Path." *Journal of Population Economics*, 24(3): 1071–1100.

Mora, Joséginés. 1996. "The demand for higher education in Spain." *European Journal of Education*, 31(3): 341–354.

Niederle, Muriel, and Lise Vesterlund. 2007. "Do women shy away from competition? Do men compete too much?" *The Quarterly Journal of Economics*, 122(3): 1067–1101.

Ohinata, Asako. 2011. "Did the US Infertility Health Insurance Mandates Affect the Timing of First Birth?" *CentER Working Paper Series.*

O'Neill, June, and Solomon Polachek. 1993. "Why the gender gap in wages narrowed in the 1980s." *Journal of Labor Economics*, 11(1, Part 1): 205–228.

Remennick, Larissa. 2010. "Between Reproductive Citizenship and Consumerism: Attitudes towards Assisted Reproductive Technologies among Jewish and Arab Israeli Women." *Kin, Gene, Community: Reproductive Technologies among Jewish Israelis*, 19: 318.

Rothman, Kenneth J., Lauren A. Wise, Henrik T. Sorensen, Anders H. Riis, Ellen M. Mikkelsen, and Elizabeth E. Hatch. 2013. "Volitional determinants and age-related decline in fecundability: a general population prospective cohort study in Denmark." *Fertility and Sterility*, 99(7): 1958 – 1964.

Ruggles, Steven, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek. 2010. "Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database]." Minneapolis, MN: Minnesota Population Center.

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

Schmidt, Lucie. 2007. "Effects of Infertility Insurance Mandates on Fertility." *Journal of Health Economics*, 26(3): 431–446.

Shavit, Yossi, Richard Arum, and Adam Gamoran. 2007. *Stratification in Higher Education: A Comparative Study.* Stanford University Press.

Simonstein, Frida. 2010. "IVF Policies with Emphasis on Israeli Practices." *Health Policy*, 97(2): 202–208.

Siow, Aloysius. 1998. "Differential Fecundity, Markets, and Gender Roles." *Journal of Political Economy*, 106(2): 334–354.

Skinner, Jonathan. 1985. "The effect of increased longevity on capital accumulation." *The American Economic Review*, 75(5): 1143–1150.

Taniguchi, Hiromi. 1999. "The Timing of Childbearing and Women's Wages." *Journal of Marriage and the Family*, 61(4): 1008–1019.

Toner, James P. 2003. "Age = egg quality, FSH level = egg quantity." *Fertility and Sterility*, 79.

Velez, MP, MP Connolly, I-J Kadoch, S Phillips, and F Bissonnette. 2014. "Universal Coverage of IVF Pays Off." *Human Reproduction*, 29(6): 1313–1319.

Volanski, Ami. 2005. *Academy in a Changing Environment.* Kotar, in Hebrew.

Wai, Jonathan, Megan Cacchio, Martha Putallaz, and Matthew C Makel. 2010. "Sex differences in the right tail of cognitive abilities: A 30year examination." *Intelligence*, 38(4): 412–423.

Wilde, Elizabeth Ty, Lily Batchelder, and David T Ellwood. 2010. "The Mommy Track Divides: The Impact of Childbearing on Wages of Women of Differing Skill Levels." *Working Paper.*

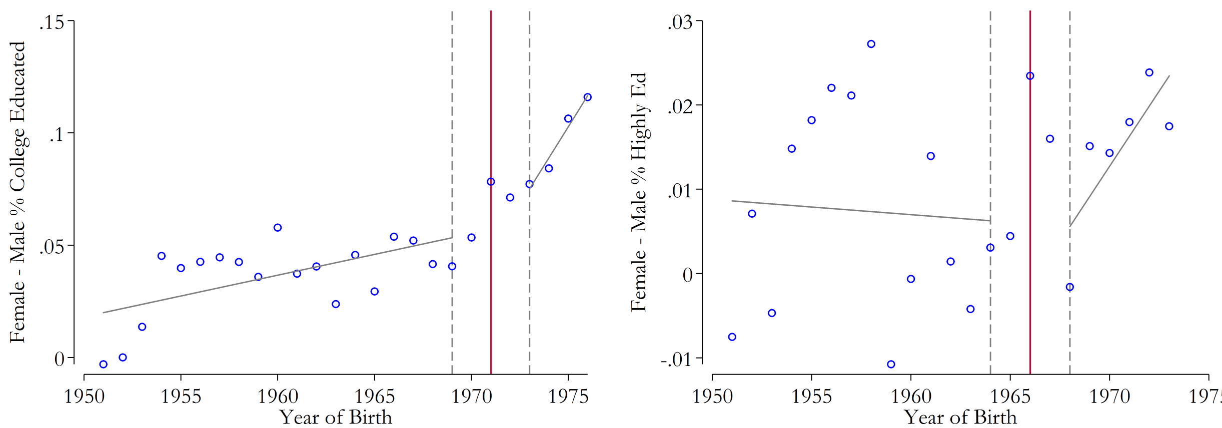
Zafar, Basit. 2013. "College major choice and the gender gap." *Journal of Human Resources*, 48(3): 545–595.

6 Appendix

Figure A1: Difference in Education Rates with Fuzzy Treatment

(a) College Education

(b) Graduate Education



Notes: These graphs show that the difference in education outcomes post-policy is still apparent with a two-year interval on either side of our “most treated” cohort. Older cohorts may have been partially treated, if some individuals had not yet made their educational decisions and thus were able to respond to the policy change, and younger cohorts may not have been fully treated, if some individuals had made their educational decisions already. Data from the 2008 Israeli population census, restricted to Israeli-born Jews.

Table A1: College Education by Percent of Cohort Treated

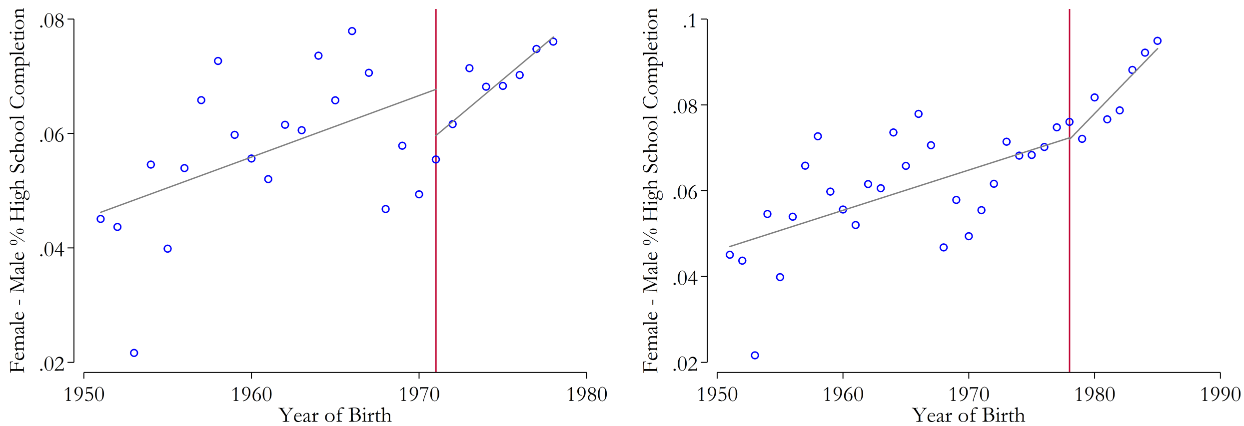
	Dependent variable: College Education Gradual Exposure			
	DiD		GLS DiD	
	(1)	(2)	(3)	(4)
fem × treated	0.090	0.051	0.090	0.055
	(0.006)***	(0.010)***	(0.010)***	(0.016)***
	[0.026]***	[0.032]		
cohort treated	-0.029	0.016	-0.036	0.002
	(0.022)	(0.019)	(0.032)	(0.030)
	[0.041]	[0.038]		
female	0.021	0.047	0.021	0.044
	(0.003)***	(0.006)***	(0.005)***	(0.010)**
	[0.021]	[0.025]*		
fem × time		0.002		0.002
		(0.000)***		(0.001)**
		[0.001]*		
Constant	0.385	0.392	0.252	0.258
	(0.013)***	(0.008)***	(0.003)***	(0.003)***
	[0.024]***	[0.024]***		
YOB FEs	YES	YES	YES	YES
Observations	173790	173790	54	54
R-Squared	0.109	0.109		

Notes: The treatment variable is a continuous variable that measures the gender-specific percent of the cohort that has yet to enter college in 1994, and hence has the potential to alter decisions due to the policy change that occurred at the same year. The percentages are based on data from the 1995 Israeli Census regarding the age distribution of 1993 college entrants. We use 1993 because it is the last pre-treatment year and thus the best measure for the college-entry age distribution both for the pre-treatment years and for 1994 in the absence of treatment (i.e. the policy change). Columns 1–2: Ordinary least-squares difference-in-differences regression using micro data, including controls for religiosity and parents' origin. Robust standard errors clustered at the gender × year level in parentheses; robust standard errors clustered at the gender × geography level in square brackets. Columns 3–4: Generalized least squares regression with data collapsed to the gender-year-of-birth level. Robust standard errors that allow for cross-sectional correlation and for panel-specific serial correlation (i.e. estimate a unique autocorrelation parameter for each group), in parentheses. Data from the 2008 Israeli population census, restricted to Israeli-born Jews.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Figure A2: High School Placebo

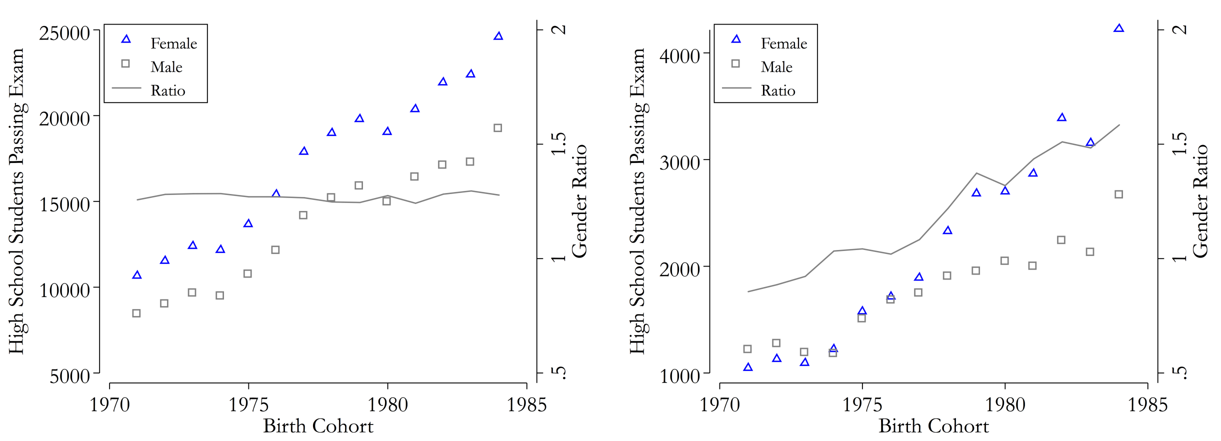
(a) 1971 Cutoff (b) 1978 Cutoff



Notes: We test for a change in high school completion rates to rule out broader increases in education driving our effects for both the 1971 cohort, which would have been affected if the shock was something that impacted all individuals born in 1971, and the 1978 cohort, which would have been affected if the shock impacted all individuals pursuing schooling in 1994. Figure (a) shows average high school completion for women and men by birth cohort. Figure (b) presents the difference in high school completion between women and men, as well as fitted lines for the pre and post periods. Data from the 2008 Israeli population census, restricted to Israeli-born Jews.

Figure A3: Female vs. Male Highschool Matriculation ("Bagrut")

(a) Jewish Students (b) Arab Students



Notes: The figures show the number of high-school students passing the matriculation exams ("Bagrut") by birth year, for men and women separately. In addition, the ratio of women to men passing the exams is calculated and measured on the right hand side y-axis. Figure (a) presents these figures for students in Hebrew speaking schools, i.e., the Jewish population of Israel. Figure (b) shows the same figures for students in Arabic schools, i.e., the Arab population of Israel. Birth years are calculated by subtracting 18 (the approximate age for students in the last year of high-school) from the year of exam. Data from the Israeli Central Bureau of Statistics year-books for the years 1989-2002.

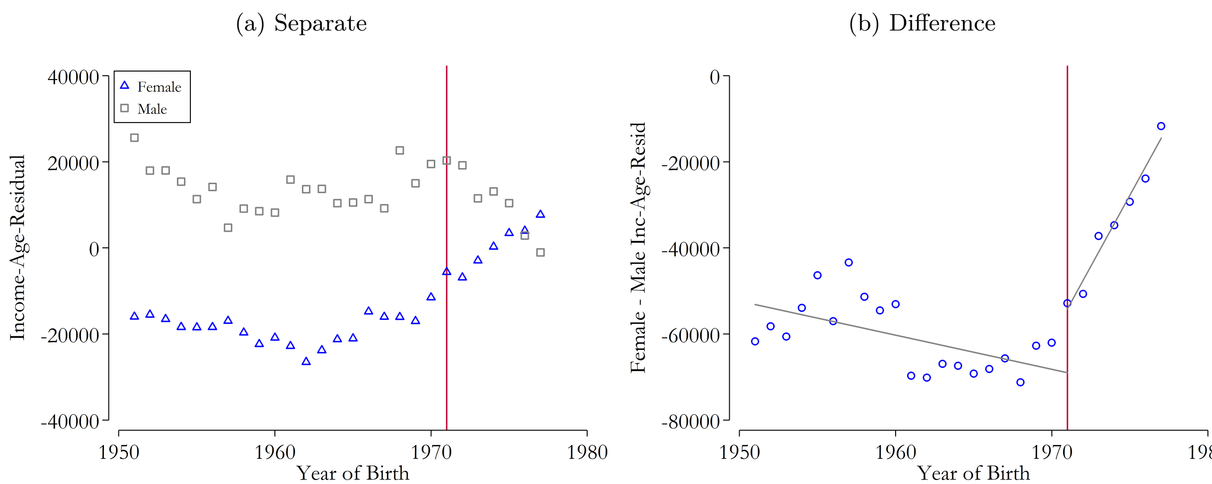
Table A2: Age at First Marriage (Arab Control)

	Dependent variable: Age First Marriage					
	DiD		Slope-Change DiD		GLS Slope-Change DiD	
	(1)	(2)	(3)	(4)	(5)	(6)
jewish × post	1.019	1.016	-0.076	-0.057	-0.033	0.488
	(0.127)***	(0.106)***	(0.152)	(0.082)	(0.150)	(0.325)
	[0.208]***	[0.208]***	[0.192]	[0.192]		
jewish × post × time			0.058	0.059	0.056	0.349
			(0.016)***	(0.010)***	(0.020)***	(0.062)***
			[0.030]*	[0.030]*		
Constant	21.984	20.049	21.688	21.626	21.692	20.172
	(0.084)***	(0.135)***	(0.036)***	(0.033)***	(0.116)***	(0.506)***
	[0.269]***	[0.216]***	[0.334]***	[0.355]***		
YOM FEs		YES		YES		YES
Observations	124744	124744	124744	124744	62	62
R-Squared	0.121	0.122	0.122	0.123		

Notes: Columns 1–4: Ordinary least-squares difference-in-differences regression using micro data (no controls included since religiosity and parents' origin controls used only apply to Jewish population). Robust standard errors clustered at the group × year level in parentheses; robust standard errors clustered at the group × geography level in square brackets. Columns 5–6: Generalized least squares regression with data collapsed to the group-year level. Robust standard errors that allow for cross-sectional correlation and for panel-specific serial correlation (i.e. estimate a unique autocorrelation parameter for each group), in parentheses. All specifications include coefficients for group, post, time, post × time, and group × time. Data from the 2008 Israeli population census, restricted to Israeli-born.

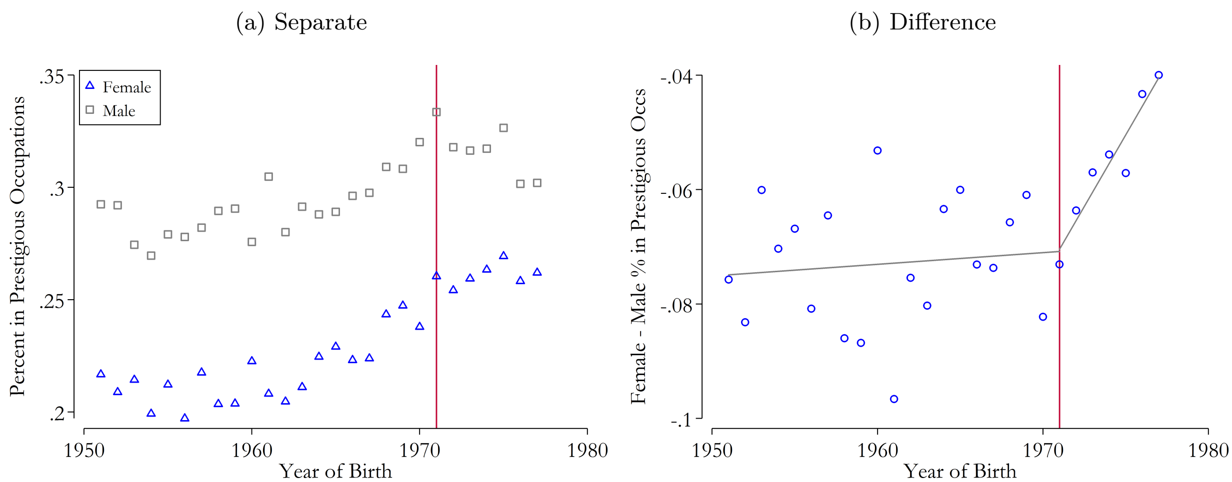
*** p<0.01, ** p<0.05, * p<0.1

Figure A4: Female vs. Male Income Residuals



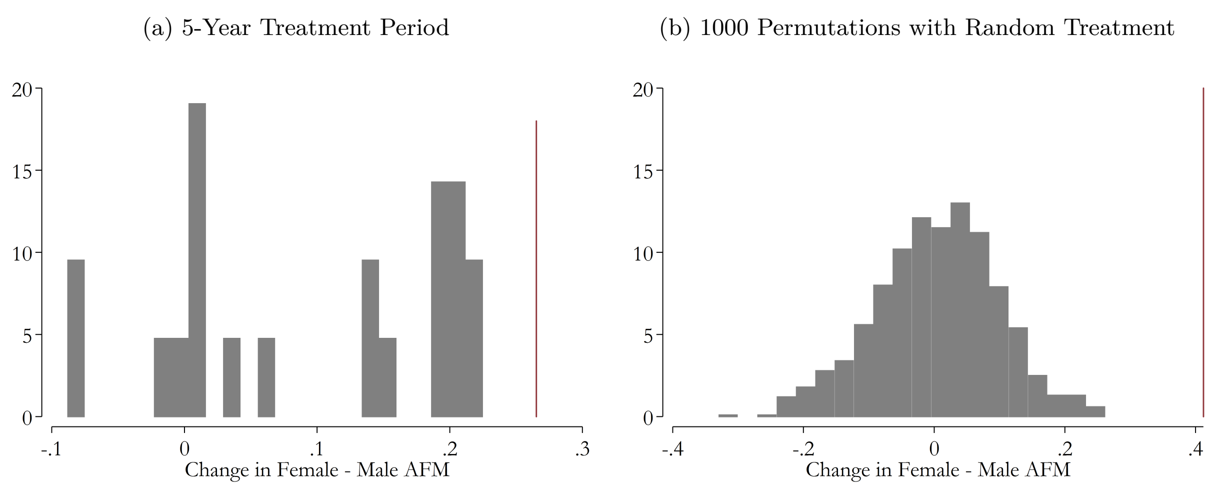
Notes: Figure (a) shows income levels, controlling for a flexible polynomial in age and a dummy for female, for women and men by birth cohort. Figure (b) presents the difference in income between women and men, as well as fitted lines for the pre (1951-1970 birth cohorts) and post (1971-1977 birth cohorts) periods (the choice of cohorts follows the college education analysis). Data from the 2008 Israeli population census, restricted to Israeli-born Jews.

Figure A5: Female vs. Male Prestigious Occupations



Notes: Figure (a) shows rates of participation in prestigious occupations, including management, law, medicine, engineering, and academia, for women and men by birth cohort. Figure (b) presents the difference in these rates between women and men, as well as fitted lines for the pre (1951-1970 birth cohorts) and post (1971-1977 birth cohorts) periods (the choice of cohorts follows the college education analysis). Data from the 2008 Israeli population census, restricted to Israeli-born Jews.

Figure A6: Permutations for Age at First Marriage



Notes: The figure on the left is created by running a similar regression as our column 1 specification, except with a ten year data period, with five years control and five years treatment, sequentially, for every possible ten year period in our data range. The red line represents the effect size of the actual treatment year, with this ten-year data period (the ten-year approach allows us to compare our actual treatment to other break points, with the same number of years before and after). The figure at right uses the same number of "treated" years as in the true model, but randomly draws them from the study period (for an example of this approach, see Agarwal et al. (2014)). We perform 1,000 such random draws.

Table A3: Spousal Income, Triple Difference

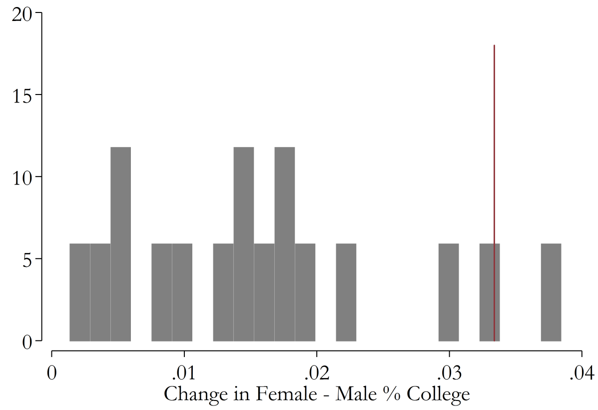
	Dependent variable: Spousal Income, Triple Difference			
	DDD		GLS DDD	
	(1)	(2)	(3)	(4)
older × post × fem	17676.189	17250.748	18046.017	17341.056
	(8597.305)**	(8155.762)**	(7929.683)**	(8522.156)**
	[8490.004]**	[8540.434]**		
older × post	1402.961	2216.055	426.693	-369.805
	(3793.243)	(3913.530)	(3185.769)	(3386.364)
	[2797.701]	[2837.764]		
older × fem	-22896.788	-22941.086	-24944.938	-23669.066
	(5939.128)***	(6183.988)***	(6829.507)***	(7471.575)***
	[13814.507]*	[13780.157]*		
fem × post	-22940.075	-21719.467	-18596.185	-17184.070
	(4659.229)***	(3876.876)***	(7385.162)**	(7832.778)**
	[4090.765]***	[3962.996]***		
post	11976.842		6300.597	
	(4795.629)**		(3028.793)**	
	[3029.632]***			
female	120102.904	119306.436	61419.868	59491.186
	(2638.721)***	(2262.463)***	(6510.743)***	(7082.276)***
	[8472.702]***	[8328.691]***		
married older	303.631	-120.452	-3394.571	-3334.257
	(2859.065)	(2384.070)	(2642.969)	(2827.988)
	[3956.800]	[3947.478]		
time	-737.461		1231.043	
	(481.807)		(227.629)***	
	[431.037]*			
Constant	727786.275	641649.432	-5745.596	-14481.554
	(580311.776)	(594258.199)	(2117.151)***	(2436.667)***
	[453091.957]	[456996.909]		
YOM FEs		YES		YES
Observations	45089	45089	72	72
R-Squared	0.183	0.184		

Notes: Columns 1–2: Ordinary least-squares triple-difference regression using micro data, including controls for religiosity and parents' origin, as well as a flexible polynomial for age. Robust standard errors clustered at the gender × age group × year level in parentheses; robust standard errors clustered at the age gender × group × geography level in square brackets. Columns 3–4: Generalized least squares regression with data collapsed to the gender-age group-year-of-marriage level. Robust standard errors that allow for cross-sectional correlation and for panel-specific serial correlation (i.e. estimate a unique autocorrelation parameter for each group), in parentheses. Data from the 2008 Israeli population census, restricted to Israeli-born Jews.

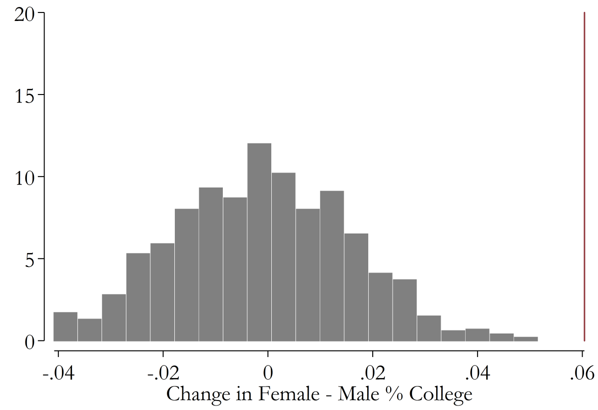
*** p<0.01, ** p<0.05, * p<0.1

Figure A7: Permutations for College Education

(a) 5-Year Treatment Period



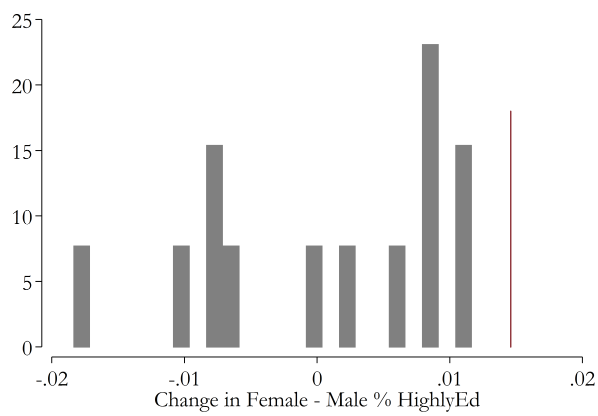
(b) 1000 Permutations with Random Treatment



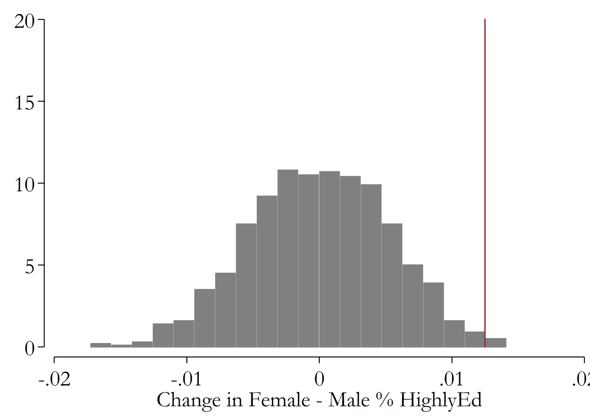
Notes: The figure on the left is created by running a similar regression as our column 1 specification, except with a ten year data period, with five years control and five years treatment, sequentially, for every possible ten year period in our data range. The red line represents the effect size of the actual treatment year, with this ten-year data period (the ten-year approach allows us to compare our actual treatment to other break points, with the same number of years before and after). The figure at right uses the same number of “treated” years as in the true model, but randomly draws them from the study period (for an example of this approach, see Agarwal et al. (2014)). We perform 1,000 such random draws.

Figure A8: Permutations for Graduate Education

(a) 5-Year Treatment Period

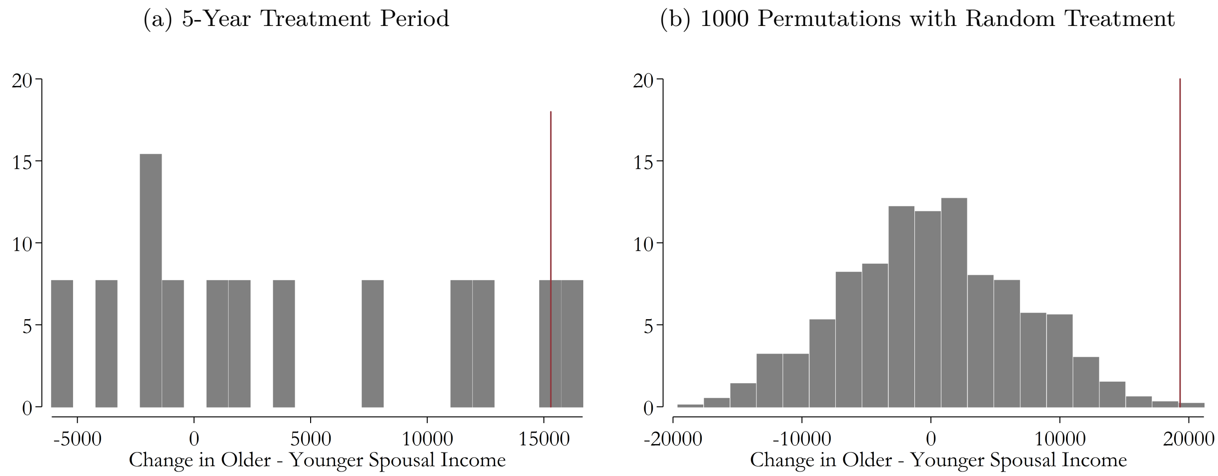


(b) 1000 Permutatins with Random Treatment



Notes: The figure on the left is created by running a similar regression as our column 1 specification, except with a ten year data period, with five years control and five years treatment, sequentially, for every possible ten year period in our data range. The red line represents the effect size of the actual treatment year, with this ten-year data period (the ten-year approach allows us to compare our actual treatment to other break points, with the same number of years before and after). The figure at right uses the same number of “treated” years as in the true model, but randomly draws them from the study period (for an example of this approach, see Agarwal et al. (2014)). We perform 1,000 such random draws.

Figure A9: Permutations for Spousal Income-Age-Residual



Notes: The figure on the left is created by running a similar regression as our column 1 specification, except with a ten year data period, with five years control and five years treatment, sequentially, for every possible ten year period in our data range. The red line represents the effect size of the actual treatment year, with this ten-year data period (the ten-year approach allows us to compare our actual treatment to other break points, with the same number of years before and after). The figure at right uses the same number of "treated" years as in the true model, but randomly draws them from the study period (for an example of this approach, see Agarwal et al. (2014)). We perform 1,000 such random draws.

Table A4: College and Graduate Graduation Rates, Men's Cohort Adjusted

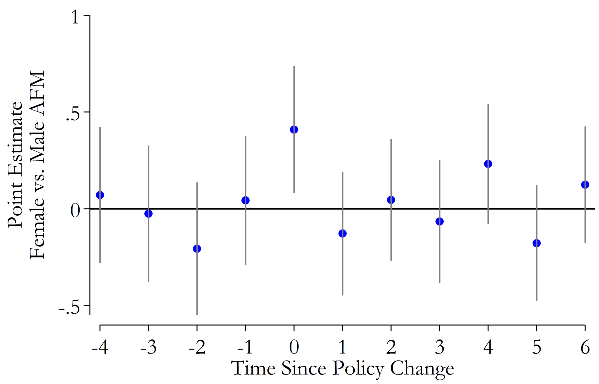
Dependent variable:	College Education		Graduate Education	
	DiD with GSTT			
	(1)	(2)	(3)	(4)
fem × post	0.024	0.018	0.014	0.009
	(0.007)***	(0.006)***	(0.006)**	(0.004)**
	[0.016]	[0.015]	[0.008]*	[0.007]
Constant	0.497	0.396	0.195	0.208
	(0.007)***	(0.012)***	(0.005)***	(0.004)***
	[0.031]***	[0.023]***	[0.016]***	[0.015]***
YOB FEs	YES		YES	
Cohort FEs		YES		YES
Observations	171617	171617	136998	136998
R-Squared	0.110	0.110	0.0468	0.0469

Notes: Columns 1–4: Ordinary least-squares difference-in-differences regression using micro data, including controls for religiosity and parents' origin. Robust standard errors clustered at the gender × cohort level in parentheses; robust standard errors clustered at the gender × geography level in square brackets. All specifications include coefficients for female, post, time, and group × time. Data from the 2008 Israeli population census, restricted to Israeli-born Jews.

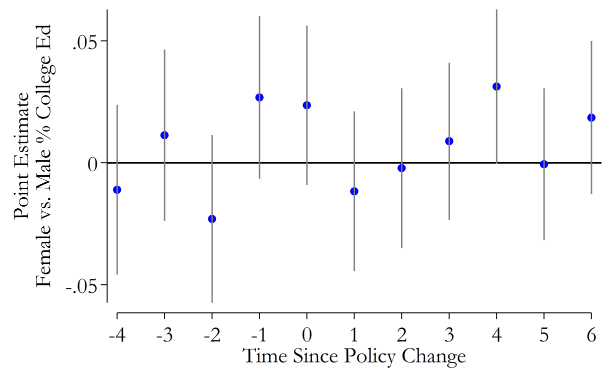
*** p<0.01, ** p<0.05, * p<0.1

Figure A10: Event Study Analysis (Distributed lags)

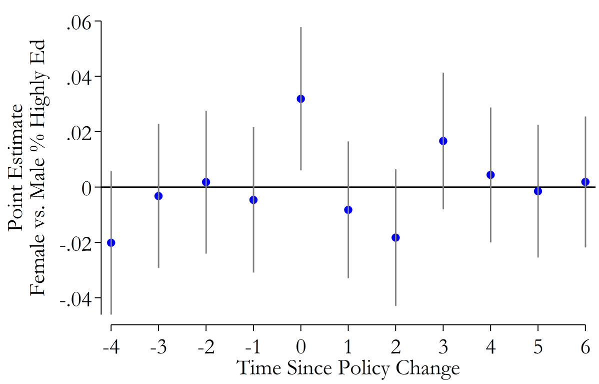
(a) Age at First Marriage



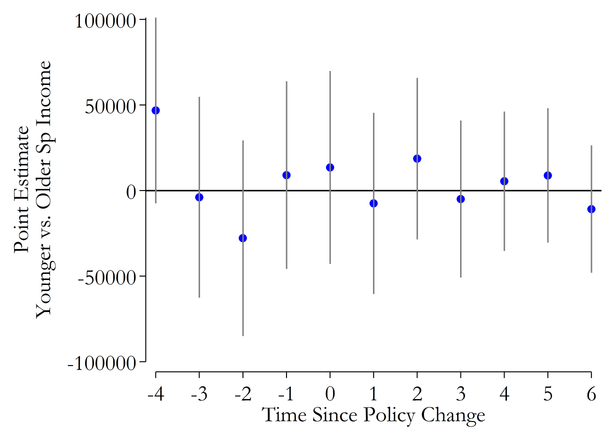
(b) College Education



(c) Graduate Education

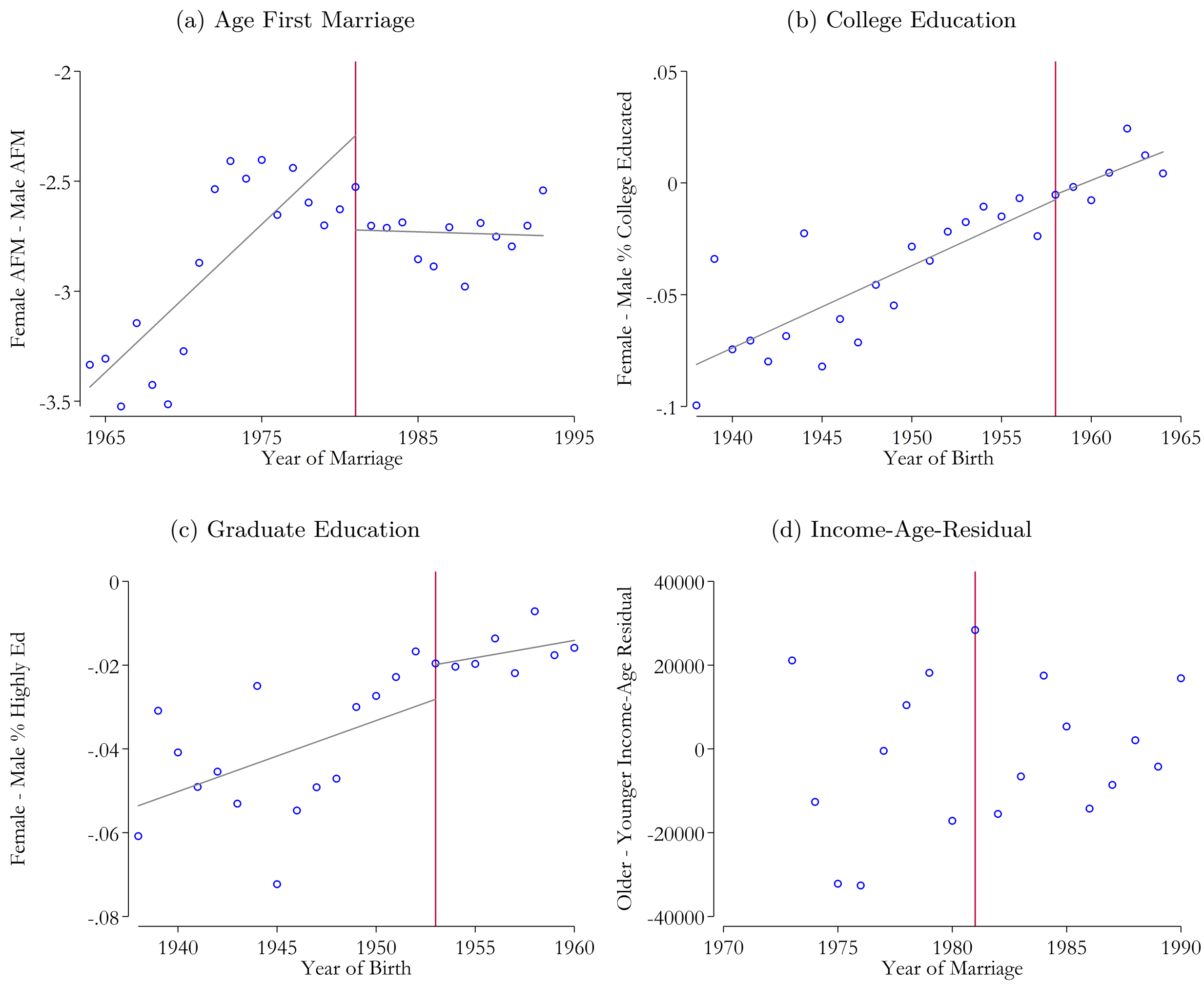


(d) Spousal Income



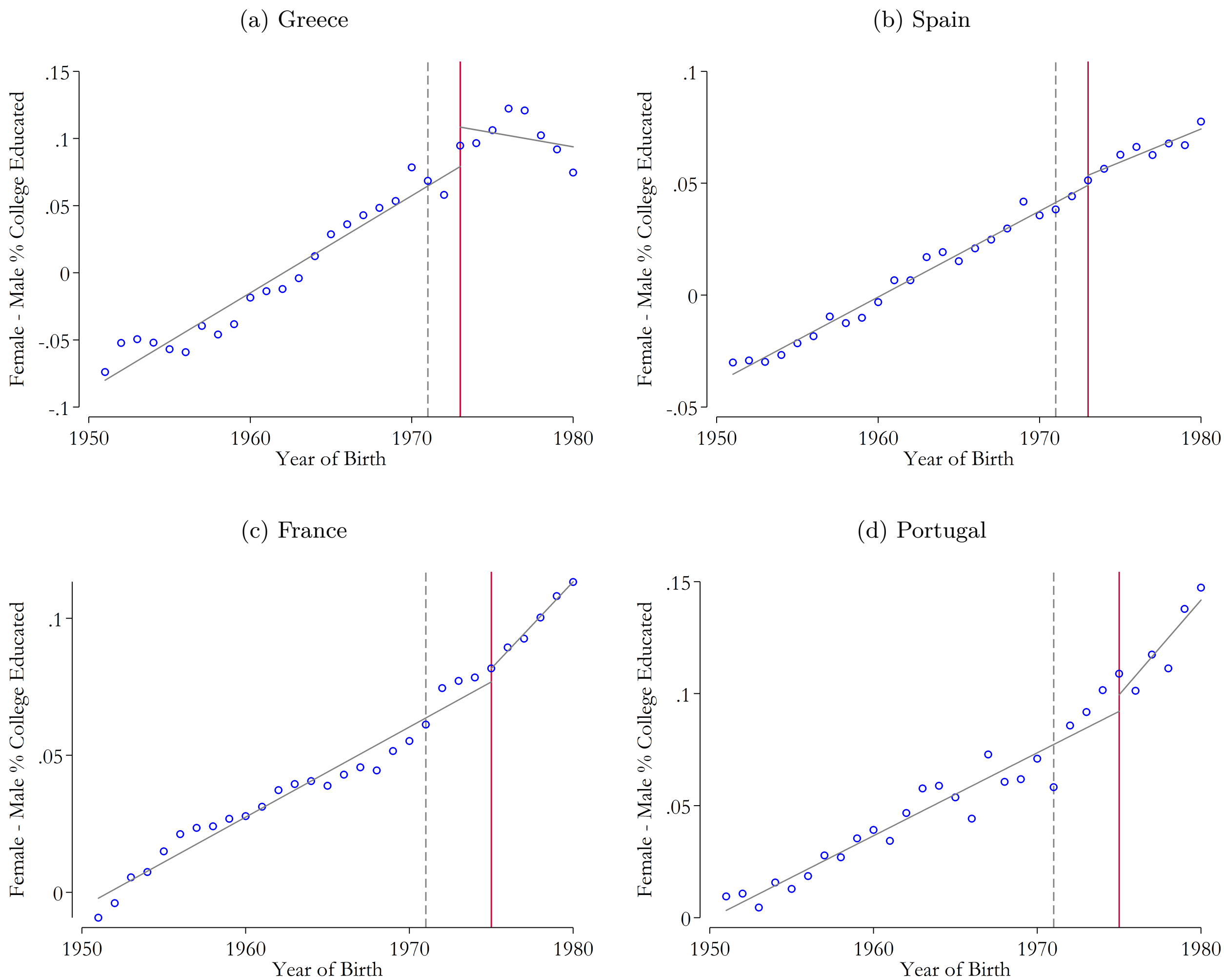
Notes: The figures present distributed lag event study analyses for the principal regression specifications for our main outcome variables. Point estimates and confidence intervals are for the coefficients on dummy for being in or greater than a given year, interacted with dummy variables for the treated group. For age at first marriage, college education, and graduate education, the treated group is women, compared to men, while for spousal income it is women over 30 at first marriage, compared to those under 30. For age at first marriage and spousal income, time is year of marriage (being married in or after the given year), with 1994 as time "zero," while it is birth cohort for college and graduate education (being born in or after the specified cohort). The regression equation includes all year-group interactions for the entire sample period, as well as a linear time trend and demographic controls. All regressions restricted to Israeli-born Jews.

Figure A11: Placebo Test using 1995 Israeli Census



Notes: Placebo test for spurious results due to retrospective analysis, using 1995 census and fictitious 1981 ”policy change”. Figure (a) presents the difference in men and women's age at first marriage by marriage year, figure (b) presents the difference in men and women's college attainment and figure (c) graduate attainment by birth year, and figure (d) presents the difference in spousal income-age-residual for women who marry at 30 or above versus those who marry younger. Data from 1995 Israeli population census, restricted to Israeli-born Jews. Note, that for age at first marriage, we use data up to 1993, to avoid the "post treatment" period (which starts at 1994).

Figure A12: College Attainment by Birth Cohort in Comparable Countries



Notes: We use countries that have censuses around the time of Israel's 2008 Census and similar GDP per capita as Israel to conduct placebo tests, showing that the cohort entering college in 1994 in other countries was not similarly affected (in each country, the break-point is shifted according to typical college entry age of students in that country, while the dotted line highlights the 1971 cohort). This would be the case if broader international shifts in the nineties were responsible for the effects that we see. We do not observe similar discontinuous increases in female versus male college attainment over time in any of the other countries. The same lack of discontinuous trends is true when looking at graduate school attainment as well, although the data is somewhat noisier. Census data for Greece from 2001. Census data from Spain, France, and Portugal from 2011. Data obtained from IPUMS.

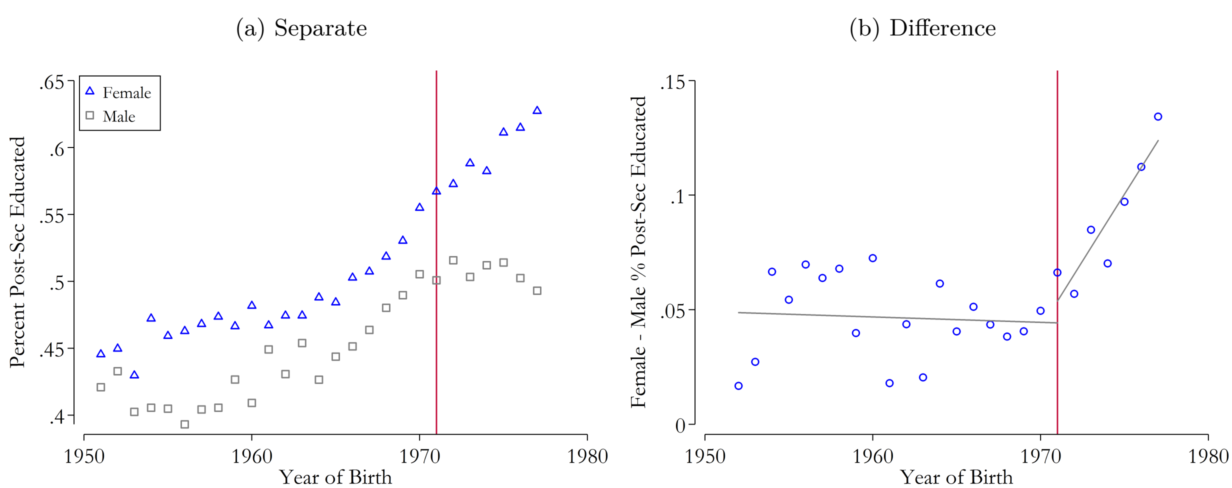
Table A5: All Outcomes, Excluding Bottom Socio-Economic Status Groups

Dependent variable:	AFM	College	Graduate	Sp Inc
	(1)	(2)	(3)	(4)
treated × post	0.213	0.004	0.015	17422.184
	(0.047)***	(0.007)	(0.007)**	(6152.163)***
	[0.143]	[0.017]	[0.009]*	[8879.802]*
treated × post × time	0.049	0.006	0.000	
	(0.001)***	(0.001)***	(0.001)	
	[0.023]**	[0.004]*	[0.002]	
Constant	23.424	0.441	0.221	-514742.422
	(0.073)***	(0.007)***	(0.006)***	(871970.673)
	[0.187]***	[0.029]***	[0.019]***	[831479.104]
FEs (YOB/YOM)	YES	YES	YES	YES
Observations	113652	122001	96788	13212
R-Squared	0.219	0.106	0.0476	0.0545

Notes: These regressions exclude the bottom 5 (of 20) socio-economic groups, in order to rule out that expansion of access to healthcare could have driven outcome measures. Only 5% of Israelis were uninsured at the time of the 1994 policy change. All columns present ordinary least-squares difference-in-differences regression using micro data, including controls for religiosity and parents' origin and time fixed effects. Robust standard errors clustered at the group × year level in parentheses (in columns 1–3 group is gender and in column 4 it is age group; in columns 2–3 year is year of birth and in columns 1 and 4 year is year of marriage); robust standard errors clustered at the group × geography level in square brackets. Columns 1–3 present slope-change DID estimates for age at first marriage, college education and graduate education. Column 4 present DID estimates for spousal income, controlling for spousal age with a flexible polynomial. All specifications include coefficients for group, post, and time. College, graduate, and AFM regressions additionally include post × time and group × time. Data from the 2008 Israeli population census, restricted to Israeli-born Jews.

*** p<0.01, ** p<0.05, * p<0.1

Figure A13: Female vs. Male Post-Secondary Education



Notes: Figure (a) shows average post-secondary educational attainment for women and men by birth cohort. Figure (b) presents the difference in post-secondary attainment between women and men, as well as fitted lines for the pre (1951-1970 birth cohorts) and post (1971-1977 birth cohorts) periods. Data from the 2008 Israeli population census, restricted to Israeli-born Jews.

Table A6: Main Outcomes, Controlling for the Number of Academic and Teachers' Colleges

Dependent variable:	AFM		College		Graduate	
	(1)	(2)	(3)	(4)	(5)	(6)
fem × post	0.376	0.381	0.026	0.026	0.033	0.033
	(0.165)**	(0.057)***	(0.010)**	(0.006)***	(0.005)***	(0.004)***
	[0.174]**	[0.172]**	[0.018]	[0.018]	[0.011]***	[0.011]***
fem × post × time	0.051	0.052	0.011	0.011	-0.003	-0.003
	(0.020)**	(0.010)***	(0.003)***	(0.002)***	(0.001)*	(0.001)**
	[0.034]	[0.034]	[0.005]**	[0.005]**	[0.003]	[0.003]
Constant	26.142	26.360	0.492	0.540	0.260	0.241
	(0.174)***	(0.106)***	(0.023)***	(0.023)***	(0.010)***	(0.010)***
	[0.305]***	[0.222]***	[0.026]***	[0.032]***	[0.018]***	[0.019]***
FEs (YOB/YOM)		YES		YES		YES
Observations	167416	167416	173790	173790	138953	138953
R-Squared	0.246	0.247	0.109	0.109	0.0465	0.0466

Notes: These regressions allow the expansion of the higher-education system to have differential effects for women versus men, by including gender-specific controls for the number of academic colleges and teachers' colleges each year. All columns present ordinary least-squares difference-in-differences regression using micro data, including controls for religiosity and parents' origin. Robust standard errors clustered at the gender × year level in parentheses (in columns 1–2 year is year-of-marriage and in columns 3–6, year-of-birth); robust standard errors clustered at the gender × geography level in square brackets. All specifications include coefficients for female, post, time, post × time, and group × time, in addition to controls for the number of academic colleges and teachers' colleges by year, interacted with gender. Data from the 2008 Israeli population census, restricted to Israeli-born Jews.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$