# The Unequal Battle Against Infertility: Theory and Evidence from IVF Success and Drop-out*

Fane Groes
Copenhagen Business School

Anna Houštecká
CERGE-EI

Daniela Iorio
University of Bologna

Raül Santaeulàlia-Llopis
New York University Abu Dhabi
UAB, BSE, and CEPR

December 4, 2024

**Abstract**

Using Danish administrative data, we show that IVF success is associated with maternal education: College-educated women have a 9% higher live birth chance than high school-educated women and 25% higher than dropouts. We exclude infertility causes, health behaviors, occupations, clinics, finances, and partner attributes as drivers. Instead, we focus on latent factors like ability and psychological traits. First, we show how proxies for these factors like Grade Point Average (GPA) shape IVF success. Second, we build a structural dynamic model of post-IVF-failure dropout where women differ in latent ability and psychological costs. Our model counterfactuals imply that ability explains 87% of the education gradient in IVF success, prompting a policy discussion.

**Keywords**: In-Vitro Fertilization, Success, Inequality, Infertility, Education, Ability, Administrative Data, Structural Dynamic Model, Endogenous Dropouts
**JEL Classification**: I00, J13, E00

*We would like to thank Ingeborg Kristiansen and Edward Sosman for excellent research assistance. We thank Jere Behrman, Michele Belot, Sonia Bhalotra, Mette Ejrnæs, Paula England, Libertad González, Hans-Peter Kohler, Fabian Lange, Veda Narasimhan, Tetyana Surovtseva and seminar participants in multiple venues. Raül Santaeulàlia-Llopis acknowledges financial support from the Spanish Ministry of Economy and Competitiveness, through the Proyectos I+D+i 2019 Retos Investigacion PID2019-110684RB-I00 Grant, Europa Excelencia EUR2021-122011 and the Severo Ochoa Programme for Centres of Excellence in R&D (CEX2019-000915-S). This output was supported by the NPO "Systemic Risk Institute" no. LX22NPO5101, funded by European Union - Next Generation EU (Ministry of Education, Youth and Sports, NPO: EXCELES).

# 1 Introduction

**Motivation** Assisted reproductive techniques such as In-Vitro Fertilization (IVF) are increasingly popular methods for childbearing. In Denmark, the proportion of children born through IVF has steadily risen since the 1990s, growing from less than 1.8% in 1995 to approximately 7.2% in 2019 (see Figure 1).[1] This increase in IVF usage aligns with the idea that women born in the 1980s and later aim to balance careers and family, as recently put forth in Goldin (2021), rather than prioritizing one over the other. A key aspect of this balance is the trade-off between delaying childbearing and preserving fertility (Doepke et al., 2023). While postponing motherhood can have positive effects on women's earnings and human capital accumulation, it also reduces natural pregnancy chances.[2] The introduction of IVF in the 1980s has potentially relaxed this biological constraint, influencing women's decisions.[3] However, the extent of this influence may vary depending on the IVF success rate (i.e. live births) and its determinants.

**Question** Unfortunately, while IVF technology is a well-established treatment for infertility, there remains limited understanding regarding the determinants of IVF success. For instance, medical studies have identified prognostic factors, including uterine receptivity, ovarian function, morphological grading of embryos, and their chromosomal competence.[4] At face value, these factors might appear to rely on purely exogenous factors and sheer "luck." But, does IVF success truly only hinge on the capriciousness of "nature"?

**Answer** Using the Danish administrative registers, which provide us with both longitudinal and detailed information on *all* IVF treatments in Denmark since 1995, we find the presence of significant education disparities in IVF success. Precisely, among childless women, college-educated patients are about 25% more likely to attain a live birth at the first trial (cycle) than women who do not have a high school degree, and 9% higher one than those with a high school degree or some college—where the average chances of a live birth in the first cycle is 0.25. Despite the fact that the education composition changes over the cycles, the education gradient in IVF success is of similar size also in subsequent IVF treatments. To the best of our knowledge, we are the first to document the positive education-IVF gradient and to underscore its significance (Doepke et al., 2023; Olivetti et al., 2024). Notably, a growing body of evidence

[1]Worldwide, there are about four million assisted reproductive techniques cycles annually, resulting in approximately one million babies born (source: International Committee for Monitoring Assisted Reproductive Technologies, ICMART, 2022). This trend mirrors previous analysis (de Mouzon et al., 2010; Connolly et al., 2010).

[2]See Caucutt et al. (2002), Gayle and Miller (2012), Adda et al. (2017), Leung et al. (2016) for assessments in the premium to delay. See also Budig and England (2001) and Kleven et al. (2019) for motherhood penalties.

[3]E.g., to study how fertility time horizons affect women's investments in human capital, age at marriage, and first birth, Gershoni and Low (2021b,a) leverage a policy reform that provided free IVF services in Israel.

[4]See the comprehensive review in van Loendersloot et al. (2010).

is validating the robustness of this gradient, as recent working papers (e.g Bensnes et al., 2023; Gallen et al., 2024), which acknowledge our precedence, have arrived at similar findings using alternative datasets.

Our results are robust to controlling for various factors, including age, year, health status, disease diagnoses, individual socioeconomic characteristics beyond education, infertility causes, and clinic fixed effects. The ability to control for clinics is particularly relevant because it rules out the possibility that the gradient in IVF results from patients sorting into different clinics. Such sorting might be influenced by unobserved clinic characteristics correlated with both individual education and IVF success. Therefore, we exploit only the variation in education across patients within clinics, who essentially share the same practitioners, protocols, equipment, and more. Another crucial aspect is financial constraints, potentially linked to education. Our analysis shows that although the education gradient is slightly reduced, it remains significant after controlling for household wealth, spouse's income, and education. Notably, Danish couples receive free fertility treatments in public clinics for the first three embryo transfer trials, and our results align when focusing on free trials within public clinics, indicating that insurance coverage has a minimal impact on the estimates.[5] Our results remain monotonously increasing in education if we include finer education specifications and is only slightly reduced and after including controls for medical degrees, and (un)healthy behaviors including BMI, smoking, and alcohol use. Further, the gradient becomes more pronounced after accounting for selection into treatment—higher participation of more-educated women. This suggests that high-school dropouts are more likely to be selected from the best part of their education-specific distribution of latent factors that affect their IVF success, relative to women with a college degree.

In addition, upon separately studying the stages of the IVF treatment, we also uncover a significantly positive education gradient in IVF success for each stage: aspiration (egg retrieval), embryo transfer, and live birth conditional on embryo transfer. For instance, among women who reached the embryo transfer stage, those with a college degree are 17.5% more likely to succeed than high school graduates and about 19.9% more likely than high school dropouts. One possible explanation is that only the best-quality embryo(s) are selected, and the quality itself might have been influenced in part by behavioral factors that are correlated with education.[6]

[5] In the United States, insurance mandates have been found to increase first birth rates for women over 35 (Schmidt, 2007) and have a positive effect on the utilization of infertility treatments, including IVF, for older and more educated women (Bitler and Schmidt, 2012). In our sample, 82% of first treatments are free of charge.

[6] The medical literature suggests that there are ways to enhance ovarian health and egg quality. Factors relevant to egg quality (and consequently embryo quality) include age, diet, BMI, hormonal issues, stress, smoking, and patient compliance with recommended gonadotropin dosages (Noorhasan et al., 2008).

**What explains the education gradient in IVF?** A useful approach to address this question is Grossman (2006), distinguishing "allocative efficiency" and "productive efficiency" mechanisms. While "allocative efficiency" suggests that more educated individuals make different health input choices due to varying prices and resource access, our analysis of IVF treatments in Denmark reveals an education gradient that persists even when examining variation *within* a clinic, where patients share the same resources (e.g. IVF protocol and equipment). This appears to counter the allocative efficiency argument. Instead, our findings support the concept of "productive efficiency," suggesting that highly educated women attain better health outcomes compared to their less-educated counterparts, even when encountering comparable prices and constraints. Here, we argue that the productive efficiency behind the education gradient in IVF success manifests in terms of latent factors, in particular, ability and psychological traits.

**Productive Efficiency** Completing an IVF cycle is undeniably challenging, demanding strict adherence to a rigorous schedule of appointments, tests, and medication (see Appendix A). In this context, highly educated individuals may excel in self-managing IVF inputs and better understand IVF technology, reflecting enhanced ability. This idea is consistent with Goldman and Smith (2002), emphasizing the crucial roles of self-management and treatment adherence in explaining the positive education gradient across health outcomes.[7] Recent research has uncovered a robust correlation between health literacy, treatment adherence, self-management, and cognitive abilities—all positively linked to Grade Point Average (GPA).[8] Hence, to delve further into our conjecture, we focus on a subset of younger high school graduates, for whom we observe their GPA. Our analysis reveals that the inclusion of GPA leads to a largely reduced and insignificant education gradient (i.e., the college coefficient), while GPA shows a significant positive effect. This suggests that GPA may serve as a more accurate proxy for measuring ability than educational achievement.[9]

[7]Previous related research has highlighted education disparities in the efficient adoption of health technologies, such as complex contraceptive methods (Rosenzweig and Schultz, 1989) or complex treatment regimens, e.g. antiretroviral therapy for HIV (Goldman and Lakdawalla, 2005). See also Cutler and Lleras-Muney (2008).

[8]See, for example, Berkman et al. (2011) and Miller (2016). More closely related with our work, ability, either proxied with education or more directly through Armed Forces Qualification Test (AFQT), has also been shown relevant in determining unintended pregnancies (Musick et al., 2009).

[9]Some works use variation in the outcome of IVF as an instrument for first birth, after controlling for education, to estimate the impact of a first child on female labor outcomes (see Lundborg et al. (2017); Bensnes et al. (2023); Gallen et al. (2024)). Our findings support that one should indeed control for education to decrease bias in the IV estimation. However, if available, using finer levels of controls for (unobserved) ability is warranted.

**Psychological Costs** Undergoing IVF is knackering and involves important psychological costs. In particular, failure in IVF can represent a crashing experience.[10] For these reasons, throughout the IVF process, it may make sense for some people to stop after failure.[11] In this regard, we further document that there is selection out of IVF based on education. Women with a college degree are less likely to discontinue treatment after failing an IVF cycle relatively to high-school dropouts. These disparities may partly reflect latent psychological traits such as grit and resilience, which are correlated with education, that allow more-educated women to manage better the stress from failure.[12]

**A Dynamic Model of IVF Dropouts with Latent Factors** To quantitatively assess the importance of the proposed latent factors, precisely, ability and psychological traits, we develop a dynamic model of women choosing when to stop IVF treatment. Women differ in ability (i.e. how well women follow the IVF protocol) and psychological traits (i.e. disutility associated with navigating the challenging IVF journey). In this framework, women utilize IVF technology in pursuit of their first child, with each treatment cycle representing a period. In the event of treatment failure, a woman faces a trade-off. She can choose to proceed with a new IVF cycle, with a positive probability of successfully achieving an IVF child. Success provides an additional stream of future utility starting in the next period. However, this success is weighed against the psychological cost associated with undergoing IVF, which reduces her current utility. Alternatively, if she decides to discontinue IVF, her life remains childless.

We structurally estimate the latent distributions of ability and psychological cost within education groups by matching the conditional success and dropout rates across education groups and treatments. The average ability within education groups is identified based on IVF success rates in the first cycle, while the variance is determined from the declining IVF success rates over cycles. This reflects the changing composition of ability, as high-ability women are more likely to succeed compared to low-ability women. Instead, the distribution of psychological cost is essentially estimated from drop-out rates following a failure across cycles. Ultimately, more-educated

[10] Assisted reproductive treatments, including IVF, are associated with increased stress, anxiety, and depression, especially after IVF failure, leading to higher psychotropic medication use (Goisis et al., 2023). Hjelmstedt et al. (2003), Joelsson et al. (2017), and Huang et al. (2019) document these treatments can deteriorate mental health.

[11] These insights are drawn from a dedicated discussion within The Weekend Intelligence podcast, focusing on the experiences of women who have undergone IVF. For example, the discussion emphasizes that IVF '*does not work for everyone and it comes at a cost and a toll, and to some people makes full sense to stop*' (quote). Source: *The Economist*'s podcast titled 'The hope and the heartbreak of IVF,' which was published in November 2023

[12] A bulk of recent studies in education and psychology show that grit and perseverance are predictors of success in a variety of areas, including college success, and of the achievement of long-term goals. Further, grit has shown incremental predictive value over and above cognitive ability for academic performance and retention in the military, the workplace, school and marriage. See Jachimowicz et al. (2018), Akos and Kretchmar (2017), Eskreis-Winkler et al. (2014), Duckworth et al. (2007) and references therein.

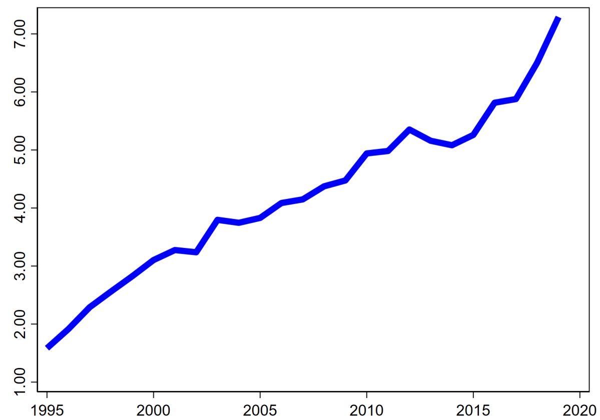
women exhibit a distribution of ability with a higher mean and lower variance compared to less-educated women. The distribution of psychological cost is more right-skewed for less-educated women. Clearly, for all education groups, the composition of types changes over the cycles due to self-selection, and both the average ability and psychological cost decline.

**Counterfactual Experiments and Policy Insights** Our counterfactual experiments imply that ability accounts for up to 87% of the difference in overall IVF success between college-educated women and women with less schooling.[13] On the other hand, the psychological cost primarily influences dropout patterns and, to a limited extent, success rates in the final treatment cycle. Due to the cost associated with continuing treatment, women with lower ability are more likely to drop out. In a counterfactual scenario where these costs decrease, women with lower productivity might opt for additional treatments. However, this could lead to declines in the overall success rate. Consequently, while policy interventions offering counseling for women undergoing IVF treatments are valuable, they may not necessarily enhance the success rate in the final treatment due to negative self-selection into further treatments.

**Further Discussion: Growing Inequality in the Battle Against Infertility** We further show a growing education gradient in IVF success over time, and assess with accounting exercises how such disparities influence the average age at first birth (AFB) and total fertility rates (TFR). In particular, highly educated women tend to have fewer children as they participate more in the labor force (Goldin and Olivetti, 2013; Olivetti, 2014; Goldin and Mitchell, 2017). In this context, we find that IVF women help mitigate the aggregate decline in TFR. Further, the fact that IVF disproportionately impacts the fecundity frontier for highly educated women, aligns well with the fact that the fertility rate gap across education groups (e.g. Jones et al., 2010) is partially reduced—in a purely accounting sense—with the rise of IVF. This raises the question of whether policies should address the IVF education gradient, given that the asymmetric battle against infertility simultaneously balances fertility rates across education groups.

Next, Section 2 describes the institutional setting and data. In Section 3 we pose our benchmark econometric specification and describe our results. We introduce the model, its estimation and counterfactuals in Section 4. Section 5 provides further discussion. Section 6 concludes.

Figure 1: IVF Children (%), Proportion of All Births: Denmark (1995-2019)




*Notes:* The proportion of IVF children (%) is computed as the number of births from the IVF register over number total births in the population in a given year.

## 2 Institutional Background and Data

First, we describe the Danish institutional setting related to the Law of Artificial Insemination, paying particular attention to couples' eligibility for subsidized treatment and their rights. Second, we discuss the Danish register panel data and provide details on the construction of our variable for live births from infertility treatments. Third, we discuss descriptive statistics of IVF in Denmark.

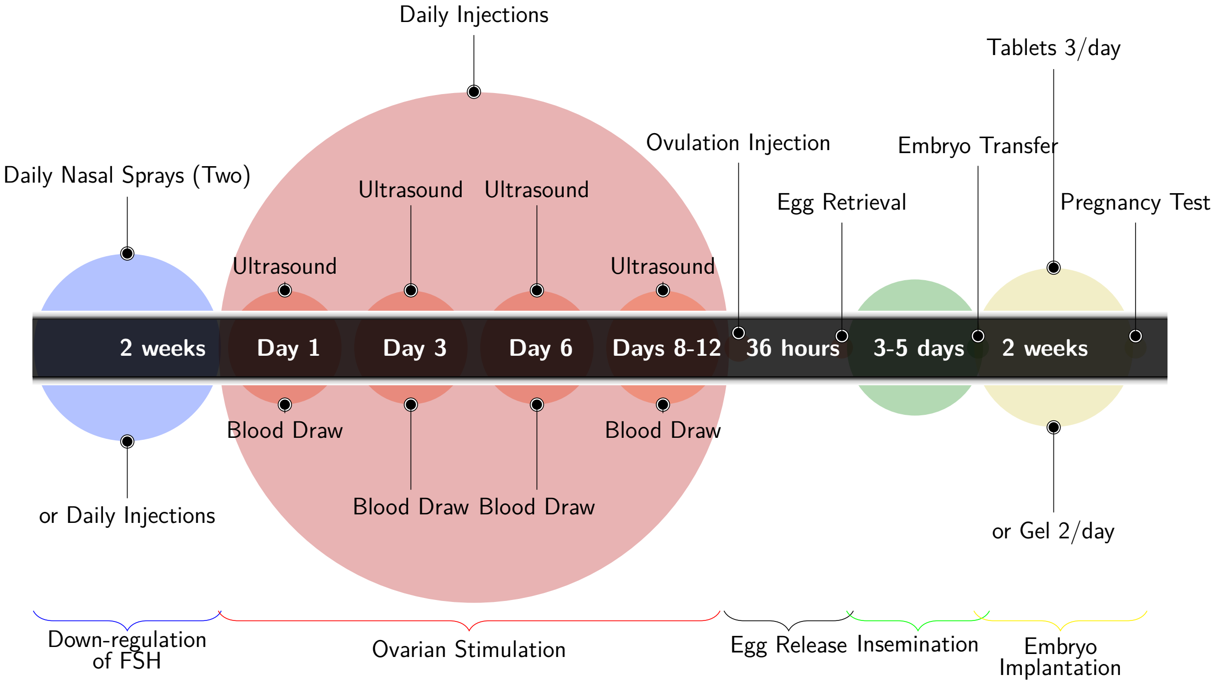
### 2.1 Institutional Background

During the entire time-span of our sample (1994-2009), Danish women had the right to artificial insemination by an in vitro method if they fulfilled the following three criteria (Ministry of Health, 1997, 2006): (i) the woman must be younger than 45 years of age at the beginning of a treatment period; (ii) the doctor needs the consent of both, the woman and the man in the couple, being treated; (iii) the couple must, in the doctor's opinion, be suitable to take care of a child, and the woman must be able (both mentally and physically) to undergo pregnancy. Further, the eligibility requirements for couples to receive *free* fertility treatments at a public hospital were the following: First, the couple was not allowed to have any joint children. Second, the couple should have attempted pregnancy naturally for at least 12 months; however, depending on the woman's age and the couple's medical history, the treatment could start earlier. Third, the woman must not be older than 40 years old at the beginning at the treatment.[14] Fourth, only the first three successful treatments were free. A treatment was considered successful if it transferred at least

[13] As we discuss in the paper, implementing mobile apps to remind patients about protocol timing and improving communication between doctors and patients can enhance IVF management and hence reduce the ability gap across educational groups. Additionally, medical advancements to reduce injections, blood draws, and frequent visits to fertility clinics can simplify the process of ovarian stimulation (Kushnir et al., 2022).

[14] Free IVF is available until age 40, with paid treatment permitted until age 45, making it illegal thereafter.

Figure 2: IVF Protocol




*Notes:* We construct this Figure to show the phases of the IVF protocol together with all the necessary treatments that are administered throughout the process based on the oldest fertility clinic in Denmark (Dansk Fertilitetsklinik, 2016). All injections displayed are self-administered by the patient. FSH is the follicle stimulating hormone.

one healthy embryo into the woman's womb. See Ministry of Health (2006, 2012, 2013) for details. In our sample, 83% of the couples who succeeded in having the first child did so within the allotted free treatments, and 76% were treated in the public sector.[15]

Treatments can terminate at any time during the IVF process for a number of reasons, including over-stimulation, cysts in the ovaries, no healthy eggs to retrieve, no fertilization of the eggs, or unsuccessful pregnancy after the embryo transfer. Sucessful IVF treatments involve not only medical interventions but an intensive amount of patient self-management, as its procedure is strict and complex. Clinics follow a standard protocol consisting of five phases: down-regulation of follicle stimulating hormone (FSH), ovarian stimulation, egg release, insemination, and embryo implantation; see Figure 2. We describe the IVF protocol in further detail in Appendix A.

[15]After the first three free treatments, additional IVF cycles must be pursued at a private fertility clinic, incurring costs. Nevertheless, the law permits the use of leftover frozen eggs from prior treatments at public hospitals for subsequent pregnancies, even exceeding the initial free trials. In certain regions, free treatment may be available when couples seek assistance for a second child (Danish Health Insurance, 2013)

## 2.2 The Administrative Danish Register Panel Data

We use unique administrative Danish register panel data from the entire Danish population from 1994 to 2009. We are interested in two specific sources of information. The first is the Danish National Board of Health, which contains detailed information about all women using in vitro techniques to achieve a pregnancy (e.g., date of cycle/treatment, reasons for undergoing treatment, if the treatment includes aspiration and/or transference, if the treatment results in a live birth). This is referred to as the IVF register, and includes the medical aspects of the individual fertility treatment histories. Precisely, for every treatment, the IVF register records when the treatment started, when the eggs were retrieved (aspiration), whether and when the embryo(s) were selected and then transferred to the woman's womb, and whether this resulted in a live birth, which is our measure of IVF success.[16] Before the start of the cycle height in centimeters and weight in kilos were recorded and women completed a questionnaire eliciting lifestyle information (e.g., number of cigarettes smoked per day and number of alcoholic beverages consumed per week). This information is available from 2006 onward. In addition, the medical records of patients at visits to any general practitioner (GP) contain information about the number of yearly services performed by the GP (e.g. consultation, blood test, vaccination, etc.), and the reimbursement (in Danish Kroner, DKK) that the GP receives from the state for the provided services. The medical records from visits to hospitals are grouped into diagnosis code by the main diagnosis of the ICD10 (10th revision of the International Statistical Classification of Diseases).

The second source is Statistics Denmark, which includes register data of annual information on socioeconomic variables (e.g., age, gender, education), income information (yearly income, earnings, and wealth), characteristics of employment (e.g., employed, self-employed, unemployed, out of the labor market), and general health information of the population. The IVF and the Statistics Denmark registers can be merged through a personal identifier. The data also includes a family-ID to link the individual to her spouse/cohabiting partner, children, parents, and other household members. Monetary variables are deflated using the Consumer Price Index. We classify patients into three mutually exclusive educational categories: (1) less than high school education, (2) high school and some college education, and (3) college and higher education.

[16] The IVF register records live births until 2005. To have a consistent measure of births over our sample period, we impute live births for the period 1995-2009 using birth records in the Statistics Denmark family register. For each fertility treatment, we calculate the expected birth date as the start date plus 280 days, checking for births within 90 days before or 21 days after. If a woman has multiple treatments less than 3 months apart, we classify the last treatment as a success and the others as failures. For the period 1995-2005 our imputed birth measure has a correlation coefficient of 0.96 with the recorded live births from the IVF register with most of the difference originating from more births being recorded in the family register.

### 2.3 IVF in Denmark: Stylized Facts

Our sample includes all Danish women 25 to 45 years of age with recorded education in our registers, who were married or cohabiting with a man, and did not have prior children before seeking the first treatment in 1994-2009. We consider this age group in order to ensure that they completed their education.[17] Since we want to follow women from their first IVF treatment, we exclude all women with treatments in 1994. With this restriction, we drop 1.7% of the treatments. This way, we can start counting initial treatments in 1995, and we classify a treatment as first if the woman did not receive any treatments in 1994.[18] In total, we follow 20,513 women who received their first treatment in both public and private clinics. Overall, there are 44,528 initiated treatments and 93% of them occur within the first five cycles. The median number is 2. The public sector performs 80% of these. We also consider a restricted sample of 16,900 women who were below 40 at the time of their first cycle at a public clinic and so qualified for free care.

Following the law, we consider a treatment to be still eligible for the free quota if the patient received less than three treatments reaching the stage of embryo implantation in a public hospital since the couple entered the sample, conditional on having no previous children. The free treatments can occur only at public clinics, and there is no reimbursement of the expenses if the couple undergoes an IVF treatment in the private sector. Note that the number of free cycles may exceed three, because this upper bound is conditional on a successful embryo transfer. After the initial free cycles, cycles should, according to the law, take place in private clinics.[19]

### 2.4 Characteristics of IVF Patients

In our sample, the majority of IVF patients have a high school degree or some college (51%), followed by college graduates or higher (35%).[20] Table 1 compares the socioeconomic and demographic characteristics of IVF patients by education groups, namely women who did not complete a high school degree (henceforth "$< HS$" or HS dropouts), high school graduates or some college (henceforth "*HS*"), and college graduates or higher (henceforth "*College*") seeking

[17] For the same reason, we also exclude individuals above 25 still in school.

[18] Our data show the stock of women in treatment during a year. To identify the first treatment, we assume if she enters a fertility treatment without a child, she either stops or continues with no break over a year.

[19] Less than 10% of all cycles in public hospitals cannot be classified as free. In most cases, these are patients who want to conceive a second child in vitro using the embryos frozen in previous cycles; see Section 2.1.

[20] In 1996, most IVF patients have a *HS* degree or some college (52%), but this declines to 43% in 2009. Conversely, in 2009, college graduates make up the majority of IVF patients (51%), a group representing just 22% in 1996. There is also a decrease in patients with less than a *HS* degree, accounting for 26% of total IVF patients in 1996 but only 6% in 2009. We include time dummies to capture this education variation across years.

Table 1: Demographic Characteristics of IVF Patients: First Treatment

| Education | < *HS* | *HS* | *College* |
|---|---|---|---|
| Age | 31.4 | 31.3 | 31.7 |
| Married (%) | 57.9 | 56.8 | 55.8 |
| Patient's income | 184,445 | 230,227 | 273,225 |
| Spousal income | 256,615 | 304,354 | 336,395 |
| Employment status (%): | | | |
| On leave | 1.8 | 0.7 | 0.3 |
| Self-employment | 2.3 | 2.7 | 1.3 |
| Employed | 73.0 | 89.3 | 94.4 |
| Out of labor force | 11.6 | 3.4 | 1.9 |
| Unemployed | 11.2 | 3.9 | 2.0 |
| Live births (%) | 21.4 | 25.5 | 27.3 |
| Sample (%) | 13.3 | 51.3 | 35.4 |
| Observations | 2,244 | 8,670 | 5,986 |

*Notes:* In terms of education groups, we denote IVF patients with less than high school as < *HS*, high school or some college as *HS*, and college or higher degree as *College*. Income is in DKK, deflated by CPI to year 2000. Employment status and income are measured the year prior to treatment. The sample consists of women below 40 in their first treatment who got treated in a public clinic ($N = 16,900$).

the first free IVF treatment ($N$ = 16,900).[21] Regardless of educational attainment, a woman starting IVF is on average a little older than 31 years old, and women in the *College* group are only a few months older when they are first treated.[22] The less educated ones were slightly more likely to be married (57.9%) than those with a *HS* degree or *College*. On the contrary, we find significant differences in both individual and spousal income across education groups. Most of the sample was employed, with employment rates ranging from 73% to 94% from the bottom to the top of the education groups, respectively.

The success rate measured by the fraction of IVF live births in the first (free) IVF treatment is 25.6% for the entire population, 21% for the < *HS* group, 25% for the *HS* group, and 27% for the *College* group. These figures are similar in the larger sample of women treated either in public or private clinics ($N$ = 20,513): 21% for the < *HS* group, 25% for the *HS* group, and 26% for the *College* group (Online Appendix Table A1). In this sample, there are relatively more college-educated women, 36.7%, versus 35.4% in the benchmark sample. Further, the educational disparities in IVF success are still present in subsequent treatments (middle panel, Online Appendix Table A2) and the difference is sharper when we look at the last treatment

[21]We choose to divide education into the three categories and specifically have a "less than high school" educational category, because 9th grade is the highest mandatory grade in Denmark. In 2009, the last year of our sample, 18.4 percent of of the Danish population (not just the ones seeking IVF), ages 25 to 45 had less than a high school degree and in 1995, the first year in our sample, this share was 28.1 percent. Thus, since we are interested in educational disparity, we want to show this relatively large share of the population separately. Further, we believe we align with the literature by using the "less than high school" as a base educational category, for example, in a review of facts about human capital, Deming (2022) summarizes results by educational categories that are presented relative to the less than high school category.

[22]On average, women with at least a master degree are 32.5 y.o., whereas college graduates are 31.4 y.o.

overall: 49%, 58%, and 62% for the < *HS*, *HS*, and *College* groups, respectively. There are also marked differences in drop-out rates after a failure (bottom panel, Online Appendix Table A2).

The majority of treatments is provided by the public sector in Denmark, about 82%, and it moderately decreases with age: 87% for women 25-29 years old, 85% for women 30-34 years old, and 77% for women 35-40 years old (see Online Appendix Table A3). This holds true for the first cycle, all eligible free treatments, for those seeking the first pregnancy (panel (a), Online Appendix Table A3). We assess IVF success rates at both the treatment and clinic levels. In the former, high-volume clinics are emphasized (panel (b), Online Appendix Table A3), while the latter assigns equal weight to all clinics (panel (c), Online Appendix Table A3). In both cases, the overall success rate is higher in the public sector than in the private sector (25.6% vs. 21.2% at the treatment level and 25.8% vs. 17.8% at the clinic level). Since older women were somewhat more likely to be treated in the private sector, we also examine differences by age. At the treatment level, we find that the success rate is similar across sectors for mid-range age groups, and higher in the public sector for younger and older women (29% vs. 26%), with the reverse true for older women (19% vs. 15%). Similar insights arise at the clinic level.

Finally, there are also educational disparities in infertility causes, medical conditions, and health behavior among women seeking the first IVF treatment. Online Appendix Table A4 reveals that <*HS* women showed a higher incidence of fallopian tube defects (37% compared to 21% in the *College* group), while *College* women were more likely to diagnosed with male causes, other medical causes, and unspecified causes, possibly influenced by their slightly higher age at entry. From 2006, our data contains information about BMI, smoking and alcohol use. The *College* group shows lower BMI and smokes fewer cigarettes in the year before the treatment compared to the lower educated groups. In contrast, the *College* group tends to consume more alcohol prior treatment than lower educated women.[23]

# 3 The IVF-Education Gradient

## 3.1 Empirical Strategy

As documented above, less and more educated women are different in a number of individual dimensions such as age, health status, socio-economic characteristics, infertility causes and, to some extent, sort into different sectors. Thus, an immediate concern is that educational disparities in IVF outcomes may partially reflect these differences. To account for the influence of these

[23] Precisely, the proportion that reports not smoking any cigarettes is 63% for the < *HS* group, 77% for the *HS* group, and 84% for the *College* group. The proportion of individuals that reported not drinking alcohol is 51% for individuals with less than a high school degree, 42% for the *HS* group and 37% for the *College* group.

systematic differences in individual and clinic characteristics, we estimate the following equation:

$$b_{ijht}^{IVF} = cons + \sum_{s>0} \alpha_s \mathbf{1}_{s_i} + \eta x_{it} + \beta_j + \gamma_h + \varepsilon_{it}, \tag{1}$$

where $b_{ijht}^{IVF}$ is a dummy variable equal to 1 if a live birth is attained with the IVF treatment $t$ for woman $i$ in the clinic $h$ at year $j$. Let $s_i$ be our measure of educational attainment for woman $i$, namely, $<$ *HS*, *HS* and *College*. Then, the coefficient $\alpha_s$ captures the education gradient.

The vector $x_{it}$ captures a benchmark set of controls which includes age dummies, time dummies, health status indicators (average number and cost of GP services, disease diagnoses), socioeconomic characteristics (marital status, log patient's income, log spousal income, employment status), infertility causes (cervical defect, ovulation defect, fallopian tube defect, male causes, other causes, unspecified causes), and clinic fixed effects. In order to avoid reverse causality issues, we use variables measured when entering the first treatment measured the calendar year before entering the first treatment, as undertaking IVF may indeed affect both employment status and income, as well as costs of GP services. We capture time-fixed effects with $\beta_t$ and clinic-fixed effects with $\gamma_h$. Last, $\varepsilon_{it}$ reflects heteroskedastic robust standard errors $N(0, \sigma^2_{\varepsilon,i})$.

While focusing on the sample of fully subsidized IVF treatments for childless couples at a fertile age helps address the potential selection problem based on education, there remains a possibility of a positive education gradient in IVF resulting from patients selecting different clinics. This could be influenced by unobserved clinic characteristics correlated with both individual education and IVF success.[24] To address this issue, we control for clinic-fixed effects and, hence, exploit only the variation of education across patients within a clinic sharing the same practitioners, protocol descriptions and equipment. Before presenting our estimation results, we assess the remaining variation in patients' educational attainment after removing clinic and year fixed effects. The mean and standard deviation are 13.5 and 2.3, respectively, and post-removal, more than 90% of the variation is retained (from 2.28 to 2.18), ensuring the precision of our estimates.

### 3.2 Results

First, we estimate a linear probability model (LPM) using information on all women 25 to 45 years of age who underwent an IVF process, as specified in equation (1). We start by reporting

[24] One concern is the presence of disparities in the amount of available resources across education groups. If the higher-educated mothers have access to better IVF technology, they will have better success in attaining a live birth simply due to higher productive efficiency. This argument can be stated in general for a wide set of medical technologies (e.g. Grossman, 1972; Kenkel, 1991; Thompson et al., 2008; Cutler and Lleras-Muney, 2012).

in Table 2 the results for the first cycle women underwent, which might be viewed as the cleanest because of possible attrition in repeated cycles.

The outcome is clear: There is a large and highly significant education gradient in IVF success rates in all specifications. Without controls, *College* women are associated with a 5.89 percentage-point (p.p) higher probability of attaining a live birth compared to the low educated ones (column 1, Table 2). When we add age, the education gradient slightly increases for *College* graduates. They might be the ones who are more likely to delay childbearing for career concerns and enter into treatment at an older age that, consistent with the medical literature, is an important determinant of IVF success in all specifications in Table 2. As women age, the probability of a successful IVF cycle decreases. As expected, the education gradient slightly decreases for both education groups when we include further controls, in addition to age: health status and year fixed effects (Column 3); and marital status, socio-economic status (i.e., both female and spousal incomes), and employment status before entering the first IVF treatment (Column 4).[25]

The Jones Institute at the Eastern Virginia Medical School, where the first US IVF baby was born in 1981, lists key factors influencing IVF outcomes: age-related ovarian reserve, uterine normalcy, semen quality, in vitro fertilization and cleavage success or failure, and the number of embryos transferred and cryopreserved, which increases the total reproductive potential of an IVF cycle.[26] Accordingly, our benchmark specification controls for infertility causes, in addition to clinic-fixed effects (column 5, Table 2). The education gradient slightly decreases with the augmented controls suggesting that the additional controls capture some of the variation in IVF success. Notably, in terms of infertility causes, a defect of the Fallopian tube significantly reduces the probability of IVF success by 2.93 percentage points; and for those patients with "other causes" of infertility this probability drops by 2.92 percentage points.[27]

To be precise, in our benchmark specification that conditions on the full set of controls (Column 5, Table 2), the *HS* group shows a 3.13 percentage-point higher probability of achieving

[25]To control for health status, we also incorporate pre-existing medical conditions, such as the number of GP services in the past year, the total medical expenditures associated with GP services, and whether the patient has been diagnosed with any serious disorder or disease before undergoing an IVF treatment (see section 2.4). Whether we used a dummy for each diagnosis given at hospitals during the year prior to treatment, or an indicator equal to 1 if the patient has been diagnosed with any disease does not alter the robustness of the results.

[26]See the information from the Jones Institute at the Easter Virginal Medical School here.

[27]Note that, as discussed in Section 2.4, while Fallopian tube as cause of infertility is more prevalent in the $<$ *HS* group, "other causes" is more prevalent in the *College* group, and might be related to the quality of the eggs because these patients are slightly older when they enter into treatment. Further, the infertility causes may also partially capture differences in underlying fecundity across education groups: Highly educated women seeking IVF may have been trying to get pregnant for shorter time than a less educated women of the same age seeking IVF treatment, resulting in the higher educated women selecting into IVF having an underlying higher fecundity than the lower educated women selecting into IVF.

a live birth, while the *College* group shows a higher probability of 5.37 percentage points. To contextualize, the average chances of live birth in the first cycle for the $< HS$ women is 21.4% (see Section 2.4). Hence, the gradient translates to a 25.1%(=5.37× 100/21.4) higher chance for college graduates and 14.6% for women with high school degrees or some college, compared to those without a high school degree. Analogously, resetting the reference group to *HS* women, the *College* group is 9.0%(=2.3× 100/25.5) more likely to succeed than the *HS* group with the two groups being statistically significantly different from each other.[28,29]

[28] See Online Appendix Table A8.

[29] In our analysis, we primarily used a linear probability model (LPM). Results from a logit model show a similar education gradient. For instance, the average marginal effect estimates are 5.39 and 3.15 for the *College* and *HS* groups, respectively (Column 5, Online Appendix Table A6).

Table 2: Education Gradient in IVF Success (Live Births)

| *IVF Live Births* | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| High School | 0.0416*** | 0.0392*** | 0.0361*** | 0.0326*** | 0.0313*** |
| | (0.00984) | (0.00981) | (0.00999) | (0.0102) | (0.0102) |
| College | 0.0589*** | 0.0619*** | 0.0589*** | 0.0545*** | 0.0537*** |
| | (0.0104) | (0.0104) | (0.0108) | (0.0111) | (0.0111) |
| Age Dummies | | ✓ | ✓ | ✓ | ✓ |
| Time Dummies | | | ✓ | ✓ | ✓ |
| Health Status: | | | | | |
| Average number of GP services | | | -0.00118 | -0.00140 | -0.00158 |
| | | | (0.00135) | (0.00135) | (0.00135) |
| Average cost of GP services | | | -7.94e-08 | -4.73e-08 | -5.54e-08 |
| | | | (1.70e-07) | (1.71e-07) | (1.71e-07) |
| Disease(s) Diagnoses | | | ✓ | ✓ | ✓ |
| Socioeconomic Characteristics: | | | | | |
| Married | | | | -0.0142** | -0.0119* |
| | | | | (0.00687) | (0.00688) |
| Log total patient's income | | | | 0.00656** | 0.00669** |
| | | | | (0.00301) | (0.00301) |
| Log total spousal income | | | | 0.00236 | 0.00244 |
| | | | | (0.00333) | (0.00330) |
| Employment status: | | | | | |
| On leave | | | | -0.0183 | -0.0199 |
| | | | | (0.0410) | (0.0413) |
| Self-employment | | | | 0.0214 | 0.0187 |
| | | | | (0.0282) | (0.0279) |
| Employed | | | | -0.00866 | -0.0113 |
| | | | | (0.0169) | (0.0169) |
| Out of labor force | | | | -0.0284 | -0.0268 |
| | | | | (0.0231) | (0.0231) |
| Infertility causes: | | | | | |
| Cervical defect | | | | | -0.00539 |
| | | | | | (0.0292) |
| Ovulation defect | | | | | -0.0141 |
| | | | | | (0.0127) |
| Fallopian Tube defect | | | | | -0.0293** |
| | | | | | (0.0118) |
| Male causes | | | | | -0.0131 |
| | | | | | (0.00954) |
| Other causes | | | | | -0.0292** |
| | | | | | (0.0118) |
| Unspecified causes | | | | | -0.0149 |
| | | | | | (0.0128) |
| Clinic Fixed Effects | | | | | ✓ |
| Constant | 0.214*** | 0.279*** | 0.279*** | 0.191*** | 0.224*** |
| | (0.00866) | (0.0173) | (0.0226) | (0.0575) | (0.0594) |
| Observations | 16,900 | 16,900 | 16,900 | 16,900 | 16,900 |
| R-squared | 0.002 | 0.010 | 0.014 | 0.015 | 0.022 |

*Notes*: Robust standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, *$p < 0.1$. HS denotes high school. Employment status reference category is "unemployed". Monetary variables are expressed at year 2000 price levels deflated using the Danish Consumer Price Index. All specifications are run on the first treatment. The omitted category is HS dropouts. The sample consists of women below 40 in their first treatment who got treated in a public clinic ($N = 16,900$) in their first treatment to get the first child.

Our results do not rely on the education measure. Even with finer educational groups or linear years of education (Online Appendix Table A5), an IVF success gradient persists. High school grads have a 3.1 p.p. higher success chance than dropouts, while college grads and Ph.D. holders show even higher advantages, 4.9 p.p. and 6.7 p.p., respectively (column 1). This implies a 22.8% higher success for college women and 32.2% for Ph.D. women compared to high school dropouts. Each additional education year associates with a 0.67 p.p. higher IVF success (column 2). The education gradient also remains significant although decreases when separately including medical degrees of IVF women and their partners (Online Appendix Table A7).

**IVF Stages within the Cycle** A successful live birth in IVF relies on the successful completion of three stages : aspiration (egg retrieval), embryo transfer (contingent on healthy embryos), and pregnancy.[30] Factors affecting each stage involve women's health behavior (prior to and during the IVF), adherence to IVF protocol, and embryo quality (only the best quality embryo(s) are selected), potentially linked to education.[31] For example, the medical literature suggests factors that are relevant for egg quality, and in turn for the embryo, are age, diet, BMI, hormonal issues, stress, alcohol consumption and smoking.[32] Then, to examine determinants of failure at specific IVF stages—lack of retrieved eggs, healthy embryo transfer, or post-transfer miscarriage——we conduct separate regressions by stage of IVF, conditional on reaching the stage. Further, our analysis of live births incorporates indicators for the number of embryos implanted.

Results in Table 3 show a positive educational gradient across all IVF stages, especially in the second and third stages. The *HS* group is 1.01 p.p. more likely to undergo egg extraction than dropouts, rising to 2.03 for the *College* group. Post-successful aspiration, the *HS* group has a 3.01 p.p higher chance of achieving a healthy embryo than dropouts, compared to 3.71 for the *College* group. Remarkably, the estimated coefficients for the live birth gradient, conditional on embryo transfer and controlling for embryos implanted (columns 3 and 4, Table 3), mirror those in our benchmark specification.[33] The probability of success, conditional on an embryo transfer, is 27.3% for the reference $<$ *HS* group—and 31.0% overall. Precisely, the *College* group is

[30]Detailed stages of the IVF treatment are in Appendix A.

[31]For instance, using data from three clinics in the greater Boston area, Mahalingaiah et al. (2011) document that patients with a graduate school education have statistically significantly higher peak estradiol levels than patients without a college degree, which in turn affect the odds of cycle cancellation before egg retrieval.

[32]Shah et al. (2011) shows that obesity is associated with fewer normally fertilized oocytes, lower estradiol levels, and lower pregnancy and live births. Rossi et al. (2011) show that four alcoholic drinks per week is associated with a decrease in IVF live births, after controlling for cycle number, cigarette use, body mass index, and age.

[33]The fact that we find an education gradient in births conditional on embryo transfer differs from the small and insignificant effect of years of education on IVF births reported in Lundborg et al. (2017), see their Online Appendix, Table 1. We have tried, and failed, to replicate their results. See our detailed discussion in Groes et al. (2024) here.

Table 3: Education Gradient in IVF Success (Live Births): Different Stages of the IVF treatment

| *Outcome*: | Aspiration (1) | Embryo Transfer (2) | Live Birth (3) | Live Birth (4) |
|---|---|---|---|---|
| High School | 0.0101* | 0.0301*** | 0.0309** | 0.0298** |
| | (0.00563) | (0.00968) | (0.0124) | (0.0124) |
| College | 0.0203*** | 0.0371*** | 0.0544*** | 0.0540*** |
| | (0.00581) | (0.0104) | (0.0134) | (0.0134) |
| Full Controls | ✓ | ✓ | ✓ | ✓ |
| 2 Embryos Trans. | | | | 0.124*** |
| | | | | (0.00887) |
| ≥3 Embryos Trans. | | | | 0.0397 |
| | | | | (0.0282) |
| Constant | 0.851*** | 0.740*** | 0.328*** | 0.235*** |
| | (0.0321) | (0.0579) | (0.0721) | (0.0727) |
| Observations | 16,900 | 16,204 | 13,783 | 13,647 |
| R-squared | 0.043 | 0.024 | 0.024 | 0.037 |

*Notes:* Robust standard errors are in parenthesis. Observations in Columns (3) and (4) are treatments that reached the embryo implantation stage. The sample consists of women below 40 in their first treatment who got treated in a public clinic ($N = 16,900$) in their first treatment to get the first child. The corresponding table with $HS$ as the baseline group can be found in the Online Appendix Table A9.

19.9%=5.44× 100/27.3 and the *HS* group is 11.3%=3.09×100/27.3 more likely to attain a live birth than the $<$ *HS* group. Child birth after embryo transfer does not seem purely idiosyncratic.[34]

## 3.3 Robustness

Here, we assess the robustness of the education gradient in IVF success to the health behavior, financial constraints, maternal occupation and spousal education.[35]

**(Un)Healthy Behavior** Behavioral information on smoking, alcohol consumption, and BMI is available only for 2006-2009. Therefore, we initially replicate our results for this time span (Column 2, Table 4). It's noteworthy that the education gradient in Column 2 is twice as large as in Column 1, indicating an increase in educational disparities over time. Comparing these with results including BMI, smoking, and alcohol, the education gradient decreases from 6.18 to 5.89 p.p. for the *HS* group and from 9.10 to 8.51 p.p. for the *College* group (Column 3).[36]

**Financial Constraints** In column 4 of Table 4, we directly control for household wealth, observing a slight decrease in the IVF education gradient. Nevertheless, the gradient remains large and positive, as wealth correlates positively with IVF success. Expanding the sample to all treat-

[34] The unsuccessful birth can either be due to the failure of the embryo implantation after transfer or a miscarriage after pregnancy.

[35] To see further results for success rate for IVF treatment for the second child see (Groes et al., 2017).

[36] All specifications include an indicator for missing information on BMI, smoking and alcohol consumption.

Table 4: Robustness

| *IVF Live Births* | (1) Benchmark | (2) 2006-2009 | (3) Behaviors 2006-2009 | (4) Wealth | (5) All women/clinics | (6) Occupations Sample | (7) Occupations | (8) Spousal educ. |
|---|---|---|---|---|---|---|---|---|
| High School | 0.0313*** | 0.0618** | 0.0588** | 0.0308*** | 0.0299*** | 0.0283*** | 0.0234** | 0.0283*** |
| | (0.0102) | (0.0243) | (0.0245) | (0.0102) | (0.00932) | (0.0107) | (0.0112) | (0.0104) |
| College | 0.0537*** | 0.0910*** | 0.0851*** | 0.0529*** | 0.0488*** | 0.0500*** | 0.0422*** | 0.0441*** |
| | (0.0111) | (0.0246) | (0.0251) | (0.0111) | (0.0100) | (0.0116) | (0.0138) | (0.0117) |
| Benchmark controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| BMI | | | ✓ | | | | | |
| Smoking | | | ✓ | | | | | |
| Alcohol | | | ✓ | | | | | |
| Wealth | | | | 1.85e-08* | | | | |
| | | | | (9.68e-09) | | | | |
| Occupation FE | | | | | | | ✓ | |
| High School, Partner | | | | | | | | 0.0136 |
| | | | | | | | | (0.00939) |
| College, Partner | | | | | | | | 0.0306*** |
| | | | | | | | | (0.0116) |
| Constant | 0.224*** | 0.314** | 0.352** | 0.232*** | 0.230*** | 0.238*** | 0.258*** | 0.210*** |
| | (0.0594) | (0.123) | (0.132) | (0.124) | (0.0517) | (0.0628) | (0.0632) | (0.0609) |
| Observations | 16,900 | 4,673 | 4,673 | 16,900 | 20,513 | 15,737 | 15,737 | 16,672 |
| R-squared | 0.022 | 0.024 | 0.025 | 0.022 | 0.028 | 0.022 | 0.024 | 0.023 |

*Notes*: Robust standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, *$p < 0.1$. HS denotes high school. Employment status reference category is "unemployed". All specifications are run on the first treatment. The omitted category is HS dropouts. The main sample considered are women below 40 treated in public clinics ($N = 16,900$) in all columns except column (5) where the sample is all women/all clinics ($N = 20,513$). The corresponding table with $HS$ as the baseline group can be found in the Online Appendix Table A10.

ments (not just those eligible for free treatment), we find a significantly positive IVF-education gradient of respectively 0.14 and 0.49 percentage point smaller magnitude for high school and college educated. This adjustment likely stems from the inclusion of ineligible women over 40, primarily highly educated, with lower fertility rates.

**Occupations** One might assume that certain working conditions, like shift work, long hours, lifting, standing, and physical exertion, could affect miscarriage rates. Hence, we replicate results for employed women (column 6; subsample of 15,737 women).[37] When controlling for occupations in column 7 (1-digit ISCO classification), the gradient slightly decreases but remains significant. On average, a College (HS) graduate sees a 4.22 p.p. (2.34) increase in IVF success.

**Spousal Education** Spousal education may impact a woman's education gradient in IVF success. When we control for spousal education, we find that the women education gradient remains significant (Column 8). The spousal education gradient is also highly significant for college graduates. On average, women with a college graduate spouse have a 3.1 p.p. higher chance of IVF success compared to those with a HS dropout spouse. This suggests that spousal education

[37]The gradient in Column 6 is slightly smaller than in Column 1.

influences outcomes beyond family income and wealth. These results hold up when considering whether the spouse holds a medical degree (Columns 4 and 5 in Online Appendix Table A7).

### 3.4 Selection into IVF Treatment

An alternative potential reasoning behind the IVF success is selection intro treatment. In particular, not all infertile women who seek to have a child decide to undertake IVF treatment, even though the first treatments are free of charge. For this reason, we study the determinants of IVF success with a standard two-stage Heckman selection model that incorporates an IVF participation equation, where we exploit the fact that there is variation in number of IVF clinics by municipalities in Denmark—*ceteris paribus*, differences in the decision to participate in IVF treatment can be generated by differential access to IVF clinics; these transit costs can be of monetary and/or psychological type, or proxy for the opportunity cost of time. Specifically, we consider the number of clinics offering IVF treatments per woman age 25 to 45 in each municipality.

Our main result is that the IVF-education gradient becomes even steeper when we take the IVF participation decision into account, see Column 1 in Table 5. IVF patients with a high school degree have a 4.47 p.p. higher probability of a live birth than high school dropouts, and IVF patients with a college degree have a 7.60 p.p. higher probability than high school dropouts. Recall that in the benchmark case these figures were 3.13 p.p. for high school graduates and 5.37 p.p. for college graduates. Further, we find an education gradient in terms of IVF participation (Column 2). As anticipated, participation increases with the presence of municipal IVF clinics relative to the number of women.[38]

[38] The number of clinics per women is significant at the 5 percent level, leading to an F-test of 5.8, which is less than the usual threshold of an F-test of 10. However, the fact that we observe an education gradient in IVF the participation aligns well with our summary statistics and the next Section 3.5 that analyses, which women continue with IVF after a fail attempt. The results from estimating our the model in section 4 also increased selection into staying with the IVF treatments of high education women, leading the education gradient in success rate to be higher if we had the same selection of women from all educational categories.

Table 5: Education Gradient in IVF Success (Live Births): Selection into Treatment

| | IVF Live Births (1) | IVF Participation (2) | Mills (3) |
|---|---|---|---|
| HS and some college | 0.0447* | 0.0804*** | |
| | (0.0243) | (0.00984) | |
| College and higher degree | 0.0760* | 0.149*** | |
| | (0.0425) | (0.0105) | |
| Age Dummies | ✓ | ✓ | |
| Time Dummies | ✓ | ✓ | |
| Health Status: | | | |
| Average number of GP services | -0.000298 | 0.0118*** | |
| | (0.00345) | (0.00101) | |
| Average cost of GP services | -4.01e-08 | -7.55e-07*** | |
| | (2.61e-07) | (1.32e-07) | |
| Disease(s) Diagnoses | ✓ | ✓ | |
| Socioeconomic Characterisitcs: | | | |
| Married | 0.0272 | 0.224*** | |
| | (0.0616) | (0.00642) | |
| Log total income | 0.00639 | 0.0115*** | |
| | (0.00445) | (0.00317) | |
| Log spousal income | 0.00739 | 0.0259*** | |
| | (0.00781) | (0.00327) | |
| Employment status: | | | |
| On leave | -0.0208 | 0.0158 | |
| | (0.0389) | (0.0389) | |
| Self-employment | -0.00131 | -0.00473 | |
| | (0.0248) | (0.0245) | |
| Employed | 0.000429 | 0.0849*** | |
| | (0.0279) | (0.0155) | |
| Out of labor force | -0.0622 | -0.114*** | |
| | (0.0386) | (0.0223) | |
| $\lambda$ | | | 0.223 |
| | | | (0.317) |
| IVF Clinics per Women (by Municipality) | | 111.6** | |
| | | (46.37) | |
| Constant | -0.529 | -3.075*** | |
| | (1.051) | (0.0572) | |
| Observations | 919,808 | 919,808 | |

*Notes:* The sample consists of the total number of childless women between ages 25 and 45 who are married or cohabitating with a spouse. Robust t-statistics in parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Table 6: Education Gradient in IVF Success (Live Births) by Treatment

| *Treatment*: | 1 | 2 | 3 | 4 | 5 | Last |
|---|---|---|---|---|---|---|
| High School | 0.0299*** | 0.0195* | 0.0223* | 0.0385** | 0.0190 | 0.0713*** |
| | (0.00932) | (0.0107) | (0.0126) | (0.0152) | (0.0206) | (0.0108) |
| College | 0.0488*** | 0.0425*** | 0.0496*** | 0.0322* | 0.0449** | 0.124*** |
| | (0.0100) | (0.0116) | (0.0138) | (0.0167) | (0.0228) | (0.0114) |
| Full controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Constant | 0.230*** | 0.107* | 0.238*** | 0.194 | 0.188 | 0.639*** |
| | (0.0517) | (0.0635) | (0.0858) | (0.147) | (0.277) | (0.0633) |
| Observations | 20,513 | 13,326 | 8,653 | 5,251 | 3,024 | 20,513 |
| R-squared | 0.028 | 0.029 | 0.033 | 0.025 | 0.048 | 0.114 |

*Notes*: Results from a linear probability model for success in each of the first 5 treatments and in the last treatment. Robust standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, *$p < 0.1$. The omitted category is HS dropouts. The sample consists of all women/all clinics ($N = 20,513$).

The steeper education gradient in IVF success, post-selection control, suggests non-random participation from each education group. For example, patients might select into IVF based on latent characteristics that influence success. In particular, high school dropouts that enter IVF treatment are likely drawn from better segments of their group-specific distribution of latent factors, relative to women with a college degree that encompass a broader range of latent factors due to higher participation probability. Without accounting for selection, the education gradient in success serves as a lower bound compared to findings when selection is considered.

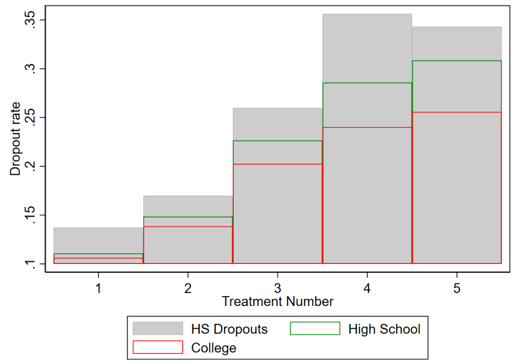
### 3.5 By Treatment Number and Overall IVF Success

One might ultimately be interested in the overall success rate, regardless of whether it occurs at the first or at a later cycle. In this pursuit, we study IVF success at cycle number $t$ given that all cycles before $t$ failed using the sample of all IVF women (i.e., our larger sample).[39] We show results up to cycle $t = 5$ in Table 6. The education gradient with respect to the $<$ *HS* group is significantly positive in each cycle and stronger for the *College* group than the *HS* group. Notably, even in the fifth cycle, the *College* group is 4.49 p.p. more likely to succeed than $<$ *HS*, a figure that is similar to that of the first treatment 4.48. That is, our results suggest that less-educated patients do not catch up after repeated cycles.

It is likely that the gradient in IVF success at each treatment $t + 1$ is affected by the selective attrition occurring in $t$—individuals exiting treatment in $t$ due to success or dropouts, as showed by the reduction of observations across cycles. Indeed, we also find that education helps determine selection out of IVF after a failure, see Figure 3. For instance, the *HS* group is 2.65 p.p. less likely

[39] The probability of success decreases with the number of cycles for all education types. Overall, this probability drops from 24.7% in the first cycle to 16.67% in the fifth cycle.

Figure 3: IVF Dropout after a failure By Education Group




*Notes:* This figure displays the probabilities of drop-out for women who do not achieve a live birth and discontinue treatment.

to drop out from IVF than the $< HS$ group after the first cycle. The corresponding figure for the *College* group is 3.1 p.p.[40] This gap is also present in all subsequent cycles, and if anything it becomes sharper. Interestingly, the disparity between high school and college women becomes more pronounced over the cycles. Hence, dropouts can endogenously affect the distribution of women—based on education and associated latent factors, that moves across cycles.[41] This idea is corroborated by the large and significant education gradient in IVF that we find by estimating the gradient using only women that either succeed or dropout in the last cycle they undertook, independently of which number it is (see "Last" Column). Precisely, the overall success in IVF is such that the *HS* group and the *College* group significantly outperform the $< HS$ group by, respectively, 7.13 p.p. and 12.4 p.p.[42]

### 3.6 What Can Explain the Remaining Gradient?

So far, we have shown that there is a strong and significant maternal educational gradient in IVF success (and dropouts) which remains significant after controlling for eligibility to free treatments, individual socio-economic and medical conditions before treatment (e.g., age, marital status, labor market status, infertility causes, etc.) and sorting into clinics. We have also ruled out a set of potential mediators as the main explanation for the gradient, such as (un)healthy behavior, financial constraints, occupation choices, spousal characteristics. Further, the gradient stands after controlling for potential selection into IVF treatment.

[40] These numbers are very similar to the ones obtained from the estimation of drop-out rate after the first failure using all baseline controls: 2.5 and 3.8. p.p., respectively, and both significant at 1%.

[41] In Section 4, we develop a dynamic model of IVF dropouts across cycles in order to assess how this endogenous attrition contributes to the educational disparities in the IVF success. Specifically, we will uses the documented education disparities in dropout rates by cycle to, partly, identify potential latent heterogeneity in psychological costs within and across education groups.

[42] It's worth noting that the sample for this overall success (Last" Column), by construction, includes all women who enter IVF treatment and exit—thus, the observations of Column 1 are identical to those of Column "Last."

In this context, one might argue that education captures individual cognitive abilities, including memory, organizational skills, and health literacy, which have been shown to affect medication adherence and self-management, thereby contributing to achieving favorable health outcomes (Goldman and Smith, 2002).[43] This argument echoes the concept of "productive efficiency," in opposition to an argument of "allocative efficiency" described by Grossman (2006). The latter refers to the notion that individuals with higher levels of education choose different health inputs because they face different prices and have access to different resources and technologies. However, in the case of IVF, we observe an education gradient even when exploiting variation within a clinic, where patients face the same IVF inputs, equipment, and a strict protocol, seemingly ruling out the allocative efficiency argument. Instead, our results point in the direction of "productive efficiency": highly educated individuals have better health outcomes than lower-educated ones, even when they all face the same prices and constraints. Furthermore, the rigors of completing an IVF treatment entail strict adherence to an intense schedule of appointments, blood tests, ultrasound tests, and procedures, as well as patient compliance with medications (see Appendix A). Therefore, a better ability to self-manage the IVF procedure can yield higher IVF "productive efficiency" among more educated individuals. Building on these considerations, we proceed to explore the relationship between a proxy for cognitive ability (potentially mixed with psychological traits), such as grade point average (GPA), and a woman's performance during her final IVF treatment, which could be any of the treatments (up to five).

GPA is recorded only for women in recent cohorts (mostly women below 35 years old) with at least a high school degree. This restriction more than halves the initial sample size and high-school graduates are our base category now. As a reference point, in Column (1) of Table 7 we first show that, despite the fact that we lose variation in education, college graduates are 5.3 p.p. more likely to ultimately achieve success than High-school graduates when considering all cohorts. In Columns (2) and (3), we restrict our sample to younger cohorts for whom we observe GPA. Note that the estimated coefficient for College shrinks in size but it is still significant in Column (2). Interestingly, the magnitude of the College coefficient halves and it is no longer significant when we include GPA in Column (3). The GPA coefficient instead is significant and positive. Specifically, when we compare women with the same educational achievement, we see that a 10% increase in GPA (i.e., about one standard deviation) is associated with a rise in the probability of success of 1.3 p.p.. Notably this finding is present not only when we consider the last cycle a woman undertakes, but also when we take into account only the first cycles that progress to the embryo transfer stage (Column 4).[44]

[43]See also, more recently, Miller (2016).

[44]Our finding that GPA absorbs education effects aligns with medical literature, indicating weakened education-health outcome relationships after accounting for cognitive ability (Serper et al., 2014).

Table 7: GPA, Pre-IVF Fixed Effects and IVF Success

| *Outcome*: | GPA: | | | | Pre-IVF FE: | | |
|---|---|---|---|---|---|---|---|
| IVF Success | Main Sample (1) | GPA Sample (2) | GPA (3) | GPA\|Embryo (4) | Main Sample (5) | Pre-IVF FE Sample (6) | Pre-IVF FE (7) |
| HS | | | | | 0.0713*** | 0.0629*** | 0.0628*** |
| | | | | | (0.0108) | (0.0125) | (0.0125) |
| College | 0.0528*** | 0.0293** | 0.0191 | 0.009 | 0.124*** | 0.111*** | 0.100*** |
| | (0.00733) | (0.0114) | (0.0120) | (0.0126) | (0.0114) | (0.0132) | (0.0135) |
| $\log(GPA)$ | | | 0.131*** | 0.0913*** | | | |
| | | | (0.0476) | (0.0507) | | | |
| Pre-IVF FE | | | | | | | 0.0736*** |
| | | | | | | | (0.0206) |
| Full Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 17,934 | 8,220 | 8,220 | 6,879 | 20,513 | 17,979 | 17,979 |

*Notes:* The table shows the results from a linear probability model for IVF success. In the left panel the baseline category are high-school graduates and it is run for the last treatment except for column (4) which for comparison purposes is run on the first treatment and corresponds to column (4) of Table 3. The right panel is run on the last treatment and the baseline category are high-school dropouts. All regressions include all controls from our baseline regression in Table 2. The different samples are subsamples of all women/clinics ($N = 20,513$).

To the extent that wages partly reflect ability, i.e., cognitive skills, an alternative way to proxy for ability is by estimating women's Fixed Effects (FE) from a wage regression using only wage data before women enter their first IVF treatment.[45] In Column (7) of Table 7, we observe that pre-IVF FE, similar to GPA, help determine IVF success. However, unlike GPA, pre-IVF FE do not absorb a significant portion of the education gradient, suggesting that the ability proxy recovered through GPAs captures better the effects of education on IVF success. In any case, both proxies for ability (GPA and pre-IVF FE) show significant explanatory power in elucidating IVF success, especially in the context of GPA and the education gradient in IVF success.

## 4 A Dynamic Model of IVF Treatments, Success and Drop-Outs

We explore mechanisms behind the education gradient developing a dynamic model of endogenous IVF treatment with two latent factors: ability and psychological costs. In particular, each woman is permanently endowed with a certain ability ($z$), that determines the degree of productive efficiency in IVF success, and a psychological cost ($b$) associated with IVF treatments following the idea that psychological traits such as grit and resilience are associated with a lower psychological cost. In our model, women use IVF technology in order to get a first child. Both the number of

[45] Specifically, we select women with completed education up to ten years before the first treatment and run the following regression on full-time log wages:

$$\begin{aligned}\log(w_{ij}) &= \beta_0 + \beta_1 experience_i + \beta_2 experience_i^2 + \beta_3 firmtenure_i + \beta_3 firmtenure_i^2 \\ &+ \beta_4 age_i^2 + \beta_5 age_i^3 + I(IND = k) + \eta_j + \alpha_i + u_{ij}\end{aligned}$$

where $i$ denotes the individual and $j$ denotes the year; and we extract the $\hat{\alpha}_i$ as the "Pre-IVF FE."

IVF treatments (hence, the IVF drop-out rates) and their IVF success rates are endogenous and depend on the individual types $s = (z, b)$. The distributions of $z$ and $b$ are specific to each of three education ($e$) groups: less than high school, high school, and college. Since each education group solves the model separately, we drop, for now, the $e$ subscripts in the model discussion.

### 4.1 The household problem

In the model, women undergo IVF treatments to conceive their first child and are followed from their initial IVF treatment ($t = 1$) to their last IVF treatment (with a maximum of $T = 5$). Each period in the model corresponds to an IVF treatment. Following the first IVF treatment, for each period $t \in \{2, T\}$, women decide whether to continue with an additional IVF treatment or to discontinue IVF. We denote this discrete choice with $d = 0, 1$, where $d = 1$ implies the continuation of IVF treatment and $d = 0$ implies otherwise. If an individual continues IVF, she faces an intertemporal trade-off between a positive chance of IVF success, which delivers an additional stream of future utility starting next period, and the psychological cost of undergoing treatment, which reduces current utility. Dropping out from IVF results in a childless life, with no re-entry option. That is, women exit IVF after success or opt to live childless.

Precisely, for treatments (i.e, periods) $t > 1$, childless individuals must choose between going through an additional IVF treatment or dropping out from IVF. That is, individuals solve:

$$V(n = 0, t; z, b) = \max \left\{ V^{d=0} (n = 0, t; z, b), V^{d=1} (n = 0, t; z, b) \right\}, \tag{2}$$

where $d \in \{0, 1\}$ is a discrete choice variable that equals one (zero) if individuals choose (not) to undergo IVF treatment number $t$ and $n \in \{0, 1\}$ is a discrete variable denoting number of children. The individual's type is permanent and defined by her ability to follow IVF procedure ($z$) and her psychological stress ($b$).

The value of dropping out from IVF treatments is a childless life following the dropout. Precisely, dropping from IVF in period $t$ delivers the following present-value utility,

$$V^{d=0}(n = 0, t; z, b) = \sum_{\tau=t}^{T} \beta^{\tau-t} \ln y,$$

which is the discounted utility in logarithmic form of the future stream of income given by a deterministic and constant $\{y\}_{t=\tau}^{T}$.

The value of undergoing IVF treatment $t$ is,

$$\begin{aligned} V^{d=1}(n=0,t;z,b) &= \ln y - bt^2 \\ &+ \beta\left[\pi(z)V(n=1,t+1;z,b) + (1-\pi(z))V(n=0,t+1;z,b)\right] \end{aligned} \tag{3}$$

subject to the IVF success probability,

$$\pi(z) = \frac{exp(1)}{exp(1) + exp(-lnz)}, \tag{4}$$

where note that the success probability depends on the individual permanent productivity in IVF success, $z > 0$. Note that the probability of IVF success increases with individual permanent productivity $z$, *ceteris paribus*. In particular, with $\Delta > 0$, $\pi(z+\Delta) > \pi(z)$. Hence, the value of undergoing IVF treatment $t$ defined in equation (3) is given by the current felicity of consuming the individual's endowment ($\ln y$) minus the psychological cost of going through IVF which we assume is increasing in the number of treatments ($bt^2$),[46] plus the expected utility of the IVF. In particular, if IVF outcome is success (i.e., an IVF child birth), which occurs with probability $\pi(z)$, then agents discontinue further IVF treatments and enjoy utility,

$$V(n=1,t+1) = \sum_{\tau=t+1}^{T} \beta^{\tau-(t+1)}\left[\ln y + \ln(1+n)\right],$$

for the rest of their lifetime where $\ln(1+n)$ captures the joy of having children. Note that the shape of the felicity function from enjoying a child ensures that not all individuals will want to have a child at all costs, i.e., $\frac{\partial \ln(1+n)}{\partial n} < \infty$. Indeed, if $n = 0$, then we are back to the present-value utility of leaving IVF treatment without a child.

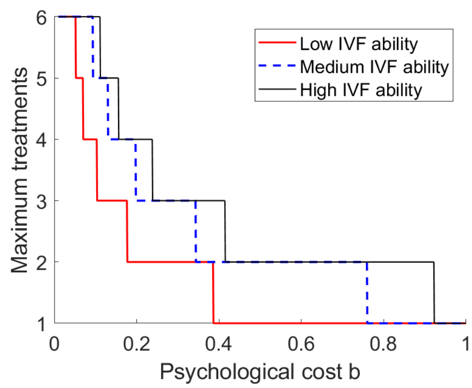
### 4.2 Solution Algorithm

We solve this model backwards for each value of $z \in [0,5]$ and $b \in [0,1]$. In each period, individuals choose whether to stay in treatment by comparing the values of being treated and dropping out. There are two parameters that are fixed in this step. We fix the discount rate to 0.99 and normalize consumption path to an exogenously given value $\{y_t\}_{t=0}^{T} = y = 5$ for all individuals which is the smallest value for $y$ such given the parameter space for $z$ and $b$ the delivered policy functions cover dropout at all treatment numbers.

Figure 4 displays the policy function for the maximum treatments, decreasing with psychological cost $b$, for three ability levels $z$. For all types, the maximum number of treatments is

[46] This captures the idea that the psychological costs increases with failure as argued in Missmer et al. (2011).

Figure 4: Drop-out Policy Function




*Note:* The figure shows the drop-out policy function after failure on the maximum number of treatments depending on $|b|$ for three different levels of ability $z$ ($z = 0$, $z = 2.5$ and $z = 5$).

decreasing in $b$. After the initial failure, a cost of 0.4 or above halts low ability types. To achieve the same for medium types, the cost must almost double to 0.75, and for high types, it must exceed 0.9. The $b$ range to stop treatment after failing the second trial is broad, from 0.19 to 0.92, depending on $z$. To stop treatment after failing the third one, the $b$ range further shrinks to about 0.12 to 0.18 for the low ability types, 0.19 to 0.36 for the medium types, and about 0.23 to 0.42 for the high type. Although the range of $b$ shrinks, it is still large because it depends on $z$. In contrast, with very low costs, ability has a minor role in deciding additional IVF treatments, as it's relatively cheap psychologically. Overall, the maximum cycles women undergo depend on disparities in $z$ and $b$, particularly at higher psychological costs.

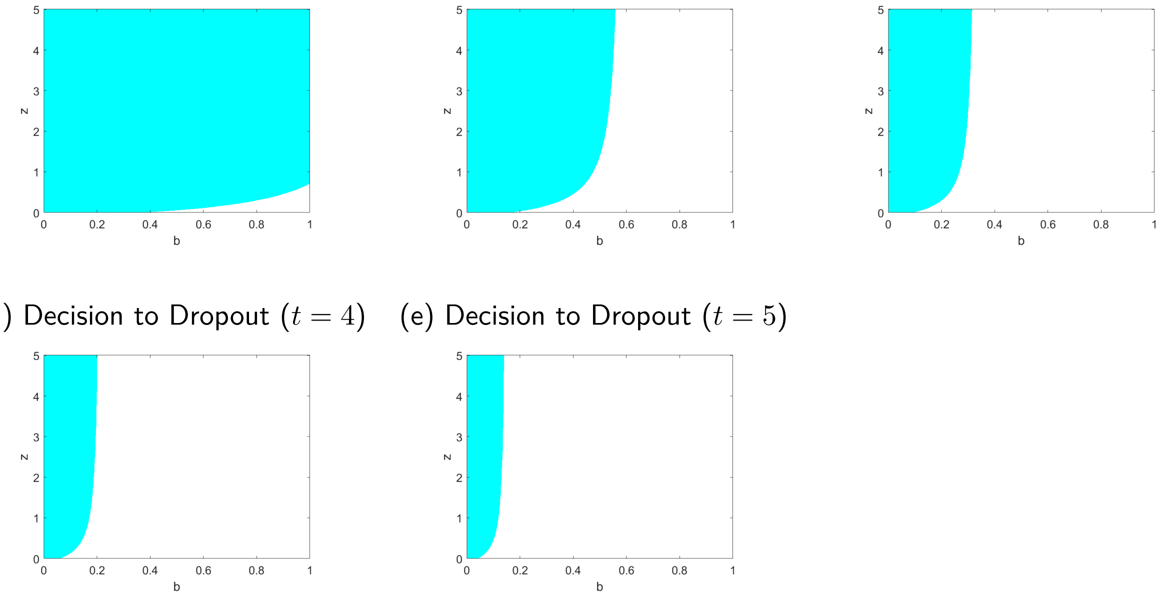
This can also be seen from Figure 5 where we plot the combination of values of ability ($z$) and psychological cost ($b$) that led individuals to undertake a new treatment (blue area), for each number of failed cycles. At each treatment, there are fewer combinations of $b$ and $z$ which lead women to remain in treatment after a failure, as a result of the increasing cost associated with higher treatments and selection in $z$ (higher types exit earlier). Moreover, across all $t$, $z$ is increasing and convex in $b$. This is due to the fact that $b$ enters quadratically in the utility function and therefore as $b$ increases higher and higher values of $z$ are needed to remain in IVF. In addition, $z$ influences the decision to stay only through the continuation value, and its role is greater in earlier cycles since there are more periods still to go.

### 4.3 Model Estimation

We use the actual patterns of success rates and drop out rates by education group and across the number treatments as targeting moments.[47] We recover the distributions of the latent ability $z$

[47] For these targeting moments we control for the interaction between age and year fixed effects. The moments are not sensitive to the inclusion of other variables from our benchmark specifications for IVF success and dropout.

Figure 5: Decision to Dropout IV: Effects of Ability $z$ and Psychological Factor $b$




(a) Decision to Dropout ($t = 1$) (b) Decision to Dropout ($t = 2$) (c) Decision to Dropout ($t = 3$)

(d) Decision to Dropout ($t = 4$) (e) Decision to Dropout ($t = 5$)

*Note:* The figures show the combinations of $b$ and $z$ of women continuing IVF (blue area) for each treatment $t = 2, ..., 5$ after failure in the previous treatment.

and psychological costs $b$, for each education group, separately. We use a log-normal distribution for $z$ to ensure that ability always takes non-negative values, and estimate its mean and variance. Instead, the distribution of $b$ is estimated non-parametrically.

The moments that we use to identify the distributions of $z$ and $b$ are IVF drop-out rates and success rates by education group and treatment (one per treatment apart from the last one where dropout is deterministic). The IVF drop-out rates for education $e$ and treatment $t$ are defined as,

$$DR(e,t) = \frac{\int_{Z \times B} \mathbf{1}_{d(z,b,t)=0} d\Phi_e(z,b,t)}{\int_{Z \times B} d\Phi_e(z,b,t)} \tag{5}$$

where $\Phi_e(z,b,t)$ is the joint distribution of $z, b$ for a given education group $e$ and period $t$, $\mathbf{1}_{d(z,b,t)=0}$ is an indicator function for each type $(z,b)$ that is equal to one if the type chooses not to continue treatment $t$. Then, the numerator in equation (5) is the endogenous population with education $e$ that at period $t$ chooses to discontinue IVF treatment (this is endogenous for each and all $t$) and the denominator is the endogenous population of education $e$ that at period $t$ must decide whether to continue IVF treatment $t$ or not (this is endogenous for $t \geq 2$).

The IVF success rates for education $e$ and treatment $t$ (one per treatment) are defined as,

$$SR(e,t) = \int_{Z \times B} \pi(z) \frac{\mathbf{1}_{d(z,b,t)=1}}{\int_{Z \times B} \mathbf{1}_{d(z,b,t)=1} d\Phi_e(z,b,t)} d\Phi_e(z,b,t) \tag{6}$$

Table 8: Estimated Distribution of Latent Ability $z$, by Education

| Education: | $<$ *HS* | *HS* | *College* |
|---|---|---|---|
| $\mu_z$ | 0.1091 | 0.1357 | 0.1404 |
| $\sigma_z^2$ | 0.9 | 0.8 | 0.7 |

*Notes:* This Table shows the estimated mean and variance of the distribution of latent ability $z$, where $\mu_z$ is the mean and $\sigma_z$ the variance of the normal distribution computed from the $<$ *HS*, *HS* and *College*. The variance was estimated on a grid varying by 0.1.

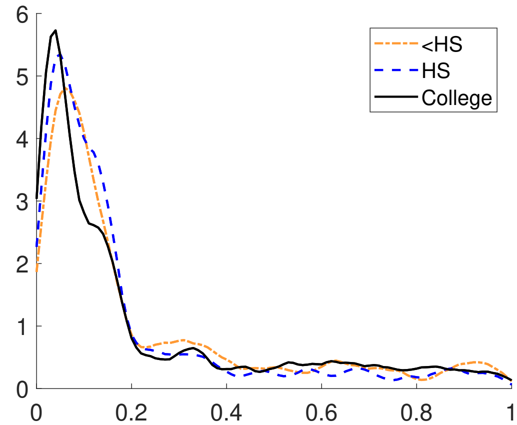
where $\pi(z)$ is the individual IVF probability of success defined in equation (4).

In order to recover the non-parametric distribution of $b$ for each education group, we first obtain the number of individuals dropping out at each treatment from the empirical conditional drop-out rates. We then identify the segments of the support of $b$ for each maximum dropout from the policy function, and populate the interval between 0 and 1 by drawing from a uniform distribution among the number of individuals who drop out at each treatment, adding mass to different segments—a total of six segments—to match the dropout and success rates.

We use the method of simulated moments to estimate the distributional parameters of the two latent factors $z$ and $b$, separately, by education. This implies identifying, per education group, the six segments of the distribution of $b$, and the parameters $\mu_z$ and $\sigma_z^2$ of the log-normal distribution of $z$. Hence, we are using ten moments (success and dropout rates for the first 5 treatments) to estimate these eight parameters and the estimation is done through a grid search, such that we obtain the global maximum.

Different moments by education help disentangle the two latent factors. First, education-specific IVF success rates help to recover the distribution of $z$ by education. The average ability can be recovered from the first cycle of IVF success rates. The variance in the ability is instead pinned down by the change in IVF success in later treatments. Within each education group, because high-ability women are more likely to succeed in the first IVF cycle and exit, women undergoing later treatments are more likely to be of lower type, and this is true for all education groups. In turn, in the model we show that the average IVF success decreases across cycles in all education groups because the ability $z$ worsens. The extent to which the probability of IVF success changes, depends on the initial range of $z$. Second, women with lower psychological cost ($b$) are more likely to continue the IVF treatment in another cycle, for a given ability. Hence, the distribution of $b$ by education groups is recovered from education specific IVF dropouts across cycles. It is important to note that the identification of $z$ and $b$ is not orthogonal. For example, recovering $b$ through the model in order to match the endogenous drop out rates takes into account that women in the model making the drop out choice conditional on the expected value of treatment, which depends on the probability of success as determined by $z$. At the same time,

Figure 6: Estimated Distribution of Latent Psychological Costs $b$, by Education




*Notes:* This Figure shows the kernel densities of the estimated distributions of latent psychological costs $b \in (0, 1)$ for the three education groups.

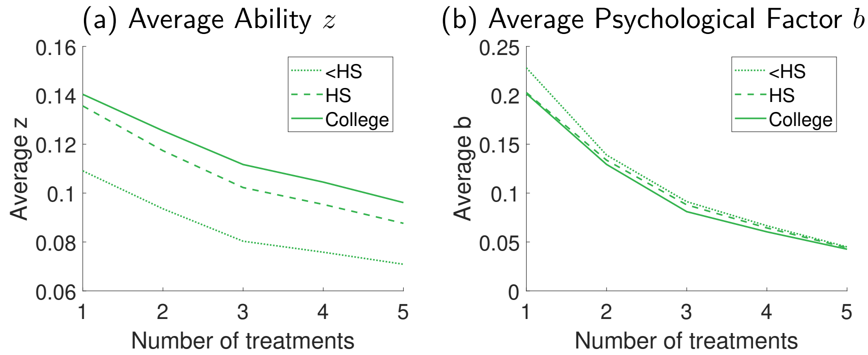
the success rates defined in (6), which are likely to determine $z$, are also a function of the selected sample that continues IVF treatment and, hence, depend on $b$.

**Estimation results** First, the estimated mean and variance of the log-normal distribution of $z$ for each education group is shown in Table 8. Average ability $z$ increases with education, by 22% when moving from < *HS* to *College*. It is closer among the *HS* and *College* groups. Further, the variance decreases with education. Consequently, the distribution of ability within $< HS$ displays a lower average and a somewhat wider distribution than $HS$ and $College$. Second, the estimated distribution of the psychological cost $b$ is most right-skewed for the *College* group, followed by the *HS* group, and then the < *HS*, see Figure 6. More college women with lower values of $b$ is consistent with lower dropouts in this group than in less educated groups.

We next examine the evolution of the average $z$ and $b$ of women with different educational attainment at each treatment to shed light on the importance of each of the latent factors in explaining the evolution of the success rates over time in each education group. The estimated average $z$ and average $b$ both show declining patterns over the treatments for all three education groups. On average, more-educated women show higher ability ($z$) and lower psychological costs ($b$) than those with less education, as depicted in Figure 7. The declining averages over treatments result from a mix of selection effects affecting the composition of women remaining in IVF.

First, there is a negative selection effect: very high $z$ women have a higher chance of achieving a live birth at the first attempt and thus successfully exit from IVF. In parallel, women paying a very high psychological cost $b$ are more likely to exit after having failed in IVF. Second, there is a potentially countervailing and more subtle effect that contributes to positive selection in $z$, because only women with an ability level high enough relative to their psychological cost do not

Figure 7: Evolution of the average ability $z$ and psychological factor $b$




*Notes:* This Figure shows average ability $z$ and average psychological cost $b$ of women in the model by education over treatments in panels (a) and (b), respectively.

lose motivation to continue IVF after a failure, hence, women with relatively high ability remain in IVF. Our estimates indicate that overall the average productivity worsens, and the first effect dominates. As the number of IVF cycles increases, we find that the women that remain in IVF become increasingly similar to one another and the selection effects become less strong. Hence, the declining trend in $z$ and $b$ somewhat flattens across cycles.

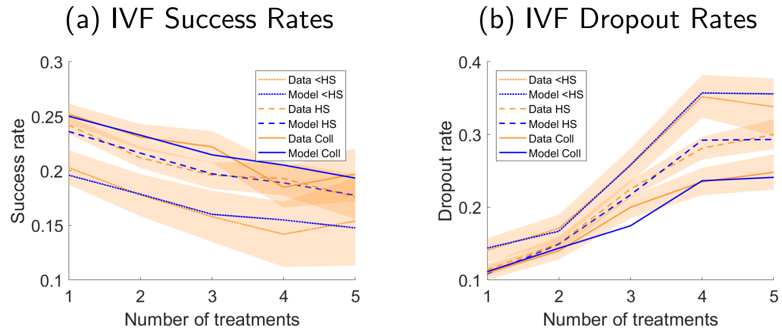
**Model Fit** Our model fits well both the decreasing pattern over treatments for success rates and the increasing pattern of dropout rates as well as the differences in both by education group, see, respectively, panel (a) and (b) in Figure 8. The decreasing pattern of success rates across treatments is similar for each education group. On the contrary, educational disparities in the conditional dropout rates increase with treatment number. This can be interpreted through the lens of our model as evidence that women with lower education face higher psychological cost related to IVF than women with higher education. While the $<$ *HS* group shows the highest conditional dropout rates for all treatments, differences between the *HS* and the *College* groups become apparent especially in later treatments, when the negative effect of the psychological cost becomes larger.

### 4.4 Counterfactual Experiments and Policy Insights

Since the estimated model accounts for the observed differences (and level) in success rates across education groups over IVF treatments, we now use it as a device to measure the extent to which the two latent factors, ability ($z$) and psychological costs ($b$), account for the education gradient.

**The role of the latent factors: ability $z$ and psychological costs $b$** First, in order to assess the explanatory power of ability $z$ in generating the education gradient in IVF, we simulate the economy attributing to all education groups the same (estimated) average productivity, $\mu_{e=C,z}$, and variance $\sigma^2_{e=C,z}$ of the college group, while maintaining the education-specific distributions

Figure 8: Model Fit, By Education Group and Treatment




*Notes:* This Figure shows the model and data moments together with the 95% confidence intervals in the data for success over treatments in panel (a) and dropout over treatments in panel (b), for each education group.

of psychological costs. We show the results of such counterfactual (CF1) in Table 9, which also shows our benchmark results to ease comparison. First, mechanically, success rate in the first treatment is equalized across education groups to 0.25. Second, the success rate in the last endogenous treatment increases from 0.46 to 0.57 in the $< HS$ group, thus closing the gradient between college success 0.59 and $< HS$ (enhanced with the ability of college educated) by $\frac{.57-.46}{.59-.46} = 87\%$. This change is the direct effect of boosting the probability of success in each treatment. In the dynamics, the overall number of treatments decreases, but only slightly because of two effects offsetting each other: on the one hand $< HS$ exit faster because they have now a higher productivity; on the other hand, the $< HS$ ones who failed stay longer in treatment, since the continuation value of remaining is higher.

Second, we simulate the model attributing to all education groups the same (estimated) distribution of psychological costs of the college group, while maintaining the education-specific ability distributions. The results of this counterfactual (CF2) are in Table 9. When facing a lower psychological cost, women educated with less than a college degree drop out less than in their benchmark counterpart. In particular, $< HS$ women increase their average number of treatments from 2.69 to 2.89. The decrease in dropouts, increases the success rate in the last treatment, although not by much, from 0.46 in the benchmark scenario to 0.48 with lower psychological costs. The reason for this small improvement is that the increase in the number of trials is offset by a worsening self-selection of who opts to remain in treatment. In fact, women with lower levels of $z$ are now more likely to continue the treatments as they face a smaller psychological cost. To put this in perspective, the gradient in the overall success rate between the bottom and the top of the education categories closes only by 13% as a result of serving the less educated with the psychological costs of the more educated.

We further re-conduct the previous two counterfactuals by giving the $< HS$ group, respectively, the distributions of ability $z$ of the $HS$ group (CF3) and the distributions of psychological

Table 9: Counterfactual Experiments and Policy

| | IVF Success Rate First Treatment | | | IVF Success Rate Last Treatment, Overall | | | Number of Treatments | | |
|---|---|---|---|---|---|---|---|---|---|
| *Education*: | <HS | HS | College | <HS | HS | College | <HS | HS | College |
| Benchmark | 0.1960 | 0.2361 | 0.2501 | 0.4634 | 0.5560 | 0.5907 | 2.6994 | 2.6786 | 2.6677 |
| Counterfactuals: | | | | | | | | | |
| CF 1 | 0.2500 | 0.2500 | - | 0.5744 | 0.5883 | - | 2.5245 | 2.6183 | - |
| CF 2 | 0.1957 | 0.2361 | - | 0.4800 | 0.5888 | - | 2.8900 | 2.7215 | - |
| CF 3 | 0.2370 | - | - | 0.5431 | - | - | 2.5810 | - | - |
| CF 4 | 0.1957 | - | - | 0.4761 | - | - | 2.8206 | - | - |

*Notes:* This table shows the benchmark and counterfactual values for 1) the IVF success rate in the first treatment, 2) the overall success rate and 3) the average number of treatments undertaken, by education group. In CF 1 (CF 2) all education groups are given the $\mu_{e=C,z}$, and variance $\sigma^2_{e=C,z}$ of the college group ($HS$ group). In CF 3 (CF 4) all educations groups are given the (estimated) distribution of psychological costs of the college group ($HS$ group). The (overall) unconditional IVF success (aggregate and by education) in the sample are shown in Table A2 of the Online Appendix: aggregate 0.5826, College 0.6170, $HS$ 0.5830, $< HS$ 0.4876.

costs $b$ of the $HS$ group (CF4); see Table 9. We find similar insights to our previous counterfactuals: Ability $z$ explains 86% of the difference in the IVF success rate across the $HS$ and $< HS$ groups, while psychological costs $b$ explains approximately 14%.

**Policy Insights** First, as per our results, we would expect that ability-boosting interventions such as nudging adherence to the IVF protocol or being more closely followed or guided through the IVF stages by a medical person (i.e. policies that increase ability $z$) can be highly successful since it can raise the number of children born through IVF without increasing the number of treatments, which would be expensive for the government that already offers them for free. Second, intuitively, offering psychological assistance and counseling during the IVF procedure (i.e. policies that reduce psychological costs $b$) would reduce the likelihood that women will discontinue treatments following a failure (drop out) due to stress and would better prepare them for the possibility that treatments will fail. However, our results show that effectively lowering the psychological costs $b$, though it helps lowering the dropout rates, it does not improve IVF success rates by much. These results align with recent review of psychological studies indicating that organized psychotherapies are successful in lowering infertility discomfort, which is extremely important, but nevertheless they cannot increase IVF success much (Hjelmstedt et al., 2003).[48]

Focusing on ability-boosting interventions, one would think that IVF patients already have a strong incentive to follow their treatment plans closely due to their intense desire to become pregnant, but medical data have shown sub-optimal medication adherence rates. Note that the IVF process itself requires considerable IVF medication use on a precise timetable, and adherence is

[48] For the medical assessment of interventions aimed at reducing infertility distress and coping with repeated treatment failures, see also (Joelsson et al., 2017; Huang et al., 2019).

Table 10: Increasingly Unequal: Non-Stationary Education Gradient In IVF Success (Live Births)

| *IVF Live Births*: | 1995-1999 (1) | 2000-2004 (2) | 2005-2009 (3) | Full Sample (4) |
|---|---|---|---|---|
| High School | 0.0386*** | 0.0118 | 0.0418** | 0.0299*** |
| | (0.0139) | (0.0165) | (0.0205) | (0.00932) |
| College | 0.0337** | 0.0418** | 0.0664*** | 0.0488*** |
| | (0.0159) | (0.0176) | (0.0208) | (0.0100) |
| Full Controls: | ✓ | ✓ | ✓ | ✓ |
| Observations | 6,725 | 7,205 | 6,583 | 20,513 |
| R-squared | 0.027 | 0.048 | 0.028 | 0.028 |

*Notes:* Results from a linear probability model for success for three different periods together with the overall. Robust standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, *$p < 0.1$. The omitted category is HS dropouts. The sample consists of all women/all clinics ($N = 20,513$).

inversely proportional to frequency of dose.[49] For many patients, IVF treatments are emotionally taxing and the terminology surrounding infertility can be confusing. Many of them have never administered drugs via self-injection before. This difficulty may result in unintentional noncompliance with the advised gonadotropin dose, an inadequate or overactive ovarian response, and possibly an unfavorable cycle outcome (Noorhasan et al., 2008; Mahoney and Russell, 2021).[50]

In practice, there are several ways to simplify and enhance the management of the IVF procedure. For instance, a mobile app that reminds patients of the protocol timing can improve medical adherence in IVF, just as it has been beneficial for other types of treatments (Lin and Wu, 2014). Another potentially effective method for enhancing patient adherence is to improve communication between the doctor and the patient (e.g. Mahoney et al., 2019; Wiecek et al., 2019). That is, without being confrontational, doctors may inquire about patients' feelings and ability to follow the recommended course of action, emphasizing the importance of protocol timing and addressing any potential barriers to adherence. If necessary, doctors should develop strategies to encourage adherence and, potentially, seek the support of the patient's spouse. Finally, medical advancements to reduce injections, blood draws, and frequent visits to fertility clinics are slowly gaining acceptance because they simplify the process of controlled ovarian stimulation, making it less invasive and time-consuming (Kushnir et al., 2022).[51]

[49] It has been shown in other contexts that patients taking medication on a schedule of four times daily achieved average adherence rates of about 50 percent (Osterberg and Blaschke, 2005).

[50] Women report barriers in medication adherence during ovarian stimulation (e.g. Mahoney et al., 2019).

[51] For example, utilizing long-acting gonadotropins or oral medications to reduce the number of injections, measuring salivary estradiol levels to decrease the need for blood draws, and promoting portable lower cost ultrasound devices that may further simplify follicular and endometrial monitoring.

## 5 Further Discussion

First, we discuss further implications of IVF for the battle against infertility in Section 5.1. In particular, we assess the non-stationarity of the education gradient in IVF success; how IVF shapes the fecundity frontier; how IVF contributes to the average age at first birth (AFB); and how IVF can help mitigate the disparities in total fertility rates (TFR) across education groups. Second, as a corollary of our results, we cannot avoid a discussion on the (mis)use of IVF success as an instrument for children in the assessment of motherhood penalties (or other outcomes) in Section 5.2.

### 5.1 An *Increasingly Unequal* Battle Against Infertility and Its Implications

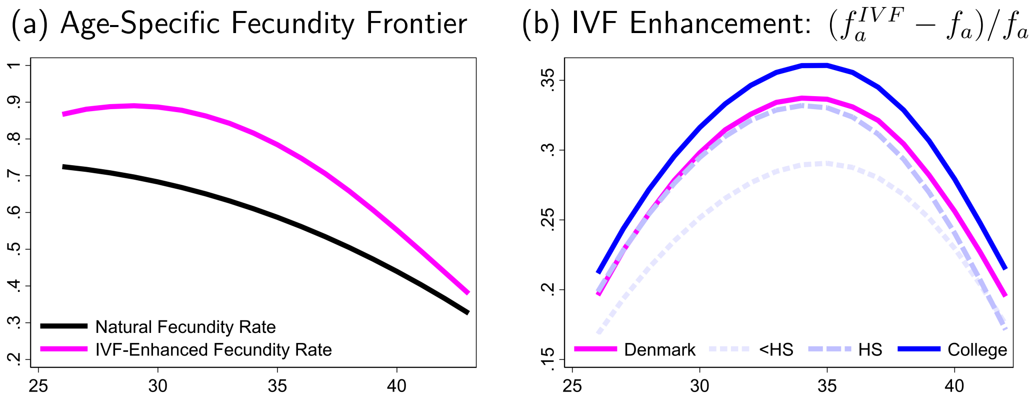
In Table 10, we present the results from re-conducting our benchmark specification, breaking the sample period into three quinquennial subperiods. In the late 1990s, college-educated women had a 3.37 p.p. higher probability of IVF success than women with less than a high-school degree; this figure increases to 6.64 in the late 2000s. That is, the education gradient in IVF success between college-educated women and women in the $< HS$ group almost doubles over a span of 10 years. Additionally, note that the difference between $HS$ and $< HS$ remains relatively constant across the entire sample period. Indeed, during the 1990s, there is no significant difference between the odds of IVF success for women with a $HS$ degree and women with a college degree. However, a gradient in IVF success emerges between $HS$ and college-educated women in the 2000s. Overall, the analysis of the behavior of the education gradient in IVF across time implies that the unequal battle against infertility is becoming increasingly unequal.

**How Much Does IVF Shape Age-Specific Fecundity?** In order to assess how the IVF technology influences women's decisions to delay motherhood, we need to understand how IVF shapes (if at all) the age-specific natural fecundity rate, defined as the ability to become pregnant conditional on seeking a child. The measurement of fecundity, however, is rather elusive for several reasons. First, the ability of women to become pregnant is conditional on the ability of their partners to impregnate. Second, it becomes necessary to control for numerous factors, particularly the latent effort put into getting pregnant. This involves controlling for variables such as the frequency and timing of intercourse, as well as the use of contraceptives.

In exploring the age-specific fecundity rate, we turn to Goldin (2021) (Chapter 7) and the results in CECOS Fédération et al. (1982).[52] This study, grounded in a natural experiment,

[52] CECOS stands for Centres d'Etude et de Conservation du Sperme Humain. The study used 11 CECOS centers with a total of 2,193 women.

Figure 9: IVF-Enhanced Age-Specific Fecundity Rate




*Notes:* In panel (a), the age-specific natural fecundity rate $f_a$ is constructed using the cumulative success after seeking a child in 12 AI cycles, as reported in Table 1 of CECOS Fédération et al. (1982). These fecundity rates are provided in 5-year age intervals, and we further use their reported sample sizes across age groups to construct separate fecundity rates for the group aged 36 to 40 and the group above 40. To plot $f_a$ in panel (a), we select the median age within each interval and estimate a cubic age polynomial. The IVF-enhanced fecundity rate, $f_a^{IVF}$, is defined by equation (7), where the IVF success rates represent the cumulative success across the first five treatments that we find for Denmark 1995-2009.

addresses, to the extent possible, the aforementioned concerns. Over two thousand French women married to azoospermic (sterile) men individually sought and underwent artificial insemination (AI) using donor sperm at varying intervals. Their difference in age but not in procedure allowed researchers to assess how age influenced their ability to conceive.[53,54] We reproduce these findings in Figure 9, where $f_a$ represents the natural fecundity rate for age group $a$ constructed using the cumulative success after seeking for a child in 12 AI cycles.[55] Assuming that all non-fecund women go through IVF, we then define the IVF-enhanced fecundity rate by age group,

$$f_a^{IVF} = f_a + (1 - f_a) \times \text{IVF Success Rate}_a, \tag{7}$$

where the IVF success rate at age $a$ is the unconditional cumulative success across the first five treatments.[56] Clearly, the IVF technology boosts the natural fecundity rate, as seen in panel

[53]All women used the same protocol developed with the Unité de Recherches Statistiques of the Institut National de la Santé et de la Recherche Médicale (INSERM).

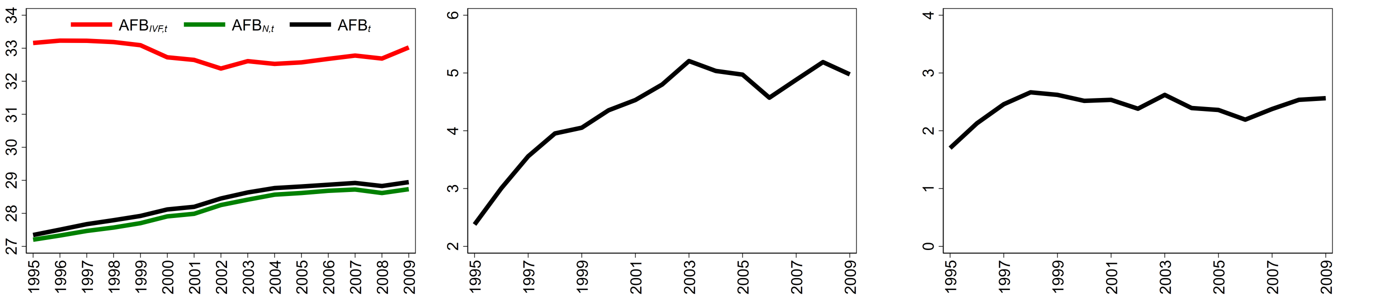
[54]The underlying identification assumption is that women pursue pregnancy with the same ability (or effort). However, drawing a parallel with our results on IVF success, one can imagine that women with different education (proxying for ability) would achieve different success rates in pregnancy even when following the same AI protocol, which would violate the underlying assumption. Although we think that this concern deserves further exploration, here, analogously to CECOS Fédération et al. (1982), we abstract from it and leave it for future work.

[55]At the time of publication, the findings in CECOS Fédération et al. (1982) were surprising. The fecundity of the studied women experienced a significant decline between the ages of thirty-one and thirty-five, much earlier than previously believed. The prevailing belief had been that fecundity started declining after the age of thirty-five, but not much before. Further, as Goldin (2021) (Chapter 7) puts it: *"Although there have been criticisms of the study and the advice offered by its authors, it remains the most scientific study of human fertility—largely because so many factors determining conception could be controlled."*

[56]We use the cumulative success rate of the first five treatments, which includes 93% of all treatments. The cap at five treatments implies the initial treatments are as close to the 2009 end period as possible.

Figure 10: IVF Technology and Average Age at First Birth (AFB), Denmark

(a) AFB with and without IVF (b) % of IVF First Births (c) Contribution of IVF to AFB (Months)




*Notes:* Panel (a) shows the average AFB for women bearing their first child through IVF, AFB$_{IVF,t}$ (solid red line), women bearing their first child naturally without IVF, AFB$_{N,t}$ (solid green line) and all women, AFB$_t$ (solid black line). Panel (b) shows the percentage of first children that are born through IVF. Panel (c) shows the difference between AFB$_t$ and AFB$_{N,t}$ in months.

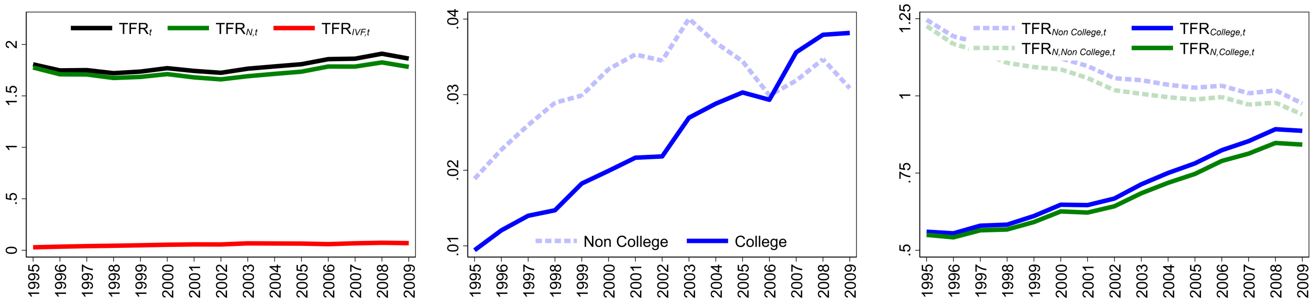
(a) of Figure 9. That is, the IVF technology shifts the fecundity frontier outwards. For women between 26 and 30, the natural fecundity rate is 70.5%, whereas the IVF-enhanced fecundity rate is 89.5%. Between ages 31 and 35, the natural fecundity rate is 62.9%, while the IVF-enhanced fecundity rate is 85.0%. In other words, the IVF-enhanced fecundity rate for women between 31 and 35 is well above the natural fecundity rate for women between 26 and 30. The ability of IVF to delay the fecundity decline decreases with age. Specifically, for women between 36 and 40, the natural fecundity rate is 50.2%, whereas the IVF-enhanced fertility rate is 65%, which is approximately the same as the natural fecundity rate for women between 31 and 35. In essence, IVF delays the decline in fecundity for five years, on average, during the second half of the 30s. The IVF enhancement gains (as a proportion of natural fecundity) over the life cycle are presented in panel (b) of Figure 9. The largest gains occur in the mid-30s, with an average gain of approximately 33%. Lastly, reflecting the ability to succeed in IVF, the IVF enhancement differs by education: 36% for college-educated women and 28% for those with less than a high school education.

Whether the shift in the fecundity frontier due to IVF translates to changes in aggregate outcomes such as the average age at first birth (AFB) or total (or completed) fertility rates is an open question. Next, we discuss how much IVF successful women affect aggregate measures of AFB and fertility in a purely accounting sense.

**Implications of IVF for AFB** In 2022, the average age at which women in Denmark had their first child was 29.9 years. This contrasts with an average age of 23.1 years in 1960. To isolate how much of the increase in the average AFB is due to successful IVF women in each period $t$, we conduct a simple accounting exercise separating women into two mutually exclusive groups $g = \{N, IVF\}$: women that have their first child naturally, $i \in N$, and women that have their first child through IVF, $i \in IVF$. We denote with $\mu_{g,t}|c$ the total amount of women that have

Figure 11: IVF Technology and Total Fertility Rate (TFR), Denmark

(a) $\text{TFR}_t$ (Aggregate) (b) $\text{TFR}_{IVF,t}$ by Education (c) $\text{TFR}_t$ and $\text{TFR}_{N,t}$ by Education




*Notes:* Panel (a) shows the total fertility rate $\text{TFR}_t$ and its decomposition between IVF children and children born naturally as per equation (9). Panel (b) shows the total fertility rate generated by IVF children by education groups as per equation (10). Panel (c) shows the total fertility rate (and that focused only on naturally born children) by education group.

a child in group $g$ in period $t$, and $\mu_t|c$ is the total amount of women that have a child in period $t$. Then, we define the average AFB in period $t$ as,

$$\overline{\text{AFB}}_t = \omega_{IVF,t}\overline{\text{AFB}}_{IVF,t} + \omega_{N,t}\overline{\text{AFB}}_{N,t} \tag{8}$$

where $\overline{\text{AFB}}_{g,t} = \frac{\sum_{i\in g}\text{AFB}_{it}}{\mu_{g,t}|c}$ is the average AFB for women in group $g$ in period $t$ and $\omega_{g,t} = \frac{\mu_{g,t}|c}{\mu_t|c}$ is the proportion of women that have their first child in period $t$ and belong to group $g$, with $\omega_N = 1 - \omega_{IVF}$. In panel (a) of Figure 10, we plot the time series of $\overline{\text{AFB}}_{N,t}$, $\overline{\text{AFB}}_{IVF,t}$ and $\overline{\text{AFB}}_t$. We note that as $\overline{\text{AFB}}_t$ increases, $\overline{\text{AFB}}_{N,t}$ increases but at a lower pace, which suggests IVF women contributes (in a purely accounting sense) to increase the average AFB over time. This is due to a combination of the fact that $\overline{\text{AFB}}_{IVF,t}$ is higher—at a relatively constant level, slightly decreasing over time—together with the fact that the share of women delivering first births through IVF increases with $\overline{\text{AFB}}_t$; panel (b) of Figure 10. The result that emerges is that IVF births contribute to increase the average AFB by 2.5 months; see panel (c) of Figure 10 showing the difference between $\overline{\text{AFB}}_t$ and $\overline{\text{AFB}}_{N,t}$. Overall, the contribution of IVF to AFB in purely accounting sense is modest. However, note that to obtain this result we need to assume that in the absence of IVF, AFB= $\overline{\text{AFB}}_{N,t}$ for all women, including those that in the data went through IVF. Therefore, of course, we cannot interpret our accounting results as causal because the absence of IVF technology may alter the behavior of women that went through IVF differently, which would violate our assumption. Further, the behavior of those women that did not go through IVF might also change in the absence of IVF.

**Implications of IVF for Fertility Rates** By altering fecundity rates, IVF technology can also influence fertility rates. In particular, a positive education gradient in IVF could contribute to mitigate the negative relationship between fertility rates and education typically present across time and space (Jones et al., 2010; Doepke et al., 2023). In Denmark, the total fertility rate (TFR) has remained relatively stable around 1.75 births per woman within our sample period from 1995 to 2009; see panel (a) in Figure 11. Further, let $b_{g,a,t}$ be the total amount of births from women of age $a$ in period $t$ in group $g$, $\mu_{g,a,t}$ the total amount of women of age $a$ in period $t$ in group $g$, with $b_{a,t} = \sum_g b_{g,a,t}$ and $\mu_{a,t} = \sum_g \mu_{g,a,t}$ where $g = \{IVF, N\}$ standing, respectively, for IVF births and natural births. We define the TFR (for women between ages 15 to 49) as,

$$\text{TFR}_t = \sum_a \frac{b_{a,t}}{\mu_{a,t}} = \sum_a \frac{b_{IVF,a,t} + b_{N,a,t}}{\mu_{a,t}} = \underbrace{\sum_a \frac{b_{IVF,a,t}}{\mu_{a,t}}}_{\text{TFR}_{IVF,t}} + \underbrace{\sum_a \frac{b_{N,a,t}}{\mu_{a,t}}}_{\text{TFR}_{N,t}} \tag{9}$$

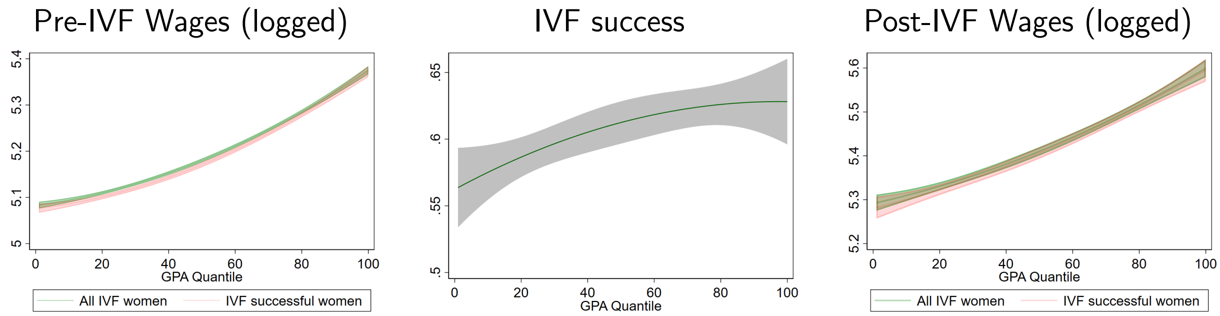
where $\text{TFR}_{N,t}$ is the total fertility rate that considers (only) natural births in the entire population of women and $\text{TFR}_{IVF,t}$ is the total fertility rate that considers (only) IVF births in the entire population of women. At the aggregate level, $\text{TFR}_{N,t}$ is larger than $\text{TFR}_{IVF,t}$ with $\text{TFR}_{N,t}$ accounting for 96% of $\text{TFR}_t$; see panel (a) in Figure 11. At the same time, throughout our sample period, there is a decline in the $\text{TFR}_{N,t}$ explained by the increase in $\text{TFR}_{IVF,t}$ from 0.03 to 0.07. A decomposition by education groups shows that the rise in $\text{TFR}_{IVF,t}$ is explained by the behavior of noncollege educated women before the mid 2000s, and by college educated women after that with $\text{TFR}_{IVF}$ being 0.038 for college women and 0.030 for noncollege women in 2009; see panel (b) of Figure 11.[57]

Finally, the convergence in TFR across education groups is substantial. Specifically, whereas TFR was 1.25 for noncollege women and 0.57 for college women in 1995, this figure is 0.98 and 0.87 for, respectively, noncollege and college women in 2009; see panel (c) in Figure 11. IVF partly contributes to this convergence in TFR across education groups. In a purely accounting sense, abstracting from the children born through IVF slows down the convergence in TFR across education groups by approximately to $(\ln \text{TFR}_{NonCollege} - \ln \text{TFR}_{College}) - (\ln \text{TFR}_{N,NonCollege} - \ln \text{TFR}_{N,College}) = 10\%$ in 2009. Again, we cannot interpret these accounting results as causal, because we do not observe the counterfactual behavior of women in the absence of IVF. Interestingly, the fact that IVF helps close the TFR gap across education groups opens the question of whether policy should aim to mitigate the education gradient in IVF (or not) given that, at the same time, the presence of the education gradient in IVF generates a more equal footing in

[57] Note that we can decompose $\text{TFR}_t$ by education groups as,

$$\text{TFR}_t = \sum_a \frac{b_{a,t}}{\mu_{a,t}} = \sum_a \frac{\sum_e (b_{IVF,a,e,t} + b_{N,a,e,t})}{\mu_{a,t}} = \sum_e \left( \underbrace{\sum_a \frac{b_{IVF,a,e,t}}{\mu_{a,t}}}_{\text{TFR}_{IVF,e,t}} + \underbrace{\sum_a \frac{b_{N,a,e,t}}{\mu_{a,t}}}_{\text{TFR}_{N,e,t}} \right) \tag{10}$$

Figure 12: GPA, IVF Success and Wages




*Notes:* This figure shows the relationship (using a quadratic fit) between GPA quantiles and pre-IVF log wages, IVF success and post-IVF log wages. The figures on log wages show the result for both all IVF women and for only women successful in IVF.

total fertility rates across education groups. The recognition of this policy tradeoff implies that policies addressing the IVF education gradient should take into consideration the concurrent role of the gradient in promoting more equal total fertility rates across education groups.

## 5.2 Ability Proxies Determine Both IVF Success and Post-IVF Outcomes

Previously we showed that IVF success is significantly influenced by GPA and, estimating our model, we find that it is primarily ability, which explains the observed education gradient. We extend these results by showing that GPA, unsurprisingly, also impacts both pre- and post-IVF wages. In Figure 12, we plot GPA percentiles (horizontal axis) against pre-IVF (average) wages (panel (a)), IVF success rates (panel (b)), and post-IVF wages two years after the last IVF treatment (panel (c)), showing a positive unconditional correlation between GPA and all these measures. Further, we observe the same patterns within the sample of women who succeeded in IVF implying that ability proxies have an effect on post-IVF wages that is independent of IVF success.[58]

In conclusion, we show that post-IVF wages are plagued by the independent effects of the same determinants—e.g. ability proxies—of IVF success.[59] Our results suggest that research using IVF success as an instrument for birth when estimating the impact of a first child on female labor outcomes should attempt to control for as fine a measure of ability as possible. Since ability is positively correlated with both IVF success and labor market outcomes, not controlling for any

[58]We find similar insights when including all of our controls; see columns (1)-(5) in the Online Appendix Table A11. Furthermore, similar insights emerge when using pre-IVF earnings fixed effects, which capture unobserved aspects of skills at the workplace that are also potentially mixed with psychological traits and other factors, as alternative ability proxy; see columns (6)-(7) in Online Appendix Table A11.

[59]In a similar vein, Bhalotra and Clarke (2019) assert that the maternal state is multifaceted and challenging to comprehensively assess and account for. This is highlighted by the fact that indicators of a mother's health and health-related behaviors are consistently positively associated with twin births, casting doubt on the validity of using twins as instrumental variables.

ability measure can create a positive bias in the estimates on child penalty, leading to smaller child penalties than what we would expect if we compare women of the same ability.

## 6 Conclusion

In this paper, we investigated the education gradient of a well-established medical technology, IVF. The unique structure of the administrative Danish register allowed us to estimate this gradient using comparisons across patients within clinics, and to control for a wide range of individual characteristics, infertility causes, number of embryos implanted, pre-treatment medical conditions, socio-economic conditions and unhealthy behaviors such as smoking and drinking.

Women with more education attain higher IVF success rates. To grasp this result, we focused on latent factors such as ability and psychological traits. First, we show that proxies for these factors, like GPA, directly shape IVF success. Second, we build a dynamic model of post-IVF failure dropout, considering women's heterogeneity in these latent factors. Our counterfactual experiments provide policy insights that, in terms of IVF success, support interventions to enhance ability—e.g. nudging adherence to the IVF protocol—over the provision of psychological help.

Further, the education gradient in IVF success increases over time. This phenomenon raises a new set of questions concerning the longer-run implications of the inequality in the battle against infertility. In particular, as IVF technology shifts the fecundity frontier more for highly educated women, the well-known fertility gap across education groups is partly mitigated by IVF. This raises a policy trade-off: should policy address the IVF education gradient, given that the asymmetric battle against infertility simultaneously balances fertility rates across education groups? We hope our analysis initiates further exploration into these questions.

## References


Adda, J., Dustmann, C., and Stevens, K. (2017). The career costs of children. *Journal of Political Economy*, 125(2):293 – 337.

Akos, P. and Kretchmar, J. (2017). Investigating grit as a non-cognitive predictor of college success. *Review of Higher Education: Journal of the Association for the Study of Higher Education*, 40(2):163–186.

Bensnes, S., Huitfeldt, I., and Leuven, E. (2023). Reconciling Estimates of the Long-Term Earnings Effect of Fertility. *IZA Discussion Paper*, (16174).

Berkman, N. D., Sheridan, S. L., Donahue, K. E., Halpern, D. J., and Crotty, K. (2011). Low health literacy and health outcomes: an updated systematic review. *Annals of internal medicine*, 155(2).

Bhalotra, S. and Clarke, D. (2019). Twin Birth and Maternal Condition. *The Review of Economics and Statistics*, 101(5):853–864.

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

Budig, M. J. and England, P. (2001). The wage penalty for motherhood. *American Sociological Review*, 66(2):204–225.

Caucutt, E., Guner, N., and Knowles, J. (2002). Why Do Women Wait? Matching, Wage Inequality, and the Incentives for Fertility Delay. *Review of Economic Dynamics*, 5(4):815–855.

CECOS Fédération, Schwartz, D., and Mayaux, M. J. (1982). Female Fecundity as a Function of Age — Results of Artificial Insemination in 2193 Nulliparous Women with Azoospermic Husbands. *New England Journal of Medicine*, 306:404–406.

Connolly, M., Hoorens, S., and Chambers, G. M. (2010). The costs and consequences of assisted reproductive technology: An economic perspective. *Human Reproduction Update*, 16(6):603–613.

Cutler, D. and Lleras-Muney, A. (2008). Education and health: Evaluating theories and evidence. In *Making Americans Healthier: Social and Economic Policy as Health Policy*. Russell Sage Foundation, New York.

Cutler, D. M. and Lleras-Muney, A. (2012). Education and health: Insights from international comparisons. NBER Working Paper 17738.

Danish Health Insurance (2013). Webpage for the danish health insurance (https://www.sygeforsikring.dk/default.aspx?id=226).

Dansk Fertilitetsklinik (2016). IVF Treatment Guide (IVF vejning).

de Mouzon, J., Goossens, V., Bhattacharya, S., Castilla, J., Ferraretti, A., Korsak, V., Kupka, M., Nygren, K., Andersen, A. N., EIM, and ESHRE (2010). Assisted reproductive technology in Europe, 2006: results generated from European registers by ESHRE. *Human Reproduction*, 00(0):1–12.

Deming, D. J. (2022). Four facts about human capital. *Journal of Economic Perspectives*, 36(3):75–102.

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

Duckworth, A. L., Peterson, C., Matthews, M. D., and Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, 92(6):1087–1101.

Eskreis-Winkler, L., Duckworth, A. L., Shulman, E. P., and Beal, S. (2014). The grit effect: predicting retention in the military, the workplace, school and marriage. *Frontiers in Psychology*, 5.

Gallen, Y., Joensen, J. S., Johansen, E. R., and Veramendi, G. F. (2024). The Labor Market Returns to Delaying Pregnancy. *IZA Discussion Paper*.

Gayle, G.-L. and Miller, R. A. (2012). Life-Cycle Fertility and Human Capital Accumulation. Technical report.

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

Gershoni, N. and Low, C. (2021b). The power of time: The impact of free IVF on Women's human capital investments. *European Economic Review*, 133:103645.

Goisis, A., Palma, M., Metsä-Simola, N., Klemetti, R., Martikainen, P., Myrskylä, M., Pelikh, A., Tosi, M., and Remes, H. (2023). Medically assisted reproduction and mental health: a 24-year longitudinal analysis using Finnish register data. *Am J Obstet Gynecol*, 228.

Goldin, C. (2021). *Career and Family: Women's Century-Long Journey toward Equity*. Cambridge University Press.

Goldin, C. and Mitchell, J. (2017). The new life cycle of women's employment: Disappearing humps, sagging middles, expanding tops. *Journal of Economic Perspectives*, 31(1):161–82.

Goldin, C. and Olivetti, C. (2013). Shocking Labor Supply: A Reassessment of the Role of World War II on Women's Labor Supply. *American Economic Review*, 103(3):257–62.

Goldman, D. and Lakdawalla, D. (2005). A Theory of Health Disparities and Medical Technology. *Contributions to Economic Analysis  Policy*, 4:1395–1395.

Goldman, D. P. and Smith, J. P. (2002). Can Patient Self-Management Help Explain the SES Health Gradient? *Proceedings of the National Academy of Sciences of the United States of America*, 99(16):10929–10934.

Groes, F., Houštecká, A., Iorio, D., and Santaeulàlia-Llopis, R. (2024). Partial replication of "Can Women Have Children and a Career? IV Evidence from IVF Treatments". Technical report, Copenhagen Business School.

Groes, F., Iorio, D., Leung, M.-Y., and Santaeuàlia-Llopis, R. (2017). Educational Disparities in the Battle Against Infertility: Evidence from IVF Success.

Grossman, M. (1972). *The Demand for Health: A Theoretical and Empirical Investigation*. New York: Columbia University Press.

Grossman, M. (2006). Education and nonmarket outcomes. volume 1 of *Handbook of the Economics of Education*, chapter 10, pages 577–633. Elsevier, 1 edition.

Hjelmstedt, A., Widström, A., Wramsby, H. M., and Collins, A. (2003). Patterns of emotional responses to pregnancy, experience of pregnancy and attitudes to parenthood among IVF couples: a longitudinal study. *Journal of Psychosomatic Obstetrics and Gynaecology*, 24(3):153–162.

Huang, M.-Z., Chien-Huei, K., Kuan-Chia, L., Jiann-Loung, H., Shuby, P., and Meei-Ling, G. (2019). Psychological Health of Women Who Have Conceived Using Assisted Reproductive Technology in Taiwan: Findings from a Longitudinal Study. *BMC Women's Health*, 19.

Jachimowicz, J. M., Wihler, A., Bailey, E. R., and Galinsky, A. D. (2018). Why grit requires perseverance and passion to positively predict performance. *Proceedings of the National Academy of Sciences of the United States of America*, 115(40):9980–9985.

Joelsson, L., Salih, Tydén, T., Wanggren, K., Georgakis, M. K., Stern, J., Berglund, A., and Alkistis, S. (2017). Anxiety and Depression Symptoms among Sub-fertile Women, Women Pregnant after Infertility Treatment, and Naturally Pregnant Women. *European Psychiatry*, 45.

Jones, L. E., Schoonbroodt, A., and Tertilt, M. (2010). Fertility Theories: Can They Explain the Negative Fertility-Income Relationship? In *Demography and the Economy*, NBER Chapters, pages 43–100. National Bureau of Economic Research, Inc.

Kenkel, D. (1991). Health behavior, health knowledge, and schooling. *Journal of Political Economy*, 99(21):287–305.

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

Kushnir, V. A., Smith, G. D., and Adashi, E. Y. (2022). The Future of IVF: The New Normal in Human Reproduction. *Reproductive Sciences*, 29(3):849–856.

Leung, M.-Y., Groes, F., and Santaeulàlia-Llopis, R. (2016). The Relationship between Age at First Birth and Mother's Lifetime Earnings: Evidence from Danish Data. *PLoS ONE*, 1(11).

Lin, H. and Wu, X. (2014). Intervention strategies for improving patient adherence to follow-up in the era of mobile information technology: a systematic review and meta-analysis. *PLoS One*, 9(8):e104266.

Lundborg, P., Plug, E., and Rasmussen, A. W. (2017). Can Women Have Children and a Career? IV Evidence from IVF Treatments. *American Economic Review*, 107(6):1611–37.

Mahalingaiah, S., Berry, Hornstein, M. D., Cramer, D. W., and Missmer, S. A. (2011). Does a woman's educational attainment influence in vitro fertilization outcomes? *Fertility and Sterility*, 95(8):2618–2620.

Mahoney, D. E. and Russell, C. L. (2021). Women's reports of barriers to and facilitators of oral medication adherence during ovarian stimulation: A mixed methods pilot study. *Journal of Reproduction & Infertility*, 22(3):184–200.

Mahoney, D. E., Russell, C. L., and Cheng, A.-L. (2019). Medication adherence among women undergoing infertility treatment: A systematic review. *International Journal of Women's Health and Reproduction Sciences*, 7(2):141–9.

Miller, T. A. (2016). Health literacy and adherence to medical treatment in chronic and acute illness: a meta-analysis. *Patient education and counseling*, 99(7):1079–1086.

Ministry of Health (1997). Law on artificial insemination (lov om kunstig befrugtning i forbindelse med lægelig behandling, diagnostik og forskning m.v. ). Law.

Ministry of Health (2006). Law on artificial insemination (lov om ændring af lov om kunstig befrugtning i forbindelse med lægelig behandling, diagnostik og forskning m.v.). Law.

Ministry of Health (2012). Official webpage for information on fertility treatment (https://www.sundhed.dk/borger/sundhed-og-livsstil/graviditet-foedsel-barsel/barnloesheds-listeside/kopi-af-behandling-barnloeshed/).

Ministry of Health (2013). *Choice of Hospital (Sygehusvalg)*. Ministeriet for Sundhed og Forebyggelse.

Missmer, S. A., Pearson, K., Ryan, L., Meeker, J., Cramer, D., and Hauser, R. (2011). Analysis of multiple-cycle data from couples undergoing in vitro fertilization: methodologic issues and statistical approaches. *Journal of Epidemiology*, 22(4):497–504.

Musick, K., England, P., Edgington, S., and Kangas, N. (2009). Education Differences in Intended and Unintended Fertility. *Social Forces*, 88(2):543–572.

Noorhasan, D. J., McCulloh, D. H., Cho, M., and McGovern, P. G. (2008). Follicle-stimulating hormone levels and medication compliance during in vitro fertilization. *Fertility and Sterility*, 90(5):2013.e1.

Olivetti, C. (2014). The Female Labor Force and Long-Run Development: The American Experience in Comparative Perspective. In *Human Capital in History: The American Record*, NBER Chapters, pages 161–197. National Bureau of Economic Research, Inc.

Olivetti, C., Pan, J., and Petrongolo, B. (2024). The evolution of gender in the labor market. Working Paper 33153, National Bureau of Economic Research.

Osterberg, L. and Blaschke, T. (2005). Adherence to medication. *The New England Journal of Medicine*, 353(5):487–497.

Rosenzweig, M. R. and Schultz, T. P. (1989). Schooling, Information and Nonmarket Productivity: Contraceptive Use and Its Effectiveness. *International Economic Review*, 30(2):457–77.

Rossi, Berry, Hornstein, Cramer, Ehrlich, and Missmer. (2011). Effect of alcohol consumption on in vitro fertilization. *Obstetrics and gynecology*, 117(1):136–142.

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

Serper, M., Patzer, R. E., Curtis, L. M., Smith, S. G., O'Conor, R., Baker, D. W., and Wolf, M. S. (2014). Health literacy, cognitive ability, and functional health status among older adults. *Health services research*, 49(4):1249–1267.

Shah, D., Missmer, S., Berry, K., Racowsky, C., and Ginsburg, E. (2011). Effect of obesity on oocyte and embryo quality in women undergoing in vitro fertilization. *Obstetrics and gynecology*, 118(1):63–70.

Thompson, V., Cavazos-Rehg, P., Tate, K., and Gaier, A. (2008). Cancer information seeking among african americans. *Journal of Cancer Education*, 23:92–101.

van Loendersloot, L. L., van Wely, M., Limpens, J., Bossuyt, P. M., Repping, S., and van der Veen, F. (2010). Predictive factors in in vitro fertilization (ivf): a systematic review and meta-analysis. *Human Reproduction Update*, 16(6):577–589.

Wiecek, E., Tonin, F. S., Torres-Robles, A., Benrimoj, S. I., Fernandez-Llimos, F., and Garcia-Cardenas, V. (2019). Temporal effectiveness of interventions to improve medication adherence: A network meta-analysis. *PLoS One*, 14(3):e0213432.

# Appendix

## A The IVF Protocol

The IVF procedure that women have to go through starts with the diagnosis of why the couple is infertile, which involves an analysis of the sperm quality and female fertility.[60] The IVF procedure itself is strict and complex as it involves both medical interventions and patient's strict adherence to treatment. It consists of the aspiration of oocytes (egg retrieval), the insemination, and the embryo transfer, which may or may not result in a pregnancy. We can break the overall procedure into five consecutive phases (Dansk Fertilitetsklinik, 2016).[61] We show the phases of the protocol and the treatments administered throughout the process in Figure 2.

Phase 1 : **Down-regulation of follicle stimulation hormones (FSH)** The down-regulation improves the control on the future ovarian stimulation (e.g., prevents premature ovulation) and requires the application of two nose sprays. One of the sprays needs to be applied twice per day and one four times per day.[62]

Phase 2 : **Ovarian Stimulation Process** The second step is the ovarian stimulation process that starts with the menstrual period. This process refers to hormone stimulation with FSH through daily injections that must be taken at the same time every evening. During this process nasal sprays are applied the same number of times (four) per day, though in less quantity. Ovarian stimulation is monitored with regular (three-to-five) blood and ultrasound tests.[63]

Phase 3 : **Egg Release** This phase consists of two steps. First, an **injection** is self-administered by the patient around ten days after the stimulation begins. This injection releases the eggs precisely thirty six hours after its application, exactly when the second step is scheduled; the **egg retrieval procedure**, performed by doctors. Thus, it is very important that the patient manages to follow procedure and takes the injection at the time determined by the doctor.

Phase 4 : **Insemination** This is a purely medical procedure without patient involvement which may lead to at least one healthy embryo to be transferred on the basis of the morphological grading of the embryos.[64] Doctors assess the quality of the embryos on the basis of some aspects of their microscopic appearance: Cell number, cell regularity, degree of fragmentation, granularity, etc. If there are healthy embryos to transfer, approximately three to five days after the eggs retrieval, the embryo transfer takes place.

Phase 5 **Embryo Implantation** In the two weeks following the embryo transfer, the embryo implants itself in the womb. During this phase, patients must place one tablet three times a day, or apply a gel twice daily, in the vagina to ensure the mucosa matures correctly, and then take a pregnancy test. Patient then must then follow a standard healthy life style conducive to a successful pregnancy.

[60] All women who are in our sample of IVF-treated women have undergone the clarification procedure.

[61] The IVF procedure can either be "long" or "short". For concreteness, we focus on the description of the "long" procedure which embeds the "short" one. We follow the description provided by the oldest fertility clinic in Denmark (Dansk Fertilitetsklinik, 2016).

[62] Alternatively, the patient injects herself with one daily Lupron injection in the belly at night.

[63] The last day of stimulation, FSH injections are complemented with a human chorionic gonadotropin (HCG) injection to support the normal development of an egg in a woman's ovary.

[64] Blastocyst culture helps select the best quality embryos for transfer and reduces multiple pregnancy risks.