On the family origins of human capital formation: Evidence from donor children.

Petter Lundborg, Erik Plug, and Astrid Würtz Rasmussen*

Forthcoming *Review of Economic Studies*

Abstract

We introduce a novel strategy to study the intergenerational transmission of human capital skills, net of genetic skill transfers. For this purpose, we use unique Danish data on children conceived through sperm and egg donation in IVF treatments to estimate the relationship between child test scores and parental years of schooling. Because the assignment of donors is not selective, these parental schooling estimates allow for a causal nurture interpretation. Once we take account of genes, we find that only the education of mothers matters: the association between father's education and child test scores (in reading and math) is insignificant and practically zero, whereas the association between mother's education and child test scores (in reading, not in math) is significant and large, and as large as the association we estimate for mothers of non-donor children.

*Plug: Amsterdam School of Economics, University of Amsterdam, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands (email: e.j.s.plug@uva.nl). Lundborg: Department of Economics, Lund University, Tycho Brahes Väg 1, 223 63 Lund, Sweden (email: petter.lundborg@nek.lu.se). Rasmussen: Department of Economics and Business Economics, Aarhus University, Fuglesangs Allé 4, 8210 Aarhus, Denmark (email: awr@econ.au.dk). Data in this article are, in part, taken from IVF registers, held by the Danish Health Data Authority (DHDA). DHDA is not responsible for the analysis or interpretation of the data presented. We thank the editor, five referees, seminar and conference participants for their comments and suggestions. Financial support from the Institute for Evaluation of Labour Market and Education Policy (IFAU), Swedish Research Council for Health, Working Life and Welfare (FORTE), and the Dutch Research Council (NWO) is gratefully acknowledged. The authors declare that they have no relevant or material financial interest that relate to the research described in this paper.

1 Introduction

Why do more skilled parents have more skilled children? This question has attracted continuous attention from social scientists for well over a century. Their polar explanations are one of nurture, and one of nature. It is nurture if more skilled parents provide a more advantageous environment for their children's skill development. It is nature if more skilled parents have certain genetic skill advantages that they pass on to their children. Of course, any intermediate explanation is possible too, and arguably more likely.

Previous adoption and twin studies on the intergenerational transmission of human capital skills (usually education) seem to settle on nurture being (somewhat) less important than nature (Taubman 1976, Plug and Vijverberg 2003, Plug 2004, Björklund et al. 2005 2006, Sacerdote 2002 2007, Cesarini 2010, Cesarini and Visser 2017). There is, however, much uncertainty about the accuracy of these nurture and nature estimates; as noted in recent literature surveys, the adoption and twin strategies used to isolate nurture from nature influences often suffer from identification problems that bias results against the nurture explanation (Björklund and Salvanes 2011, Black and Devereux 2011, Holmlund et al. 2011, and Sacerdote 2011).[1]

In this paper, we introduce a novel identification strategy to more credibly identify the nurture effect in the intergenerational transmission of human capital skills. In particular, we exploit that some children are genetically unrelated to one of their rearing parents because they are conceived through sperm or egg donors in in vitro fertilization (IVF) treatments. Sperm donation refers to fertilization of the mother's egg with the sperm of an anonymous donor man. Resulting children are genetically related to the mother but not to the father. Egg donation is like sperm donation in that the children are genetically related to the father but not to the mother. Within this IVF context, we take an education-oriented human capital perspective and identify the nurture effect by estimating how the educational outcomes of donor children relate to the educational outcomes of their genetically unrelated parents. Because the assignment of donors is not selective, we can give the corresponding intergenerational mobility estimates a causal nurture effect interpretation, one that captures the nurturing effect of both prenatal and postnatal environments. As such, the donor strategy we propose closely resembles the ideal experiment in which children as embryos with different genetic make up are randomly assigned to parents with different human capital levels.

[1]The most common reasons for underestimating the nurture influence include positive sibling spillovers, treatment differentials (identical twins are treated more similarly than non-identical twins, adoptees are treated less favorably than non-adoptees), and adoptees arriving at a later age in their adoptive families.

We use several administrative registers to compile our primary sample of IVF children born in Denmark (between 1994-2007) with information on their donor status (conceived through sperm or egg donation), various test score outcomes from nationwide standardized tests taken throughout their primary and lower secondary school years, and the educational and labor market characteristics of their parents. This sample allows us to estimate intergenerational associations for donor and non-donor children. In addition, we use the same registers to compile two validity samples of children: one of adopted children, and one of all other children. These samples together allow us to compare intergenerational associations between donor and adopted children (to better assess the role of prenatal and postnatal environments) and between donor and all other children born in Denmark in this period (to better assess the generalizability of our findings).

To preview our main results, we find no evidence that the education of fathers matters for their children's test scores, once we take account of their genes. In donor families where children are genetically unrelated to the father, the associations between paternal education and child test scores (in reading and math) are statistically insignificant and close to zero. By contrast, we find strong evidence that the education of mothers matters, also net of their genes. In donor families where children are genetically unrelated to the mother, we find that children with more educated mothers perform much better on standardized achievement tests, particulary in reading. The associations between maternal education and child reading test scores are significant and large, and as large as the associations we find for mothers and genetically related children. The associations between maternal education and child math test scores, on the other hand, are notably smaller but imprecise.

When we compare the nurture estimates in donor families with those in adoptive families (that closely resemble donor families), we find that the nurture effect estimates are the same for adoptive fathers but different for adoptive mothers. In particular, the associations between maternal education and child test scores in adoptive families are small, and notably smaller than those obtained in donor families. One possible explanation is that prenatal (and very early childhood) conditions are important for the development of child skills. It explains why nurture effects in donor families are stronger for mothers than for fathers: unlike fathers, mothers are pregnant, carry children, give birth, and spend observably more time taking care of children in early childhood. It also explains why nurture effects are stronger for donor children than for adopted children: unlike adopted children, donor children benefit from their mother's exposure in pregnancy and very early childhood. We provide some additional evidence, albeit mixed, for this prenatal condition explanation. When we look at maternal smoking during pregnancy, which is considered one of the primary prenatal risk factors associated with poor health and cognitive development of children (Almond and Currie 2011), we find that more educated donor mothers (pregnant of genetically unrelated babies) smoke less. When we look at standard infant health measures at birth, however, we do not find that the same mothers give birth to (measurably) healthier children. It is possible, though, that the combination of small samples and limited variation in our infant health measures prevents us from detecting significant effects.

Another possible explanation is that donor families are simply different. If these differences affect how human capital skills transfer from parents to children, the estimated nurture effects gathered in donor families may not generalize to more representative families. We provide several tests for this. We first leverage the legal setting in Denmark, where donor children in donor families must be genetically related to one of their parents. We find no evidence that the overall intergenerational skill associations between parents and their genetically related children are different in donor and representative families. We next explore whether donor parents treat their genetically unrelated children differently. When we look at parental leave take up, labor supply, and divorce risk, which serve as crude proxies for parental involvement and investments in infancy and early childhood, we find no systematic differences between genetically related and unrelated children. Without direct information on parental involvement and investments, however, it is still possible that donor parents are differently involved in the upbringing of their genetically unrelated children. In particular, we cannot rule out that more educated mothers in egg-donor families are also more involved mothers driving their nurturing impact up, and reversely, that fathers in sperm-donor families are less involved fathers driving their nurturing impact down.

The rest of the paper unfolds as follows. Section 2 provides the literature background, and lists the main contributions of our study. Section 3 describes the institutional context of IVF treatments in Denmark and the administrative data. Section 4 describes how donors are assigned to IVF treated families. Section 5 introduces the novel strategy to identify the nurture influences of parental education in intergenerational mobility models using donor and nondonor IVF children. Section 6 presents our main set of results. Section 7 compares our results to those obtained using representative and adoption samples. Section 8 concludes.

2 Previous Literature

There is an active literature concerned with estimating the nurturing effect of parental education on the skills and educational outcomes of children. See Björklund and Salvanes (2011), Black and Devereux (2011), Holmlund et al. (2011), and Sacerdote (2011) for recent literature reviews on the topic. In isolating nurture from nature influences, most studies estimate education associations of either sibling pairs or parent-child pairs using genetically informative samples of twins and adoptees. In this section, we first summarize the nurture findings from previous sibling and intergenerational studies, discuss the main channels by which parental education shapes the nurturing environment for children (to do well in school), and conclude with the main contributions of our study.

Sibling Studies

One line of studies relies on sibling associations. With different sibling pairs sharing different combinations of common genes and environment, researchers have used behavioral genetic models to decompose the overall educational outcome variation into nature and nurture components. The nurture component represents the impact of some latent family background component that captures all the environmental factors shared by siblings.

Studies that use twins identify the nature component from sibling association differences between identical and fraternal twins, which is then used to recover the nurture component. Studies that use adoptees identify the nurture component from sibling associations of either adopted siblings or adopted and non-adopted siblings. These studies generally find that nurture matters, explaining about 10 to 45 percent of the overall variation in the educational attainment.[2] These nurture estimates, however, rely on rather controversial model assumptions regarding the representativeness of twins and adoptees, siblings (not) affecting one another, gene-environment independency, similarity in (parental) treatments, random partner choice (in case of twins), and random assignment to families (in case of adoptees). Any conclusion based on such nurture estimates should therefore be treated cautiously.

[2]We have taken sibling associations in education reported elsewhere and constructed a comprehensive set of nurture estimates. With twins, the nurture estimates range from 10 percent (Miller et al. 1995), 25-35 percent (Taubman 1976, Jencks and Brown 1977, Lykken et al. 1990, Isacsson 1999, Cesarini 2010), to 45 percent (Ashenfelter and Krueger 1994). With adoptees, the nurture estimates equally vary and range from 10 percent (Scarr and Weinberg 1994), 20 percent (Lichtenstein et al. 1992, Cesarini 2010), to 45 percent (Teasdale and Owen 1984). Recent adoption studies that account for random assignment of adoptees to families find that nurture explains about 15 percent of the overall education variability (Sacerdote 2007, Fagereng et al. 2021).

Intergenerational Studies

Another line of studies relies on intergenerational associations. With different parents providing different combinations of genes and environment to their children, researchers have used regression models to link the educational characteristics of parents and children, after taking account of genetic skill transfers. The corresponding estimate expresses the nurturing impact of parental education, which captures the causal effect of parental education and any other environmental factor that is correlated with it.

Studies that use identical twin parents identify the nurturing effect from within-twin regressions linking the educational differences of twin parents to the educational differences of their children.[3] Studies that use adoptees identify the nurturing effect from simple cross-sectional regressions linking the educational outcomes of adoptive parents to the educational outcomes of their adopted children. Like the sibling association studies, these studies also find that nurture matters, explaining about 30 percent of the overall intergenerational association in education for mothers, and about 60 percent for fathers.[4]

While these regression models provide nurture estimates that are easier to interpret than those provided by behavioral genetic decomposition models, there are limitations that may bias these nurture effect estimates downwards. The main limitation in twin studies relates to the non-heritable traits that twins share. If the differencing of twins not only differences out all the heritable traits but also some of the non-heritable traits that twins share, as Griliches (1979) and others have argued, the within-twin nurture effect estimates likely understate the nurturing influence of parental education. And similarly, the main limitation in adoption studies relates to the early childhood environment that adoptees may miss in their adoptive family. If there is an important role for prenatal and early childhood conditions in explaining child outcomes, as Heckman (2007) and co-authors repeatedly emphasize (see also the references in online Appendix A), the nurture effect estimates taken from adoptees will also understate the true impact of nurture on

[3]These within-twin studies typically aim at identifying the causal impact of parental education. We nonetheless interpret the corresponding within-twin estimates as nurture effect estimates because they also capture the impact of all the unshared non-heritable abilities of twin parents that are correlated with parental education, as Griliches (1979) indicated some time ago.

[4]Analogous to the sibling associations we summarized earlier, we have collected intergenerational associations in education from twin and adoption studies and expressed the corresponding nurture effect estimates as fractions of the overall (cross-sectional) intergenerational associations in education. With twin parents, we find that most fraction estimates range from 25 to 40 percent for mothers, and from 30 to 75 percent for fathers (Behrman and Rosenzweig 2002, Bingley et al. 2009, Holmlund et al. 2011, Pronzato 2012, Amin et al. 2015). With adoptees, these fraction estimates are somewhat higher, ranging from 50 to 65 percent for mothers, and from 65 to 85 percent for fathers (Dearden et al. 1997, Sacerdote 2000, Plug 2004, Holmlund et al. 2011). Recent adoption studies that more carefully account for the random assignment of adoptees to families report smaller fraction estimates of about 30 to 45 percent for mothers and fathers (Björklund et al. 2006, Sacerdote 2007, Adermon et al. 2021, Black et al. 2022).

those outcomes. Other limitations relate to the larger impact of measurement error on within-twin estimates, the non-random assignment of adopted children into adoptive families, and the lack of representativeness of samples of twins and adoptees.[5]

Mechanisms

A last line of studies explores how parental education contributes to a stable, supportive and stimulating environment for children to do well in school. Black and Devereux (2011) list the main mechanisms through which parental education can improve the skills of children: more educated parents are richer parents who can more easily allocate money to skill-enhancing activities (Becker and Tomes 1986); more educated parents are more informed/better decision makers who can better allocate their time and money to activities that enhance child skills (Cutler and Lleras-Muney 2006); and more educated parents are more productive/better parents who can generate more child skills out of the same amount of time and money allocated to child development (Cunha and Heckman 2009). Establishing whether any of these mechanisms are causal has proved difficult. Evidence on causal mechanisms is, for this reason, scarce.[6]

Our Contributions

Our study on children conceived through sperm and egg donations, when viewed as embryo adoptions, most closely relates to the recent adoption studies, which account for non-random assignment and investigate how educational and wealth outcomes of adopted children relate to the educational outcomes of their rearing parents (Björklund et al. 2006, Sacerdote 2007, Hægeland et al. 2010, Holmlund et al. 2011, Fagereng et al. 2021, Black et al. 2022). Our study is complementary to these adoption studies in two important ways. First, these adoption studies estimate nurture effects on samples of children who are adopted during early childhood (up to six to eighteen months). Our study estimates nurture effects on samples of donor children who are transferred

[5]Recent genome-wide association studies (GWAS), which estimate the nature component directly from genetic information, find much smaller nature estimates than those reported in twin and adoption studies (Lee et al. 2018, Okbay et al. 2022). Based on millions of genotyped individuals, these studies find that the polygenetic scores alone can explain about 10 to 15 percent of all the variance in education measured in years of schooling. With additional genetic information on siblings, parents, or a clever combination thereof, it is also possible to identify the causal effect of these polygenetic scores (Lee et al. 2018, Okbay et al. 2022, Young et al. 2022). Based on much smaller samples, these studies find that the polygenetic score effect sizes get considerably smaller and together explain only a quarter to a half of all the polygenetic-score-induced variation in educational attainment. GWAS designs, however, fail to capture all the relevant genetic variation, so the corresponding nature estimates likely underestimate the role of nature (Young 2019).

[6]Online Appendix A provides a more detailed discussion on the existing evidence on causal mechanisms and how our study relates to this literature.

into the womb 3 to 5 days after a successful fertilization. This means that, unlike the adoption studies, our nurture effect estimates capture prenatal and very early childhood influences. Second, the adoption studies rely on parents and children that bear little resemblance to any other sample of (representative) parents and children.[7] Our study uses samples of donor children born and raised by IVF parents treated with either eggs from other IVF-treated mothers or sperm from Danish sperm donors (with arguably comparable traits).[8] This means that our study has a somewhat greater representation than the adoption studies.

Compared to these adoption studies, however, our study also comes with some limitations. First, IVF treatments based on donors are quite rare which means that we work with relatively small samples. Second, IVF treatments based on donors are quite recent interventions which means that most donor children are too young to measure their performance in terms of realized educational attainment and labor market outcomes. Instead, we work with intermediate school outcomes and measure the children's performance in terms of test score outcomes from national tests taken in primary and lower secondary education. Comparable test scores taken in primary education are often found to be strong predictors for outcomes that are realized later in life such as final exam scores, educational attainment and labor market earnings (Beuchert and Nandrup 2018, Woessmann 2018).[9] Third, IVF treatments in Denmark are based on either egg or sperm donors, but never both, which means that the empirical strategy we propose only takes account of the genes of one parent. While this is a disadvantage whenever parents match nonrandomly, later in the paper we show that the estimated nurture effect estimates do not suffer much from assortative mating bias (see online Appendix D). Instead, we take advantage of this disadvantage; that is, we can test the external validity of this donor experiment by comparing the intergenerational skill associations obtained with children in donor families and their genetically related parents to those obtained with more representative children and parents.

Our approach is somewhat related to Rice et al. (2009). They survey IVF-treated mothers with and without donor children treated in several UK and US fertility clinics.

[7]It is difficult to compare adoption and non-adoption families for a number of reasons. Adoptees are less comparable in that they are separated from their birth parents, possibly with traumatizing effects (Brodzinski 1987). Adoptees are also less comparable in that they are assigned to rearing parents that are very different from themselves. International adoptees (including Korean-born adoptees) look distinctively different from their rearing parents. National adoptees may look more similar to their parents, but often have distinctively different backgrounds. Björklund et al. (2006), for instance, document that Swedish-born adoptees are mostly born in less-advantaged families but placed in more-advantaged families.

[8]Nonexperimental evidence suggests that donor-treated parents under less restrictive donor assignment rules tend to choose donors who resemble the (infertile) partner (Nielsen et al. 1995).

[9]When we focus on the children in our sample who are old enough to be in college, we find that the test scores we use are indeed strong predictors of college attendance (see also online Appendix D).

They examine how prenatal smoking affects child outcomes and find that prenatal smoking reduces birth weight in genetically unrelated and related children. Their approach has two potential problems: the assignment of donors to recipient mothers may not be random, and survey response may be selective. Our study differs from theirs in focus and approach. We are not only asking a different question, but we are also answering it with a more convincing empirical approach and better data; that is, we exploit the quasi-random assignment of egg donors to more credibly identify the mother-child skill relationship, net of genetic skill transfers, and we use an administrative sample with information on all donor treated mothers in Denmark.

3 IVF in Denmark: Institutions and Data

In this section, we describe the institutional setting for IVF treatments in Denmark, with emphasis on the use of donor eggs and sperm, and discuss how we construct our data from several administrative registers.

IVF Institutions

Danish couples who experience fertility problems typically visit their general practitioner for medical advice and fertility testing. When childless couples are medically diagnosed as infertile, their general practitioner can refer them to a fertility clinic or hospital. In case the women in infertile couples are below the age of 40, they are entitled to three IVF treatments at no cost.[10] Each year, about 2,500 couples start an IVF-treatment and the average success rate per IVF treatment is 25-30 percent. Most couples undergo about 3 treatments, on average, leading to an overall success rate of 70-75 percent.

The standard IVF procedure works by collecting eggs, fertilizing eggs with sperm in a laboratory environment, and implanting the most promising embryo(s) back into the womb. Most IVF treatments involve the couples' own eggs and sperm. Some IVF treatments, however, involve either donor eggs or donor sperm. Over the period we consider, Danish law prohibits the fertilization of donor eggs with donor sperm, the argument being that the child should be genetically related to at least one of the parents.

[10]While Danish law has set 46 as the maximum treatment age, public clinics generally use 40 as the threshold. The annual costs for IVF-related medication is born by the couple and amounts to about 4,000 DKK annually (which corresponds to US$640 in 2016). Free IVF treatments applies to first-born children only. In our study, single women were not allowed to undergo IVF treatments. A law change in 2007, however, made it possible for single women to undergo IVF treatment with donor sperm.

This implies that children conceived with donor eggs are genetically unrelated to their mother but genetically related to their father, and vice versa.

The process of using donors in Denmark is highly regulated. Again, over the period we consider, egg and sperm donations were anonymous. Donor recipients (as well as the children born with donated eggs or sperm) were not informed about the identity of their donor, and vice versa, donors were not informed about the identity of their donor recipient. Donors were neither informed about the outcome of the treatment (that is, whether their donated eggs or sperm resulted in a pregnancy) nor could they claim legal parenthood over the children born with their donated eggs or sperm. Parents of donor children are encouraged to tell their children about their donor status around the age of five but are not legally obliged to do so. Survey evidence suggests parent-child relationships of lower quality in donor families that did not tell their children about their donor conception (Golombok et al. 2011).

Danish law also regulates who can become an egg donor. Up to 2007, which is also the study period we consider, only women who underwent IVF treatment themselves were allowed to donate eggs. Donor eggs were *surplus eggs* from IVF-treated women who produced more eggs than needed for their own IVF treatment. To ensure donor egg quality, egg-donor candidates were medically screened before they could donate eggs and, once approved, egg donors had to be younger than 35 when they donated eggs. Monetary compensation was not allowed. Only few women volunteer to donate eggs; over the 1994-2007 period, there was a donor egg shortage (donor egg demand exceeded donor egg supply). From 2007 onwards, Danish law allows also other women to donate eggs to address the donor egg shortage.

Danish law is less restrictive for sperm donors. While the law dictated that, in IVF treatments, only clinics could buy donor sperm from anonymous donors, most men were allowed to donate sperm. Like egg-donor candidates, sperm-donor candidates were medically screened. Sperm-donor candidates with a family history of serious hereditary mental and physical disorders were rejected. Once approved, candidates were repeatedly tested for infectious diseases for the full duration of the donation period. For sperm donations, a small monetary compensation was allowed. Many men donate sperm; over the 1994-2007 period, sperm banks held enough sperm to treat all infertile couples in need of donor sperm (donor sperm supply exceeded donor sperm demand).

IVF Register

In our empirical analyses, we exploit data from the Danish IVF register, currently held by the Danish Health Data Authority (Sundhedsdatastyrelsen).[11] The register contains information on IVF treatments taking place in public and private fertility clinics and hospitals in Denmark from 1994 onwards. We focus on couples who underwent IVF treatment in the period 1994-2007. Due to reforms in donor legislation (described earlier), it will be harder for us to examine (and ensure) quasi-random assignment of egg donors and egg-donor recipients from 2007 onwards.

The register covers information on the main reason for infertility, the mode of treatment, the use of donor eggs and donor sperm, the number of eggs retrieved from the womb, the number of fertilized eggs transferred back, the date of treatment and clinic identifiers. It also records whether treated women agreed to donate eggs and, if so, how many. We have merged the IVF register to other administrative registers to get longitudinal information on standard demographic variables (including birth year, gender, immigrant status, marital status, number of children, and education) running from 1991 to 2016 and standard labor market variables (including labor force status and annual earnings) running from 1991 to 2012.

To study intergenerational mobility patterns in education, we require educational outcomes of IVF-treated parents and their children. We use data from the Danish Education Register, which holds records on educational achievement in primary, lower and upper secondary, and tertiary education from the early 1970s onwards. For parents, we observe realized educational outcomes and take years of schooling as our main parental outcome. For their children, we observe test scores taken from multiple nationwide tests (including 4 tests in reading and 2 tests in math) that were introduced in Danish primary and lower secondary schools in 2010. Most children in our sample window have taken 3 to 4 tests. For each test, we calculate the standardized test score based on the test score mean and standard deviation in the representative sample of children that take the test (which we discuss in more detail below). Our main child outcome is the average of all available standardized test scores in reading and math.[12]

Our main analysis sample is restricted to those IVF children for whom we observe the score of at least one nationwide test. We select 19,509 children in the IVF register,

[11]Lundborg et al. (2017) analyze the IVF register in another context: they exploit IVF treatment success at the first IVF treatment as a natural experiment to estimate the causal effect of having children on the career of women.

[12]While we also examine test score performance for each reading and math test separately, we prefer the overall average of multiple test scores as the main child outcome for two reasons. First, and most importantly, it raises precision and reduces the influence of outliers. Second, it accounts for parental influences that possibly spill over across the different tests.

Table 1
Summary statistics

	IVF non-donor children	IVF sperm-donor children	IVF egg-donor children	all other children	all adopted children	adopted children in IVF-treated families
	(1)	(2)	(3)	(4)	(5)	(6)
	Child characteristics and outcomes:					
standardized test score (combined)	0.12	0.26	0.17	0.01	–0.03	–0.05
	0.86	*0.86*	*0.86*	*0.90*	*0.91*	*0.89*
male	0.51	0.53	0.56	0.51	0.44	0.40
birth year	2001.86	2001.41	2000.18	2001.05	2001.75	2002.62
number of children	19,509	820	157	650,930	2,674	1,196
number of multiple births	7,301	336	48	26,066	0	0
	Parent characteristics and outcomes:					
Pre-treatment characteristics						
years of schooling father	12.90	12.71	13.04	12.59	13.16	13.44
	2.39	*2.34*	*2.55*	*2.38*	*2.53*	*2.44*
college education father (1/0)	0.30	0.29	0.36	0.25	0.37	0.39
pre-treatment earnings father	32.60	32.37	32.09	26.06	31.97	37.90
(in 10,000 DKK)	*19.24*	*16.87*	*19.20*	*16.47*	*20.40*	*20.50*
immigrant status father (1/0)	0.05	0.03	0.05	0.11	0.03	0.02
birth year father	1966.44	1964.87	1962.82	1968.25	1965.99	1964.95
years of schooling mother	13.28	13.16	12.75	12.98	13.54	13.76
	2.29	*2.29*	*2.27*	*2.35*	*2.37*	*2.20*
college education mother (1/0)	0.41	0.39	0.32	0.37	0.48	0.50
pre-treatment earnings mother	23.87	23.52	23.68	16.78	23.30	27.24
(in 10,000 DKK)	*12.55*	*11.56*	*11.95*	*11.61*	*14.09*	*13.61*
immigrant status mother (1/0)	0.06	0.05	0.03	0.11	0.03	0.03
birth year mother	1968.95	1968.79	1964.13	1970.83	1967.64	1966.19
years of schooling egg donor	–	–	12.88	–	–	–
	–	–	*2.18*	–	–	–
college education egg donor	–	–	0.32	–	–	–
pre-treatment earnings egg donor	–	–	22.43	–	–	–
(in 10,000 DKK)	–	–	*10.06*	–	–	–
immigrant status egg donor (1/0)	–	–	0.03	–	–	–
birth year egg donor	–	–	1968.30	–	–	–
missing information egg donor	–	–	0.36	–	–	–
number of IVF attempts	2.75	2.83	3.04	–	–	–
Post-treatment outcomes						
(first 5 years following child birth)						
relative change in annual earnings father	0.07	–0.00	0.10	–	–	–
parental leave days father	19.08	16.67	12.15	16.27	27.36	28.72
relative change in annual earnings mother	–0.10	–0.17	–0.09	–	–	–
parental leave days mother	332.06	298.59	298.10	269.23	251.52	267.28
divorce	0.17	0.21	0.11	0.27	0.34	0.11
number of children	19,509	820	157	650,930	2,674	1,196
number of mothers	14,200	617	127	406,109	2,328	1,000

Note–The table shows means (with standard deviations in italics) for different intergenerational samples of children with test scores: (i) all non-donor IVF children; (ii) all sperm-donor IVF children; (iii) all egg-donor IVF children; (iv) a representative sample of all other children; (v) a representative sample of adopted children; and (vi) a representative sample of IVF-treated families with adopted children. Online Appendix Table B.1 contains the definition of all variables who were not conceived through either donor sperm or donor eggs, 820 sperm-donor children, and 157 egg-donor children. Of these egg-donor children, we are able to match (with some certainty) 97 to their egg donors (see Section 4 for details).

For comparison purposes, we construct two additional samples. The first sample is the representative sample, which is a 30 percent random sample of families with non-adopted children born in similar years as the IVF children. The second sample is our adoption sample, which is a 30 percent random sample of families with children adopted from abroad. We observe 650,930 non-adopted and non-IVF children and 2,674 adopted children. Of these adoptees, 1,196 were adopted in families who experienced failed IVF treatments.[13]

Table 1 provides sample means for the intergenerational samples. We make three informative comparisons. First, IVF children, and donor children in particular, perform much better in nationwide tests than most other children. Compared to children in the representative sample, for instance, we find that IVF children have 0.12-0.26 standard deviation higher test scores. Second, IVF-treated parents and representative parents tend to be different. IVF-treated parents are older and in more stable relationships than the representative parents. When we look at education, which is the parental characteristic at the center of this study, the differences are less pronounced. Compared to the educational levels of parents in the representative sample, IVF-treated fathers (but not IVF-treated mothers) are more educated, but not by much. Compared to the education levels of parents in the adoption samples, however, IVF-treated parents and representative parents appear much more similar than different. And third, we find that, among IVF treatments with egg donors, donor recipients are much older than egg donors, which is consistent with the age restriction of 35 imposed on egg donors. But when we look at their educational levels, we find that IVF-treated women and their egg donors are nearly identical.

4 Assignment of Donors to Families

In this section we document how the sperm and eggs of donors were assigned to Danish couples in IVF treatments. We are particularly interested in the extent to which donor assignment occurs randomly.

[13]Over the period we consider, prospective adoptive parents have to meet formal criteria (regarding age, marital stability, income, housing, mental and physical health, and crime history) and follow a pre-adoption counseling course before they are qualified to adopt (see Danish adoption law BEK nr. 283 af 20/03/2019).

The Donor Assignment Process

As we mentioned above, donor assignment is bound by strict rules on donor anonymity. While donors are strictly anonymous, prospective parents can state their preferences for donors on five dimensions: skin color, hair color, eye color, weight and height. These stated preferences for donor characteristics are expressed in a donor market with an excess supply of sperm donors and excess demand for egg donors, which implies that these preferences are likely met for sperm donors, but not for egg donors.

With an excess supply of sperm donors, sperm donors are assigned based on the prospective parents' preferences, which makes donor assignment in principle random conditional on these stated preferences. The IVF register does not record these preferences. In our case, the intergenerational results for fathers are not consistent with any strong selection based on preferences for these five donor characteristics, as we will show later.

With an excess demand for egg donors, prospective parents in need of an egg donor are placed on a waiting list. Fertility clinics organize their own waiting lists. There are in total 21 fertility clinics (including public clinics, private clinics, and hospitals that offer IVF treatments). Prospective parents choose one fertility clinic which, together with the shortage of eggs, means that donor assignment depends on the position on the clinic-specific waiting list rather than on preferences for donor characteristics. To substantiate the claim, we quote from the guidelines of one of the largest IVF clinics in Denmark (Ciconia): "Because of the long waiting time, it is not possible to match physical characteristics. You are offered donor eggs in the same order as you have been put on the waiting list" (Ciconia 2015).[14] The IVF register contains detailed records of the date and place of the IVF treatment. If we take the date of the first donor treatment (measured in calendar months) and a full set of clinic indicators to accurately proxy the clinic-specific waiting list order, the assignment of donor eggs to prospective parents should be as good as random conditional on the calendar month of first donor treatment and clinic-fixed effects.

Is Donor Assignment Conditionally Random (for Women)?

With the IVF register at hand, we can identify (with some certainty) the recipients' egg donor and empirically assess whether egg donor assignment is conditionally random.

[14]We do not want to a priori rule out the possibility that clinics allow for some rudimentary form of matching on the ethnic origin of donors and recipients. But its possibility is small (only 3 to 5 percent of all IVF treatments involve immigrant couples) and accounted for in the analysis (our intergenerational regressions that follow always control for recipient immigrant status). We can therefore treat the waiting list principle as the primary matching mechanism for donor eggs.

Table 2
Relationship between egg-donor characteristics and donor-recipient education

	donor years of education		**donor age**		**donor immigrant**		**donor earnings**	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
recipient years of education	–0.002	–0.006	–0.012	0.003	–0.002	–0.003	0.185	0.206
	0.031	*0.038*	*0.058*	*0.062*	*0.002*	*0.003*	*0.216*	*0.234*
recipient birth year		–0.016		0.013		–0.001		–0.002
		0.022		*0.029*		*0.001*		*0.140*
recipient immigrant status		0.103		0.018		0.028		0.327
		0.467		*0.545*		*0.022*		*2.089*
partner years of education		0.020		–0.042		0.003		–0.029
		0.049		*0.038*		*0.003*		*0.253*
partner birth year		–0.011		–0.029		–0.0003		0.071
		0.022		*0.028*		*0.002*		*0.106*
partner immigrant status		0.072		0.288		0.018		–0.710
		0.622		*0.592*		*0.021*		*3.126*
R-squared	0.008	0.025	0.005	0.008	0.003	0.019	0.009	0.010
joint p-value		0.128		0.459		0.115		0.221
joint p-value (w/o immigrant status)		0.272		0.759		0.455		0.161
number of observations	533	533	533	533	533	533	517	517
number of egg donors	419	419	419	419	419	419	408	408
number of donor recipients	364	364	364	364	364	364	342	342
clinic FE and treatment date controls	✓	✓	✓	✓	✓	✓	✓	✓

Note–The dependent variables are egg-donor years of education, age, immigrant status, and pretreatment earnings. In columns 1, 3, 5 and 7, we regress donor characteristics on recipient years of education. In columns 2, 4, 6 and 8, we add recipient birth year and immigrant status, and recipient's partner years of schooling, birth year and immigrant status (or partner averages of recipients treated in the same year and same clinic if partner information is missing). All regressions control for the first donor treatment date (in calendar months) and a full set of clinic indicators. The p-values for joint F-test indicate whether the recipient characteristics are jointly significant (all values indicate statistically insignificance). Standard errors are clustered by recipients and shown in italics; * indicates significance at 10 percent level, ** indicates significance at 5 percent level, and *** at 1 percent level.

In our IVF context where the nurturing influence of parental education is the treatment of interest, the natural verification test is to link the educational attainment of donor recipients to the educational attainment (or any other pre-assignment characteristic) of their donors. Conditional random assignment would predict zero associations, that is, after taking account of the assignment control variables.

The IVF register contains information on women who provide donor eggs and women who receive fertilized donor egg implants. For egg donors, the register records the number of donor eggs extracted and the extraction date. For donor egg recipients, the register records the number of donor eggs implanted and the treatment date, which may represent either the egg-implantation date (which occurs after the donors' extraction date) or the preparation-for-implantation date (which occurs before the donors' extraction date). With the treatment histories, we can link donor egg recipients to their egg donors based on the (correct) premise that fertility clinics take their egg donors from the women they treat and predominantly use fresh (fertilized) donor eggs as egg implants.[15] We define a match (with some certainty) if donor-egg recipients and egg donors are treated at the same clinic and the recipients' treatment date occurs within one week after the donors' extraction date (when recipients receive the embryo implants) or within seven weeks before the donors' extraction date (when recipients receive medication to prepare their uterus for pregnancy and egg implants). The matched pairs constitute our verification sample. Over the 1994-2007 period, we are able to identify 533 matches treated at the same clinic at the same period. With matched-pair observations, egg donors and egg-donor recipients may enter the sample several times. If egg donors produce enough viable donor eggs, they can serve multiple donor recipients. If previous IVF attempts with donor eggs failed, egg-donor recipients may be treated with eggs from multiple donors (in multiple treatments). There are, in total, 364 different donor recipients and 419 different egg donors.[16]

For a first glimpse at the donor assignment procedure, we calculate a raw correlation of 0.026 between the donors' and recipients' years of schooling. This number is quite low and already suggests that there is no systematic assignment procedure. Table 2

[15]Over the period we study, most treatments involve fresh embryo transfers. For the years 1994-1995, Westergaard et al. (2000) report that 90 percent of all embryo transfers were fresh embryo transfers. For the years 2006-2011, when we have information on fresh and frozen embryo transfers in IVF treatments, 86 percent of egg-donor treatments made use of fresh (fertilized) eggs.

[16]The egg donor recipients in the verification sample only partially overlap with those in the intergenerational sample. The verification sample, for instance, contains more egg-donor recipients because we sampled all egg-donor recipients including those whose donor-egg implants did not result in children and those whose children had no available test scores. Of the 533 donor treatments in the verification sample, only 118 were successful and lead to children. We observe test scores for 97 children. The intergenerational sample, on the other hand, contains more egg-donor recipients with children because we could not match all egg-donor recipients to their donors.

presents the more formal test results. In particular, we estimate how pre-assignment characteristics of egg donors (education, age, immigrant status, and pre-treatment earnings) relate to the educational attainment of egg-donor recipients after controlling for the waiting list variables (month of first donor treatment and clinic-fixed effects). We measure pre-treatment earnings by taking averages over the annual earnings (in 10,000 DKK) observed in the 4 years preceding the treatment year. In columns 1, 3, 5 and 7, we regress the four donor outcomes on recipient years of education and waiting list controls. In columns 2, 4, 6 and 8, we augment the regression with recipient birth year and immigrant status, and the recipient's partner years of education, birth year and immigrant status. This specification most closely resembles the intergenerational mobility specification we use throughout this study. At the bottom of the table, we present p-values for joint F-tests for whether the recipient (and recipient partner) characteristics are jointly statistically significant. Because there is some uncertainty about immigrant status being an appropriate characteristic to test for random donor assignment, as we mentioned earlier in footnote 14, we also present joint test results without the immigrant status measures.

We find that, conditional on the recipient's position on the waiting list, there is no relationship between the four donor characteristics and the recipient's education. All the education estimates are statistically insignificant and small. We also find that there is no relationship between the four donor characteristics and all other characteristics of the recipient and her partner. Test results are, in general, as one would expect with conditional random assignment of donor eggs to donor recipients.[17]

5 Empirical Strategy

In our main analysis, we take the sample of IVF-treated families and distinguish parents of non-donor children from parents of donor children. We begin by presenting a simple reduced-form intergenerational mobility model where both parents influence

[17]Even if donor assignment is conditionally random at each match, assignment need not be random for each match that leads to a successful pregnancy. We find no evidence, however, that assignment turns selective once we condition on successful treatments. We provide three pieces of evidence. First, we find that our estimates for recipients and partners remain statistically insignificant and close to zero when we add a successful treatment dummy (which is 1 if the recipient got pregnant and had a child, and 0 otherwise) to the assignment regressions (online Appendix Table B.2, odd-numbered columns). Second, we find again no statistical relationship between any of the donor and recipient characteristics when we restrict the sample to matches where recipients got pregnant and had a child (online Appendix Table B.2, even-numbered columns). We recognize, though, that the smaller sample of successful matches makes it also less likely to uncover statistically significant estimates. And third, we find that the intergenerational associations do not change when we control for egg donor education in our intergenerational skill transmission regressions (Table 3, columns 3 and 4).

the educational achievement of their non-donor children

$$Y_i^c = \alpha^c + \alpha^m Y_i^m + \alpha^f Y_i^f + \delta' X_i + e_i^c. \tag{1}$$

In this regression model Y_i^c represents an intermediate educational achievement outcome (measured in nationwide achievement tests in primary and lower secondary education) of child c born and raised in family i with mother m and father f, Y_i^m and Y_i^f represent the educational achievement of the child's mother and father (measured in completed years of schooling), X_i represents a set of pre-determined child, family, and treatment variables (including the child's gender, birth year and multiple-birth status, the mother's and father's birth year and immigrant status, the first treatment date and a full set of clinic indicators), and e_i^c represents exogenous child-specific characteristics. We measure birth year in continuous years and first treatment date in continuous calendar months because of sample size considerations. We control for multiple-birth status because many IVF treatments result in twin (or higher-order) births. The intergenerational coefficients α^f and α^m measure the intergenerational association between the educational achievement of genetically related children and parents and represent an unknown blend of nurture and nature influences. With samples of non-donor children, ordinary least squares (OLS) estimation of (1) yields estimates of α^f and α^m.

Our data on donor children allow us to isolate the nurturing component from the intergenerational coefficients β^m and β^f. The intuition for how we identify the nurturing effects follows from a simple stylized intergenerational transmission model, akin to the transmission model of Björklund et al. (2006), where all parents and donors can influence the child's education

$$Y_{ijk}^c = \beta^c + \beta^m Y_i^m + \gamma^m y_j^m + \beta^f Y_i^f + \gamma^f y_k^f + \theta' X_i + e_{ijk}^c, \tag{2}$$

where the subscripts i, j and k stand for child c raised and born in family i but conceived by egg donor j and sperm donor k, the superscripts m and f stand for the child's mother and father, Y_i^m and Y_i^f represent the observable measures of the educational achievement of the child's rearing (and genetically unrelated) mothers and fathers, y_j^m and y_k^f represent the unobservable measures of the educational achievement of the child's genetically related egg- and sperm-donor providers, X_i represents a set of pre-determined child, family, and treatment variables (as defined in (1)), and e_{ijk}^c represents unobserved child-specific characteristics.[18]

[18]While this is clearly an oversimplified description of the real world (it ignores, for instance, that test scores of children can be affected by possible interactions between nature and nurture), Björklund et al. (2006) show that this very simple model characterizes the intergenerational transmission of education surprisingly

The intergenerational coefficients γ^f and γ^m measure the intergenerational associations between the educational achievement of the child and donors and represent the nature effects. The intergenerational coefficients β^f and β^m measure the intergenerational associations between the educational achievement of the child and genetically unrelated parents and represent the nurture effects. These nurture effects, which are the prime targets of estimation, must be interpreted broadly and capture the causal influence of parental education and any other unobserved parenting/nurturing skill that is correlated with it; that is, the nurture effects capture the influence of those parenting/nurturing skills that can be both the cause and consequence of parental education.

In case of egg donor children raised in family i with genetically unrelated mother i, egg donor j, and genetically related father i (with identical Y^f and y^f), donor assignment follows a clinic-specific waiting list principle. With controls for first treatment date (measured in calendar months) and a full set of clinic indicators to account for the position mothers take on the waiting list, we can ignore the influence of the donor's genes, as if γ^m is zero, and rewrite the intergenerational transmission model as

$$Y_{ij}^c = \beta^c + \beta^m Y_i^m + (\beta^f + \gamma^f) Y_i^f + \theta' X_i + e_{ij}^c. \tag{3}$$

With samples of egg-donor children, direct estimation of (3) gives us an unbiased nurture estimate of β^m.[19]

In case of sperm-donor children raised in family i with genetically related mother i (with identical Y^m and y^m), genetically unrelated father i, and sperm donor k (with unknown y_k^f), donor assignment is guided by preferences for the five donor characteristics (skin color, hair color, eye color, weight and height). Adding these variables to X in the intergenerational transmission model would analogously eliminate the genetic influences of the sperm donor

$$Y_{ik}^c = \beta^c + (\beta^m + \gamma^m) Y_i^m + \beta^f Y_i^f + \theta' X_i + e_{ik}^c. \tag{4}$$

The problem is that we do not observe these preference variables. With the preferred donor characteristics excluded from (4), the estimate of β^f may be biased and capture not only the nurture effect of the rearing father but also part of the nature effect of the donor, that is, if the educational outcomes of rearing fathers and sperm donors are somehow related through the omitted preferences. There are, however, good a priori

well, at least for Swedish adoptees born in 1962-1966.

[19] Our data on matched egg donors make it possible to estimate the same intergenerational models with the educational attainment of the matched egg donor as additional regressor. If assignment is conditionally random, the estimated nurture effect should not change when we include the matched egg donor's educational attainment. This is indeed what we document later (in Table 3).

reasons to believe that these omitted preferences cause only little (upward) bias, given that donors are anonymous and parents can only choose out of five donor traits that are at best crude proxies for the donor's educational attainment.

Equations (3) and (4) relate the educational attainment of children to the educational attainment of *both* rearing parents. We purposefully do so to take account of confounding assortative mating effects. Would we exclude the genetically related father from (3), for example, the estimate of β^m captures not only the direct influence of the mother's education (representing the nurture effect) but also the indirect influence of the father's education (representing a mixture of nurture and nature effects) because more educated mothers tend to marry more educated fathers. More specifically, we do not want that the father genes confound the nurture effect of mothers, and vice versa, that the mother genes confound the nurture effect of fathers. With the educational attainment of both rearing parents in the same specification, we can better separate out the direct effects of the genetically unrelated parent from the indirect effects of the other genetically related parent. In our sensitivity analysis (see online Appendix D), we provide some evidence that confounding assortative mating effects are of little concern.[20]

6 Results

Table 3 presents the main intergenerational transmission estimates for education. We report estimates of regressions of our standardized test scores (for reading and math combined) on the years of education of both parents, with controls for the child's gender, birth year and multiple-birth status, the parents' immigrant status and birth year, and the calendar month of first IVF treatment and clinic-fixed effects. These control estimates are not reported. We run separate regressions on samples of IVF-treated parents with non-donor, sperm-donor, and egg-donor children.[21]

In column 1 we begin with the intergenerational mobility associations obtained from the sample of non-donor children. For both parents, we find that the estimated associations between parental schooling and child test scores are large, positive, and statistically significant indicating that higher educated parents have on average children with higher test scores. The overall magnitudes of these associations, which differ only

[20]Collado et al. (2023) make a similar point when they model and quantify intergenerational transmission of education using extended family members (defined by common great grandparents) and their spouses. They find, like us, that partners hardly sort on genetic factors that drive their educational attainment.

[21]For the sample of egg-donor children, it is essential to control for treatment date and clinic-fixed effects to ensure that the assignment of donor eggs is conditionally random. For the other samples, however, it is not. Online Appendix Table B.3 shows that the intergenerational transmission estimates for non-donor and sperm-donor children do not change when we remove treatment date controls and clinic-fixed effects from the regressions.

Table 3
Regressions of child test scores on parental education using IVF children

	non-donor child	sperm-donor child	egg-donor child	
	(1)	(2)	(3)	(4)
years of education mother	0.076	0.066	0.072	0.073
	*0.004****	*0.019****	*0.026***	*0.022***
years of education father	0.064	–0.008	0.063	0.062
	*0.003****	*0.015*	*0.024***	*0.023***
years of education egg donor				0.008
				0.042
R-squared	0.132	0.090	0.258	0.261
number of observations	19,509	820	157	157

Note–The dependent variables are the averaged standardized achievement test scores for reading and math. The independent variables are the parents' educational attainment measured in the nominal years spent in school. All specifications control for the gender, multiple-birth status and birth year of children, immigrant status and birth year of rearing mothers and fathers, and the date of first treatment (in calendar months) and a full set of clinic indicators. These intergenerational mobility regressions are run on three different samples of IVF-treated parents with non-donor children (in column 1), sperm-donor children (in column 2), and egg-donor children (in columns 3 and 4). In the last specification (column 4), we have added a control for educational attainment of the egg donor and an imputation indicator. In case the egg donor is not identified (60 out of 157 observations), we have imputed the educational attainment of the egg donor with the average educational attainment of the donors treated in the same year and same clinic. Standard errors are clustered at the clinic level and shown in italics; * indicates significance at 10 percent level, ** indicates significance at 5 percent level, and *** at 1 percent level.

a little for mothers and fathers, tell us that four more years of parental education of either parent are associated with children having about 0.26-0.30 standard deviation higher test scores. These estimates are comparable to those estimated in previous intergenerational mobility studies (Hægeland et al. 2010).

In column 2 we run the same intergenerational mobility regressions on the sample of sperm-donor children. Here the intergenerational associations are supposed to take account of the father's genes. For fathers of sperm-donor children, we find that the estimate gets much smaller. In fact, the estimated association, which we interpret as the nurture effect, is close to zero, statistically insignificant, and (statistically) significantly smaller than the one obtained for fathers of non-donor children.[22] For mothers of sperm-donor children, however, we find that the intergenerational association between the educational attainment of mothers and child test scores remains large, positive, and statistically significant, and statistically similar to the one obtained for mothers of

[22]A critical concern for our nurture effect interpretation is that the assumption of almost random sperm donations may not hold in regressions without donor preference controls. Would the assignment of sperm donors be selective, part of what we interpret as the nurture effect may in fact be genetic. With nurture estimates close to zero, however, we consider such a bias unlikely. Online Appendix C provides a more formal intergenerational transmission model with selective donor assignment and shows that the nurture effects we estimate are most consistent with a quasi-random donor assignment procedure.

non-donor children.

In column 3 we switch to the sample of egg-donor children and report intergenerational associations that take account of the mother's genes. For mothers of egg-donor children, we find that the intergenerational association is large, positive, and statistically significant, and as large as the one obtained for mothers of non-donor children. For fathers of egg-donor children, we also find a positive, sizable, and statistically significant association that is practically identical to the one obtained for fathers of non-donor children. In column 4 we expand the baseline regression model for egg-donor children with the matched egg donor's educational attainment (in which we replaced the missing observations with the average education of egg donors treated in the same year and clinic). We expect similar intergenerational association for rearing mothers because the assignment of donor eggs is essentially random. This is indeed what we find. The intergenerational associations for the rearing parents hardly change. The intergenerational association for the matched donors is small, although imprecisely estimated.[23]

Two tentative conclusions can be drawn from these findings. The first is that the education of genetically unrelated fathers is of little help to their children's test scores. Based on the estimates obtained in donor families, we find that the nurture effect of paternal education on child test scores is close to zero. The second is that the education of genetically unrelated mothers does matter. Based on the estimates obtained in donor families, we find that the nurture effect of maternal education on child test scores is large, and as large as the overall intergenerational effects for genetically related mothers. Taken at face value, these nurture effects are consistent with the common notion that mothers are the primary caretakers of children, being much more involved in their children's primary (and lower secondary) schooling than fathers.

Separate Results by Field and Grade

We measure the educational achievement of children by overall test performance averaging the standardized reading and math test scores over the different grades. One concern is that the test score gains we estimate may depend on test score subject and

[23]In our intergenerational mobility sample of rearing mothers of egg-donor children, we are able to identify the likely donor for 64 percent of all donor children. We have replaced missing education with the average education of the donors treated in the same year and clinic, and added a missing dummy for those donor children without donor information to the intergenerational mobility model. Excluding the missing observations does not affect our results. When we estimate the intergenerational model in column 4 using the subsample deleting the missing observations, the coefficients on the mother's education, father's education, and the matched donor's education are 0.062 (with a standard error of 0.043), 0.072 (with a standard error of 0.038), and -0.001 (with a standard error of 0.043), respectively. While less precise, these estimates are broadly similar to those reported for the full sample.

grade. Table 4 therefore presents the intergenerational associations for each separate reading and math test taken in different grades for our samples of non-donor children, sperm-donor children, and egg-donor children.

For IVF parents of non-donor children (in columns 1 and 4), the intergenerational associations (capturing the overall impact of nurture and nature influences) are all positive, statistically significant, and nearly identical across the different grades and fields. For fathers of sperm-donor children (in columns 2 and 5), the estimated nurture effects are also very similar and almost all statistically insignificant and close to zero, regardless of field and grade. We have no clear explanation for the negative and significant nurture impact on math test scores in grade 8. What we can say, though, is that all our estimates are consistent with more educated fathers being of little help to their children's test scores. For mothers of egg-donor children (in columns 3 and 6), we see that sample sizes get very small. We find nonetheless significant, positive, and nearly identical nurture effects for reading test scores across most grades. The estimated nurture effects for math in grades 3 and 8 are notably lower, but neither effect is precisely estimated. In particular, the latter effect estimates for math are not precise enough to rule out substantial positive effect sizes as large as 0.067 in grade 3 and 0.168 in grade 8. So statistically speaking, we cannot exclude that the maternal nurture effects for reading and math are broadly similar.

On the whole, we find only minor differences in the nurture effects across the primary and lower secondary school years. Since little seems lost by focusing on the average of all available test scores in both reading and math, we will take the combined reading-math test scores as the main child outcome in the analysis to follow. One has to keep in mind, though, that the positive and significant nurture effects we estimate for rearing mothers are primarily driven by their impact on child test scores in reading (and not in math).

Table 4
Regressions of child test scores on parental education using IVF children by field and grade

	non-donor child	**sperm-donor child**	**egg-donor child**	**non-donor child**	**sperm-donor child**	**egg-donor child**
	(1)	(2)	(3)	(4)	(5)	(6)
	Reading test scores in grade 2			**Reading test scores in grade 6**		
education mother	0.077	0.075	0.060	0.078	0.043	0.120
	*0.005****	*0.027****	*0.047**	*0.004****	*0.027*	*0.035****
education father	0.054	–0.017	0.111	0.066	0.015	–0.007
	*0.005****	*0.021*	*0.056**	*0.004****	*0.023*	*0.039*
R-squared	0.108	0.157	0.397	0.119	0.107	0.356
observations	12,198	428	69	9,777	380	96
	Reading test scores in grade 4			**Reading test scores in grade 8**		
education mother	0.073	0.086	0.099	0.076	0.043	0.126
	*0.005****	*0.022****	*0.056**	*0.006****	*0.033*	*0.035****
education father	0.063	–0.018	0.128	0.064	0.001	–0.004
	*0.005****	*0.027*	*0.053***	*0.006****	*0.029*	*0.039*
R-squared	0.112	0.126	0.489	0.117	0.095	0.320
observations	11,381	360	88	7,489	377	89
	Math test scores in grade 3			**Math test scores in grade 8**		
education mother	0.068	0.046	0.001	0.077	0.080	0.042
	*0.004****	*0.021***	*0.033*	*0.008****	*0.035***	*0.063*
education father	0.063	0.016	0.094	0.071	–0.061	0.110
	*0.004****	*0.021*	*0.057*	*0.005****	*0.036**	*0.065*
R-squared	0.088	0.121	0.195	0.102	0.119	0.267
observations	11,811	351	77	9,759	383	94

Note–The dependent variables are the standardized achievement test scores by grade and field. Each panel considers test scores from one particular nationwide test taken in one particular grade. The independent variables are the parents' educational attainment measured in the nominal years spent in school. All specifications control for the gender, multiple-birth status and birth year of children, immigrant status and birth year of rearing mothers and fathers, and the date of first treatment (in calendar months) and a full set of clinic indicators. These intergenerational mobility regressions are run on three different samples of IVF-treated parents with non-donor children (in columns 1 and 4), sperm-donor children (in columns 2 and 5), and egg-donor children (in columns 3 and 6). Standard errors are clustered at the clinic level and shown in italics; * indicates significance at 10 percent level, ** indicates significance at 5 percent level, and *** at 1 percent level.

Results for Parental Earnings

In our study on the intergenerational transmission of human capital, we follow (most of) the existing literature and use parental years of schooling as our primary human capital measure. We recognize, though, that human capital is multidimensional and that the parents' years spent in school capture only one of the human capital dimensions.

To address this, we consider a human capital measure that covers a wider set of skills. If we think of human capital as those skills that make (in our case) parents more productive and generate returns, parental earnings can serve as such an alternative human capital measure. In particular, we consider pre-child earnings (which is arguably more exogenous than post-child earnings) and examine how our results change when we replace the parents' years of schooling with the parents' pre-child earnings. In constructing this earnings-oriented human capital measure, we take averages of annual earnings (measured in 10,000 DKK) observed in the 4 years before the parents were successfully treated and had a child.

Table 5 (panel A, columns 1 to 3) contains the estimates for the parents' pre-child earnings. We find that the nurture effect estimates mirror those for years of schooling; that is, the estimate for the father's pre-child earnings in the sperm-donor sample is insignificant and small, whereas the estimate for the mother's pre-child earnings in the egg donor sample is significant and large. When we run regression models with years of schooling and pre-child earnings as complementary skill measures, we find the same nurture effect estimates as before (panel B, columns 4, 5 and 6). For fathers of sperm-donor children, the estimates for schooling and earnings together are close to zero and statistically insignificant. For mothers of egg-donor children, the estimates for schooling and earnings are both positive, sizable, statistically significant for schooling (but not for earnings), and do not differ much from those taken from regressions where schooling and earnings enter as the single skill control. When we express the estimated nurture effects in terms of standard deviation changes in maternal schooling (with a standard deviation of 2.27) and pre-child earnings (with a standard deviation of 11.92 DKK), the nurture effect sizes are 0.168 for schooling and 0.107 for earnings. Both magnitudes are sizable, suggesting that parental years of schooling and pre-child earnings capture relevant but different dimensions of human capital. We continue to conclude though that the intergenerational transmission of human capital skills remains qualitatively similar for education-oriented and earnings-oriented skill measures; that is, once we take account of genes, we find that only the human capital skills of mothers matter.

Table 5
Regressions of child test scores on parental earnings and education using IVF children

	non-donor child	sperm-donor child	egg-donor child	non-donor child	sperm-donor child	egg-donor child
	(1)	(2)	(3)	(4)	(5)	(6)
	A: parental earnings			**B: parental earnings and education**		
earnings mother	0.009	0.011	0.013	0.005	0.009	0.009
	*0.001****	*0.003****	*0.005****	*0.001****	*0.003****	*0.006*
earnings father	0.003	–0.001	0.007	0.002	–0.001	0.006
	*0.0004****	*0.002*	*0.004***	*0.0003****	*0.002*	*0.004*
education mother				0.070	0.056	0.072
				*0.003****	*0.019****	*0.028***
education father				0.061	–0.010	0.037
				*0.003****	*0.017*	*0.035*
R-squared	0.062	0.084	0.236	0.140	0.100	0.280
observations	19,509	820	157	19,509	820	157

Note–These estimates are taken from specifications that deviate from the baseline regression models (reported in Table 3). In panel A, we replace parental education for parental pre-child earnings as the main independent variable. In panel B, we include parental pre-child earnings as additional independent variable. Standard errors are clustered at the clinic level and shown in italics; * indicates significance at 10 percent level, ** indicates significance at 5 percent level, and *** at 1 percent level.

Other Specification Checks

We consider several specification issues that are common to the analysis of intergenerational skill transmission. With the overall test score average as our main child outcome, we examine how assortative mating, nonlinearities in intergenerational skill transfers (including nature-nurture interactions), twin births, parental health, and the number of IVF attempts (as proxy for the demand for children), and alternative parental and child skill measures affect our nurture effect estimates.

Online Appendix D presents and discusses these specification issues in detail. Online Appendix Table D.1 (panels A to L) contains the corresponding nurture estimates, which we briefly summarize here. In panels A and B we examine the sensitivity to partner sorting on genes. Whether we include controls for parent and partner income, health, and employment status (measured over the four years leading up to child birth) or exclude all partner characteristics, the nurture effects remain unaffected. In panels C and D we test for nature-nurture interactions. Whether we add education interactions between the genetically-unrelated parent and genetically-related partner or between the genetically-unrelated mother and genetically-related egg donor, we find little evidence that interactions matter much for the child test score outcomes. In panels E and F we examine the sensitivity to twin births (common to IVF treatments). Whether we combine twin observations or delete all twin observations from the samples, the nurture effect estimates do not really change. In panels G, H, and I, we examine how parental health (measured by fertility-related health disparities) and the number of IVF attempts affect our estimates. Whether we restrict the samples to men and women whose fertility problem is the main reason to seek treatment or include the number of IVF attempts as additional control, our intergenerational transmission estimates remain largely similar. In panels J, K and L we examine the sensitivity to alternative skill measures. Whether we replace parental education with a parental college indicator, express parental education and child test scores in ranks, or take predicted college based on test scores as child outcome, we continue to find the same nurture patterns as before.[24]

7 Generalizability of Results

An important question is whether the nurture effect estimates taken from parents of donor children are (informative about and) generalizable to other parents of non-donor children. In an attempt to answer this external validity question, we compare intergenerational mobility patterns across different samples of parents and children.

IVF Families

We start by comparing IVF families with and without donor children. Within these families, there are two reasons why the nurture effects may be different. One is that donor children are different from non-donor children (either because of donor selection or because of experiences specific to donor children that affect their development). Another reason is that parents of donor children are different from parents of non-donor children (either because of self-selection or because of experiences specific to nongenetic parents that affect their parenting). We perform several tests to detect such nurture differences.

We first compare donor and non-donor families and test whether the overall intergenerational transmission of skills (taking the nurture and nature effects together) is different in donor and non-donor families. Because donor children in donor families must be genetically related to one of their parents (by Danish law), we can estimate how the educational attainment of rearing fathers and mothers relates to the test scores of their genetically related children in both donor and non-donor families. If there are some

[24]While our study takes an intergenerational perspective, it is possible to contrast our intergenerational associations with sibling associations using different combinations of donor and non-donor sibling pairs. Online Appendix E takes a behavioral genetics perspective and decomposes the test score variance into a nurture and nature component using the different sibling associations.

unobserved family-specific characteristics in donor families that lead to weaker nurture effect for fathers in sperm-donor families, and stronger nurture effect for mothers in egg-donor families, we should see that such unobserved family-specific characteristics also lead to weaker intergenerational associations for mothers in sperm-donor families, and stronger associations for fathers in egg-donor families. Table 3 displays the main intergenerational transmission estimates. For mothers of sperm-donor and non-donor children, who bring both nature and nurture, we find that a one year increase in maternal education is associated with standard deviation gains of 0.070 and 0.076 in child test scores (see Table 3, columns 1 and 2). For fathers of egg-donor and non-donor children, we find that one more year of education leads to comparable standard deviation gains of 0.063 and 0.064 in child test scores (see Table 3, columns 1 and 3). With the overall intergenerational associations being nearly identical, we find no evidence that human capital skills are transmitted differently between parents and their genetically related children in donor and non-donor families.

We next compare donor and non-donor children and test whether they respond differently to their parents' upbringing. As suggested by Golombek (2021), we explore whether the zero nurture effects for fathers arise because donor children (and sperm-donor children in particular) turn less receptive to their (unrelated) parents' attention once they go through adolescence and struggle with their unknown biological background. When we examine how the nurture (and nature) effects evolve over the different school years (as reported earlier in Table 4), however, there is little evidence that nurture effects get weaker when donor children go through adolescence. For the fathers of sperm-donor children, we find that the estimated nurture effects are already close to zero in the earlier grades of primary education. And for mothers of egg-donor children, we find that the estimated nurture effects for reading test scores are sizable and positive throughout primary and lower-secondary education.

And lastly, we compare donor and non-donor parents and test whether they treat their children differently. As we do not have information on parental involvement and investments in child schooling, we cannot directly test for treatment differences. If we instead assume that more involved parents spend more time with their children (which we proxy with parental leave take up and labor supply) and invest more in a stable family environment (which we proxy with parental divorce risk), we can indirectly test whether donor and non-donor parents treat their children differently by comparing their parental leave take up, labor supply, and divorce risk during the child's preschool years. Again, our concern with fathers of sperm-donor children is that the zero nurture effects may arise if unrelated fathers are less involved fathers and, as a result, take up less parental leave, work longer hours, and face a higher divorce risk. And reversely, our

Table 6
Relationship between parents' time spent with children and donor type in IVF-treated families

	days parental leave^{a} (0-2 yrs)		**days parental leave^{a} (0-5 yrs)**		**labor supply change (0-5 yrs)**		**divorce (0-5 yrs)**
	mother (1)	father (2)	mother (3)	father (4)	mother (5)	father (6)	couple (7)
non-donor families (ref.)							
sperm-donor families	–0.476	–1.200	–1.599	2.651	–0.042	0.004	0.034
	5.793	*1.245*	*11.018*	*1.463**	*0.048*	*0.028*	*0.022*
egg-donor families	15.161	–4.098	–10.422	–4.461	0.036	0.054	–0.056
	10.266	*1.094****	*14.595*	*1.242****	*0.061*	*0.067*	*0.022***
R-squared	0.167	0.028	0.030	0.129	0.026	0.028	0.013
number of observations	15,387	15,387	11,577	11,557	14,535	14,229	15,387
mean (non-donor families)	263.881	14.275	346.985	18.557	–0.076	0.133	0.168

Note–The dependent variables are the parents' parental leave take up measured as the number of registered parental leave days during the first two and five years following child birth (columns 1 to 4), labor supply response measured as the percentage change between the average labor earnings in the four years before child birth and the average labor earnings in the first five years after child birth (columns 5 and 6), and divorce measured as indicator for whether married/cohabitating parents divorce/break up during the first five years following child birth (column 7). The independent variables of interest are indicators for egg- and sperm-donor families. The non-donor families serve as the reference group. The samples treat each donor family as a separate observation. All specifications control for the child's gender, multiple-birth status and birth year, the parents' immigrant status and birth year, the first treatment date (in calendar months) and a full set of clinic indicators. Robust standard errors are shown in italics; * indicates significance at 10 percent level, ** indicates significance at 5 percent level, and *** at 1 percent level.
Note a–Before 2002, the paid parental leave policies (in our study period) allowed parents to take up 90 days maternity leave, 20 days paternity leave and 50 days shared leave. After 2002, these paid parental leave periods were 90 days maternity leave, 10 days paternity leave, and 160 days shared leave.

concern with mothers of egg-donor children is that the positive nurture effects may arise if unrelated mothers are more involved mothers and take up more parental leave, work fewer hours, and invest more in their marriage.[25]

Table 6 reports estimates from least-squares regressions on the sample of all IVF-treated families, with four different dependent variables: parental leave take up measured as the number of registered parental leave days taken by the parents during the first two and five years following child birth, labor supply response of parents measured as the percentage change between the average labor earnings in the four years before child birth and the average labor earnings in the first five years after child birth, and divorce measured as an indicator for whether married/cohabitating parents divorce/break up during the first five years following child birth. The independent variables of interest are indicators for egg- and sperm-donor families. As before, we include (but not report on) pre-determined controls for the child's gender, multiple-birth status and birth year, and the parents' immigrant status and birth year in our regressions. Table 6 also reports summary statistics of parental leave take up, change in labor supply, and divorce rates. The sample sizes vary because of missing earnings averages and because of unavailable 5-year parental leave measures for the youngest cohorts of children.

Three things become clear from this table. First, the parental leave statistics show that all mothers spend much more time with children than fathers do. Second, none of the parental leave and labor supply estimates suggest that unrelated mothers spend more time and unrelated fathers spend less time with their donor children. The estimates are mostly statistically insignificant, and the few that are have the opposite sign.[26] And third, the divorce estimates, on the other hand, are sizable and have signs in the hypothesized directions; that is, compared to non-donor IVF-treated couples, we find that egg-donor couples are 6 percentage points less likely to divorce, whereas sperm-donor couples are 3 percentage points more likely to divorce.[27]

We perform two more tests to explore whether the different divorce rates observed in donor families can explain why the nurture effects are stronger for mothers than for

[25]Unrelated parents may be differently involved into child rearing for different reasons (Dawkins 2006, Holmlund et al. 2011). In our setting, unrelated fathers may be less involved fathers because they miss some evolutionary drive. Unrelated mothers, on the other hand, may be more involved mothers because of self selection, as they have put in much more effort than most other mothers to get pregnant and have children.

[26]We do not know why fathers of egg-donor children take up less parental leave days than other fathers do. What we do know, however, is that these negative parental leave estimates are not so relevant on their own given that fathers in general take up so few parental days.

[27]These results do not change when we additionally control for pre-determined socioeconomic characteristics, including pre-child earnings (average earnings in years t-1 to t-4), labor market participation (0 if parents do not work for pay in years t-1 to t-4), sick leave (average number of sick leave days in years t-1 to t-4), leisure time (average number of non-working days in years t-1 to t-4, 0 if parents work full-time), and a multiple birth indicator (see online Appendix Table B.4).

fathers. We first run our intergenerational mobility regression and replace child test scores with the couples' divorce status (or break-up status for cohabitating couples) in the sample of egg-donor children. With an estimated association between the mothers' education and divorce status (with standard errors in parentheses) of -0.003 (0.018), we do not find that more educated mothers face lower divorce risks. We second run our intergenerational mobility regression and additionally control for the couples' divorce status. With estimated nurture effects of -0.011 (0.016) for fathers of sperm-donor children, and 0.072 (0.026) for mothers of egg-donor children, we continue to find we the same nurture effects as before. This all suggests that divorce alone cannot explain the nurturing impacts observed in this study.

In sum, we recognize that there are several possible mechanisms explaining why the nurture estimates obtained in donor families may not generalize to non-donor families. Of those we examine, we find little evidence that donor and non-donor parents are different in how they transmit their human capital skills to their children.

Representative Families

We also compare IVF families and representative families. In particular, we compare intergenerational mobility patterns between IVF families with genetically related children and families drawn from the full population of families with children born in the same research time window. In both types of families, the intergenerational transmission estimates represent an overall blend of nature and nurture influences. If we would get the same intergenerational transmission estimates, we conjecture that the process of skill transmission is comparable in the two types of families, and that the intergenerational transmission estimates obtained with IVF families have a wider generalizability.

Table 7 presents the intergenerational transmission estimates for education for the different samples. In column 1 we report estimates for the sample of representative families. In column 2 we reproduce our baseline estimates for the sample of non-donor IVF families (as reported in table 3) for ease of comparison. We find that the estimated associations between parental schooling and child test scores are all large and positive. Although they are not identical, the intergenerational associations differ only a little across the two types of families.[28]

[28]We also explore whether the intergenerational associations vary by the child's gender in the samples of representative and non-donor families. Online Appendix Table B.5 contains these estimates. For the children in the representative sample, the gender differences are statistically significant but small and of little consequence. For the non-donor children in the IVF sample, we find that the gender differences are equally small and statistically insignificant. These results suggest, as before, that representative families and non-donor IVF families are comparable in terms of intergenerational skill transmission.

Table 7
External validity regressions of child test scores on their parents' education

	all other children	**IVF non-donor children**	**all adopted children**	**adopted children in families with IVF history**	**sperm-donor children**	**egg-donor children**
	(1)	(2)	(3)	(4)	(5)	(6)
years of education mother	0.085	0.076	0.019	–0.014	0.066	0.072
	*0.001******	*0.004******	*0.008******	*0.012*	*0.019******	*0.026****
years of education father	0.072	0.064	0.026	0.011	–0.008	0.063
	*0.001******	*0.003******	*0.008******	*0.011*	*0.015*	*0.024****
R-squared	0.162	0.132	0.127	0.155	0.090	0.258
number of children	650,930	19,509	2,674	1,196	820	157
country-of-origin FE			✓	✓		
clinic FE, calendar month		✓			✓	✓

Note–The dependent variables are averaged standardized test scores. The independent variables are the parents' educational attainment measured in the nominal years spent in school. All specifications control for the gender, multiple birth status and birth year of children, and the birth year and immigrant status of mothers and fathers. The specification in columns 3 and 4 further controls for country-of-origin fixed effect (for the 20 most popular countries of origin). The specifications in column 2, 5 and 6 further controls for clinic-fixed effects and calendar month of first donor treatment. Standard errors are clustered at the clinic level and shown in italics; * indicates significance at 10 percent level, ** indicates significance at 5 percent level, and *** at 1 percent level.

Adoptive Families

And lastly, we compare intergenerational mobility patterns in families with donor children and families with adopted children. In both types of families, we can estimate nurture effects, that is, when children are genetically unrelated to their rearing parents. The adoptees are suitable for identifying (and comparing) nurture effects: adoptees are adopted at infancy (the vast majority of adoptees is adopted below the age of 2) to ensure that the nurture effect estimates capture most of the early childhood influences; and adoptees are foreign born to ensure that the assignment process is fairly random.[29]

In column 3 we report the intergenerational transmission estimates for all foreign-born adoptees. In columns 5 and 6 we reproduce our baseline estimates for the sample of donor children (as reported in Table 3) for ease of comparison. We find small, positive, and statistically significant nurture effects indicating that higher educated parents provide a better nurturing environment for adopted children to perform well in school. The associations are similar for mothers and fathers and imply that four more years of parental education of either parent are associated with children having about 0.10 standard deviation higher test scores. In column 4 we estimate the intergenerational transmission models in families who after (a sequence of) unsuccessful IVF treatments ended up adopting a child. Because these adoptive families are arguably more comparable to IVF families in terms of observable and unobservable characteristics, we believe that the corresponding nurture effect estimates should in principle be more comparable to those obtained in donor families. When we compare these nurture effect estimates to those obtained in the full sample of foreign-born adoptees, we get smaller and statistically insignificant nurture effects.[30]

How are we to interpret the difference between the nurture effect estimates taken from adopted children to those taken from donor children? If we think of the nurture effect estimates obtained with adoptive parents and foreign-born adoptees as representative for all other parents and children (including IVF parents and donor children), the difference should capture the nurturing influence of parental education on the pre-

[29]As discussed in Holmlund et al. (2011), most adoptive parents know little, if anything, about the biological background of foreign-born adoptees. They know, like we do, the adoptees' gender, age, and country of origin. We therefore run our intergenerational mobility regressions with additional controls for the adoptees' country of birth (measured by indicator variables for the 20 most popular countries of birth) and assume that assignment of adoptees to families is conditionally random.

[30]As an alternative strategy, we also constructed more comparable adoption samples based on propensity score matching. The adoptees are reweighted to match either the sample of egg-donor children or the sample of sperm-donor children. The weights are based on the estimated propensity scores taken from probit regressions on the likelihood of either being an egg-donor or a sperm-donor child on the child's gender and birth year, their parents' years of schooling, immigrant status, and birth years. Whether we match on sperm-donor or egg-donor characteristics, we get nurture effect estimates that are almost identical to those obtained for all adoptees (see Appendix Table B.6, column 3 and 4).

natal and very early childhood environment. We find the only substantial difference for rearing mothers, that is, the intergenerational transmission estimate for mothers of egg-donor children is at least three times as high as the estimate for mothers of adopted children. When we test for differential impacts (by pooling the egg-donor and adoption samples and estimate columns 3 and 6 with a fully interacted regression model), we find that the nurture impact of egg-donor mothers is significantly larger than that of adoptive mothers (the estimated difference is equal to 0.053 and comes with a standard error of 0.023). These nurture effect estimates appear consistent with the view that more educated mothers are better able in creating a prenatal and early childhood environment that improves child test scores.

Prenatal Conditions

With data on risky behavior of pregnant women and child health outcomes measured at birth, we can take a closer look at the relationship between prenatal environment and the education of (prospective) mothers. For this purpose, we draw two supplementary samples from the IVF register. For the IVF-treated women, the IVF register contains information on self-reported smoking behavior during pregnancy for the years 2006 to 2011, which enables us to construct a prenatal smoking indicator (ever smoked during pregnancy). We note here that we consider all treated women, including unsuccessfully treated women, to account for possible prenatal smoking influences on the success of treatment (having a livebirth).[31] For the IVF children, the IVF register contains information on their birth weight, gestation length, and APGAR scores.[32] We construct three standard indicators of adverse child health at birth: low birth weight (less than 2500 grams), pre-term birth (less than 37 completed weeks of gestation), and low APGAR scores (less than 10). We then regress these prenatal outcomes on the years of education of both partners (in the treated couple), with varying sets of control variables. For the prenatal smoking of IVF-treated women (including unsuccessfully treated women), we control for the immigrant status and birth year of treated women and their partners, and the date of first treatment (in calendar months) and a full set of clinic indicators. For the health outcomes of IVF children, we additionally control for the child gender,

[31]Smoking during pregnancy is widely considered harmful to the developing fetus, with adverse effects on both early- and later-life outcomes, ranging from infant health and cognition, adolescent test scores and schooling, to adult earnings (Bharadwaj et al. 2013, Currie and Hyson 1999, Härkonen et al. 2012, Simon 2016).

[32]APGAR scores represent a composite index of a child's health at birth based on muscle tone, heart rate, reflex irritability, skin coloration, and breathing rate and effort. APGAR scores run from 0 to 10. The conventional indicator for poor prenatal health uses less than 7 as the cutoff. With too few children scoring less than 7, however, this does not work in our sample of IVF children. Instead, we consider less than perfect prenatal health and use less than 10 as the cutoff (as in Razaz et al. 2019).

Table 8
Regressions of prenatal behavior and outcomes on education of IVF-treated women

	non-donor treatment	**sperm-donor treatment**	**egg-donor treatment**
	(1)	(2)	(3)
A. Prenatal smoking (0/1)			
years of education treated woman	–0.020	–0.012	–0.014
	*0.002******	*0.006**	*0.007**
R-squared	0.039	0.024	0.031
number of treated women	49,856	1,121	589
mean prenatal smoking	0.110	0.097	0.093
B. Low birth weight (0/1)			
years of education mother	–0.007	–0.009	0.011
	*0.002******	*0.007*	*0.018*
R-squared	0.225	0.247	0.448
number of mothers	20,714	776	188
mean low birth weight	0.204	0.204	0.277
C. Premature birth (0/1)			
years of education mother	–0.006	–0.020	–0.003
	*0.002******	*0.010****	*0.022*
R-squared	0.214	0.229	0.325
number of treated women	20,909	785	191
mean pre-term birth	0.225	0.217	0.309
D. Low APGAR scores (0/1)			
years of education mother	–0.003	–0.016	0.006
	*0.001****	*0.006******	*0.011*
R-squared	0.015	0.054	0.143
number of mothers	20,597	770	186
mean low APGAR scores	0.114	0.126	0.156

Note–The dependent variables are prenatal smoking, low birth weight (less than 2,500 grams), pre-term birth (less than 37 completed gestation weeks), and low APGAR scores (less than 10). The main independent variable is the (prospective) mother's educational attainment measured in the nominal years spent in school. In the prenatal smoking sample, the regression specification further controls for the partner's educational attainment, immigrant status and birth year of both partners, first IVF treatment date controls (measured in calender months) and a full set of clinic indicators. In the infant health sample, the regression specification additionally controls for infant birth year, gender, and multiple-birth status. Standard errors are clustered at the clinic level shown in italics; * indicates significance at 10 percent level, ** indicates significance at 5 percent level, and *** at 1 percent level.

multiple-birth status, and birth year. If more educated women indeed offer a better in utero environment for their children, we should find that more educated women smoke less during pregnancy, and have children with healthier birth outcomes.

Table 8 reports the least-squares regression results, together with sample means for maternal smoking, low birth weight, prematurity, and low APGAR scores.[33] For maternal smoking during pregnancy (in panel A), we find that the coefficients on the treated woman's years of education are all negative, significant, and comparable across the three samples of IVF-treated women. These estimates are large in magnitude, indicating that one more year of education is associated with a 12-18 percent reduction in prenatal smoking. For infant health (in panels B, C and D), we also find that the coefficients on the mother's years of schooling are negative and (mostly) statistically significant, but only in the larger samples of IVF mothers with non-donor and sperm-donor children. Compared to the prenatal smoking estimates, these infant health estimates are notably smaller in magnitude, though, indicating that one more year of maternal education reduces the risk of having low weight infants by 4 percent, premature births by 3-9 percent, and low APGAR scores by 3-13 percent. In the sample of IVF mothers with egg-donor children, none of the infant health coefficients are significant, which is not so surprising given the difficulty to uncover smaller maternal influences in smaller samples.

As another test for the prenatal condition story, we have included birth outcomes as (bad) controls in our intergenerational regressions. If in utero conditions are important for child test scores, we should see that the nurture effect for mothers gets substantially smaller when we take account of these birth outcomes. This is not what we observe. When we restrict the egg-donor sample to non-missing birth indicators for low birth weight, pre-term birth and low APGAR scores (we are left with 144 observations), the mother's nurture effect (with standard error in parentheses) with and without birth indicator controls are 0.064 (0.006) and 0.071 (0.006), respectively. Of importance here is that the associations between the mother's education and child test scores do not change either when we control for these birth indicators in the larger samples where the associations between the mother's education and child birth outcomes were statistically significant. We find that the intergenerational associations between the mother's education and child test scores are 0.076 and 0.076 (with and without birth indicator controls) in the non-donor sample, and 0.067 and 0.066 (with and without birth indicator controls) in the sperm-donor sample.

[33]The prematurity, low birth weight, and low APGAR score rates (reported at the bottom of panels B, C, and D) are quite high because of the high twin rates that are common in IVF treatments. This is accounted for by controlling for multiple-birth status in the infant health regressions.

On the whole, our empirical evidence for the prenatal story is mixed. When we consider maternal smoking as a proxy for a poor and harmful prenatal environment, we find large effects of schooling among donor mothers. When we consider the women who carry and give birth to genetically unrelated babies, we do not find that more educated mothers have healthier babies. With only limited variation in the three infant health measures (and no infant cognition measures) in small donor samples, however, it is statistically challenging to uncover strong prenatal influences.

Interpreting the Results

When we interpret the estimates for donor children as nurture effects, which we can credibly do, our findings suggest that the rearing effects are much stronger for mothers than for fathers. With mothers as primary caretakers of young children, this is not surprising at all. When we compare the estimates for non-donor and donor children and interpret the difference as nature effects, which we can only do under strict assumptions about the representativeness of donor families and independence of nurture and nature influences, our findings suggest that the genetic effects are stronger for fathers than for mothers. With what we know from biology, this is highly surprising and simply impossible. In an attempt to reconcile our findings with the biological fact that genes are passed on to children from fathers and mothers alike, we probe three possible explanations, two behavioral (that relax the strict assumptions) and one statistical.

One possible explanation is that nature-nurture interactions matter for child test scores, and that the different nurture effects reflect, among other things, innate differences among the different donor children. In this case, we should see that sperm-donor children are children with genetically disadvantaged backgrounds (who benefit less from an improved environment), and reversely, egg-donor children are children with genetically advantaged backgrounds (who benefit more from an improved environment). We find no evidence for this: the raw test score differences we observe are inconsistent with the innate skill differences we hypothesize for donor children (see Table 1); in addition, we find small and insignificant interaction estimates in intergenerational mobility regression models that allow for nature-nurture interactions (see online Appendix Table D.1, panels C and D).

Another possible explanation is that donor parents are just different in how they raise their children, and that the different nurture effects reflect, among other things, different levels of parental involvement and effort into child rearing. In this case, we should see that sperm-donor children have less involved fathers (who exert less effort in raising their children), and reversely, egg-donor children have more involved mothers (who have better child rearing skills, exert more effort in providing a more favorable environment for their children to do well, or both). We find little evidence that unrelated parents are differently involved with their donor children (see Table 6), but we readily admit that parental leave, labor supply, and divorce may not be the best measures of parental involvement into child rearing. Exploring these parenting differences further is an important task, but lies beyond the scope of this study (without more appropriate measures of parental involvement and investments).

The third and last explanation is that our donor samples are just too small and the nurture effect estimates too imprecise to statistically uncover that fathers impact child test scores through environment and mothers through genes. A closer look at the nurture effect estimates (and their confidence intervals) reveal indeed that the nurture effects for fathers and mothers may not be so different from each other. The near zero nurture effect we estimate for fathers in our sample of 820 sperm-donor children, for example, is precise enough to statistically rule out nurture effect sizes larger than 0.023 (which is about one third of the overall intergenerational association we estimate for other fathers). The positive nurture effect for mothers we estimate in our sample of 157 egg-donor children, on the other hand, is large enough to be statistically significantly different from zero but not precise enough to rule out nurture effects sizes as small as 0.023 (which is again about one third of the overall intergenerational association we estimate for other mothers).

8 Discussion and Conclusion

In this paper we investigate the intergenerational persistence in human capital skills, net of genetic skill transfers, using a novel strategy based on Danish children conceived through sperm and egg donations in IVF treatments. We measure intergenerational persistence in human capital skills by estimating the relationship between parental education and child test performance on nationwide tests taken in primary and lower secondary education using samples of donor children. By looking at donor children, we eliminate the genetic connection between children and one of their parents; that is, children from sperm donors are genetically related to their mother but not their father, and children from egg donors are genetically related to their father but not their mother. Moreover, by following donor children from embryo implants onwards, we capture the parents' education-related prenatal and early childhood influences on child test scores. We find a strong and significant relationship between the mothers' education and egg-donor child test scores, but no relationship between the fathers' education and sperm-donor child test scores.

We realize that our empirical strategy only works if donor assignment is either random or related to variables that we observe and control for. In the Danish context, we know exactly how donors are assigned to prospective parents. For rearing mothers of egg-donor children, the assignment of eggs is based on the position mothers take on the clinic's waiting list. With regression models that account for the clinic queue order, we show that donor assignment is as good as random. For rearing fathers of sperm-donor children, the assignment of donor sperm is based on unobserved preferences for a limited number of donor characteristics. If the assignment of sperm donors is selective, our primary concern would be that part of what we interpret as the nurture effect may in fact be genetic. With nurture estimates for fathers close to zero, however, we consider such a selection bias unlikely. Because the assignment of donors to parents is not selective, we can interpret the nurture effect estimates (as presented for mothers and fathers) in a causal way and conclude that it is the nurturing impact of the education of mothers, and not fathers, that matters most for the test scores of children.

Why are the nurture effects so much stronger for mothers than for fathers? There are several enforcing mechanisms. In terms of prenatal influences, mothers are (unlike fathers) pregnant, carry children, and give birth. We find mixed evidence that more educated mothers improve prenatal conditions (measured by reduced smoking in pregnancy). In terms of postnatal influences, we find clear evidence that mothers spend (unlike fathers) much more time with their children in early childhood (measured by parental leave take up). If more educated mothers, as primary caretakers, are also more involved and effective in child rearing, it is quite plausible that mothers matter most for their children's test scores.

We also realize, of course, that the nurture effects we estimate are population, outcome, and context specific. It may therefore be worthwhile to speculate how our nurture effect estimates for fathers and mothers might change when we move beyond the population of donor families, look at other outcomes than test scores in primary and lower secondary education, and consider countries outside Denmark. We first study donor families, which are clearly not representative; donor parents are, on average, richer, more educated, older, and possibly raise their donor children differently from non-donor children. With more representative fathers, we may find stronger nurture effects if donor fathers are less involved in child rearing (possibly because they bond less well with unrelated children and as a result exert less effort in raising them). With more representative mothers, on the other hand, we may find weaker nurture effects if donor mother are more involved in child rearing (possibly because they had to put in much more effort than most other mothers to get pregnant and have children). Second, we study test score outcomes in primary and lower secondary education when children are young and time intensive to raise. With different outcomes, we may find stronger nurture effects for fathers if more educated fathers matter more for later-life outcomes that are more financially intensive and less time intensive, such as attending university, labor market outcomes, or portfolio holdings. And third, we study intergenerational persistency in human capital skills in Denmark where education is heavily subsidized, skill returns are low, and parental leave arrangements are generous. With children growing up in countries with more costly education and higher skill returns, we may find larger nurture effects for fathers if more educated fathers (who would otherwise see little return to their child investments) devote more time and money on children. With children growing up in countries with less generous parental leave arrangements, we may get weaker nurture effects if more educated mothers (who would otherwise spend more time with their children) take up less maternity leave. Clearly, much more work needs to be done to explore these possibilities for fathers and mothers; this is a priority for our future research.

Data Availability Statement

The data and code underlying this research are available in the Zenodo data repository, at https://doi.org/10.5281/zenodo.10608471.

References

Adermon Adrian, Mikael Lindahl, and Mårten Palme. 2021. "Dynastic Human Capital, Inequality, and Intergenerational Mobility," *American Economic Review,* 111:5 pp. 1523-1548.

Almond, Douglas, and Janet Currie. 2011. "Human Capital Development Before Age Five," in *Handbook of Labor Economics,* Orley Ashenfelter and David Card, eds. Vol. 4, pp.1315-1486, Elsevier.

Amin, Vikesh, Petter Lundborg, and Dan-Olof Rooth. 2015. "The Intergenerational Transmission of Schooling: Are Mothers Really Less Important Than Fathers?" *Economics of Education Review,* 47, pp. 100-117.

Ashenfelter, Orley, and Alan Krueger. 1994. "Estimates of the Returns to Schooling from a New Sample of Twins," *American Economic Review,* 84:5, pp. 1157-1173.

Becker, Gary, and Nigel Tomes. 1986. "Human Capital and the Rise and Fall of Families," *Journal of Labor Economics,* 4:3, Part 2, S1-S39.

Behrman, Jere, and Mark Rosenzweig. 2002. “Does Increasing Women’s Schooling Raise the Schooling of the Next Generation?" *American Economic Review,* 92:1, pp. 323-334.

Beuchert, Louise, and Anne Nandrup. 2018. “The Danish National Tests at a Glance," *Nationaloekonomisk Tidsskrift,* 1, pp. 1-37.

Bharadwaj, Prashant, Julian Johnsen, and Katrine Løken. 2014. “Smoking Bans, Maternal Smoking and Birth Outcomes," *Journal of Public Economics,* 115, pp. 72-93.

Bingley, Paul, Kaare Christensen, and Vibeke Myrup Jensen. 2009. “Parental Schooling and Child Development: Learning from Twin Parents," Danish National Centre for Social Research Social Policy and Welfare Services Working Paper 07:2009, Copenhagen.

Björklund, Anders, Markus Jäntti and Gary Solon. 2005. "Influences of Nature and Nurture on Earnings Variation: A Report on a Study of Various Sibling Types in Sweden," in *Unequal Chances: Family Background and Economic Success*, Samuel Bowles, Herbert Gintis and Melissa Osborne, eds. Princeton University Press.

Björklund, Anders, Mikael Lindahl, and Erik Plug. 2006. “The Origins of Intergenerational Associations: Lessons from Swedish Adoption Data," *Quarterly Journal of Economics*, 121:3, pp. 999-1028.

Björklund, Anders and Kjell Salvanes. 2011. “Education and Family Background: Mechanisms and Policies," in *Handbook of the Economics of Education.* Eric Hanushek, Stephen Machin, and Ludger Woessmann, eds. Vol. 3, pp. 201-247. Elsevier.

Black, Sandra and Paul Devereux. 2011. “Recent Developments in Intergenerational Mobility," in *Handbook of Labor Economics.* Orley Ashenfelter and David Card, eds. Vol. 4, pp. 1487-1541. Elsevier.

Black, Sandra, Paul Devereux, Petter Lundborg, and Kaveh Majlesi. 2022. “Poor Little Rich Kids? The Role of Nature versus Nurture in Wealth and Other Economic Outcomes and Behaviors," *Review of Economic Studies*, 89:2, pp. 876-908.

Brodzinsky, David. 1987. “Adjustment to Adoption: A Psychosocial Perspective." *Clinical Psychology Review,* 7:1, pp. 25-47.

Cesarini, David. 2010. *Essays on Genetic Variation and Economic Behavior*. PhD Thesis. Massachusetts Institute of Technology, Boston, US.

Cesarini, David and Peter Visscher. 2017. “Genetics and Educational Attainment," *NPJ Science of Learning,* 2.1: 4.

Ciconia (2015). *Patient information (Egg donation).* Aarhus Fertilitetsklinik.

Collado, Dolores, Ignatio Ortuño-Ortìn, and Jan Stuhler. 2023. “Estimating Intergenerational and Assortative Processes in Extended Family Data," *Review of Economic Studies,* 90:3, pp. 1195-1227.

Cunha, Flavio, and James Heckman. 2009. “The Economics and Psychology of Inequality and Human Development," *Journal of the European Economic Association,* 7:2, pp. 320-364.

Currie, Janet, and Rosemary Hyson. 1999. “Is the Impact of Health Shocks Cushioned by Socioeconomic Status? The Case of Low Birthweight.” American Economic Review 89 (2): 245–250

Cutler, Davis, and Adriana Lleras-Muney. 2006. Education and Health: Evaluating Theories and Evidence. NBER Working Paper No. 12352, Cambridge Massachusetts, US.

Dawkins, Richard. 2006. *The Selfish Gene: With a New Introduction by the Author.* UK: Oxford University Press.(Originally published in 1976).

Danish Adoption Law, BEK nr. 283 af 20/03/2019. https://www.retsinformation.dk/eli/lta/2019/283.

Dearden, Lorraine, Stephen Machin, and Howard Reed. 1997. “Intergenerational Mobility in Britain," *Economic Journal,* 110:440, pp. 47-64.

Fagereng, Andreas, Magne Mogstad, and Marte Rønning. 2021. “Why Do Wealthy Parents Have Wealthy Children?" *Journal of Political Economy,* 129:3, pp.703-756.

Golombok, Susan. 2021. “Love and Truth: What Really Matters for Children Born Through Third-Party Assisted Reproduction." *Child Development Perspectives,* 15:2, pp 103-109.

Golombok, Susan, Jennifer Readings, Lucy Blake, Polly Casey, Laura Mellish, Alex Marks, and Vasanti Jadva. 2011. “Children Conceived by Gamete Donation: Psychological Adjustment and Mother-Child Relationships at Age 7." *Journal of Family Psychology,* 25:2 230-239.

Griliches, Zvi. 1979. “Sibling Models and Data in Economics: Beginnings of a Survey," *Journal of Political Economy,* 87:5, pp. 37-64.

Härkönen, Juho, Hande Kaymakçalan, Pirjo Mäki, and Anja Taanila. 2012. “Prenatal Health, Educational Attainment, and Intergenerational Inequality: The Northern Finland Birth Cohort 1996 Study," *Demography,* 49:2, pp. 525–52

Heckman, James. 2007. "The Economics, Technology, and Neuroscience of Human capability Formation," *Proceedings of the National Academy of Sciences,* 104:33, pp. 13250-13255.

Holmlund, Helena, Mikael Lindahl and Erik Plug. 2011. “The Causal Effect of Parents’ Schooling on Children’s Schooling: A Comparison of Estimation Methods," *Journal of Economic Literature,* 49:3, pp. 615-651.

Hægeland, Torbjørn, Lars Kirkeboen, Oddbjørn Raaum and Kjell Salvanes. 2010. Why Children of College Graduates Outperform their Schoolmates: A Study of Cousins and Adoptees. IZA Discussion Paper No. 5369, Bonn, Germany.

Jencks, Cristopher, and Marsha Brown. 1977. “Genes and Social Stratification," in *Kinometrics: The Determinants of Economic Success within and between Families.* Paul Taubman, eds. Amsterdam: North-Holland Publishing Company, pp. 169-231.

Isacsson, Gunnar. 1999. “Estimates of the Return to Schooling in Sweden from a Large Sample of Twins," *Labour Economics,* 6:4, pp. 471-489

Okbay, Aysu, et al. 2022. “Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals," *Nature Genetics*, 54:4, pp. 437-449.

Lee, James and et al. 2018. “Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals," *Nature Genetics,* 50:8, pp. 1112-1121.

Lichtenstein, Paul, Nancy Pedersen, and Gerald McClearn. 1992. “The Origins of Individual Differences in Occupational Status and Educational Level: A Study of Twins Reared Apart and Together," *Acta Sociologica,* 35:1, pp. 13-31.

Lundborg, Petter, Erik Plug and Astrid Würtz, Rasmussen. 2017. “Can Women have Children and a Career? IV Evidence from IVF Treatments," *American Economic Review,* 107:6, pp. 1611-1637.

Lykken, David, Thomas Bouchard, Matt McGue, and Auke Tellegen. 1990. “The Minnesota Twin Family Registry: Some Initial Findings," *Acta Geneticae Medicae et Gemellologiae: Twin Research,* 39:1, pp. 35-70.

Miller, Paul, Charles Mulvey, and Nick Martin. 1995. “What Do Twins Studies Reveal about the Economic Returns to Education? A Comparison of Australian and US Findings," *American Economic Review,* 85:3, pp. 586-599.

Nielsen, Anders Faurskov, Bjørn Pedersen, and Jørgen Glenn Lauritsen. 1995. “Psychosocial Aspects of Donor Insemination: Attitudes and Opinions of Danish and Swedish Donor Insemination Patients to Psychosocial Information Being Supplied to Offspring and Relatives," *Acta Obstetricia et Gynecologica Scandinavica,* 74:1, pp. 45-50.

Plug, Erik, and Wim Vijverberg. 2003. “Schooling, Family Background, and Adoption: Is it Nature or is it Nurture?" *Journal of Political Economy,* 111:3, pp. 611-641.

Plug, Erik. 2004. “Estimating the Effect of Mother’s Schooling on Children’s Schooling Using a Sample of Adoptees," *American Economic Review,* 94:1, pp. 358-368.

Pronzato, Chiara. 2012. “An Examination of Paternal and Maternal Intergenerational Transmission of Schooling," *Journal of Population Economics,* 25:2, pp. 591-608.

Razaz, Neda, Sven Cnattingius, and KS Joseph. 2019. Association between APGAR Scores of 7 to 9 and Neonatal Mortality and Morbidity: Population Based Cohort Study of Term Infants in Sweden," *British Medical Journal,* 365: l1656.

Rice, Frances, Gordon Harold, Jacky Boivin, Dale Hay, Marianne van den Bree, Anita Thapar, and Gene Robinson. 2009. “Disentangling Prenatal and Inherited Influences in Humans with an Experimental Design." *Proceedings of the National Academy of Sciences*, 106:7, pp. 2464-2467.

Sacerdote, Bruce. 2000. The Nature and Nurture of Economic Outcomes. NBER Working Paper No. 7949, Cambridge Massachusetts, US.

Sacerdote, Bruce. 2002. “The Nature and Nurture of Economic Outcomes," *American Economic Review (Papers and Proceedings),* 92:2, pp. 344-348.

Sacerdote, Bruce. 2007. “How Large Are the Effects from Changes in Family Environment? A Study of Korean American Adoptees," *Quarterly Journal of Economics,* 122:1, pp. 119-157.

Sacerdote, Bruce. 2011. “Nature and Nurture Effects on Children’s Outcomes: What Have We Learned from Studies of Twins and Adoptees?" in *Handbook of Social*

Economics. Jess Benhabib, Alberto Bisin, and Matthew Jackson, eds. Vol. 1, pp. 1-30.

Scarr, Sandra and Richard Weinberg. 1994. "Educational and Occupational Achievements of Brothers and Sisters in Adoptive and Biologically Related Families," *Behavior Genetics,* 24:4, pp. 301-325.

Simon, David. 2016. "Does Early Life Exposure to Cigarette Smoke Permanently Harm Childhood Welfare? Evidence from Cigarette Tax Hikes," *American Economic Journal: Applied Economics,* 8:4, pp. 128-159.

Taubman, Paul. 1976. "The Determinants of Earnings: Genetics, Family, and Other Environments: A Study of White Male Twins," *American Economic Review,* 66:5, pp. 858-870.

Teasdale, TW, and David Owen. 1984. "Heredity and Familial Environment in Intelligence and Educational Level — A Sibling Study," *Nature,* 309:5969, pp. 620.

Westergaard, Hanne Brix, Anne Mette Tranberg, Karin Erb, and Anders Nyboe Andersen. 2000. "Danish National IVF Registry 1994 and 1995. Treatment, Pregnancy Outcome and Complications During Pregnancy," *Acta Obstetricia et Gynecologica Scandinavica,* 79:5, pp. 384-389.

Woessmann, Ludger. 2018. "Central Exit Exams Improve Student Outcomes," *IZA World of Labor,* 419.

Young, Alexander. 2019. "Solving the Missing Heritability Problem," *PLoS Genetics*, 15:6, pmid: e1008222.

Young, Alexander, and et al. 2022. "Mendelian imputation of parental genotypes improves estimates of direct genetic effects," *Nature Genetics,* 54:6, pp. 897-905.