HELSINKI GSE DISCUSSION PAPERS 3 · 2023

Parental Education and Invention: The Finnish Enigma

Philippe Aghion
Ufuk Akcigit
Ari Hyytinen
Otto Toivanen

Helsinki GSE Discussion Papers

Helsinki GSE Discussion Papers 3 · 2023

Philippe Aghion, Ufuk Akcigit, Ari Hyytinen, and Otto Toivanen:
Parental Education and Invention: The Finnish Enigma

ISBN 978-952-7543-02-3 (PDF)
ISSN 2954-1492

Helsinki GSE Discussion Papers:
https://www.helsinkigse.fi/discussion-papers

Helsinki Graduate School of Economics
PO BOX 21210
FI-00076 AALTO
FINLAND

Helsinki, February 2023

Parental Education and Invention: The Finnish Enigma*

Philippe Aghion Ufuk Akcigit Ari Hyytinen Otto Toivanen

Abstract

Why is invention strongly positively correlated with parental income not only in the US but also in Finland which displays low income inequality and high social mobility? Using data on 1.45M Finnish individuals and their parents, we find that: (i) the positive association between parental income and off-spring probability of inventing is greatly reduced when controlling for parental education; (ii) instrumenting for the parents having a MSc-degree using distance to nearest university reveals a large causal effect of parental education on offspring probability of inventing; and (iii) the causal effect of parental education has been markedly weakened by the introduction in the early 1970s of a comprehensive schooling reform.

JEL codes: O31, O35, I25, I26

Keywords: Education, Inventors, Innovation, Patents, Parents' education

*Addresses: Aghion: College de France, London School of Economics and Insead (P.Aghion@lse.ac.uk). Akcigit: University of Chicago (uakcigit@uchicago.edu). Hyytinen: Hanken School of Economics and Helsinki Graduate School of Economics (ari.hyytinen@hanken.fi). Toivanen: Aalto University School of Business, Helsinki Graduate School of Economics and KU Leuven (otto.toivanen@aalto.fi). This paper was prepared for the 2022 Klein Lecture, University of Pennsylvania. We are most grateful to the IER editor, Dirk Krueger, who encouraged us to prepare this draft for the Review. We would like to thank Atte Pudas, Sohvi Kupila and Sonja Marttinen for excellent research assistance, the ERC (grant agreement No. 786587), the Yrjö Jahnsson foundation and FORTE for financial support, Jan Jääskeläinen for help with the distance data, and Heidi Williams for very helpful discussion on an earlier version of our paper. All errors are ours.

1 Introduction

Invention is a human activity and a major source of economic growth, but everybody cannot become an inventor.[1] Whether one becomes an inventor likely depends on innate ability and the social environment, including family resources and parental education. Figure 1 depicts the relationship between the probability of off-spring becoming an inventor and parental income, using recent and historical US as well as Finnish data. In both countries, the probability increases with parental income and the increase is particularly steep at the highest levels of parental income. The probability of inventing for the off-spring of high income parents is about ten times larger than the corresponding probability at the bottom of the income distribution. We view the similarity of Finland to the US as an enigma because, unlike the US, Finland displays low income inequality and high social mobility (see, e.g., Black and Devereux 2011, Jäntti and Jenkins 2015) and has offered, for cohorts born since the mid-1960s, free education up to and including university.[2,3] The enigma is, why the relation between the parental income and the probability of offspring inventing is not less pronounced in Finland despite its more egalitarian society and equitable educational system. To resolve the Finnish enigma, we ask whether parental education explains it, and quantify the causal effect of parents' education on the probability of their offspring becoming inventors on which there is very little evidence.

Figure 1 here

While the socio-economic background of inventors is of interest in itself (e.g., Bell et al. 2019), parental education arguably stands out in its policy relevance: The educational system of a country is malleable and a public policy instrument. It is therefore notable that Figure 1 hides another striking relation: Figure 2 displays the probability that a parent has an MSc-degree, conditional on her income percentile. The parental income percentile and the probability that the parent has an MSc have an equally stark convex relationship as that in Figure 1. While inter-generational transmission of endowments is often considered universal, combining Figures 1 and 2 suggests that the reason why parents' income predicts their offspring inventing is that it mirrors how parental education and offspring's inventions are related.[4] How strongly parental education affects offspring invention is important both because the policy relevance of any particular inter-generational transmission depends on its magnitude and also because of its ramifications for inter-generational mobility.

Figure 2 here

To study the relation between parental education and invention, we merge four data sets: (i) individual data from Statistics Finland for individuals born between 1953 and 1981 and their parents; (ii) the parents' distances to the nearest university at age 19 with help of data from the National Land Survey of Finland; (iii) individual-level patenting data from the European Patent Office; and (iv) for a sub-sample, IQ data from the Finnish Defense Force. Our main explanatory variable is a dummy for at least one parent having an MSc. We follow a literature starting with Card (1995) and use as the

[1] There has been a rising interest over the past years in the process whereby new ideas come about and translate into new patents; see e.g. Bloom et al. (2017) and Gordon (2017). At population level, patenting by an individual is a rare but important event, not least because it is a concrete manifestation of the person's inventiveness.

[2] University tuition has always been low in Finland. However, essentially all parents in our data attended the old school system based on tracking where some schools had fees. We return to this later.

[3] In our data 54% of the individuals - not parents - belong to the cohorts where everybody attended comprehensive school.

[4] We show the relation between parental education and the probability to invent in Figure 3.

instrument parents' distance to the nearest university at the age of 19. We discuss in Section 4.1 how we alleviate omitted variable problems linked to this instrument (e.g. Carneiro and Heckman 2002); suffice it to say here that we have a rich vector of new municipality-cohort-specific observables at our disposal, and make use of new universities being established.[5] To alleviate endogeneity concerns further, we also study the individuals for whom we have IQ data.

We find that parental income is positively associated with the probability of becoming an inventor and that the effect is greatly diminished once parental education is controlled for. Coupled with the fact that parental education is unevenly distributed, this finding sheds light on the Finnish enigma. Moreover, as shown in Figure 2, higher parental income is positively correlated with parental education. Similarly, a white-collar parent earns more than a blue-collar parent (see Aghion et al. 2017). These positive correlations between several socio-economic characteristics of parents and the probability of offspring invention likely mask several causal mechanisms, calling for an investigation of how an intergenerational supply side measure, parental education, affects invention.[6]

Our second finding comes from instrumental variables (IV) estimations, showing that parental university education has a large, positive local average treatment effect (LATE) on the probability of a child becoming an inventor.[7] The estimated magnitude of LATE depends on whether the maternal distance, paternal distance or both are used as the instrument. Third, while the causal impact of parental education on sons is higher than that on daughters, the impact *relative to* the baseline is larger for daughters. When we use the sub-sample of males for which IQ data are available, we obtain qualitatively similar estimates. Our third finding is that the average treatment effects on the treated (ATTs) are similar to LATEs, but those on the untreated are roughly one third lower.

We also dig deeper into the Finnish enigma. Our evidence suggests that the education reform implemented in the late 1960s has reduced the causal impact of parental education and income on the probability of inventing: in other words, this reform has reduced the number of "lost Einsteins and Marie Curies" in Finland.

The remainder of the paper is organized as follows. We provide a short review of the relevant literature in Section 2. We explain the data sources and present descriptive statistics in Section 3. In this section we also present OLS regressions to explore the association between parental income and the probability of becoming an inventor with and without controlling for parental education. In Section 4 we describe our IV approach and our instrument(s). We analyze the causal effect of parental education on the probability of off-spring inventing in Section 5. There we also report results from a Roy model. In Section 6 we report a series of robustness tests. Section 7 includes a discussion of potential mechanisms at work and an initial analysis of how the causal effect of parental education may be intermediated by the schooling system the individual was exposed to. Section 8 concludes.

[5] In the interest of brevity, we refer the reader to Toivanen and Väänänen (2016) and Suhonen and Karhunen (2019) who discuss how the locations of these universities were decided upon. Briefly, there was a lot of documented randomness in the decision process.

[6] See, e.g., Takalo and Toivanen (2015); Bloom et al. (2019) for a review of demand and supply-side innovation policies.

[7] In the Appendix we report results using other dependent variables and other ways of measuring parental education. These results are in line with our main results.

2 Literature

This paper relates first to the literature on innovation-based growth, in particular to Schumpeterian growth models (e.g see Aghion and Howitt 1992; Aghion et al. 2014; and Akcigit and Nicholas 2019). In the spirit of Nelson and Phelps (1966), education is in these models not so much a production input as an investment that increases individuals' ability to both catch up with the technological frontier and to innovate at the frontier.

Our analysis relates specifically to the recent work that has merged individual data with patenting data and that has provided empirical evidence on the characteristics of inventors. Early studies in this line of research include Toivanen and Väänänen (2012), studying the returns to Finnish inventors, and Toivanen and Väänänen (2016), looking at the causal effect of own education on the probability of becoming an inventor. Our findings complement Toivanen and Väänänen (2016) who found that better own education increases the supply of inventors.[8] Bell et al. (2019) merge US individual fiscal and test-score data with USPTO patent data to look at the life-cycle of inventors. They find that parental income, occupation and sector of activity, race, gender, and childhood neighborhood are important determinants of the probability of becoming an inventor. Celik (2015) matches inventors' surnames with socioeconomic background information using US census data in 1930. He finds that individuals from richer backgrounds are more likely to become inventors. Akcigit et al. (2017) merge historical patent and individual census records from the US and look at how demography, geography, social origins, and the cultural environment, affect innovation. They show that the probability of becoming an inventor around 1940s was highly correlated with father's income but this strong relationship disappears once child's education is controlled for. Jaravel et al. (2018) merge US individual tax and patenting data to quantify the impact of coauthors on inventors' careers, finding large spillover effects.[9]

In a related strand of the literature, the interest has been in how innovation rents are shared and how innovation affects innovators' coworkers; see Van Reenen (1996), Aghion et al. (2018), Kline et al. (2019), and Aghion et al. (2022). For example, Aghion et al. (2022) explore how the arrival of an invention affects the outcomes of an inventor's coworkers, such as their wage returns and their probability of moving out of employment, and how these effects vary with the coworkers' distance to the human capital frontier.

We also build on the literature on the effects of parental education, surveyed in Holmlund et al. (2011). Lundborg et al. (2014) report for Sweden that maternal education has a positive causal impact on sons' cognitive ability and health, but no evidence of paternal education effects.[10] Hoisl et al. (2022) utilize offspring gender composition to study the parental impact on offspring invention and education, stressing the influence of parents being inventors and how this varies by gender of both the parent and the offspring. We contribute by uncovering the causal effect of parental education on the individual's probability of becoming an inventor, and by arguing that at least in equitable Nordic countries like

[8] We differ from Toivanen and Väänänen (2016) in several ways: They focus on individuals' own education while we focus on inter-generational aspects and parental education; they look at USPTO inventors whereas we use EPO data; they study a shorter time period than we and did not have access to data on municipality-cohort-level controls or IQ. Related to our work is also Suhonen and Karhunen (2019). They use largely the same educational data as we to study the causal impact of parental university education on off-spring education, finding a significant positive effect.

[9] Jaravel et al. (2018) build on seminal work by Azoulay et al. (2010) which examines the effect of the premature death of 112 scientists on their co-authors, providing the first convincing evidence of the effect of exposure to human capital on the production of new ideas.

[10] They use an overall measure of cognitive ability whereas we use the visuospatial IQ (which arguably is a less malleable sub-component of cognitive skills).

Finland, this effect may largely account for the observed relation between parental income and the probability of becoming an inventor.

3 Data and descriptive statistics

3.1 Data sources and sample construction

The data used in this paper come from the databases of Statistics Finland (SF), the National Land Survey of Finland (NLSF), the European Patent Office (EPO) and the Finnish Defense Force (FDF). SF is our source of individuals' and parents' characteristics, and information on the distances between municipalities and universities is from NLSF. EPO data allow us to identify Finnish inventors, while the source of the IQ data is FDF. Taking each of these in turn:

SF: We exploit SF's Finnish longitudinal employer-employee data (FLEED) for 1988-2012 and the 1970, 1975, 1980 and 1985 population censi. FLEED is a standard administrative register-based data, collected and maintained by SF. It covers the whole working age Finnish population. We utilize information on individual age, location, language and education. We use FLEED from 1988 (its first year) to 2013. Information on parent characteristics is drawn from the population census records and FLEED. We use the 1970, 1975, 1980 and 1985 censi for parental education, income and place of birth.

NLSF: We used ArcGis to calculate municipal distances, using a dataset depicting the municipal division in the whole of Finland. Our distance variable is the natural logarithm of the distance to the nearest university, measured from the birth-municipality of the parent in the year the parent turns 19. We set the distance to 1km for those born in a university town. There is cross-sectional and over-time variation; the former due to geography, the latter due to the opening of new universities, listed in Table A-4 in the Appendix. University openings were collected from Eskola (2002).

FDF: The FDF data contain the IQ information for conscripts from the 1982 conscript cohort onwards. Conscription in Finland applied and still applies to the whole male cohort; female military service was not yet possible during our observation period. Choosing unarmed service has been an option since 1930s, but doing so became easier in mid 1980s (i.e., for the birth cohorts from late 1960s onwards). However, only a relatively small fraction of each cohort chooses it. Moreover, a fraction of each cohort is dismissed for physical or psychological reasons either before or early on during the service. The FDF's IQ-test is similar to commonly used IQ tests, containing sub-components for analytic (numerical), verbal and visuospatial skills; those entering the service (a large majority - over 75% - of each cohort) take the test during the first weeks of service; most conscripts take their military service at the age of +/-20.[11] We use the visuospatial IQ score (IQ henceforth), as it is considered in the literature to be more strongly predetermined than the other two measures.[12] Following common practice, we normalize IQ to have mean 100 and standard deviation 15. We do this by the year of entering military

[11] Jokela et al. (2017) provide more detail on the test and the sub-components and presents evidence that selection is unlikely to bias the sample. Jokela et al. (2017) describe the visuospatial test as follows: "The visuospatial reasoning task is a set of matrices containing a pattern problem with one removed part, and the participant needs to decide which of the given alternative figures completes the matrix; it is similar to Raven's Progressive Matrices".

[12] Using similar IQ information from the Swedish Arm Forces, Dal Bó et al. (2017) argue that the IQ score is a good measure of general intelligence and cognitive ability. While the results of Lundborg et al. (2014) may suggest that IQ is malleable and thus potentially a bad control, the issue is not clear-cut as they use overall IQ. The results of Pekkarinen et al. (2009) suggest that the Finnish comprehensive school reform had no effect on visuospatial IQ, a marginally significant effect on analytic IQ, and a positive impact on verbal IQ.

service to account for the Flynn effect (i.e., for the gradual increase in IQ test scores over time). We take the percentiles of this IQ measure.

EPO: The EPO data provide information on inventor names and applicant names. Our data cover all EPO patents applications with at least one inventor with a Finnish address up to and including 2013. The data originates with PATSTAT, but SF has used the OECD REGPAT database built on PATSTAT. In the raw patent data, we have a total of 25,711 patents and 17,566 inventors. The mean and median number of inventors per patent is 2; the largest number of inventors per patent is 14. For each patent, we observe all the inventors, their names and address, the patentee and its address, and the number of citations in the first 5 years.

Data matching: SF's FLEED contains unique but anonymized individual identifiers, which are based on unique social security numbers that everybody in Finland has. These identifiers, together with SF's table linking parents to their children and municipal identifiers, allow us to link the different datasets from SF, NLSF and FDF and to create a merged data where we observe individual characteristics, the characteristics of the parents as well as where they lived.

The EPO data, in contrast, does not contain individual identifiers. Linking of patent data to individuals was done by a civil servant of SF, using the information on individual name (first and surname), employer name, individual address and/or employer's address (postcode, street name street number), and year of patent application. These were used in different combinations, also varying the year of the match to be before or after the year of application (e.g., matching a patent applied for in 1999 with the street address of the firm from the registry taken in 1998 or 2000). The match rate is 90% when calculated for the patents applied for in the years 1988-2013. The procedure follows that used in Aghion et al. (2018).

Sample: Our estimation sample contains all individuals born after 1953, whose parents were born 1901 or later, for whom we were able to match all the data sets. We exclude individuals born after 1981 as they are unlikely to have invented by 2013. The resulting cross-sectional sample contains around 1.45M individuals and 9 844 inventors.

3.2 Descriptive statistics on inventors and their parents

Our main outcome variable is a binary indicator identifying inventors, taking value one for the individuals obtaining at least one patent and taking value zero otherwise. We use as alternative outcome variables the number of patents obtained by the individual and the number of forward citations obtained by the inventor for all her patents. The former obviously measures the quantity aspect of invention, the latter in turn the quality dimension. In our sample, 0.6% of the individuals are inventors; the inventors' share among women is 0.15% and among men 1.2%. Inventors hold on average 3.2 patents (female mean 2.4; male mean 3.3) and their patents have obtained on average 4.3 citations within the first 5-year period (female mean 3.2; male mean 4.4).

Our main variable for parental education is a binary indicator taking value one if at least one parent has an MSc or higher degree (measured at age 35) in any field (STEM or non-STEM) and zero otherwise.[13] We focus on this educational level for two reasons. First, it allows us to shed light on how the university system contributes to the supply-side of innovations in the long-term. Second, obtaining a higher university degree was still relatively rare among the parental cohorts in our sample, making it a discriminating background factor among the offspring that we study and containing thus policy

[13]STEM fields are science, technology, engineering and mathematics.

potential.[14] As a robustness test we use the count of parents with an MSc, and alternatively, a binary indicator taking value one if at least one parent has an BSc or higher degree (measured at age 35) and zero otherwise.

Figure 3 allows taking a first look at the association between parental education level and field (STEM, non-STEM) and the probability of daughters and sons to become inventors. Inventors' parents are better educated: Fixing a parent's (mother, father) education field, the probability of the off-spring becoming inventors is increasing in the level of the parent's education.[15] The association is universally stronger for sons than daughters and usually stronger for STEM than non-STEM educated parents, though the latter pattern is not visible for the lower levels of parental education.

Figure 3 here

To shed further light on how parents' education and the probability of their offspring becoming inventors are related, we compare the education of inventors' parents with parental education in the general population. In our data, 6% of the individuals have at least one parent with an MSc. The corresponding number for female inventors is as high as 24%, whereas for male inventors, it is 19%. Zooming then at mothers' education, an individual's mother has only a base education (e.g., elementary school) for 59% of the individuals. The corresponding numbers for female and male inventors are quite a bit lower, 35% and 41%, respectively. Echoing this, the mother has an MSc for 2% of the individuals in the data, but among the female (male) inventors, the share is 11% (7%). The respective figures for fathers are (base education) 58% and (MSc) 5% (female inventors: 33% and 21%; male inventors: 37% and 17%).[16]

We study how the probabilities of an individual being an inventor, of an individual having at least one parent with an MSc, and of either parent obtaining an MSc are related to the parental distance to the nearest university in Appendix A-1.2. As expected, we find that the correlations between the probabilities and the distances are negative (though small in absolute value) and that e.g. the share of offspring with at least one parent with an MSc is decreasing with both parents' distance to university.

3.3 Descriptive regressions

Our descriptive analyses have revealed that parental income, parental education and the probability of offspring becoming inventors are clearly correlated with each other. Specifically, Figure 1 and 2 suggested a strong association both between the probability of becoming an inventor and parental income (see also Bell et al. 2019) and the probability of a parent having an MSc and parental income. Figure 3 showed, in turn, that inventors' parents are better educated. These pairwise correlations raise the question of whether the strong positive and convex relation between parental income and off-spring becoming an inventor remains once parental education is accounted for.

To explore this, we resort to descriptive OLS regressions. The regression equation for these estimations is:

$$y_i = \boldsymbol{X_i'}\boldsymbol{\beta} + f(income_{p,i}, \boldsymbol{\theta}) + g(Educ_{p,i}, \gamma) + \epsilon_i \quad (1)$$

[14] The median individual's mother was born in 1941 and father in 1939.

[15] The one clear violation of the monotonic increase in parental education is between mothers with an MSc and mothers with a PhD; keep in mind though that in our data, the fraction of individuals with a PhD-educated mother is only 0.07%.

[16] For detailed descriptive statistics, see Tables A-1, A-2 and A-3 in the Appendix.

where y_i is the binary indicator for individual i being an inventor; $\boldsymbol{X_i'\beta}$ are control variables and the associated parameter vector; $f(income_{p,i}, \boldsymbol{\theta})$ is a fifth order polynomial of income of the parent of type p ($p = mother,\ father$), with $\boldsymbol{\theta}$ being the associated parameter vector; function $g(\boldsymbol{Educ_{p,i}}, \gamma)$ includes a vector of field (STEM, non-STEM) and level (secondary, college, masters, PhD level, with base-level being omitted) of education indicators $Educ_{p,i}$ of parent of type p, with $\boldsymbol{\gamma}$ being the associated parameter vector; and ϵ_i is the error term. The vector $\boldsymbol{X}_i$ contains the full set of maternal and paternal year-of-birth indicators, and an indicator for the mother tongue of individual i not being Finnish.

We estimate equation (1) for daughters and sons separately using maternal and paternal income and with and without controlling for parental education. These regressions allow us to study how the probability of becoming an inventor is associated with parental income, and how this association changes when parental education is controlled for.

Figure 4 here

We report our key findings graphically in Figure 4. In the upper left panel of the figure we display the relation between the probability that a daughter becomes an inventor and maternal income (percentiles).[17] The blue curve displays the estimated $f(income_{p,i}, \hat{\boldsymbol{\theta}})$ function when not controlling for maternal education (replicating Figure 1), and the red curve the same function controlling for maternal education. Two pronounced changes are interesting: The strong convexity of the blue curve at high maternal income levels has become dramatically flatter, and the whole curve has clearly shifted down. The first change suggests that the strong convexity of the blue curve may be due to the strong association with maternal income and education displayed in Figure 2. The second change suggests that the association between maternal income and the probability of an individual becoming an inventor may be largely driven by maternal education which covaries strongly with maternal income. The upper right panel of Figure 4 displays a similar development regarding the association of a daughter's probability of becoming an inventor and paternal income, though the changes after controlling for paternal income are smaller.

The lower panel of Figure 4 displays the same associations for sons' probability of becoming inventors. The graphs display a remarkably similar picture as those for daughters, suggesting that the likely mechanisms at work are the same.

These descriptive regressions confirm the importance of parental education for the next generation's inventor outcomes in our data. While these regression results are not meant to be conclusive, they clearly suggest that parental education is a key force that drives the relation of parental income with offspring inventions: Parental education goes a long way in explaining "away" the strong relation between parental income and off-spring becoming an inventor.

4 Instrumental variables approach

We now turn to studying whether the relation between parental education and the individual's probability to invent is causal using an instrumental variable approach. Quantifying the long-term effects of parental education on the supply of inventors allows us to shed light on how higher education systems contribute to innovation policy (e.g., Takalo and Toivanen 2015; Bloom et al. 2019).

[17]We measure parental income percentiles using income at age 35. The percentiles are calculated over the whole cohort.

4.1 The instrumental variable

Our instruments are based on the (log of) distance to the nearest university from the birth municipality of the mother (father) in the year she (he) turned 19.[18] The use of distance to college as an instrument for education was introduced by Card (1995). It rests on the idea that this distance increases the costs of obtaining university education and thereby decreases the probability of attending college without affecting the outcome of interest directly. The instrument has been used successfully by e.g. Currie and Moretti (2003), Cameron and Taber (2004), Carneiro et al. (2011), Eisenhauer et al. (2015), Toivanen and Väänänen (2016), Heckman et al. (2018) and Suhonen and Karhunen (2019).

As has been noted in many earlier studies, a potential problem with the distance instrument is that parents choosing to have children close to a university are different from parents choosing to have children further away.[19] Such selection could induce a correlation between the instrument and unobservables that affect the outcome variable, invalidating the exclusion restriction. This is not just a theoretical concern: Carneiro and Heckman (2002) (see their Table 1) demonstrate that distance to college measured in the NLSY79 for white males is negatively correlated with the AFQT ability test. The literature has offered several ways to inspect the severity of the problem and to ameliorate it. We implement some of these inspections and take steps to address this potential problem.

First, as in e.g. Currie and Moretti (2003) and Toivanen and Väänänen (2016), the main source of exogenous variation in our data is the establishment of new universities.[20] Because of the randomness in the political decision process regarding the location of the new universities as well as the timing of their eventual opening (see Toivanen and Väänänen 2016 and Suhonen and Karhunen 2019), this variation is plausibly exogenous. The variation induced by new universities affects our instrument by changing distance to yet-to-be-established universities for all those parents 1) who are born at most 19 years prior to the establishment of a new university and 2) for whom the to-be-established university changes the distance from the birth municipality to the nearest university. As we show in Appendix A-1.2, a significant fraction of parents in our data were affected by the opening of new universities (Figure A-2).

Figure 5 illustrates the identifying variation in our data, showing how individuals' distance to the nearest university changes as a result of the establishment of the new universities. The figure displays the mean distance to the nearest university for a given cohort of parents (reported in the year they turn 19) and the 25^{th} and the 75^{th} percentiles. The red vertical lines show the years when a university was established in a new location, thus changing distance to the nearest university for at least some of the parents. As can be seen from the Figure, there is variation in the distance to university that is independent of the establishment of universities in new locations, and some changes in this over the years. The reason for the changes is variation in where the (future) parents were born. It comes across strongly however that the establishment of universities in new locations resulted in large drops in both the mean and the two reported percentiles of the distance-to-university distribution. Since Finland is a large country and was relatively sparsely populated at the time when the new universities were established, it is plausible that geography mattered for the choices of the affected individuals to pursue

[18] We use the distance between an individual's birth municipality and her/his nearest university at age 19, rather than distance between the municipality where the parent lived at age 19 and the nearest university at that age because of lack of residential data at age 19 and because the birth municipality is more likely to be exogenous.

[19] E.g. Carneiro and Heckman (2002), Currie and Moretti (2003), Toivanen and Väänänen (2016), and Suhonen and Karhunen (2019).

[20] See Appendix A-1.2, where Table A-4 shows a list and Figure A-1 a map of the new universities.

either a BSc or MSc in a university.

Figure 5 here

Second, following e.g. Carneiro et al. (2011), we introduce additional controls that allow holding constant factors that may induce a correlation between parents' distance to the nearest university and offspring outcomes. These controls are: 1) the size of the parents' municipal birth cohort; 2) the share of the parents' municipal birth-cohort obtaining an MSc by age 35; 3) the share of the municipal birth-cohort earning above median income, measured for the same birth-cohort but nationally, at age 35; and 4) the share of individuals of the birth-cohort in their birth municipality earning an income in the top 10% of the birth-cohort, calculated as in 3). We exclude the parent in question when calculating these variables. These controls are to our knowledge new to the literature and are designed to capture the potential longer-term impacts of the differences in the rearing environment of the parents.

We study the association between these controls and parental distances to nearest university in detail in Appendix A-1.2, but report the main features here: The size of the municipal birth cohort is negatively correlated with distance to university; Table A-5 shows the correlation coefficients and Figure A-6 shows the relationships. The fraction of the municipal birth cohort that obtains an MSc degree is strongly correlated with distance to university (correlation coefficients -0.26 for maternal and -0.55 for paternal distance; see also Figure A-7). Similarly, the fractions of the municipal birth cohort obtaining above median or top decile income are strongly correlated with parental distance to university (-0.23 and -0.42 for above median, -0.21 and -0.45 for top decile; see also Figures A-8 and A-9). These strong correlations - all highly statistically significant - suggest that using these control variables indeed can strengthen the exclusion restriction we need for the causal interpretation of our results. The additional controls allow us to rule out e.g. the possibility that variation in income or social status that has origins in the parents' birth location would compromise the exclusion restriction.

In an extension, we also study a sub-sample of individuals (not parents) for which we have IQ data. This sub-sample contains the male cohorts starting from birth year 1961, allowing us to explore whether the unobserved ability of the offspring drives our findings.

4.2 Estimation equation and the first stage

Our goal is to estimate the causal effect of parental education on an offspring's probability to invent. Our estimation equation is

$$y_i = \boldsymbol{X_i'}\boldsymbol{\beta} + \delta D_i + \epsilon_i \tag{2}$$

where y_i identifies the off-spring inventors (alternative outcome variables are considered in the robustness checks) and takes value 1 if individual i is an inventor and is zero otherwise, $\boldsymbol{X_i}$ is a vector of controls (maternal and paternal year-of-birth dummies, the indicator for mother tongue not being Finnish, and the controls for the birth municipalities of both parents discussed above);[21] $\boldsymbol{\beta}$ is the associated coefficient vector; D_i is the parental education indicator taking value 1 if individual i has at least one parent with at least an MSc and 0 otherwise (other measures of parental education are used in the robustness checks); δ is the causal parameter of interest and ϵ_i is an error term. In line with the

[21] It is probably worth noting here that we do not control for parental income or wealth. These would be so-called bad controls as they themselves would be affected by parental education. See the discussion in Section 7.1.

literature, we use linear probability models for ease of interpretation, though note that we also estimate a Roy model in Section 5.2.

The worry regarding the identification of the causal effect is that D_i (i.e., the binary indicator for at least one parent having an MSc) and the error term are correlated. When instrumenting D_i in the first stage, we follow Heckman et al. (2006) and use a propensity score as the eventual instrument. Using the propensity score as the eventual instrument guarantees positive weights when integrating the marginal treatment effect (Heckman et al. 2006).

The propensity score is estimated by projecting D_i on our distance instruments and the controls. We employ a 3^{rd} order polynomial of log parental distance and use either only maternal or paternal distance to the nearest university, or both parental distances to the nearest university. Besides the distance instruments, these regressions include our base controls, i.e., the full set of maternal and paternal year-of-birth dummies and a dummy for mother tongue not being Finnish, and our full set of additional municipality controls, i.e., the number of children born in the parental birth municipality in the year of maternal / paternal birth; the fraction of the parental municipal birth cohort that have obtained an MSc by age 35; the fraction of the parental municipal birth cohort that had above median income at age 35, where the median is calculated over the whole national birth cohort; and similarly, the fraction of the parental municipal birth cohort that had an income in the top percentile of the national cohort at age 35.

We display the details of the first stage results in Appendix A-2. The distance variables are individually and jointly highly significant, suggesting that the instruments are strong and relevant (see also below the F-test values in Table 1). The additional controls for the characteristics of the parental birth municipality are also jointly highly significant.

5 The causal effect of parental education on the probability of becoming an inventor

5.1 Baseline IV estimates

Table 1 contains our main results. The top panel presents results using all the data; the middle panel results using data on daughters; and the bottom panel results using data on sons. Starting from the top panel, we find from Column (1) a statistically significant OLS coefficient of 0.02, suggesting a relatively strong positive association (recall that the sample mean probability of inventing is 0.0067) with having at least one MSc parent and the probability of an offspring becoming an inventor. In Column (2) we use maternal distance as the instrument. The resulting IV-estimate of δ is 0.05 (s.e. 0.01), i.e., a five percentage point increase in the probability of becoming an inventor. The results in Column (3) are produced using only paternal distance as the instrument: The resulting estimate of δ is 0.03 (s.e. 0.01). Finally, we use distances of both parents as instruments in Column (4). Doing so, the dummy for having at least one MSc parent carries a coefficient of 0.03 (s.e. 0.01).

Table 1 here

These results suggest the following: First, there is a positive causal impact of parental (university) education on the probability that an offspring becomes an inventor. The effect is sizeable as all IV estimates are 5-8 times the sample mean (0.0067) of the probability of becoming an inventor.[22] Moreover, the size of the point estimates varies with the instrument. This variation suggests that the treatment effect is heterogenous, i.e., that we identify a local average treatment effect (LATE): According to Heckman et al. (2010), using two different (sets of) instruments is one way to test for the heterogeneity of the treatment effect. The LATE interpretation means in our context that the estimated treatment effect applies to those individuals whose either parent was induced to obtain an MSc due to a change (reduction) in distance to nearest university. Second, the IV estimates are all larger than the OLS estimates, suggesting a negative correlation between the unobservables and parental education. While a downward OLS bias is a frequent finding in the returns to education literature (reported already in Griliches 1977), in our inter-generational context the downward bias needs to be interpreted carefully. It suggests that the individuals, who on the basis of their unobservables are more likely to invent, are less likely to have educated parents. While there can be various channels and mechanisms at work, we can nevertheless say that had the new universities not been established, the complying parents' would not have gone to university and it would have been less likely that their offspring would have become inventors. The IV estimate could thus be larger e.g. because the establishment of the new universities made it easier for higher-ability parents to study in a university and because such induced educational choice (and the associated accumulation of parental human capital) then enhanced offspring's skills and human capital in a way that eventually supports their inventiveness later in life.

Turning to the middle panel we find a smaller but statistically significant OLS coefficient for daughters at 0.01. Using only maternal distance as the instrument we obtain a noisy coefficient of 0.01. Using paternal distance we in contrast find a statistically significant IV-estimate of 0.02. Using both parents' distances the point estimate decreases slightly, but is highly statistically significant. These results suggest that parental education, in particular paternal education and that of both parents, has a positive causal impact on those daughters whose parents were induced to obtain an MSc due to (changes in) the distance to the nearest university. Comparing these estimates to the daughters' sample mean probability of inventing shows that these effects are large, increasing the probability of inventing at between 6 - 13 fold (sample mean 0.0016).

The estimations on sons, reported in the lowest panel, produce larger point estimates. The other important difference to daughters is that using only maternal distance as the instrument produces a very large and statistically significant coefficient (0.09), but using paternal distance produces a much smaller albeit still statistically significant coefficient (0.04) - the reverse pattern from that observed with daughters.[23] Just like for daughters, using both parents' distances leads to a statistically significant estimate (0.05). A comparison to the son's sample mean reveals that these are large effects, indicating a roughly 4-9 fold increase in the probability of inventing (sample mean 0.0118).

Comparing the results on daughters and sons in terms of economic significance reveals that the point estimates for sons are larger than those for daughters. However, when one relates the point estimates to sample averages, the ordering is reversed when either paternal or parental distances are used as the instrument.[24] For example, the ratio of the coefficient to the sample mean for daughters (0.0016) is 13 using the paternal distance IV. The same ratio for sons (sample mean 0.0118) is 3.7.

[22] Toivanen and Väänänen (2016) report an IV coefficient of 0.20 for own MSc university education.

[23] As we discuss later, the observed gender differences in how paternal and maternal distances to the nearest university affect the probability of offspring invention provide us with clues of the likely (and unlikely) mechanisms at work.

[24] When using maternal distance as instrument the ratios are close to each other.

5.2 Average treatment effects on all, the treated and the untreated

The above LATE estimates are policy relevant as the main variation producing them is caused by a policy decision - university openings. It is nonetheless of policy interest to uncover average treatment effects for example because they allow an easier quantification of the benefits of the policy in question. To do so we adopt the approach of e.g. Heckman et al. (2006) and estimate a generalized Roy model.[25] The Roy model consists of three equations where the first one (equation 3) defines the potential outcomes for individual i conditional on treatment status j ($j = 0$ no parent with an MSc; $j = 1$ at least one parent with an MSc); the second the observed outcome for individual i (equation 4); while the third equation determines the treatment status of individual i (equation 5):

$$y_{ji} = \boldsymbol{X_i'}\boldsymbol{\beta_j} + \epsilon_{ji}, \quad j = 0, 1 \tag{3}$$

$$y_i = D_i y_{1i} + (1 - D_i) y_{0i} \tag{4}$$

$$D_i = \mathbb{1}[\boldsymbol{X_i'}\boldsymbol{\theta_1} + \boldsymbol{Z_i'}\boldsymbol{\theta_2} + \nu_i] \tag{5}$$

where y_{ji} is the treatment specific outcome, y_i the observed outcome (i.e., takes value one if individual i is an inventor and zero otherwise), D_i the treatment status of individual i (i.e., takes value one if individual i has at least one parent with an MSc and zero otherwise), $\mathbb{1}[.]$ the indicator function, $\boldsymbol{X_i}$ the vector of control variables discussed above, $\boldsymbol{Z_i}$ the vector of instruments (i.e., parental distances to nearest university) and $\boldsymbol{\beta_j}$, $\boldsymbol{\theta_1}$ and $\boldsymbol{\theta_2}$ are parameter vectors.

For simplicity, we estimate a parametric version of (3), (4) and (5), specifying that control variables enter in a linear and additive fashion and the error terms ϵ_{ji} and ν_i are jointly normally distributed. We use the third order polynomial of both parents' distance to the nearest university as our instrument vector and replace the parental year-of-birth dummies with year-of-birth variables.[26]

Table 2 here

Table 2 reveals that LATE estimates are slightly smaller than those in Column (3) of Table 1. The estimated average treatment effect (ATE), i.e., the effect of assigning at least one MSc parent to a randomly chosen individual, is 0.007 for daughters and 0.024 for sons.[27] Estimated average treatment effects on the treated (ATTs) are close to LATEs and larger than ATEs for all the three samples.[28] The ATTs suggest a significant impact of parental education on the offspring: e.g., the probability of a son becoming an inventor increases by a factor of four compared to the sample average. While there can be several specific mechanisms at work, the LATE estimates capture the causal effect of parental education for the offspring of those complying parents who went to a university because one was built nearby. Estimated average treatment effects on the untreated (ATUTs) are close to the estimated ATEs.[29] Also,

[25] We use the code in Andresen (2018).

[26] This is done to speed up the bootstrap; the point estimates are only marginally affected.

[27] ATE can be interpreted as the average change in outcomes that would be experienced if all individuals were required to get the treatment compared to the case in which they are not allowed to get the treatment.

[28] ATT in our context is the average causal effect on the probability of inventing for those individuals in our data who actually have at least one MSc parent. It measures the difference for those individuals between the outcome observed in the data where they had at least one MSc parent to the counterfactual outcome where neither of their parents had obtained an MSc. More generally, ATT refers to the average change in outcomes that would be experienced by the treated group if it switched from a regime in which the treatment is optional to a regime that forbids the treatment.

[29] ATUT in our context is the average causal effect on the probability of inventing for those individuals in our data whose

we cannot reject the null hypotheses of no observed and no unobserved heterogeneity for daughters, but can reject them for sons.

6 Robustness tests

To investigate the robustness of our results, we consider alternative IV estimations in which we either change the way parental education is measured or alter the main outcome variable for the offspring's inventiveness (for details, see Appendix A-3).

First, we measure parental education using an indicator for the parents having obtained a BSc (instead of an MSc). The motivation for this is that obtaining a BSc was more frequent among the parents than obtaining an MSc. This changed in Finland in the 1970s with a change in the degree structure which made the MSc the first official degree in most disciplines. As our second robustness test, we use the count of MSc parents as the measure of parental education.

So far, our main outcome variable has been a binary indicator for the offspring ever becoming an inventor. For the third robustness test, we use the number of patents in which the offspring is mentioned as one of the inventors as the outcome variable. Fourth, we change the outcome to be the total number of forward citations to all the patents of an individual. This outcome captures both the quantity and quality of the individual's inventions.

Robustness test 1: We re-estimate our model using parents obtaining a BSc as our measure of parental education, i.e., a dummy taking value one if at least one of the parents has a bachelor degree. The results are in line with our baseline findings (see Table A-7), with slightly lower point estimates in all three samples. To give an example, the coefficients of $D(BSc\ parents)$ using the sample on daughters are 0.01 and statistically significant at 5% level or better using either paternal or both parents' distance as instruments. When we use only maternal distance as instrument, the coefficient is 0.005 and not distinguishable from zero at conventional levels. These coefficients are roughly 50% lower than those obtained using $D(MSc\ parents)$ as the measure of parental education. The coefficients for sons are also up to 50% lower than those reported in Table 1, with two significant at 1% level and one (using paternal distance only as instrument) at 10% level. These results give further confirmation that parental education matters for off-spring invention, and furthermore suggest that there is a difference in the size of the causal effect of parental education at different levels of university education, triggered by a change in the distance to nearest university.

Robustness test 2: Using the number of parents with an MSc as the measure of parental education we obtain results that echo our main results (see Table A-8). The point estimates are somewhat smaller, but the pattern of statistical significance is the same. An inbuilt characteristic of this model is that the causal effect doubles for those individuals who have two parents with at least an MSc.

Robustness test 3: Using the patent count as the outcome variable and maternal distance to university as an instrument, we obtain a coefficient of 0.16 (s.e 0.001), larger than the 0.05 OLS estimate (s.e. 0.001); see Table A-9 for details. Toivanen and Väänänen (2016) report an IV estimate of 0.20 using an MSc-dummy for own education and controlling for paternal education (and an OLS coefficient of 0.04). Using the subsample of daughters, the only IV estimate that is statistically significant is the one using

neither parent has an MSc. It measures the difference between the counterfactual outcome where these individuals would have had at least one parent with an MSc to the outcome observed in the data where neither parent had an MSc. More generally, ATUT refers to the average change in outcomes that would be experienced by the control group if the treatment was made mandatory.

both parents' distances as instruments (0.04, s.e. 0.010). For sons, using maternal distance as instrument produces a coefficient of 0.26 (s.e. 0.080); using paternal distance produces a noisy 0.09 coefficient; using both parental distances a marginally significant point estimate of 0.11.

Robustness test 4: Turning to citations-regressions we find results close to those obtained using the patent count, but the coefficients are larger and more precisely estimated (see Table A-10). This result confirms that our baseline findings are robust to using an inventiveness measure that reflects the quality of offspring inventions.

7 Discussion and extensions

7.1 Potential mechanisms at work

What is captured by the estimated causal effect of parental education on the likelihood of offspring invention? Answering this question conclusively is tricky, because the literature suggests that better parental education may improve offspring outcomes not only through several causal pathways but also along several dimensions (see e.g. Lundborg et al. 2014 and Lundborg et al. 2018).

One way to dive deeper into the mechanisms at work would be to introduce for example parents' occupation, income or socioeconomic status (e.g. whether they are blue- or white-collar workers), or alternatively offspring health, fertility decisions, or educational choices (e.g. years of schooling, or STEM vs. non-STEM -field) into our analysis. However, as is well-known, adding such intermediate outcomes to the regression often leads to biases (Rosenbaum 1984, Wooldridge 2005, Angrist and Pischke 2009). Identifying and quantifying the underlying natural direct and indirect mechanisms would require a detailed mediation analysis (e.g., Heckman and Pinto 2015, and VanderWeele 2015). Implementing such an analysis would require further assumptions and data. The data currently available to us is rich, but not detailed enough to allow us to decompose the total effect of parental education that we have estimated into the relevant natural direct and indirect effects (see, e.g., VanderWeele 2015, Chapter 2).

The literature on the intergenerational effects of parental human capital suggests that the total causal effect of parental education on the likelihood of offspring invention can mirror how parents' greater earnings power (economic resources) and enhanced human capital improve the children's human capital (e.g., their educational choices, skills and health) as well as change their tastes (Lundborg et al. 2014, Lundborg et al. 2018, and Holmlund et al. 2011). We next consider whether some mechanisms are more plausible.

Let us first consider parental income. As it does elsewhere (Card, 2001; Carneiro et al., 2011), better education tends to lead to higher incomes also in Finland (Uusitalo 1999), including the parents in our data. Our descriptive regressions showed, however, that accounting for parental education greatly weakens the association between parental income and off-spring becoming an inventor. While suggestive, this result, combined with our finding that parental education has a positive causal effect on the probability of offspring becoming an inventor, goes a long way in explaining away the Finnish enigma, i.e., the strong positive and convex relation between parental incomes and the likelihood of offspring invention that we observe in the Finnish data despite the country's lower income inequality, more equitable schooling system, and higher social mobility than in many other countries (e.g., Black and Devereux 2011, Jäntti and Jenkins 2015).[30]

[30] Whether and if so, to what extent, the same holds for the much less egalitarian US remains an open question. It also bears on the policy implications of the earlier findings on the importance of family endowments and social environment for offspring

Second, parental education may change children's tastes for pursuing different opportunities (e.g., a field of education, or a career) that affect the likelihood of becoming an inventor (e.g. Hoisl et al. 2022). A potential mechanism at work is parental role modeling, which can be interpreted as a type of gender-based within-family homophily (e.g., Lindquist et al. 2015 and Boucher 2015, Brenøe and Epper 2022), leading potentially to more intense information sharing and transfer of tastes between fathers and sons, and between mothers and daughters (see also Hoisl et al. 2022). Another related mechanism is more approving and supportive behavior of parents toward same-sex children (e.g. Thomas 1994).[31] If these channels were at work, we would expect a stronger same sex effect of parental human capital on children's outcomes. Yet, our IV results do not support the empirical importance of these channels, since they suggest that using maternal (paternal) distance to the nearest university as the instrument produces a larger causal effect for sons (daughters). [32]

Third, a small number of studies suggest that an exogenous increase in parental education enhances children's education, although the magnitude of the effect is often small and varies depending on whether fathers' or mothers' education is considered; see Black et al. (2005) for Norwegian, Holmlund et al. (2011) for Swedish and Suhonen and Karhunen (2019) for Finnish evidence. Consistent with these patterns of intergenerational transmission of human capital, Lundborg et al. (2014) uses Swedish data and IV estimations to show that, besides improving their sons' health and and non-cognitive skills, mothers' additional schooling enhances the cognitive skills of their sons.[33]

This strand of the literature suggests that our IV estimates of the effect of parental education on the probability of offspring inventing may reflect the fact that (high-ability) mothers' and fathers' university education induces their (high-ability) children to acquire more and better cognitive skills (through e.g. their educational choices) which, in turn, make them more likely to become inventors. To see why that might be the case, it is useful to note that our IV estimator attaches more weight to the marginal effects at the higher end of the education (and thus skill) distribution because it utilises exogenous variation related to the establishment of new universities (and not e.g. due to a reform of compulsory schooling). Specifically, had the new universities not been established, the complying (high-ability) parents' would not have gone to the university, making it less likely that their (high-ability) offspring would have become inventors. This line of reasoning is consistent with Card (2001), arguing that marginal returns to education among those typically affected "by supply-side innovations tend to be relatively high, reflecting their high marginal costs of schooling, rather than low ability that limits their return to education". In our context, the returns to education are intergenerational and come about as new inventions.

Our IV estimate may therefore mirror the fact that (i) the establishment of the new universities made it easier for higher-ability parents to study in a university and that (ii) this enhanced both the affected parents' and their offspring's malleable human capital and skill formation in a way that supports the inventiveness of the offspring. These observations raise, however, the obvious question of whether our IV estimate is biased because of unobserved genetically heritable skills that parents and their children

inventiveness (Bell et al. 2019).

[31] Consistent with more supportive behavior of parents toward their same-sex children, Thomas (1994) finds for example that mothers (fathers) allocate more resources toward their daughters (sons).

[32] The findings of Hoisl et al. (2022) suggest that the gender-based mechanisms, if at work, may also depend on the precise gender composition and even birth order of a family's children.

[33] See also Lundborg et al. (2018) who provide related evidence using twins and adopted children and who discuss the reasons why the IV and twin/adoption estimates may differ. It is worth pointing out that Pekkala Kerr et al. (2013) find no effect on cognitive test scores of the affected individuals when the Finnish school system was reformed.

share. If so, our IV estimations would suffer from an omitted variable bias. We therefore turn to it next.

7.2 Role of IQ for offspring inventiveness

To rule out the alternative explanation that genetically heritable innate ability (IQ) is driving our causal findings, we study a sub-sample of individuals (not parents) for which we have IQ data. This sub-sample contains the male cohorts starting from birth year 1961.

To start with, the association between own ability and inventing is strong: Figure 6 plots the probability to invent against IQ percentiles. We use the visuo-spatial IQ, which arguably is less malleable than the other IQ components. The probability to invent has an increasing and convex association with IQ. Comparing individuals at the extreme right tail of the IQ distribution to those in the middle shows that the former are five to six times more likely to invent than the latter.[34]

Figure 6 here

Second, to explore the concern that our IV estimations might suffer from an omitted variable bias, we measure the correlation between parental distance to university and IQ of the offspring. In line with Carneiro and Heckman (2002), we find a negative correlation but in contrast to them, our measure refers to an inter-generational correlation. It is also notable that these correlations, at -0.064 and -0.066 (both significant at 1% level), are weaker than those of our birth municipality characteristics (but similar to what Carneiro and Heckman 2002 report).[35]

These two empirical observations suggest that not controlling for IQ may cause omitted variable bias and hence a robustness check is in order. To explore the importance of measurable IQ for our findings, we use a 4^{th} order polynomial IQ of individual i (not of parents) as an additional control in the IV estimations. The IQ sample consists of all those individuals in our main estimation sample for whom we have IQ data from the FDF. In practice, this means the large majority of the males starting from the 1961 birth cohort.

To be able to analyze the effect of including the IQ as a control we present results using the same sample, and estimating the model with and without the IQ controls. We also estimate the propensity scores separately using and not using IQ as a control. Otherwise, the specification is as in Column 3 of Table 1, i.e., we use a 3^{rd} order polynomial of both parental distances to university. Results are reported in Table 3.

Table 3 here

Comparing the results in the upper panel of Table 3 to those in the bottom panel of Table 1 we find that two coefficients have decreased somewhat in size and two have increased. Specifically, comparing the results with and without the IQ controls we find that introducing IQ reduces the OLS coefficient from 0.03 to 0.02, suggesting that IQ affects the probability of inventing, and omitting it may lead to upward bias; the IQ variables are jointly statistically significant. Turning to the IV estimates, we find that they also decrease when IQ is added. With the coefficients varying between 0.03 and 0.06, the change is between 2/3 (OLS) and 2 percentage points or 20 - 37%. The coefficients using maternal or

[34]In an unreported graph we find that using overall IQ produces an even more pronounced increase in the probability to invent at the right tail of the distribution.

[35]Figure A-10 in Appendix A-1.2 provides additional details.

paternal distances as instruments are statistically significant at 6% level, the one using both parents' distances at 2% level.

Adding the offspring IQ to the model is not necessarily the only or even primary reason for why the IV estimates decrease somewhat in size. The first birth cohort of 1961 for which IQ data is available also happens to be the one with which the transition to comprehensive schooling started in Finland (as we explain in more detail shortly). In line with what we report below, this reform may also weaken how parental schooling affects offspring outcomes.

Figure 6 shows that high-IQ individuals are much more likely to invent than those at the middle of the IQ ditribution. That observation raises the question whether parental education has a differential effect on high-IQ individuals. To study this question, we interacted the dummy for at least one parent having an MSc with a dummy indicating whether individual i is in the top-decile of the (visuospatial) IQ distribution. In an OLS regression the interaction obtains a positive and significant coefficient. However, once we instrument the interaction using interactions with our instruments and the high-IQ dummy the interaction obtains a negative and statistically insignificant coefficient while the dummy for at least one parent having an MSc carries essentially the same coefficient as in the specification without the interaction. In an additional analysis, we redefined high-IQ individuals to be those in the top-5% of the (visuospatial) IQ distribution. The results did not change.

We conclude that genetically heritable innate ability is not driving our causal findings. Moreover, we cannot find support for the view that parental education would differentially improve the probability of inventing in the top of the IQ distribution once the endogeneity of parental education is accounted for.

7.3 Implications of the Finnish comprehensive school reform

Why should parental education matter for the offspring outcomes in a country where education is free, comprehensive, and produces high test-scores in international comparisons? The answer may be that the move towards an inclusive education system in Finland was relatively recent (Pekkala Kerr et al. 2013). It was not until the 1970s that the education system became comprehensive. Prior to the reform, a two-track system channelled a limited fraction of students at age 11 to general secondary education; others continued on the vocational track. The reform postponed the choice between general and vocational education until age 16.[36] Pekkarinen et al. (2009) find that the Finnish comprehensive school reform reduced the intergenerational income elasticity by 23%.

The reform was implemented in 1972-1977 in a geographically gradual manner, starting from Northern Finland: students above the fifth grade continued in the old system, younger cohorts transferred into the new system (see, e.g., Pekkala Kerr et al. 2013). The 1961 cohort is the oldest affected by the reform and the 1966 cohort the first to completely enter comprehensive school. As almost all of the parents in our data were born before the 1960s, the reform cannot have affected their education; however, more than half of the individuals in our data were born in 1966 or later, and it is conceivable that the reform affected the relation between parental education and the probability of inventing.

How the educational system that the children face interacts with the educational background of their parents requires a thorough analysis that is beyond the scope of this paper. To provide a first step, we studied the last cohorts not affected by the reform (1956–1960) and the first cohorts fully

[36]The reform also changed the primary school curriculum, emphasizing STEM subjects. The level of teaching between ages 11-16 was adjusted for higher student heterogeneity.

affected by the reform (1966–1970). We first compare by parental income percentile the probability of offspring invention by dividing $P(Inventor)$ of the 1966-1970 cohorts by the $P(Inventor)$ of the 1956-1960 cohorts by parental income percentile. This ratio is depicted in Figure 7 separately using maternal and paternal income percentiles. What we find is that the probability of becoming an inventor has in general increased going from the 1956-1960 to the 1966-1970 cohorts; a ratio above one indicates this. We find in addition that the increase has been the largest at the middle of the income distribution. This suggests that in the later cohorts, the relation between parental income and $P(Inventor)$ has weakened.

Figure 7 here

To study this more formally we return to our regression framework but limit the sample to the 1956-1960 and 1966-1970 cohorts. We estimate our base model separately for the two subsamples. The coefficient of interest is that of $D(MSc\ parent)$.

Table 4 here

In Table 4 we present the results using data on the 1956-1960 (pre) and 1966-1970 (post) cohorts, i.e., those born in the 5 last years before the implementation of the comprehensive school reform and the 5 first cohorts for which comprehensive school covered the whole country. We find that the OLS coefficient increases from 0.015 to 0.022 when we move from the 1956-1960 to the 1966-1970 cohorts. Using maternal distance to university as instrument, we find a statistically significant $D(MSc\ parent$- coefficient of 0.10 for the earlier cohorts, but a small (0.01) and imprecise coefficient for the younger cohorts. When instrumentation is based on paternal distance to university, the point estimate again decreases (from 0.026 to 0.014) moving from the earlier to the later cohort, but neither is statistically significant. When both parental distances are used as instruments, the LATE estimate is again larger for the earlier (0.045) than for the younger (0.033) cohort.

Taken together, these results strongly suggest that the comprehensive school reform has weakened the effect of parental education on the probability of inventing. This pattern is what we would expect if the new system is more equitable and egalitarian. However, we see this analysis as a first step since there are several potential channels through which such a system-wide reform may have affected individuals' outcomes.

8 Conclusion

Using data on Finnish individuals born 1953-1981 and their parents, we find that: (i) parental income is positively associated with the probability of inventing, but the association is greatly diminished once parental education is controlled for; (ii) instrumenting parents' education with their distance at age 19 to the nearest university we find a large local average treatment effect of parental education on children's probability of inventing; (iii) the causal impact of parental education on sons is three times the impact on daughters in absolute terms, but typically smaller in relative terms; and (iv) the estimated average treatment effects on both the treated and untreated are positive.

These results are robust to a variety of tests, including using the IQ sub-sample, and apparently not driven e.g. by intergenerational transfer of tastes. We conclude that the establishment of the new universities made it easier for higher-ability parents to study in a university and that this improved access to higher education enhanced both the parents' and their children's human capital and skill formation in a way that increased the capacity of the offspring to invent.

We also find evidence that the causal effect of parental education on the probability to invent is intermediated by the features of the schooling system that the individual is subjected to. This finding is also in line with the view that the establishment of new universities had an intergenerational causal impact on children's outcomes, especially those who did not (yet) have the opportunity to benefit from the more egalitarian schooling system.

Our analysis has interesting policy implications. In particular, it suggests that by massively and persistently investing in education up to (STEM) PhD level, a country can significantly increase its aggregate innovation potential while making innovation-led growth also more inclusive.

More generally, our findings in this paper have implications for the debate on the optimal model of capitalism. Should we follow the more innovative US capitalism or German/Scandinavian capitalism, which is more inclusive and protective? Are we bound to an "either/or" choice between these two forms of capitalism? Our analysis in this paper leads us to depart from the "either/or" view: indeed, it suggests that there are policies which can help move capitalism both towards more innovativeness and towards more protection or inclusiveness. In particular, we have reasons to believe that the Finnish education reform in the early 1970s both stimulated aggregate innovation and made growth more inclusive by allowing more talented individuals with low-educated parents to become innovators, i.e., it reduced the number of "lost Einsteins and Marie Curies".

References

Aghion, P., U. Akcigit, and P. Howitt: 2014, 'What Do We Learn From Schumpeterian Growth Theory?'. In: P. Aghion and S. N. Durlauf (eds.): *Handbook of Economic Growth, Volume 2*. pp. 515–563.

Aghion, P., U. Akcigit, A. Hyytinen, and O. Toivanen: 2017, 'The Social Origins of Inventors'. National Bureau of Economic Research WP 24110.

Aghion, P., U. Akcigit, A. Hyytinen, and O. Toivanen: 2018, 'On the returns to invention within firms: Evidence from Finland'. In: *AEA Papers and Proceedings*, Vol. 108. pp. 208–12.

Aghion, P., U. Akcigit, A. Hyytinen, and O. Toivanen: 2022, 'A Year Older, A Year Wiser (and Farther from Frontier): Invention Rents and Human Capital Depreciation'. *The Review of Economics and Statistics* pp. 1–31.

Aghion, P. and P. Howitt: 1992, 'A Model of Growth through Creative Destruction'. *Econometrica* **60**(2), 323–351.

Akcigit, U., J. Grigsby, and T. Nicholas: 2017, 'The Rise of American Ingenuity: Innovation and Inventors of the Golden Age'. National Bureau of Economic Research WP 23047.

Akcigit, U. and T. Nicholas: 2019, 'History, Micro Data, and Endogenous Growth'. *Annual Review of Economics* **11**, 615–633.

Andresen, M.: 2018, 'Exploring marginal treatment effects: Flexible estimation using Stata'. *Stata Journal* **18**, 118–158.

Angrist, J. D. and J.-S. Pischke: 2009, *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Azoulay, P., J. S. Graff Zivin, and J. Wang: 2010, 'Superstar extinction'. *The Quarterly Journal of Economics* **125**(2), 549–589.

Bell, A., R. Chetty, X. Jaravel, N. Petkova, and J. Van Reenen: 2019, 'Who Becomes an Inventor in America? The Importance of Exposure to Innovation'. *Quarterly Journal of Economics* **134**(2), 647–713.

Black, S. and P. Devereux: 2011, 'Recent developments in intergenerational mobility'. In: O. Ashenfelter and D. Card (eds.): *Handbook of Labour Economics*, Vol. 4B. Elsevier, pp. 1487–1541.

Black, S. E., P. J. Devereux, and K. G. Salvanes: 2005, 'Why the apple doesn't fall far: Understanding intergenerational transmission of human capital'. *American Economic Review* **95**(1), 437–449.

Bloom, N., J. Van Reenen, and H. Williams: 2019, 'A toolkit of policies to promote innovation'. *Journal of Economic Perspectives* **33**(3), 163–184.

Boucher, V.: 2015, 'Structural homophily'. *International Economic Review* **56**(1), 235–264.

Brenøe, A. A. and T. Epper: 2022, 'Parenting values and the intergenerational transmission of time preferences'. *European Economic Review* **148**, 104–208.

Cameron, S. and C. Taber: 2004, 'Estimation of educational borrowing constraints using returns to schooling'. *Journal of Political Economicy* **112**(1), 132–182.

Card, D.: 1995, 'Using geographic variation in college proximity to estimate the return to schooling'. In: L. Christofides, K. Grant, and R. Swidinsky (eds.): *Aspects of Labour Market Behaviour: Essays in Honour of John Vanderkamp*. Toronto: University of Toronto Press.

Card, D.: 2001, 'Estimating the return to schooling: Progress on some persistent econometric problems'. *Econometrica* **69**(5), 1127–1160.

Carneiro, P. and J. Heckman: 2002, 'The evidence on credit constraints in post-secondary schooling'. *Economic Journal* **112**, 705–734.

Carneiro, P., J. Heckman, and E. Vytlacil: 2011, 'Estimating marginal returns to education'. *American Economic Review* **101**, 2754–2781.

Celik, M. A.: 2015, 'Does the Cream Always Rise to the Top? The Misallocation of Talent and Innovation'. *Unpublished, University of Toronto*.

Currie, J. and E. Moretti: 2003, 'Mother's education and the intergenerational transmission of human capital: Evidence from college openings'. *Quarterly Journal of Economics* **118**, 1495–1532.

Dal Bó, E., F. Finan, O. Folke, T. Persson, and J. Rickne: 2017, 'Who Becomes a Politician?'. *Quarterly Journal of Economics*. forthcoming.

Eisenhauer, P., J. Heckman, and E. Vytlacil: 2015, 'The generalized Roy model and the cost-benefit analysis of social programs'. *Journal of Political Economy* **123**(2), 413–443.

Eskola, S.: 2002, 'Korkeakouluja Helsingin ja Turun ulkopuolelle (Universities outside Helsinki and Turku")'. In: P. Tommila and A. Tiitta (eds.): *Suomen Tieteen Historia, osa 4, Tieteen ja Tutkimuksen Yleinen Historia 1880–luvulta Lahtien (The history of Finnish science, part 4, The general history of science and research from 1880s onwards)*, Vol. 4. Helsinki: WSOY.

Griliches, Z.: 1977, 'Estimating the returns to schooling: some econometric problems'. *Econometrica* **45**(1), 1–22.

Heckman, J., J. E. Humphries, and G. Veramend: 2018, 'Returns to education: The caual effects of education on earnings, health and smoking'. *Journal of Political Economy* **126**(S1), S197–S246.

Heckman, J., D. Schmierer, and S. Urzua: 2010, 'Testing the Correlated Random Coefficient Model'. *Journal of Econometrics*.

Heckman, J., S. Urzua, and E. Vytlacil: 2006, 'Understanding instrumental variables in models with essential heterogeneity'. *Review of Economics and Statistics* **88**(3), 389–432.

Heckman, J. J. and R. Pinto: 2015, 'Econometric mediation analyses: Identifying the sources of treatment effects from experimentally estimated production technologies with unmeasured and mismeasured inputs'. *Econometric Reviews* **34**(1–2), 6–31.

Hoisl, K., H.-C. Kongsted, and M. Myriam: 2022, 'Lost Marie Curies: Parental Impact on the Probability of Becoming an Inventor'. *Management Science*.

Holmlund, H., M. Lindahl, and E. Plug: 2011, 'The causal effect of parent's schooling on children's schooling: A comparison of estimation methods'. *Journal of Economic Literature* **49**(3), 615–651.

Jäntti, M. and S. Jenkins: 2015, 'Income mobility'. In: A. Bourguignon (ed.): *Handbook of Income Distribution*, Vol. 2. Elsevier, pp. 807–935.

Jaravel, X., N. Petkova, and A. Bell: 2018, 'Team-specific capital and innovation'. *American Economic Review* **108**(4-5), 1034–73.

Jokela, M., T. Pekkarinen, M. Sarvimäki, M. Terviö, and R. Uusitalo: 2017, 'Secular rise in economically valuable personality traits'. *Proceedings of the National Academy of Sciences* **114**(25), 6527–6532.

Kline, P., N. Petkova, H. Williams, and O. Zidar: 2019, 'Who Profits from Patents? Rent-Sharing at Innovative Firms'. *Quarterly Journal of Economics* **134**(2), 843–894.

Lindquist, M. J., J. Sol, and M. Van Praag: 2015, 'Why Do Entrepreneurial Parents Have Entrepreneurial Children?'. *Journal of Labor Economics* **33**(2), 269–296.

Lundborg, P., A. Nilsson, and D.-O. Rooth: 2014, 'Parental education and offspring outcomes: Evidence from the Swedish compulsory school reform'. *American Economic Journal: Applied Economics* **6**(1), 253–278.

Lundborg, P., M. N. Nordin, and D. O. Rooth: 2018, 'The intergenerational transmission of human capital: the role of skills and health'. *Journal of Population Economics* **31**, 1035–1065.

Nelson, R. R. and E. S. Phelps: 1966, 'Investment in Humans, Technological Diffusion and Economic Growth'. *American Economic Review*.

Pekkala Kerr, S., T. Pekkarinen, and R. Uusitalo: 2013, 'School Tracking and Development of Cognitive Skills'. *Journal of Labor Economics* **31**(3), 577–602.

Pekkarinen, T., R. Uusitalo, and S. Kerr: 2009, 'School tracking and intergenerational income mobility: Evidence from the Finnish comprehensive school reform'. *Journal of Public Economics* **93**(7), 965–973.

Rosenbaum, P. R.: 1984, 'The consequences of adjustment for a concomitant variable that has been affected by the treatment'. *Journal of the Royal Statistical Society: Series A (General)* **147**(5), 656–666.

Suhonen, T. and H. Karhunen: 2019, 'The intergenerational effects of parental higher education: Evidence from changes in university accessibility'. *Journal of Public Economics* **176**, 195–217.

Takalo, T. and O. Toivanen: 2015, 'Economics of Innovation Policy'. *Nordic Economic Policy Review* **2**, 65–90.

Thomas, D.: 1994, 'Like Father, like Son; Like Mother, like Daughter: Parental Resources and Child Height'. *Journal of Human Resources* **29**(4).

Toivanen, O. and L. Väänänen: 2012, 'Returns to Inventors'. *Review of Economics and Statistics* **94**(4), 1173–1190.

Toivanen, O. and L. Väänänen: 2016, 'Education and Invention'. *Review of Economics and Statistics* **98**(2), 382–396.

Uusitalo, R.: 1999, 'Return to education in Finland'. *Labour Economics* **6**(4), 569–580.

Van Reenen, J.: 1996, 'The Creation and Capture of Rents: Wages and Innovation in a Panel of U.K. Companies'. *Quarterly Journal of Economics* **111**(1), 195–226.

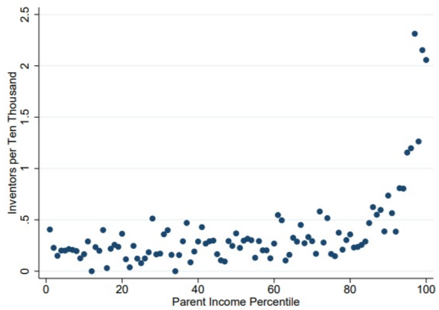
VanderWeele, T. J.: 2015, *Explanation in Causal Inference: Methods for Mediation and Interaction*. Oxford University Press.

Wooldridge, J. M.: 2005, 'Violating ignorability of treatment by controlling for too many factors'. *Econometric Theory* **21**(5), 1026–1028.

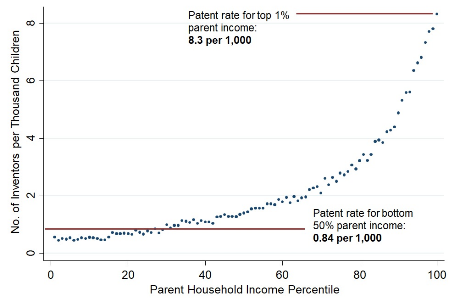
Figures and Tables

Figure 1: Parental income and Prob(*Inventor*)

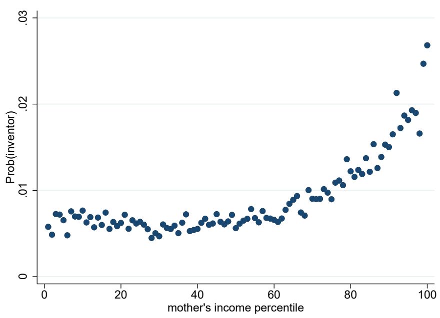
1A. 1930s U.S.



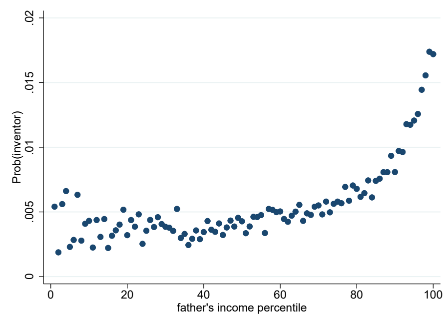
1B. 1980s U.S



1C. Finland: Maternal income



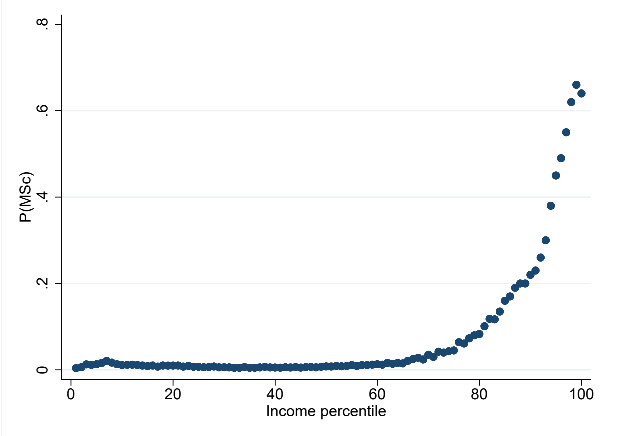
1D. Finland: Paternal income



Sources: Figure 1A: Akcigit et al. (2017), Figure 1B: Bell et al. (2019), Figures 1C & 1D: own calculations.

Figure 2: Parental income and Prob(MSc_{parent})

2A. Maternal income & Prob(MSc_{mother})



2B. Paternal income & Prob(MSc_{father})

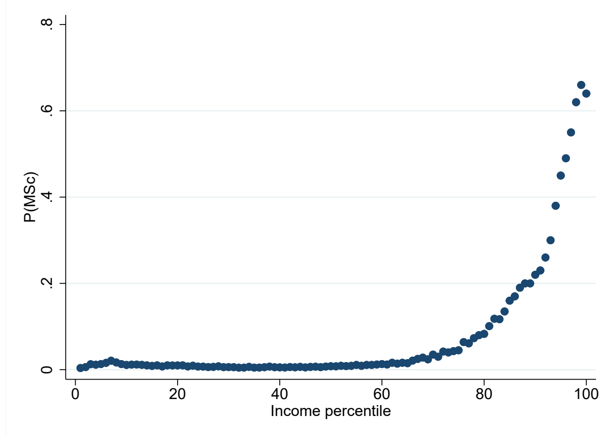
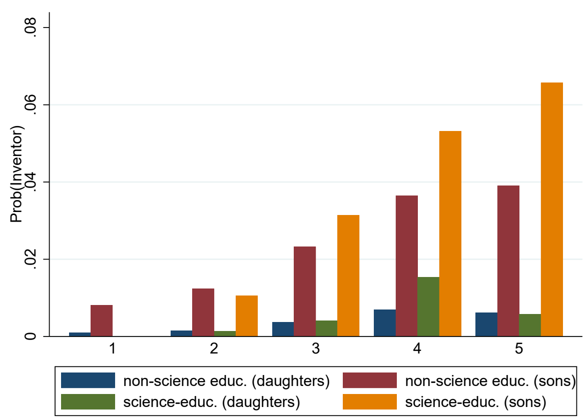
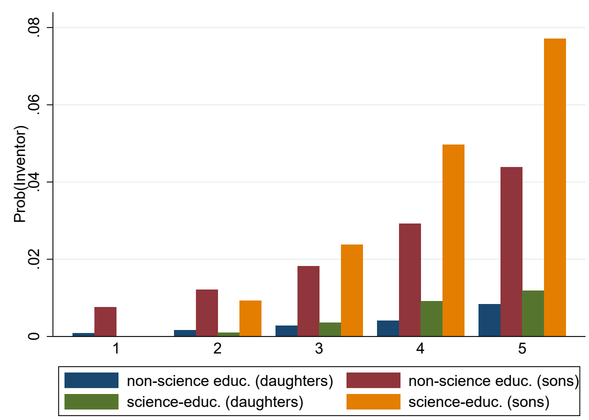


Figure 3: Prob($Inventor$) and parental education

Mothers

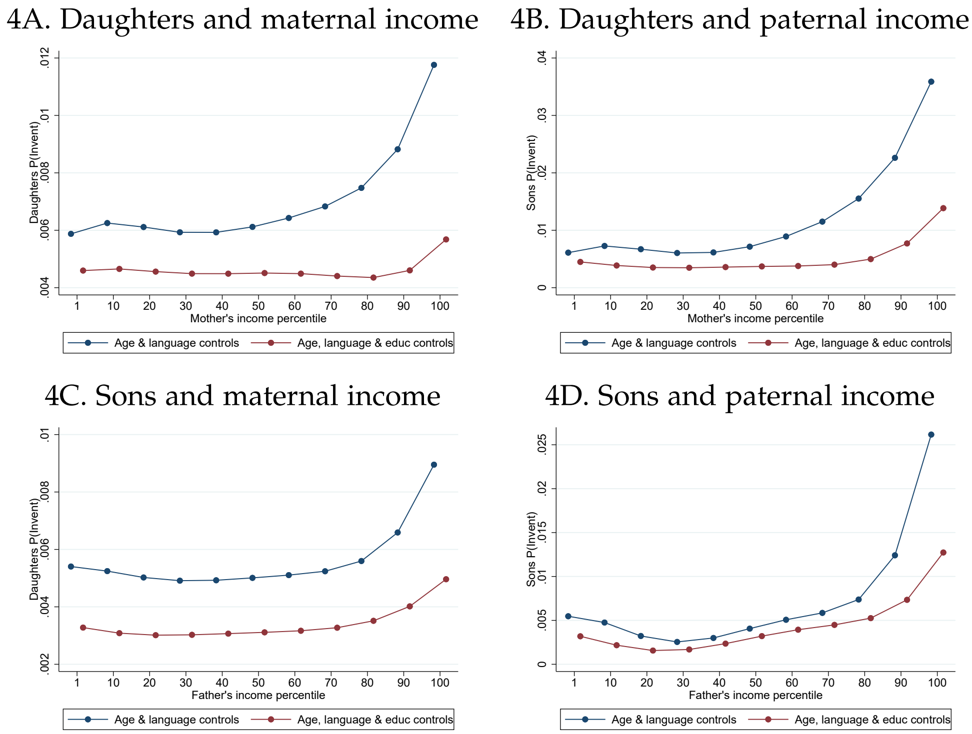


Fathers



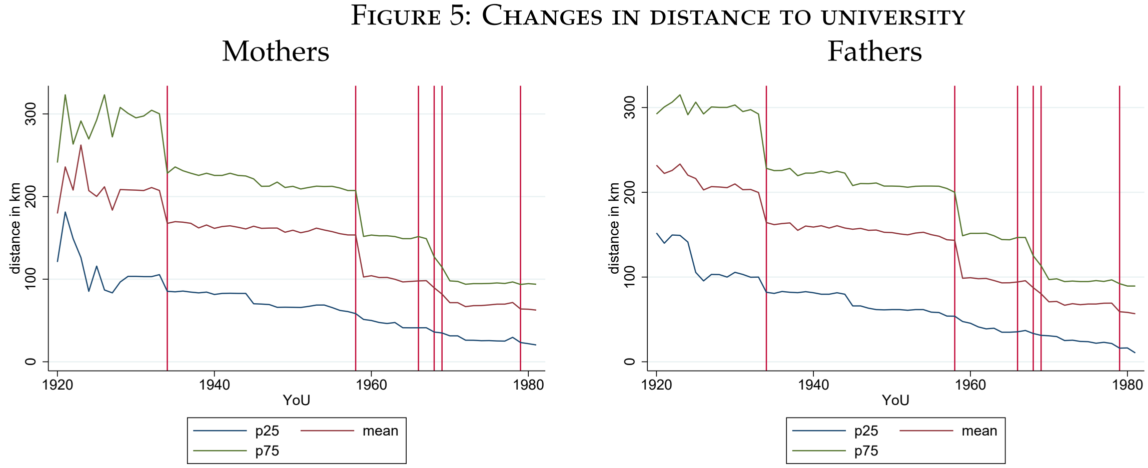
Notes: 1 = base education; 2 = secondary education; 3 = BSc; 4 = MSc; 5 = PhD. Non-science and science refer to the field of education of the parent.

Figure 4: Parental income and Prob(*Inventor*): regression-based relationships



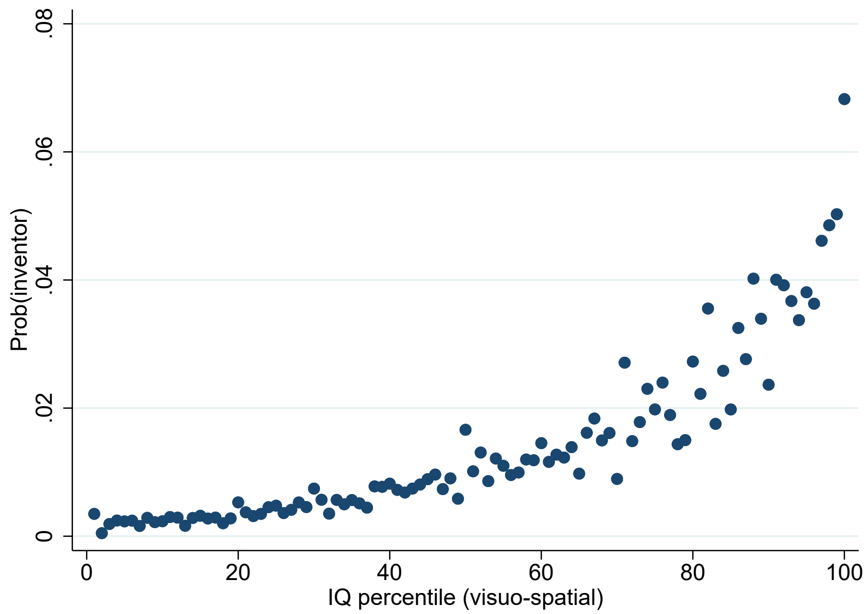
Notes: All regressions include full set of maternal and paternal year-of-birth dummies and a dummy for mother-tongue not being Finnish.

Figure 5: Changes in distance to university



Notes: YoU = **Y**ear **o**f **U**niversity, i.e., year when parent turns 19. The red vertical bars denote those years when distance to university changes due to a new university being established.

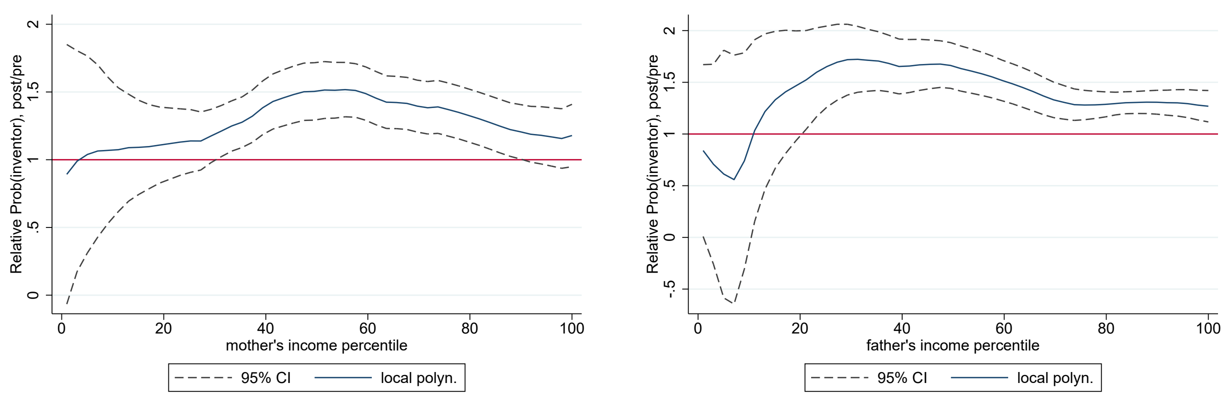
Figure 6: IQ percentile and Prob($Inventor$)



Notes: x-axis shows the percentiles of the visuospatial IQ distribution.

Figure 7: Relative $P(Inventor)$ 1966-1970 compared to 1956-1960

Maternal income   Paternal income



Notes: On the x-axis is the maternal (left hand figure) or paternal (right hand figure) income percentile. On the y-axis is the parental income percentile-specific ratio of $P(Inventor)$ calculated for the 1966-1970 cohorts and divided by $P(Inventor)$ calculated for the 1966-1960 cohorts. The curve shows a local polynomial estimate and the shaded are the 95% confidence interval. A ratio of one indicates no change in $P(Inventor)$, a ratio above one an increase and a ratio of less than one a decrease from 1956-1960 to 1966-1970.

Table 1: Estimation results

	(1)	(2)	(3)	(4)
Panel A. All Children				
	OLS	IV	IV	IV
$D(MScparents)$	0.0159***	0.0506***	0.0328***	0.0327***
	(0.00132)	(0.0110)	(0.009)	(0.0049)
F	-	55.73	140.73	508.87
Nobs	1 450 789			
Dep. var. mean	0.0067			
Panel B. Daughters				
$D(MScparents)$	0.0049***	0.0100	0.0203**	0.0160***
	(0.0005)	(0.0085)	(0.0086)	(0.0034)
F	-	47.25	75.87	326.80
Nobs	709 117			
Dep. var. mean	0.0016			
Panel C. Sons				
$D(MScparents)$	0.0261***	0.0866***	0.0430**	0.0487***
	(0.0023)	(0.0193)	(0.0205)	(0.0092)
F	-	35.95	94.12	264.76
Nobs	741 671			
Dep. var. mean	0.0118			
Instruments				
Maternal dist.	NO	YES	NO	YES
Paternal dist	NO	NO	YES	YES

Standard errors in parentheses are clustered at the year-of-birth level. Instrument is the propensity score estimated using LPM and a 3^{rd} order polynomial of the (logs) of the parental distances marked YES in the two last rows of the table. All specifications include a full set of maternal and paternal year-of-birth dummies, a dummy for mother tongue not being Finnish and the municipal controls explained in the text. F is the value of an F-test of all the instruments in the regression of the at least one MSc parent-dummy on instruments and controls. Dep. var. mean is the mean of the dependent variable for the sample in question (all, daughthers, sons), i.e., the sample probability of inventing.

Table 2: ATE, ATT and ATUT results

	(1)	(2)	(3)
	All	Daughters	Sons
LATE	0.0236***	0.0092***	0.0369***
	(0.0018)	(0.0014)	(0.0031)
ATE	0.0153***	0.0066***	0.0240***
	(0.0032)	(0.0027)	(0.0061)
ATT	0.0242***	0.0091***	0.0378***
	(0.0018)	(0.0027)	(0.0031)
ATUT	0.0147***	0.0064***	0.0231***
	(0.0034)	(0.0029)	(0.0065)
Nobs	1 450 789	709 117	741 671
Obs. het. p-value	0.0000	0.0938	0.0002
Unobs. het. p-value	0.0013	0.2548	0.0061

Boostrapped standard errors (100 rounds) in parentheses. Estimates are based on a parametric Roy model using a probit specification for the propensity score. All specifications include a full set of maternal and paternal year-of-birth dummies , a dummy for mother tongue not being Finnish and the municipal controls explained in the text. Similar to Column (3) in Table 1, a 3^{rd} order polynomial of the (logs of) parental distances to university is used as the vector of instruments.

Table 3: Estimation results using the IQ subsample

	Panel A. No IQ variables			
	(1)	(2)	(3)	(4)
	OLS	IV	IV	IV
$D(MScparents)$	0.0294***	0.0746***	0.0572**	0.0463***
	(0.0027)	(0.0266)	(0.0218)	(0.0115)
F	-	34.74	51.28	258.51
	Panel B. IQ variables			
$D(MScparents)$	0.0228***	0.0550*	0.0454*	0.0291**
	(0.0022)	(0.0274)	(0.0233)	(0.0121)
F	-	31.25	55.21	262.18
F_{IQ}		217.28	219.87	210.87
Nobs	421 729			
Maternal dist.	NO	YES	NO	YES
Paternal dist	NO	NO	YES	YES

The specifications reported in the upper panel are the same as in Column (3) of Table 1; the specifications in the lower panel include a 4^{th} order polynomial of visuospatial IQ as controls (in both the 1^{st} and 2^{nd} stage). Standard errors in parentheses are clustered at year-of-birth level. Instrument is the propensity score estimated using LPM and a 3^{rd} order polynomial of the (logs) of the parental distances marked YES in the two last rows of the table. All specifications include a full set of maternal and paternal year-of-birth dummies, a dummy for mother tongue not being Finnish and the municipal controls explained in the text. F is the value of an F-test of all the instruments in the regression of the measure of parental education on instruments and controls. F_{IQ} is the value of an F-test of the joint significance of the IQ variables.

Table 4: Estimation results using pre- and post- comprehensive school samples

	Panel A. Pre, 1956-1960			
	(1)	(2)	(3)	(4)
	OLS	IV	IV	IV
$D(MScparents)$	0.0151***	0.100***	0.0256	0.0445***
	(0.0017)	(0.0316)	(0.0507)	(0.0123)
F	-	43.16	25.19	65.44
Nobs	234 685			
	Panel B: Post, 1966 - 1970			
$D(MScparents)$	0.0221***	0.0116	0.0141	0.0333**
	(0.0016)	(0.0263)	(0.0323)	(0.0125)
F	-	54.60	44.40	79.15
Nobs	203 923			
Maternal dist.	NO	YES	NO	YES
Paternal dist	NO	NO	YES	YES

The specifications reported in are the same as in Table 1. Standard errors in parentheses are clustered at year-of-birth level. Instrument is the propensity score estimated using LPM and a 3^{rd} order polynomial of the (logs) of the parental distances marked YES in the two last rows of the table. All specifications include a full set of maternal and paternal year-of-birth dummies, a dummy for mother tongue not being Finnish and the municipal controls explained in the text. F is the value of an F-test of all the instruments in the regression of the measure of parental education on instruments and controls.

Appendix

A-1 Data description

In this part of the Appendix, we report descriptive statistics for our data and information on the establishment of new universities and how they are associated with parents' distance to the nearest university and other variables.

A-1.1 Descriptive statistics

We first show the descriptive statistics for the whole data, daughters, and sons. For each of these, we condition on inventor status, and provide t-tests for the difference between inventors and non-inventors.

Other variable names and subscripts ought to be self-explanatory, but i) $Educ\_levx$ is the educational level, with $x \in \{1, ..., 5\}$ referring respectively to base ($x = 1$), secondary (2), lower tertiary (BSc, 3), MSc (4) and PhD (5) level of education. These are measured separately for STEM and non-STEM educations, except base education, for which no such distinction exists because it refers to completing just an elementary school, or equivalent. We measure education at age 35. ii) $Income\_px_{muni,p}$, with $x \in \{., 50, 90\}$, refer to the average income percentile the birth-cohort of parent p ($p \in \{mo, fa\}$) in question in his/her birth municipality and the fractions of individuals in the birth cohort of the parent in the parent's birth municipality, that are in at least the x^{th} income percentile of the national cohort, measured at age 35. For the two latter measures, we exclude the parent in question. Thus for example the mean income percentile of the maternal birth-cohort in the mother's birth municipalities is 48.58; and on average 50% (10%) of the maternal municipal birth cohort has above median (above 90^{th} percentile) income at age 35 (the percentiles are calculated across the whole cohort, i.e., men and women).

Table A-1: Descriptive statistics

	All individuals									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Variable	All			Non-inventors			Inventors			
	Mean	p50	sd	Mean	p50	sd	Mean	p50	sd	p-value
inventor	0.0068	0.0000	0.0820	0.0000	0.0000	0.0000	1.0000	1.0000	0.0000	
patent_count	0.0214	0.0000	0.5566	0.0000	0.0000	0.0000	3.1648	1.0000	5.9821	0.00
patent_count$_{cens}$	0.0175	0.0000	0.2972	0.0000	0.0000	0.0000	2.5778	1.0000	2.5364	0.00
citations	0.0291	0.0000	1.0539	0.0000	0.0000	0.0000	4.2882	1.0000	12.0699	0.00
citations_cens	0.0183	0.0000	0.3160	0.0000	0.0000	0.0000	2.7069	1.0000	2.7313	0.00
MSc$_{atleastone}$	0.0593	0.0000	0.2361	0.0583	0.0000	0.2344	0.1948	0.0000	0.3961	0.00
#MSc	0.07	0.00	0.31	0.07	0.00	0.30	0.25	0.00	0.55	0.00
YoB	1967.47	1967.00	8.34	1967.46	1967.00	8.34	1967.62	1968.00	7.46	0.07
Finnish	0.95	1.00	0.22	0.95	1.00	0.22	0.95	1.00	0.22	0.43
Income	51.79	52.00	28.51	51.57	52.00	28.42	83.84	93.00	22.25	0.00
MSc	0.13	0.00	0.34	0.13	0.00	0.34	0.63	1.00	0.48	0.00
Dist_uni	54.85	52.70	49.86	54.91	52.70	49.88	46.27	41.24	46.78	0.00
YoB$_{mo}$	1940.19	1941.00	10.59	1940.19	1941.00	10.60	1940.05	1941.00	9.36	0.19
Income$_{mo}$	30.76	28.00	25.38	30.72	28.00	25.35	36.78	33.00	29.19	0.00
MSc$_{mo}$	0.02	0.00	0.14	0.02	0.00	0.14	0.08	0.00	0.27	0.00
Dist_uni$_{mo}$	120.40	96.74	104.59	120.46	97.01	104.64	110.96	94.64	97.80	0.00
MSc$_{muni,mo}$	0.02	0.01	0.02	0.02	0.01	0.02	0.02	0.02	0.03	0.00
Income_p$_{muni,mo}$	48.58	48.43	6.08	48.57	48.43	6.08	49.53	49.29	6.03	0.00
Income_p50$_{muni,mo}$	0.50	0.49	0.09	0.50	0.49	0.09	0.51	0.51	0.09	0.00
Income_p90$_{muni,mo}$	0.10	0.09	0.05	0.10	0.09	0.05	0.11	0.10	0.06	0.00
Educ_lev1$_{mo}$	0.59	1.00	0.49	0.59	1.00	0.49	0.40	0.00	0.49	0.00
Educ_lev2$_{nonstem,mo}$	0.20	0.00	0.40	0.20	0.00	0.40	0.21	0.00	0.41	0.02
Educ_lev2$_{stem,mo}$	0.06	0.00	0.24	0.06	0.00	0.24	0.05	0.00	0.22	0.01
Educ_lev3$_{nonstem,mo}$	0.37	0.00	0.98	0.36	0.00	0.98	0.74	0.00	1.30	0.00
Educ_lev3$_{stem,mo}$	0.00	0.00	0.06	0.00	0.00	0.06	0.01	0.00	0.10	0.00
Educ_lev4$_{nonstem,mo}$	0.02	0.00	0.13	0.02	0.00	0.13	0.06	0.00	0.23	0.00
Educ_lev4$_{stem,mo}$	0.00	0.00	0.06	0.00	0.00	0.06	0.02	0.00	0.13	0.00
Educ_lev5$_{nonstem,mo}$	0.00	0.00	0.02	0.00	0.00	0.02	0.00	0.00	0.04	0.00
Educ_lev5$_{stem,mo}$	0.00	0.00	0.02	0.00	0.00	0.02	0.00	0.00	0.04	0.00
YoB$_{fa}$	1937.67	1939.00	11.17	1937.67	1939.00	11.18	1937.90	1939.00	9.98	0.05
Income$_{fa}$	66.07	71.00	25.22	66.01	71.00	25.22	75.95	84.00	23.92	0.00
MSc$_{fa}$	0.05	0.00	0.22	0.05	0.00	0.22	0.18	0.00	0.38	0.00
Dist_uni$_{fa}$	123.35	100.73	106.08	123.41	100.89	106.11	114.32	95.29	99.93	0.00
MSc$_{muni,fa}$	0.05	0.04	0.05	0.05	0.04	0.05	0.06	0.05	0.05	0.00
Income_p$_{muni,fa}$	48.83	48.69	6.36	48.82	48.67	6.36	49.82	49.65	6.28	0.00
Income_p50$_{muni,fa}$	0.50	0.50	0.09	0.50	0.50	0.09	0.51	0.51	0.09	0.00
Income_p90$_{muni,fa}$	0.11	0.10	0.06	0.11	0.10	0.06	0.11	0.10	0.06	0.00
Educ_lev1$_{fa}$	0.58	1.00	0.49	0.58	1.00	0.49	0.37	0.00	0.48	0.00
Educ_lev2$_{nonstem,fa}$	0.08	0.00	0.28	0.08	0.00	0.28	0.09	0.00	0.28	0.34
Educ_lev2$_{stem,fa}$	0.15	0.00	0.36	0.15	0.00	0.36	0.12	0.00	0.32	0.00
Educ_lev3$_{nonstem,fa}$	0.19	0.00	0.74	0.19	0.00	0.74	0.31	0.00	0.91	0.00
Educ_lev3$_{stem,fa}$	0.07	0.00	0.26	0.07	0.00	0.26	0.15	0.00	0.36	0.00
Educ_lev4$_{nonstem,fa}$	0.03	0.00	0.17	0.03	0.00	0.17	0.07	0.00	0.26	0.00
Educ_lev4$_{stem,fa}$	0.02	0.00	0.13	0.02	0.00	0.13	0.08	0.00	0.27	0.00
Educ_lev5$_{nonstem,fa}$	0.00	0.00	0.05	0.00	0.00	0.05	0.01	0.00	0.11	0.00
Educ_lev5$_{stem,fa}$	0.00	0.00	0.05	0.00	0.00	0.05	0.02	0.00	0.13	0.00

Table A-2: Descriptive statistics

	Daughters									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Variable	All			Non-inventors			Inventors			
	Mean	p50	sd	Mean	p50	sd	Mean	p50	sd	p-value
inventor	0.0016	0.0000	0.0396	0.0000	0.0000	0.0000	1.0000	1.0000	0.0000	
patent_count	0.0038	0.0000	0.1702	0.0000	0.0000	0.0000	2.4170	1.0000	3.5525	0.00
patent_count$_{cens}$	0.0035	0.0000	0.1197	0.0000	0.0000	0.0000	2.1973	1.0000	2.0759	0.00
citations	0.0050	0.0000	0.3237	0.0000	0.0000	0.0000	3.1578	1.0000	7.5374	0.00
citations_cens	0.0038	0.0000	0.1370	0.0000	0.0000	0.0000	2.4251	1.0000	2.4682	0.00
MSc$_{atleastone}$	0.0586	0.0000	0.2348	0.0583	0.0000	0.2343	0.2377	0.0000	0.4258	0.00
#MSc	0.0720	0.0000	0.3063	0.0717	0.0000	0.3054	0.3166	0.0000	0.6120	0.00
YoB	1967.47	1967.00	8.33	1967.47	1967.00	8.33	1968.32	1969.00	7.52	0.00
Finnish	0.95	1.00	0.22	0.95	1.00	0.22	0.95	1.00	0.21	0.52
Income	43.61	41.00	25.48	43.56	41.00	25.45	73.44	84.00	25.55	0.00
MSc	0.15	0.00	0.36	0.15	0.00	0.35	0.83	1.00	0.38	0.00
Dist_uni	54.97	52.70	49.84	54.99	52.70	49.84	44.45	36.05	47.60	0.00
YoB$_{mo}$	1940.20	1941.00	10.60	1940.20	1941.00	10.60	1940.61	1942.00	8.98	0.19
Income$_{mo}$	30.81	28.00	25.32	30.80	28.00	25.32	38.43	35.00	29.74	0.00
MSc$_{mo}$	0.02	0.00	0.14	0.02	0.00	0.14	0.11	0.00	0.31	0.00
Dist_uni$_{mo}$	120.46	97.01	104.55	120.48	97.01	104.55	107.22	85.42	101.00	0.00
MSc$_{muni,mo}$	0.02	0.01	0.02	0.02	0.01	0.02	0.03	0.02	0.03	0.00
Income_p$_{muni,mo}$	48.57	48.43	6.08	48.57	48.43	6.08	49.94	49.63	5.94	0.00
Income_p50$_{muni,mo}$	0.50	0.49	0.09	0.50	0.49	0.09	0.51	0.51	0.09	0.00
Income_p90$_{muni,mo}$	0.10	0.09	0.05	0.10	0.09	0.05	0.11	0.10	0.06	0.00
Educ_lev1$_{mo}$	0.60	1.00	0.49	0.60	1.00	0.49	0.35	0.00	0.48	0.00
Educ_lev2$_{nonstem,mo}$	0.20	0.00	0.40	0.20	0.00	0.40	0.19	0.00	0.39	0.54
Educ_lev2$_{stem,mo}$	0.06	0.00	0.24	0.06	0.00	0.24	0.05	0.00	0.22	0.19
Educ_lev3$_{nonstem,mo}$	0.36	0.00	0.98	0.36	0.00	0.98	0.86	0.00	1.36	0.00
Educ_lev3$_{stem,mo}$	0.00	0.00	0.06	0.00	0.00	0.06	0.01	0.00	0.09	0.00
Educ_lev4$_{nonstem,mo}$	0.02	0.00	0.13	0.02	0.00	0.13	0.08	0.00	0.26	0.00
Educ_lev4$_{stem,mo}$	0.00	0.00	0.06	0.00	0.00	0.06	0.03	0.00	0.18	0.00
Educ_lev5$_{nonstem,mo}$	0.00	0.00	0.02	0.00	0.00	0.02	0.00	0.00	0.04	0.04
Educ_lev5$_{stem,mo}$	0.00	0.00	0.02	0.00	0.00	0.02	0.00	0.00	0.03	0.16
YoB$_{fa}$	1937.67	1939.00	11.18	1937.67	1939.00	11.18	1938.73	1940.00	9.63	0.00
Income$_{fa}$	66.00	71.00	25.23	65.98	71.00	25.22	78.23	87.00	23.62	0.00
MSc$_{fa}$	0.05	0.00	0.22	0.05	0.00	0.22	0.21	0.00	0.40	0.00
Dist_uni$_{fa}$	123.51	100.89	106.08	123.53	100.89	106.08	110.60	92.16	103.37	0.00
MSc$_{muni,fa}$	0.05	0.04	0.05	0.05	0.04	0.05	0.06	0.05	0.05	0.00
Income_p$_{muni,fa}$	48.81	48.65	6.36	48.81	48.65	6.36	50.20	50.13	6.33	0.00
Income_p50$_{muni,fa}$	0.50	0.50	0.09	0.50	0.50	0.09	0.52	0.52	0.09	0.00
Income_p90$_{muni,fa}$	0.11	0.10	0.06	0.11	0.10	0.06	0.12	0.11	0.06	0.00
Educ_lev1$_{fa}$	0.58	1.00	0.49	0.58	1.00	0.49	0.33	0.00	0.47	0.00
Educ_lev2$_{nonstem,fa}$	0.08	0.00	0.27	0.08	0.00	0.27	0.08	0.00	0.28	0.85
Educ_lev2$_{stem,fa}$	0.15	0.00	0.36	0.15	0.00	0.36	0.10	0.00	0.30	0.00
Educ_lev3$_{nonstem,fa}$	0.19	0.00	0.74	0.19	0.00	0.74	0.34	0.00	0.96	0.00
Educ_lev3$_{stem,fa}$	0.07	0.00	0.26	0.07	0.00	0.26	0.17	0.00	0.37	0.00
Educ_lev4$_{nonstem,fa}$	0.03	0.00	0.17	0.03	0.00	0.17	0.07	0.00	0.26	0.00
Educ_lev4$_{stem,fa}$	0.02	0.00	0.13	0.02	0.00	0.13	0.10	0.00	0.30	0.00
Educ_lev5$_{nonstem,fa}$	0.00	0.00	0.05	0.00	0.00	0.05	0.02	0.00	0.12	0.00
Educ_lev5$_{stem,fa}$	0.00	0.00	0.05	0.00	0.00	0.05	0.02	0.00	0.13	0.00

TABLE A-3: DESCRIPTIVE STATISTICS

	Sons									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Variable	All			Non-inventors			Inventors			
	Mean	p50	sd	Mean	p50	sd	Mean	p50	sd	p-value
inventor	0.0118	0.0000	0.1078	0.0000	0.0000	0.0000	1.0000	1.0000	0.0000	
patent_count	0.0383	0.0000	0.7601	0.0000	0.0000	0.0000	3.2603	1.0000	6.2182	0.00
patent_count$_{cens}$	0.0309	0.0000	0.3983	0.0000	0.0000	0.0000	2.6264	1.0000	2.5854	0.00
citations	0.0521	0.0000	1.4393	0.0000	0.0000	0.0000	4.4326	1.0000	12.5244	0.00
citations_cens	0.0322	0.0000	0.4207	0.0000	0.0000	0.0000	2.7429	1.0000	2.7612	0.00
MSc$_{atleastone}$	0.0599	0.0000	0.2373	0.0584	0.0000	0.2344	0.1894	0.0000	0.3918	0.00
#MSc	0.0735	0.0000	0.3088	0.0715	0.0000	0.3043	0.2460	0.0000	0.5465	0.00
YoB	1967.46	1967.00	8.34	1967.46	1967.00	8.35	1967.53	1968.00	7.44	0.46
Finnish	0.95	1.00	0.22	0.95	1.00	0.22	0.95	1.00	0.22	0.37
Income	59.60	65.00	29.06	59.30	65.00	29.00	85.17	93.00	21.43	0.00
MSc	0.12	0.00	0.32	0.11	0.00	0.31	0.61	1.00	0.49	0.00
Dist_uni	54.74	52.65	49.89	54.84	52.65	49.92	46.50	41.24	46.67	0.00
YoB$_{mo}$	1940.19	1941.00	10.58	1940.19	1941.00	10.59	1939.98	1941.00	9.40	0.06
Income$_{mo}$	30.72	28.00	25.43	30.65	28.00	25.38	36.57	33.00	29.11	0.00
MSc$_{mo}$	0.02	0.00	0.15	0.02	0.00	0.14	0.07	0.00	0.26	0.00
Dist_uni$_{mo}$	120.34	96.74	104.64	120.44	96.74	104.72	111.43	94.83	97.38	0.00
MSc$_{muni,mo}$	0.02	0.01	0.02	0.02	0.01	0.02	0.02	0.02	0.02	0.00
Income_p$_{muni,mo}$	48.58	48.44	6.09	48.57	48.43	6.09	49.48	49.23	6.04	0.00
Income_p50$_{muni,mo}$	0.50	0.49	0.09	0.50	0.49	0.09	0.51	0.51	0.09	0.00
Income_p90$_{muni,mo}$	0.10	0.09	0.05	0.10	0.09	0.05	0.11	0.10	0.06	0.00
Educ_lev1$_{mo}$	0.59	1.00	0.49	0.59	1.00	0.49	0.41	0.00	0.49	0.00
Educ_lev2$_{nonstem,mo}$	0.20	0.00	0.40	0.20	0.00	0.40	0.21	0.00	0.41	0.01
Educ_lev2$_{stem,mo}$	0.06	0.00	0.24	0.06	0.00	0.24	0.05	0.00	0.23	0.02
Educ_lev3$_{nonstem,mo}$	0.37	0.00	0.98	0.36	0.00	0.98	0.73	0.00	1.29	0.00
Educ_lev3$_{stem,mo}$	0.00	0.00	0.06	0.00	0.00	0.06	0.01	0.00	0.10	0.00
Educ_lev4$_{nonstem,mo}$	0.02	0.00	0.13	0.02	0.00	0.13	0.05	0.00	0.23	0.00
Educ_lev4$_{stem,mo}$	0.00	0.00	0.06	0.00	0.00	0.06	0.02	0.00	0.12	0.00
Educ_lev5$_{nonstem,mo}$	0.00	0.00	0.02	0.00	0.00	0.02	0.00	0.00	0.04	0.00
Educ_lev5$_{stem,mo}$	0.00	0.00	0.02	0.00	0.00	0.02	0.00	0.00	0.04	0.00
YoB$_{fa}$	1937.68	1939.00	11.16	1937.68	1939.00	11.17	1937.79	1939.00	10.02	0.35
Income$_{fa}$	66.15	71.00	25.22	66.04	71.00	25.22	75.66	83.00	23.95	0.00
MSc$_{fa}$	0.05	0.00	0.22	0.05	0.00	0.22	0.17	0.00	0.38	0.00
Dist_uni$_{fa}$	123.19	100.72	106.07	123.29	100.73	106.15	114.79	95.64	99.48	0.00
MSc$_{muni,fa}$	0.05	0.04	0.05	0.05	0.04	0.05	0.06	0.04	0.05	0.00
Income_p$_{muni,fa}$	48.85	48.71	6.37	48.84	48.70	6.37	49.77	49.61	6.28	0.00
Income_p50$_{muni,fa}$	0.50	0.50	0.09	0.50	0.50	0.09	0.51	0.51	0.09	0.00
Income_p90$_{muni,fa}$	0.11	0.10	0.06	0.11	0.10	0.06	0.11	0.10	0.06	0.00
Educ_lev1$_{fa}$	0.58	1.00	0.49	0.58	1.00	0.49	0.37	0.00	0.48	0.00
Educ_lev2$_{nonstem,fa}$	0.08	0.00	0.28	0.08	0.00	0.28	0.09	0.00	0.28	0.44
Educ_lev2$_{stem,fa}$	0.15	0.00	0.36	0.15	0.00	0.36	0.12	0.00	0.32	0.00
Educ_lev3$_{nonstem,fa}$	0.19	0.00	0.74	0.19	0.00	0.74	0.30	0.00	0.90	0.00
Educ_lev3$_{stem,fa}$	0.07	0.00	0.26	0.07	0.00	0.26	0.15	0.00	0.35	0.00
Educ_lev4$_{nonstem,fa}$	0.03	0.00	0.17	0.03	0.00	0.17	0.07	0.00	0.26	0.00
Educ_lev4$_{stem,fa}$	0.02	0.00	0.13	0.02	0.00	0.13	0.08	0.00	0.26	0.00
Educ_lev5$_{nonstem,fa}$	0.00	0.00	0.05	0.00	0.00	0.05	0.01	0.00	0.10	0.00
Educ_lev5$_{stem,fa}$	0.00	0.00	0.05	0.00	0.00	0.05	0.02	0.00	0.13	0.00

A-1.2 Establishment of new universities

Establishment of new universities and distance to university. The following table reports the years, locations and year of establishment of universities that were established by the end of our observation period in Finland. There have been several changes in the geography and organization of Finnish universities since. Notice that some the establishments listed in the table had no effect on distance to nearest university as there already was a university in the same location (e.g. U. of Turku, Technical U. of Helsinki). We display the locations (and year of establishment and distance to Helsinki) of those universities the establishment of which changed distance to university to at least some municipality in Figure A-1.

Table A-4: Establishment of universities in Finland up to end of 1970s

University	Year of establishment	Location
U of Helsinki	1640	Originally in Turku; in Helsinki since 1827
Technical U. of Helsinki	1849	Originally in Helsinki; in neighboring Espoo since 1950s
Hanken School of Economics	1909	Helsinki
Åbo Akademi	1918	Turku
U. of Turku	1920	Turku
Helsinki School of Economics	1911	Helsinki
U. of Jyväskylä	1934	Jyväskylä
U. of Tampere	1925	Tampere
U. of Oulu	1959	Oulu
Technical U. of Tampere	1965	Tampere
U. of Vaasa	1968	Vaasa
Lappeenranta U. of Technology	1969	Lappeenranta
U. of Joensuu	1970	Joensuu
U. of Kuopio	1972	Kuopio
U. of Lapland	1979	Rovaniemi

Notes: Technical U. of Tampere started as an off-shoot of Helsinki U. of Technology and was formally established in 1972. Helsinki U. of Technology moved its activities to Espoo over several years. Neither of these changes affects our distance measures.

The variation in distance to university is cross-sectional on the one hand, and comes from the opening of new universities on the other hand. To get an idea of what fraction of our data is affected by the opening of new universities we display in Figure A-2 the distribution of parents by their "Year of University" (YoU), i.e., the year they turn 19. In the histograms, the red vertical lines mark those years when a new university that changes distances to university is established. As can be seen, a significant fraction of both mothers and fathers in our data turn 19 during the time that many universities (that change distance, i.e., are opened in locations that did not have a university previously) are established.

Relation between distance to university and other variables. We analyze the relationship between our instruments and the inventor - dummy; the relationship between our instruments and a dummy for the mother or father obtaining an MSc; and the relationship between distance to university and our controls for the quality of the birth-municipality: the number of children born in the parental birth municipality in the year of maternal / paternal birth; the fraction of the parental municipal birth cohort that have obtained an MSc by age 35; the fraction of the parental municipal birth cohort that had above median income at age 35, where the median is calculated over the whole national birth cohort; and similarly, the fraction of the parental municipal birth cohort that had an income in the top percentile of the national cohort at age 35. We also include (visuospatial) IQ into the analysis.

We first display the correlations, and then graphically the relation between (log) distance and the above listed variables. The graphs are based on a projecting the variable in question on a fourth order polynomical of log distance (with no control variables). The median distance is round 100km (4.5 in logs) for both parents; the 90^{th} percentile round 250km (5.5 in logs).

The correlation (see Table A-5) between the offspring inventor -dummy and maternal and paternal distance to university are -0.01 and -0.02 with only the latter significant at standard levels. The dummy for having at least one MSc parent ($D(MSc\ parent)$) exhibits correlations of -0.04 with maternal and -0.01 with paternal distances to university; only the first is statistically significant. The correlation between the mother or father having an MSc and distance to university are -0.02 and -0.01; both are highly statistically significant. The size of the birth cohort is negatively correlated with distance. The share of the cohort (excluding the parent in question) of the municipal birth - cohort that obtain an MSc, have above median income within the national cohort or are in the top decile of the national

Figure A-1: Map of Finnish university establishments 1918 - 1979

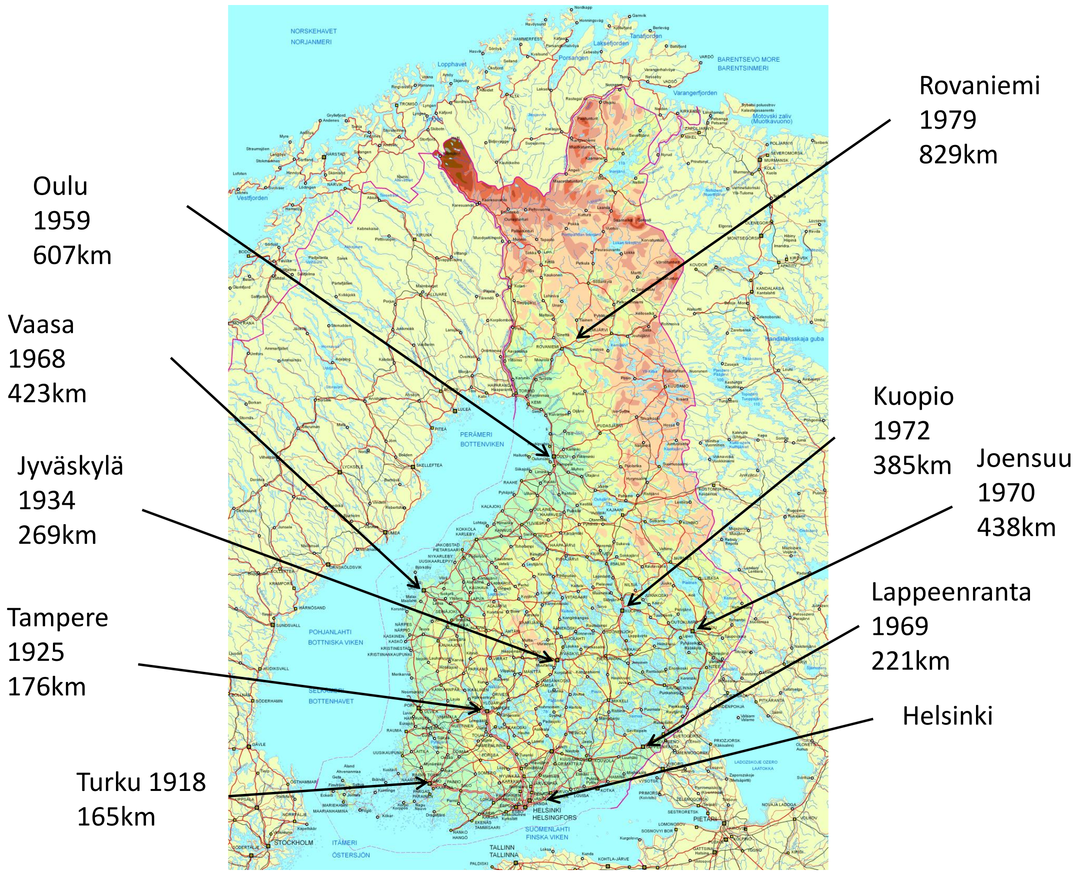
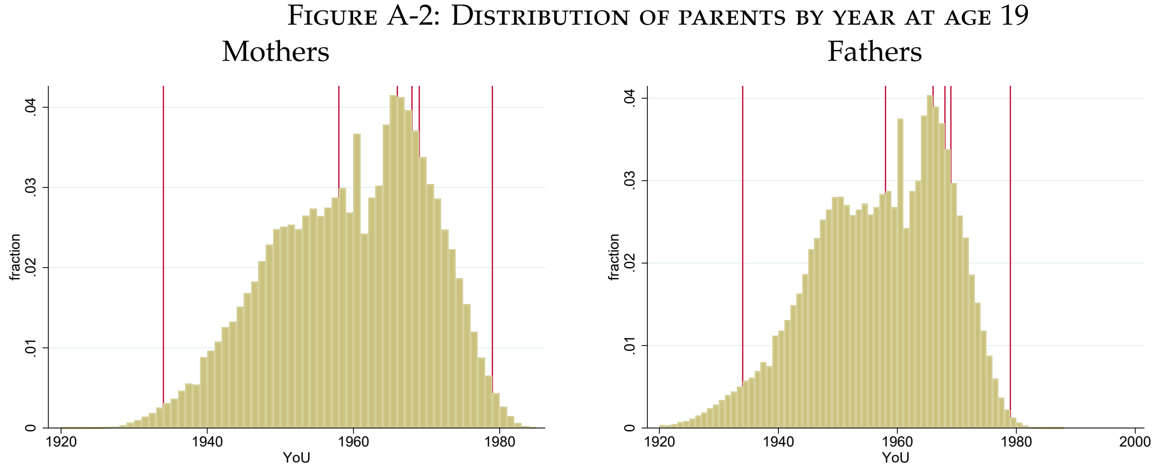


Figure A-2: Distribution of parents by year at age 19



Notes: YoU = **Y**ear **o**f **U**niversity, i.e., year when parent turns 19. The red vertical bars denote those years when distance to university changes due to a new university being established.

cohort (all measured at age 35) are all negatively and statistically significantly correlated with distance. Finally, reminiscent of Carneiro and Heckman (2002), IQ of the sons is also negatively correlated with distance. Notice though that the absolute value of the correlation coefficient is smaller than those of the municipal characteristics we control for.

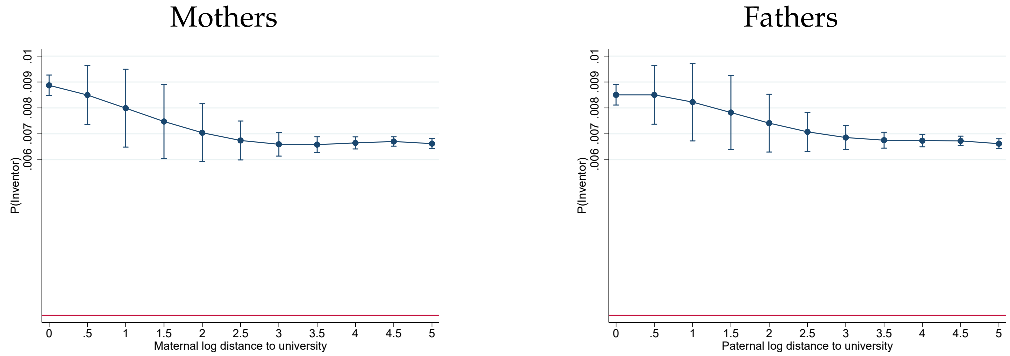
We visualize and provide a more precise picture of the correlations between the aforementioned variables and (log) distance to the nearest university by regressing each variable on a 4^{th} order polynomial of (the log of 1+) parental distance to nearest university. The relation between the probability to invent and parental distance to university is decreasing in parental distance before flattening out (Figure A-6). The relation between $D(MSc\ parent)$ and distance is almost monotonically decreasing for both parental distances to university; the relation is initially steep before flattening (Figure A-4). The relation with the parent him or herself having an MSc, and distance to university displays a similar pattern (Figure A-5). Figure A-6 shows a similar steeply declining relation. Figures A-7, A-8 and A-9 depict the relation between the characteristics of the birth cohort in a given municipality. The farther from a university, the lower the share of the cohort that obtain an MSc; the share with above median income is non-monotone in distance, first increasing and then decreasing. The share obtaining top income (=top decile of the national cohort) is not quite monotonic either, but in general more strongly decreasing in distance than that for above median income. In contrast, the mean IQ has its maximum of slightly above 102 at zero distance to university, i.e., in university towns. The mean then declines quickly, reaching the sample mean (or at least becoming insignificantly different from it) at a very small distance (a few kilometers or 1 to 1.5 log kilometers). Some 12% of the individuals in our sample have parents born this close to a university. The relationship between IQ and parental distance to university is non-monotone, but after the initial decline the changes are small.

Table A-5: Distance correlations

Parent	$P(inventor)$	D(MSc parents)	MSc_p	Count	MSc_{cohort}	p50	p90	IQ
Maternal	-0.0094	-0.0616	-0.0382	-0.2482	-0.2634	-0.2313	-0.2058	-0.0640
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Paternal	-0.0094	-0.1095	-0.1021	-0.6140	-0.5459	-0.4161	-0.4486	-0.0662
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)

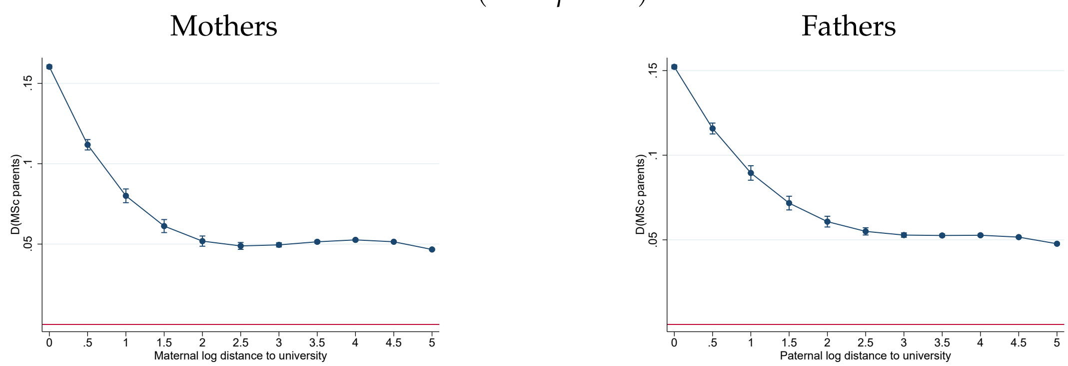
Notes: The reported numbers are the correlation coefficient between log of maternal or paternal distance (on rows) and the column variable in question and, in parentheses, its p-value. $P(inventor)$ is a dummy for the individual being an inventor; D(MSc parents) a dummy for at least one of the parents having an MSc; MSc_p is a dummy for parent of type p (depending on the row) having an MSc; Count is the size of the parental municipal birth cohort; MSc_{cohort} is the fraction of the parental municipal cohort having and MSc (excluding the parent in question) at age 35; and p50 and p90 are the fraction of the parental municipal cohort having above median or top decile income of the entire national cohort, measured at age 35. IQ is the visuospatial IQ of the individual, measured by FDF. The observation unit is a municipal birth-cohort otherwise, but an individual for $P(inventor)$ and IQ.

FIGURE A-3: $P(invent)$ AND DISTANCE TO UNIVERSITY



Notes: Figures based on a 4^{th} order polynomial of (log) distance. The observation unit is an individual.

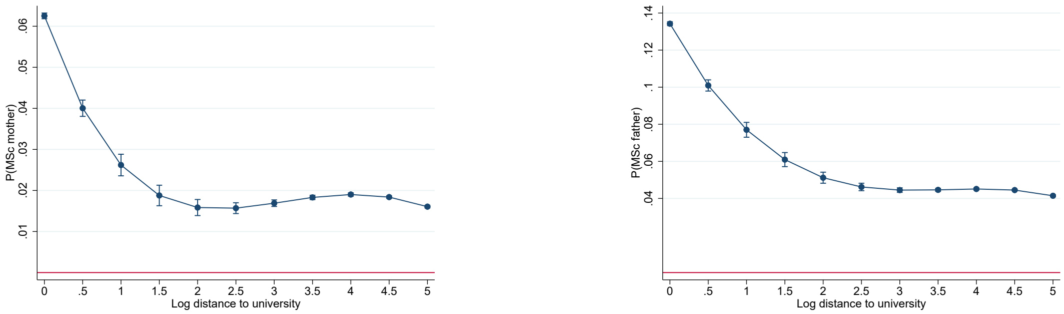
FIGURE A-4: $D(MSc\ parent)$ AND DISTANCE TO UNIVERSITY



Notes: Figures based on a 4^{th} order polynomial of (log) distance. The observation unit is a municipal birth-cohort.

Figure A-5: Prob(MSc_{parent}) and distance to university

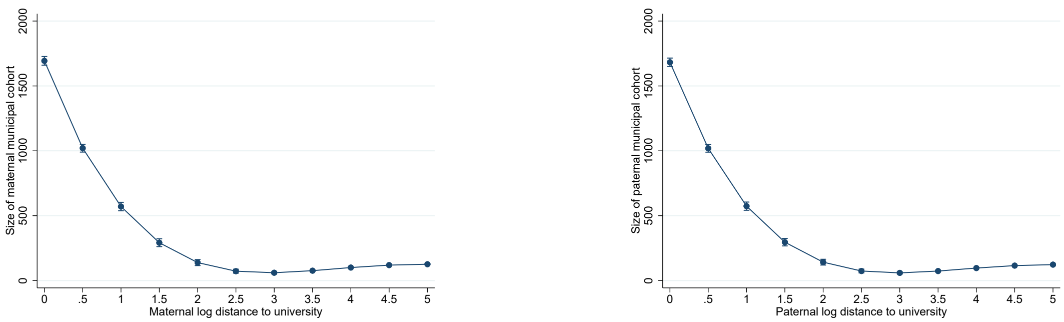
Mothers Fathers



Notes: Figures based on a 4^{th} order polynomial of (log) distance. The observation unit is a municipal birth-cohort.

Figure A-6: Size of municipal cohort and distance to university

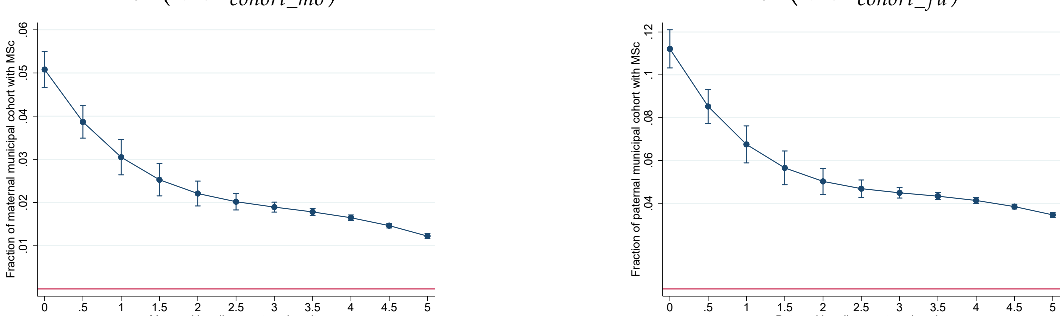
Mothers Fathers



Notes: Figures based on a 4^{th} order polynomial of (log) distance. The observation unit is a municipal birth-cohort.

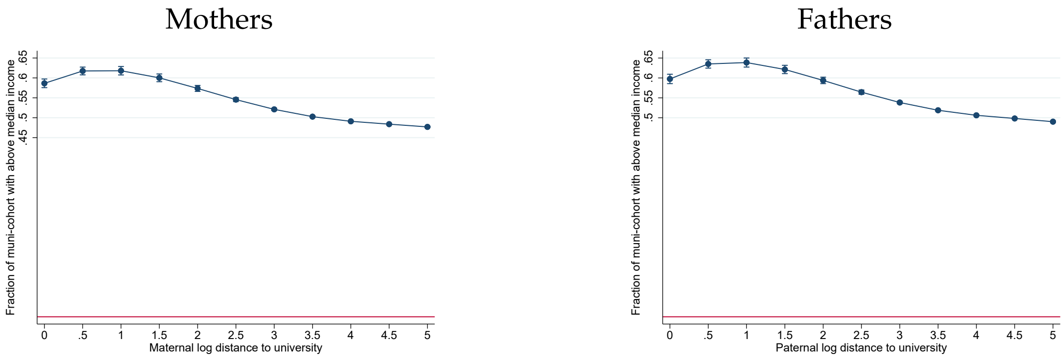
Figure A-7: Fraction of municipal cohort having an MSc and distance to university

Prob(MSc_{cohort_mo}) Prob(MSc_{cohort_fa})



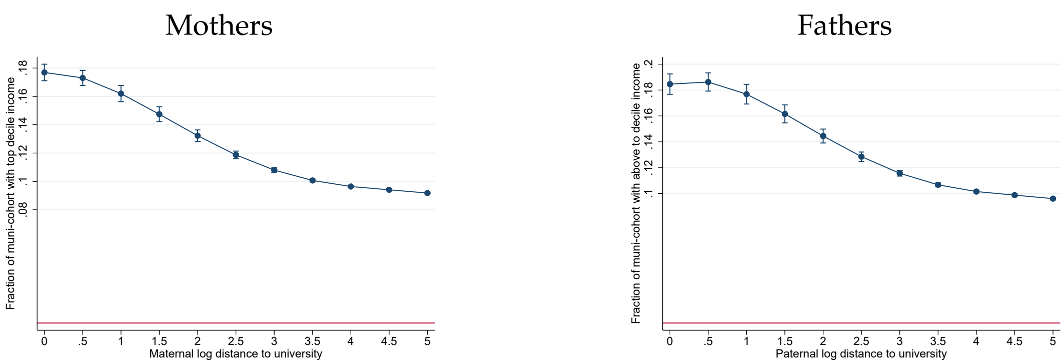
Notes: Figures based on a 4^{th} order polynomial of (log) distance. The observation unit is a municipal birth-cohort.

FIGURE A-8: ABOVE MEDIAN INCOME AND DISTANCE TO UNIVERSITY



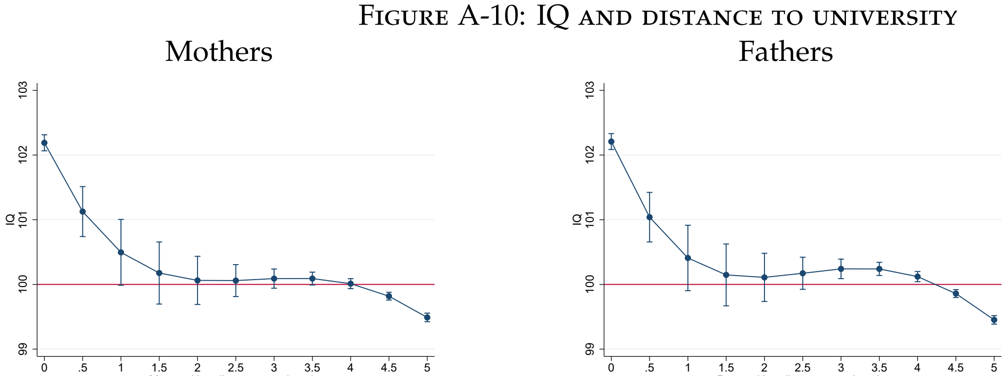
Notes: Figures based on a 4^{th} order polynomial of (log) distance. The observation unit is a municipal birth-cohort.

FIGURE A-9: TOP DECILE INCOME AND DISTANCE TO UNIVERSITY



Notes: Figures based on a 4^{th} order polynomial of (log) distance. The observation unit is a municipal birth-cohort.

FIGURE A-10: IQ AND DISTANCE TO UNIVERSITY



Notes: Figures based on a 4^{th} order polynomial of (log) distance. The observation unit is an individual.

A-2 First stage results

Here we show the first stage results for our main results. The dependent variable is the dummy for the individual having at least one parent with an MSc. The vector of instruments is a third order polynomial of (the logs of) 1) maternal distance to nearest university; 2) paternal distance to the nearest university; or 3) both parental distances to nearest university. In the last case we include the full set of interactions (of the powers of distance).

All specifications include our base controls, i.e., the full set of maternal and paternal year-of-birth dummies and a dummy for mother tongue not being Finnish. All specifications also include our full set of municipality controls: the number of children born in the parental birth municipality in the year of maternal / paternal birth; the fraction of the parental municipal birth cohort that have obtained an MSc by age 35; the fraction of the parental municipal birth cohort that had above median income at age 35, where the median is calculated over the whole national birth cohort; and similarly, the fraction of the parental municipal birth cohort that had an income in the top percentile of the national cohort at age 35.

The table also reports the p-values of F-tests for the joint significance of the instruments and the additional municipal controls. The corresponding F-test values for the instruments are reported in the main text (see Table 1).

Table A-6: First stage estimation results

	All	Daughters	Sons	All	Daughters	Sons	All	Daughters	Sons
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
lndist_uni$_{mo}$	-0.0343***	-0.0351***	-0.0334***				-0.0430***	-0.0454***	-0.0408***
	(0.00333)	(0.00391)	(0.00391)				(0.00336)	(0.00463)	(0.00418)
lndist_uni$^2_{mo}$	0.00940***	0.00978***	0.00903***				0.0114***	0.0124***	0.0106***
	(0.00105)	(0.00126)	(0.00127)				(0.00115)	(0.00156)	(0.00154)
lndist_uni$^3_{mo}$	-0.000772***	-0.000819***	-0.000727***				-0.000253**	-0.000347**	-0.000164
	(9.66e-05)	(0.000118)	(0.000119)				(0.000119)	(0.000156)	(0.000163)
lndist_uni$_{fa}$				-0.0264***	-0.0267***	-0.0261***	-0.0313***	-0.0329***	-0.0298***
				(0.00208)	(0.00309)	(0.00225)	(0.00265)	(0.00450)	(0.00305)
lndist_uni$^2_{fa}$				0.00694***	0.00722***	0.00666***	0.00692***	0.00770***	0.00616***
				(0.000746)	(0.00109)	(0.000810)	(0.000894)	(0.00159)	(0.00112)
lndist_uni$^3_{fa}$				-0.000509***	-0.000547***	-0.000472***	0.000276***	0.000202	0.000349***
				(7.14e-05)	(0.000105)	(7.88e-05)	(8.68e-05)	(0.000154)	(0.000118)
lndist_uni$_{mofa}$							0.00573***	0.00598***	0.00549***
							(0.000410)	(0.000603)	(0.000500)
lndist_uni$^2_{mofa}$							-0.000569***	-0.000593***	-0.000545***
							(2.65e-05)	(3.56e-05)	(3.32e-05)
lndist_uni$^3_{mofa}$							5.97e-06***	6.40e-06***	5.56e-06***
							(4.61e-07)	(5.77e-07)	(5.69e-07)
Count$_{muni,mo}$	-6.96e-07	-5.94e-07	-7.95e-07	3.45e-06***	3.61e-06***	3.29e-06***	5.58e-07	5.78e-07	5.36e-07
	(8.67e-07)	(9.00e-07)	(1.02e-06)	(5.80e-07)	(5.78e-07)	(7.22e-07)	(8.56e-07)	(8.92e-07)	(1.01e-06)
MSc$_{muni,mo}$	0.462***	0.447***	0.477***	0.549***	0.536***	0.562***	0.453***	0.438***	0.468***
	(0.0296)	(0.0300)	(0.0323)	(0.0278)	(0.0264)	(0.0329)	(0.0294)	(0.0297)	(0.0321)
Income_p50$_{muni,mo}$	0.0403***	0.0341***	0.0461***	0.0425***	0.0371***	0.0477***	0.0384***	0.0325***	0.0440***
	(0.00580)	(0.00589)	(0.00740)	(0.00645)	(0.00614)	(0.00803)	(0.00587)	(0.00596)	(0.00742)
Income_p90$_{muni,mo}$	0.173***	0.185***	0.162***	0.191***	0.203***	0.180***	0.167***	0.178***	0.157***
	(0.0114)	(0.0137)	(0.0138)	(0.0137)	(0.0158)	(0.0157)	(0.0117)	(0.0139)	(0.0140)
Count$_{muni,fa}$	1.79e-06***	1.67e-06***	1.90e-06***	-1.40e-06**	-1.43e-06**	-1.38e-06**	-2.83e-07	-2.51e-07	-3.11e-07
	(4.52e-07)	(4.89e-07)	(5.64e-07)	(5.64e-07)	(6.70e-07)	(5.57e-07)	(5.83e-07)	(6.85e-07)	(5.80e-07)
MSc$_{muni,fa}$	0.530***	0.532***	0.527***	0.468***	0.472***	0.464***	0.466***	0.469***	0.463***
	(0.0272)	(0.0268)	(0.0320)	(0.0260)	(0.0250)	(0.0321)	(0.0261)	(0.0249)	(0.0324)
Income_p50$_{muni,fa}$	0.0154**	0.00929	0.0213***	0.0220***	0.0160**	0.0277***	0.0160***	0.0104*	0.0213***
	(0.00605)	(0.00655)	(0.00701)	(0.00571)	(0.00630)	(0.00678)	(0.00572)	(0.00607)	(0.00692)
Income_p90$_{muni,fa}$	0.146***	0.156***	0.136***	0.132***	0.143***	0.122***	0.121***	0.131***	0.111***
	(0.0102)	(0.0117)	(0.0119)	(0.00932)	(0.0109)	(0.0112)	(0.00988)	(0.0114)	(0.0118)
Observations	1,450,789	709,117	741,671	1,450,789	709,117	741,671	1,450,789	709,117	741,671
R-squared	0.056	0.056	0.055	0.055	0.056	0.055	0.060	0.061	0.060
Instruments	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Municipal controls	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Standard errors in parentheses are clustered at year-of-birth level. Instrument is the propensity score estimated using LPM and a 3^{rd} order polynomial of the mentioned variable(s). Variables with subscript mo are maternal, those with subscript fa paternal, and those with subscript $mofa$ are interactions. Instruments and Municipal controls report the p-value of an F-test of joint significance. All specifications include a full set of maternal and paternal year-of-birth dummies and a dummies for mother tongue not being Finnish.

A-3 Robustness tests

In this section we report the results from our different robustness tests. We first briefly list the robustness tests and then display the associated result tables.

Robustness test 1: We re-estimate our model using parents obtaining a BSc as our measure of parental education, i.e., the parental education dummy takes values one if at least one of the parents has a BSc and is zero otherwise. The results are reported in Table A-7.

Robustness test 2: We use the number of parents with an MSc instead of a dummy for having at least one parent with an MSc as the treatment variable. The results are reported in Table A-8.

Robustness test 3: We use the same specification as for the main results, but change the outcome variable: instead of the inventor - dummy, we use the number of patents. To alleviate the potential effects of the long right tail of the number of patents - distribution, we truncate the distribution at 10 patents, i.e., at the 99.96^{th}percentile. The results are reported in Table A-9.

Robustness test 4: We use the same specification as for the main results, but change the outcome variable to be the number of citations. The number of citations we use is the number of forward citations in the first 5 years of a patent's life, summed over all patents of an inventor. To alleviate the potential effects of the long right tail of the number of citations -distribution, we truncate the distribution at 10 citations, i.e., at the 99.96^{th} percentile. As there are inventors with zero citations in our data, we amend the number of citations so that it is zero for non-inventors, one for inventors with at least one patent but no citations as well as for those with one citation regardless of the number of patents they have, and the number of citations for all those inventors with at least two citations in total. The results are reported in Table A-10.

A-3.1 Using a BSc as measure of parental education

Table A-7: Estimation results using $D(BSc\ parents)$

	Panel A. All Children			
	(1)	(2)	(3)	(4)
	OLS	IV	IV	IV
$D(BSc\ parents)$	0.0126***	0.0272***	0.0164***	0.0213***
	(0.00102)	(0.00607)	(0.00525)	(0.00295)
F	-	70.72	238.55	799.88
Nobs	1 450 789			
	Panel B. Daughters			
$D(BSc\ parents)$	0.00360***	0.00544	0.0107**	0.0102***
	(0.000318)	(0.00476)	(0.00463)	(0.00227)
F	-	50.78	114.82	349.69
Nobs	709 117			
	Panel C. Sons			
$D(BSc\ parents)$	0.0210***	0.0483***	0.0207*	0.0317***
	(0.00175)	(0.0110)	(0.0115)	(0.00556)
F	-	64.67	154.62	614.45
Nobs	741 671			
	Instruments			
Maternal dist.	NO	YES	NO	YES
Paternal dist	NO	NO	YES	YES

Standard errors in parentheses are clustered at year-of-birth level. Instrument is the propensity score estimated using LPM and a 3^{rd} order polynomial of the (logs) of the parental distances marked YES in the two last rows of the table. All specifications include a full set of maternal and paternal year-of-birth dummies, a dummy for mother tongue not being Finnish and the municipal controls explained in the text. F is the value of an F-test of all the instruments in the regression of the measure of parental education on instruments and controls.

A-3.2 Using number of parents with an MSc as measure of parental education

Table A-8: Estimation results using # parents with MSc

	(1)	(2)	(3)	(4)
	OLS	IV	IV	IV
Panel A. All Children				
#MSc parents	0.0125***	0.0407***	0.0248***	0.0251***
	(0.00108)	(0.0087)	(0.0072)	(0.0038)
F	-	202.03	190.72	427.41
Nobs	1 450 789			
Panel B. Daughters				
#MSc parents	0.00392***	0.0080	0.0155**	0.0124***
	(0.0004)	(0.0068)	(0.0068)	(0.0027)
F	-	100.96	85.40	213.70
Nobs	709 117			
Panel C. Sons				
#MSc parents	0.0206***	0.0698***	0.0324*	0.0371***
	(0.0019)	(0.0153)	(0.0158)	(0.0071)
F	-	102.35	105.79	214.56
Nobs	741 671			
Instruments				
Maternal dist.	NO	YES	NO	YES
Paternal dist	NO	NO	YES	YES

Standard errors in parentheses are clustered at year-of-birth level. Instrument is the propensity score (predicted number of parents with MSc) estimated using LPM and a 3^{rd} order polynomial of the (logs) of the parental distances marked YES in the two last rows of the table. All specifications include a full set of maternal and paternal year-of-birth dummies, a dummy for mother tongue not being Finnish and the municipal controls explained in the text. F is the value of an F-test of all the instruments in the regression of the measure of parental education on instruments and controls.

A-3.3 Using number of patents as the outcome variable

Table A-9: Estimation results using # patents as outcome

	(1)	(2)	(3)	(4)
	OLS	IV	IV	IV
Panel A. All Children				
$D(MSc\ parents)$	0.0493***	0.158***	0.0603	0.0723**
	(0.0055)	(0.0429)	(0.0344)	(0.0205)
F	-	55.73	140.73	508.87
Nobs	1 450 789			
Panel B. Daughters				
$D(MSc\ parents)$	0.0115***	0.0381	0.0269	0.0377**
	(0.0017)	(0.0227)	(0.0316)	(0.0099)
F	-	47.25	75.87	326.80
Nobs	709 117			
Panel C. Sons				
$D(MSc\ parents)$	0.0847***	0.262***	0.0856	0.1050*
	(0.0100)	(0.0795)	(0.0658)	(0.0390)
F	-	35.9504	94.12	264.76
Nobs	741 671			
Instruments				
Maternal dist.	NO	YES	NO	YES
Paternal dist	NO	NO	YES	YES

Standard errors in parentheses are clustered at year-of-birth level. Instrument is the propensity score estimated using LPM and a 3^{rd} order polynomial of the (logs) of the parental distances marked YES in the two last rows of the table. All specifications include a full set of maternal and paternal year-of-birth dummy, a dummy for mother tongue not being Finnish and the municipal controls explained in the text. F is the value of an F-test of all the instruments in the regression of the measure of parental education on instruments and controls.

A-3.4 Using number of citations as the outcome variable

Table A-10: Estimation results using # citations as outcome

	Panel A. All Children			
	(1)	(2)	(3)	(4)
	OLS	IV	IV	IV
D(MSc parents)	0.0475***	0.150***	0.0985*	0.0798**
	(0.0058)	(0.0413)	(0.0386)	(0.0213)
F	-	55.73	140.73	508.87
Nobs	1 450 789			
	Panel B. Daughters			
D(MSc parents)	0.0119***	0.0117	0.0503	0.0345***
	(0.0017)	(0.0189)	(0.0341)	(0.0113)
F	-	47.25	75.87	326.80
Nobs	709 117			
	Panel C. Sons			
D(MSc parents)	0.0808***	0.274***	0.1390*	0.1230***
	(0.0103)	(0.0794)	(0.0735)	(0.0376)
F	-	35.9504	94.12	264.76
Nobs	741 671			
	Instruments			
Maternal dist.	NO	YES	NO	YES
Paternal dist	NO	NO	YES	YES

Standard errors in parentheses are clustered at year-of-birth level. Instrument is the propensity score estimated using LPM and a 3^{rd} order polynomial of the (logs) of the parental distances marked YES in the two last rows of the table. All specifications include a full set of maternal and paternal year-of-birth dummies, a dummy for mother tongue not being Finnish and the municipal controls explained in the text. *F* is the value of an F-test of all the instruments in the regression of the measure of parental education on instruments and controls.