NBER WORKING PAPER SERIES

THE SOCIAL ORIGINS OF INVENTORS

Philippe Aghion
Ufuk Akcigit
Ari Hyytinen
Otto Toivanen

Working Paper 24110
http://www.nber.org/papers/w24110

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2017

This project owes a lot to early discussions with Raj Chetty, Xavier Jaravel and John Van Reenen when we were embarking on two parallel projects, them on US inventors and us on Finnish inventors. We also benefitted from very helpful comments and suggestions from our discussants at the NBER Summer Institute and at the AEA meetings, namely Pierre Azoulay, Chad Jones, and Heidi Williams respectively, and also from seminar participants at the University of Chicago, the Kaufman Foundation, the Paris School of Economics, the London School of Economics, the IOG group at the Canadian Institute for Advanced Research, the Institute of Fiscal Studies, HECER, the University of Maastricht, Copenhagen Business School, EIEF and College de France. Akcigit gratefully acknowledges the National Science Foundation, the Alfred P. Sloan Foundation, and the Ewing Marion Kauffman Foundation for financial support. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed a financial relationship of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w24110.ack

The Social Origins of Inventors
Philippe Aghion, Ufuk Akcigit, Ari Hyytinen, and Otto Toivanen
NBER Working Paper No. 24110
December 2017
JEL No. I24,J18,J24,O31

ABSTRACT

In this paper we merge three datasets - individual income data, patenting data, and IQ data - to analyze the determinants of an individual's probability of inventing. We find that: (i) parental income matters even after controlling for other background variables and for IQ, yet the estimated impact of parental income is greatly diminished once parental education and the individual's IQ are controlled for; (ii) IQ has both a direct effect on the probability of inventing an indirect impact through education. The effect of IQ is larger for inventors than for medical doctors or lawyers. The impact of IQ is robust to controlling for unobserved family characteristics by focusing on potential inventors with brothers close in age. We also provide evidence on the importance of social family interactions, by looking at biological versus non-biological parents. Finally, we find a positive and significant interaction effect between IQ and father income, which suggests a misallocation of talents to innovation.

Philippe Aghion
College de France
3 Rue D'Ulm
75005 Paris
France
paghion@fas.harvard.edu

Ufuk Akcigit
Department of Economics
University of Chicago
1126 East 59th Street
Saieh Hall, Office 403
Chicago, IL 60637
and NBER
uakcigit@uchicago.edu

Ari Hyytinen
University of Jyväskylä
PO Box 35
40014 University of Jyväskylä
Finland
ari.t.hyytinen@jyu.fi

Otto Toivanen
Professor Otto Toivanen
Department of Managerial Economics, Strategy and I
Faculty of Business and Economics
Katholieke Universiteit Leuven
Naamsestraat 69, B-3000 Leuven, Belgium
HOG 03.145
otto.toivanen@econ.kuleuven.be

1 Introduction

Who becomes an inventor? Does innovation attract the most talented individuals or is there misallocation of talents into innovation? In practice, not everybody can become an innovator: whether one becomes an innovator or not is likely to depend upon the social environment (e.g., parental resources and education) and upon ability. To the extent that both parental inputs and ability are unevenly distributed across individuals, the innovative potential of the economy may be underutilized due to misallocation of talent. Misallocation means that a positive fraction of potential inventors are not performing as well as they could due to inadequate or insufficient parental support. Inadequate parental support may be especially harmful for highly talented individuals, eroding even further the utilization of the innovative potential of the economy.

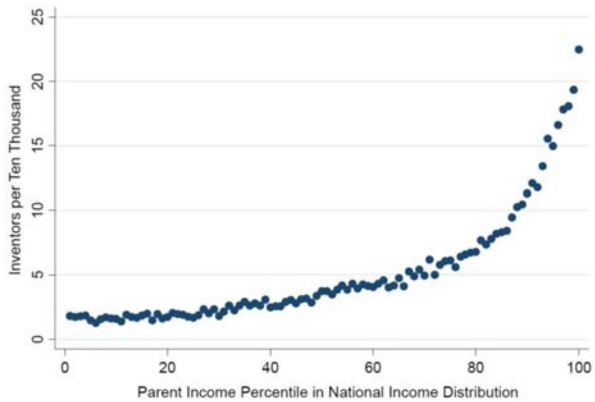
In this paper we merge three comprehensive Finnish datasets to analyze the determinants of an individual's probability to become an inventor, and to investigate whether and to which extent there is a misallocation of talents to innovation.[1] We thus consider in detail the role of: (i) family resources, parental education and social background; (ii) innate ability as proxied by visuospatial IQ; (iii) the interaction between parental background and ability.

The following striking fact motivates our analysis. Figure 1 depicts the relationship between an individual's probability of becoming an inventor and his father's income, respectively on the basis of recent US data (Figure 1A is drawn from Bell et al., 2017), on the basis of US historical data (Figure 1B is drawn from Akcigit et al., 2017), and for Finland (Figure 1C is based on our own data). We see that in all three cases the individual's probability of becoming an inventor increases with father's income, and that the effect is highly non-linear, being particularly steep at the highest levels of father's income. We also see that the probability of innovating for an individual whose father is at the very top of the income distribution is about ten times larger than the corresponding probability for an individual with a father at the bottom end of the income distribution. The similarity between Finland and the US is all the more remarkable given that, unlike the US, Finland displays low income inequality and high social mobility in international comparison (see Figure 2) and offers free education up to and including university. What lies behind the steep relationship in Figure 1C between father income and the probability of becoming an inventor in Finland?

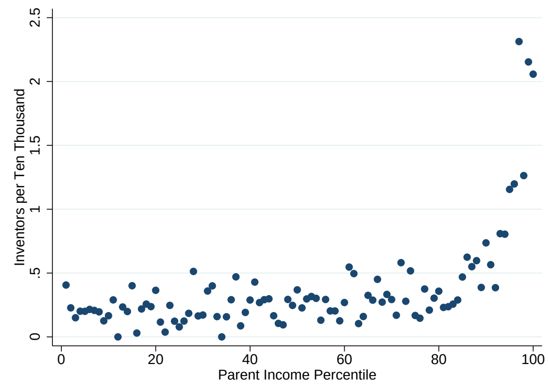
To explore this enigma, we merge three Finnish data sets: (i) individual data on income, education and other characteristics from Statistics Finland for individuals born

[1] In a companion paper (see Aghion et al., 2017), we analyze the return to innovation.

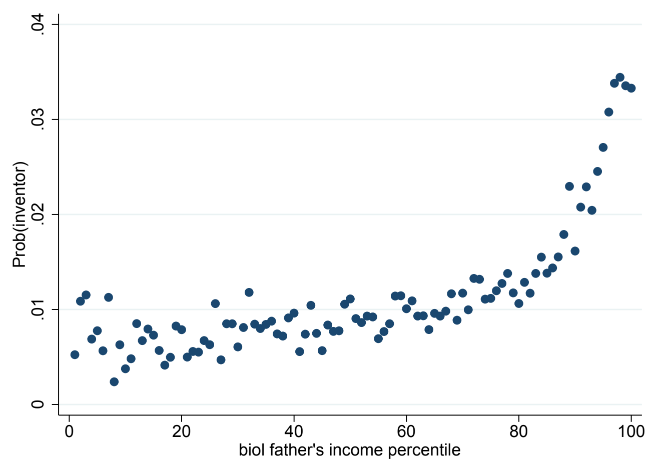
Figure 1: Parental Income and Becoming an Inventor



A. Source: Bell et al. (2017)



B. Source: Akcigit et al. (2017)

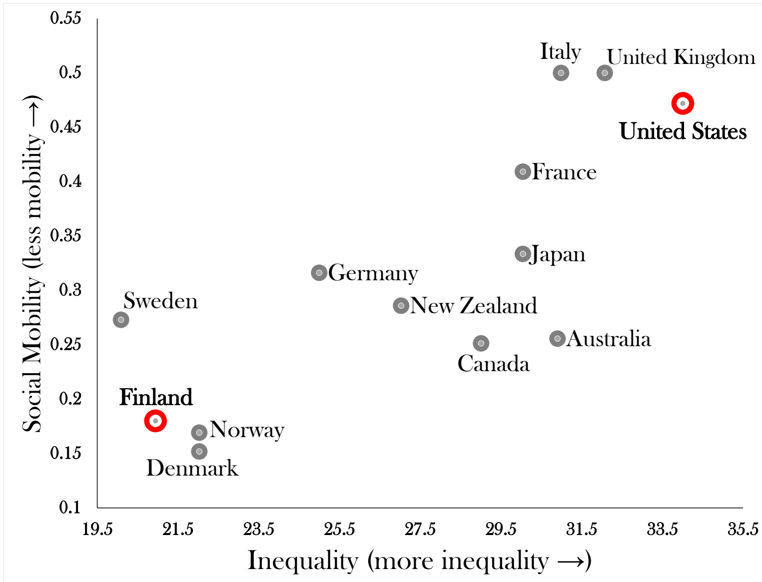


C. Source: This paper.

Notes: The figure displays the probability to invent as a function of father's / parents' income percentile. Figure 1A comes from Bell et al. (2017) and Figure 1B from Akcigit et al. (2017). In Figure 1C, father's income is measured in 1975 for individuals born 1961–1975, and in 1985 for individuals born in 1976-1985. We calculate the percentile ranks using residuals from a regression of the natural log of deflated income of fathers and mothers on year-of-birth dummies.

between 1961 and 1984, and their parents; (ii) individual patenting data from the European Patent Office; (iii) IQ data from the Finnish Defence Forces. Given that conscription only affects males in Finland, we concentrate on the male workforce in this paper. We analyze how family resources, social and educational background, own ability, and the interaction between background and ability impact on an individual's probability of inventing.

FIGURE 2: THE GREAT GATSBY CURVE



SOURCE: CORAK (2004)

Our main findings can be summarized as follows. First, parental income matters for the probability of becoming an inventor, and it does so even after controlling for other background variables and for IQ. Second, the estimated impact of parental income is greatly diminished once parental socioeconomic status and education, and the individual's IQ are controlled for, dropping by 2/3rds. Third, IQ has both a direct effect on the probability of inventing which is almost five times as large as that of having a high-income father, and an indirect effect through education. Finally, the impact of IQ is larger and more convex for inventors than for medical doctors or lawyers.

To address the potential endogeneity of IQ, we zoom our analysis on potential inventors with brothers close in age. This allows us to control for family-specific time-invariant unobservables. The effect of visuospatial IQ on the probability of inventing is robust to adding these controls.

Next, we provide evidence on the importance of family structure by comparing individuals brought up by their biological parents with individuals with a missing biological parent and/or individuals with a step parent. We find that parental divorce decreases the probability of becoming an inventor and that the income of biological parents matters only when the child lives with them, but that the effect of parental education is not dependent on living together.

We then explore the potential complementarity between IQ and family background. We find a positive and significant interaction between the individual's IQ and his father's income, which in turn points to a potential source of misallocation: namely, a positive fraction of individuals with very high IQ will underperform as potential innovators due to inadequate parental background.

The paper relates to several strands of literature. There is first a theoretical and empirical literature on innovation incentives.[2] However this literature does not look at the effects of social background on the probability of inventing, nor do they analyze the social mobility of inventors and co-workers.

Second, there is a recent literature on misallocation and economic growth. In particular Hsieh et al. (2013) look at how much of the increase in aggregate GDP per capita between 1960 and 2010 is due to an improved allocation of talents to tasks in the US, and in particular to an improved access of talented women or black men to high occupation tasks. A key assumption in their analysis is that the distribution of innate ability is identical across groups, and they point to the importance of labor market discrimination and of financial barriers to human capital investments as being key drivers of the misallocation of talents in the US. Here, we do not make any prior assumption on the distribution of ability across socioeconomic groups, and we focus on a country - Finland - with a priori no or little labor market discrimination and where schooling is entirely free. Yet, we find a significant misallocation of talent even in Finland, affecting high IQ individuals in particular.

Third, our paper relates to the literature on innovation and social mobility. Balkin et al. (2000) finds that innovation increases CEO pay in high-tech industries. Aghion et al. (2017) use the same data that we use in the current paper to show that innovation increases an individual innovator's probability to make it to the higher income brackets, and that innovation has an even larger effect on firm owners' income. Bell et al. (2017) find that the most successful innovators see a sharp rise in income. Akcigit et al. (2017) merge historical census and patenting data across US states over the past 150 years and find a positive correlation between patenting intensity and social mobility. Finally, building on Chetty et al. (2014), Aghion et al. (2015) look at the relationship between innovation, inequality and social mobility using *aggregate* cross-state and cross-commuting-zone data. They show that innovation measured by the flow or quality of patents is positively correlated with the top 1% income share of income, is uncorrelated with broader measures of income inequality, and is positively correlated with social mobility (measured as in Chetty et al., 2014). In this paper our focus is on the relationship

[2]In particular, see Holmstrom (1989), Aghion and Tirole (1994), Pakes and Nitzan (1983), Scotchmer (2004), Giuri et al. (2007), Franco and Mitchell (2008), Azoulay et al. (2010), Manso (2011), Akcigit et al. (2016), and Depalo and Di Addario (2015).

between parental education and income, IQ and the individual's probability of inventing.[3]

Closer to our analysis in this paper is a recent literature merging individual income data with individual patenting data. First, Toivanen and Väänänen (2012) use Finnish patent and income data to study the return to inventors of US patents. They find strong and long-lasting impacts, especially for the inventors of highly cited patents. Toivanen and Väänänen (2016) look at the effect of education on the probability of becoming an inventor and they find a positive and significant treatment effect, suggesting the one may increase innovation through education policy. Second, Celik (2015) matches inventors' surnames with socioeconomic background information inferred from those surnames by looking at the US census data back in 1930. His main finding is that individuals from richer backgrounds are far more likely to become inventors. Akcigit et al. (2017) merge historical patent and individual census records and show that probability of becoming an inventor around 1940s was very highly correlated with father's income but this strong relationship disappears once child's education is controlled for. Finally, Jaravel et al. (2015) merge US individual tax data and individual patenting data to quantify the impact of coauthors in the career of inventors, finding evidence of large spillover effects.[4]

Most closely related to the present paper, is Bell et al. (2017) who merge US individual fiscal and test score data with US patenting data over a recent period to look at the lifecycle of inventors and the returns to invention. These authors find that parental income, parental occupation and sector of activity, race, gender, and childhood neighborhood are important determinants of the probability of becoming an inventor. Our analysis complements theirs, as on the one hand, we do not have as detailed information as they do on parents' inventive activity or on geographical origins, but on the other hand we have information on parental socioeconomic status and education, individual IQ, and family structure.

The information on parental socioeconomic status and education allows us to show that to a large extent in Finland the relationship between parental income and the probability of becoming an inventor is driven by parental education both directly and through

[3]Our paper intersects with the psychology literature and the debate on whether IQ tests are linked to genetics or to the social environment (e.g. see Mc Gue et al., 1993; Ceci, 2001; Pinker, 2005; and Plomin and Spinath, 2004). And our emphasis on (visuospatial) IQ as a measure of cognitive ability is shared by recent work by Dal Bó et al. (2017) who use similar IQ test information from the Swedish Arm Forces to analyze the selection of municipal politicians in Sweden.

[4]Jaravel's work builds on prior seminal work by Azoulay et al. (2010) which examines the effect of the premature death of 112 eminent scientists on their co-authors. This work provides the first convincing evidence of exposure to human capital (particularly at the high end) on the production of new scientific ideas, using the exogenous passing of elite scientists as an empirical lever.

its impact on the child's education. The information on IQ allows us to show that IQ impacts both directly and indirectly through education on the probability of becoming an inventor, and that it is more important than parental characteristics. Finally, the information on family structure allows us to identify how the effects of income and education of biological parents on the probability of inventing is affected by (not) living with them, and what effect the income of genetically unrelated step parents has on the probability of becoming an inventor.

The remainder of the paper is organized as follows. Section 2 presents the data and shows some descriptive statistics. Section 3 analyzes the determinants of becoming an inventor. Section 4 focuses on potential innovators with close brothers to address the concern that IQ is endogenous. Section 5 looks at the effect of family structure. Section 6 looks at the interaction between IQ and family background. Section 7 analyzes the role of education in becoming an inventor. Finally, Section 8 concludes by drawing some policy conclusions and by suggesting avenues for future research. Appendix A contains additional tables. Appendices B–G, which are online, present various supplementary materials.

2 Data and descriptive statistics

2.1 Data sources

Our data come from the following sources.

First data source: Statistics Finland (SF). We exploit two data sets from SF: (i) The Finnish Linked Employer-Employee Data (FLEED) for the period 1988-2012 and (ii) the population census 1975 and 1985.

FLEED covers the whole working age (15 and older) population of Finland. This annual panel is constructed from administrative registers of individuals, firms and establishments, maintained by SF. It includes information on individuals' labor market status, salaries and other sources of income extracted from tax and other administrative registers. It also includes information on other individual characteristics. We utilize information on individual age, location of residence, language, education (degree and field) and socioeconomic status. We use FLEED data from 1988, the first year it is available, to 2013, the year our patent data ends. Because only a small minority of inventors are women and because we only have IQ data for men, we exclude women from our sample.

Information on parent characteristics is drawn from the population census. More specifically, we use the population censuses of 1975 and 1985 for information about parental education, socioeconomic status and income of biological and social parents. The majority of individuals in our data have fathers born either in the 1940s (36%) or 1950s (25%). 37% of the individuals have mothers born in the 1940s and 30% mothers born in the 1950s. For 45% of these individuals, the father has only a base education (max. 9 years), while 6% have a Master's degree or higher. The respective figures for their mothers are 44% and 4%.

Second data source: the European Patent Office (EPO). EPO data provide information on characteristics such as the inventor names and applicant names.[5] We have collected information on all patent applications to EPO with at least one inventor who registers Finland as his or her place of residence. The data cover all EPO patent applications (starting in 1978) with an inventor with a Finnish address up to and including 2013. The data originates with PATSTAT, but Statistics Finland has used the OECD REGPAT database built on PATSTAT. In the raw patent data, we have a total of 25,711 patent applications and 17,566 inventors. The mean and median number of inventors per patent is 2; the largest number of inventors per patent is 14.[6] For each patent, we observe all the inventors, their name and address, the patentee and its address, the number of citations in the first 5 years, and the size of the patent family (i.e., the number of countries in which the patent exists).

Third data source: the Finnish Defence Forces. The Finnish Defense Forces (FDF) provided us with information on IQ test results for conscripts who did their military service in 1982 or later. These data contains the raw test scores of visuospatial, verbal and quantitative IQ tests. The IQ tests are a 2-hour multiple choice tests containing sections for verbal, arithmetic and visuospatial reasoning. The latter is similar to the widely used Raven's Progressive Matrices – test. Overall, the Finnish Defense Force IQ test is similar to the commonly used IQ tests; moreover, a large majority (over 75%) of each male cohort performs the military service and therefore takes the test: most conscripts take their military service around the age of 20. All conscripts take the IQ test in the early stages of the service (see Jokela et al., 2017, for more detail).

We use the deciles in visuospatial IQ score (IQ henceforth for brevity), as it is considered in the IQ literature to be more strongly predetermined than the other two measures.

---

[5] Here we want to thank the research project "Radical and Incremental Innovation in Industrial Renewal" by the VTT Research Centre (Hannes Toivanen, Olof Ejermo and Olavi Lehtoranta) for granting us access to the patent-inventor data they compiled.

[6] As is typical, the distribution of the number of patents, and citations per patent (we use the number of citations to the highest cited patent of an individual, measured over the first five years of the patent's life), are highly skewed - see Figure B1 and B2 in the Appendix.

As is standard for IQ data, we normalize the raw test scores to have mean 100 and standard deviation of 15. We do this by the year of entering military service to avoid the so-called Flynn effect. In robustness tests we use also the verbal and analytic IQ scores.[7]

Data matching: The linking of all other data but the patent data was done using individual identifiers. The linking of patent data to individuals was done 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 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 lies above 90%. We provide further details on data matching in Appendix B.1.

Sample: Our estimation sample contains all individuals for whom we were able to match all four data sets (EPO, FLEED, parental data, IQ). This means that besides women we exclude men born before 1961, as IQ data are not available for them. We further exclude individuals born after 1984 as they are unlikely to have invented by 2012. The resulting cross-sectional sample comprises of around 350,000 individuals and contains 4,754 inventors.

2.2 The institutional environment

In this section, we highlight some central features related to Finnish institutional environment. A more detailed discussion is provided in Appendix B.2.

Overall economic environment in 1988-2012. During our observation period from 1988 to 2012, Finland's gross domestic product (GDP) grew on average 2.1% per year. Research and development (R&D) is a major investment item in Finland. R&D expenditures reached their peak in 2011 when the total R&D expenditure by business sector and public sector amounted to 3.8% of the GDP. Based on its Global Competitiveness Index, World Economic Forum has quite consistently ranked Finland to be one of the ten most competitive countries in the world.

Educational system. A key characteristic of the Finnish education system is that it is (essentially) free of charge at all levels, up to and including university studies (to

[7] Using similar IQ test information from the Swedish Arm Forces to analyze the selection of municipal politicians in Sweden, Dal Bó et al. (2017) argue that these IQ scores are good measures of general intelligence and cognitive ability. The question remains as to whether IQ tests are linked to genetics or to the social environment. 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.

a PhD). Applicants to most field-specific degree programs of the Finnish polytechnics and universities are required to take an entrance examination. Entry into degree programs in law and medicine, as well certain fields of science, technology and business, is competitive.

Wage setting. Wage setting is a mixture of collective and individual mechanisms in Finland. As Uusitalo and Vartiainen (2009) stress, for most employees in the manufacturing sector the minimum wages rarely bind. These features of the Finnish labor market mean that relative wages have largely been set by market forces and that wage bargaining is to a significant extent local. Moreover, various firm-specific arrangements and performance-related pay components became more widespread in the 1990s.

Remuneration of inventors. A specific law governs innovations made by employees ("Act on the Right in Employee Inventions", originally given in 1967, augmented in 2000). The provisions of the act apply to inventions (potentially) patentable in Finland. In sum, the act assigns the right to ownership of an employee invention, but it does not directly determine the amount firms have to pay if they exercise the right. Rather, the determination of the amount of compensation is largely left to the market forces.

Economic inequality. In an international comparison, income inequality is relatively low in Finland. Using disposable cash income (excl. capital gains) as the income measure, the Gini-coefficient has ranged from 20.5 in 1992 to 26.4 in 2007 (Income distribution statistics, Statistics Finland). On average, it has been 23.6 during our sample period. Consistent with the relatively low income inequality, intergenerational mobility is in Finland - like in other Nordic countries - quite high, exceeding that of the UK and US (see, e.g., Black and Devereux, 2011). In line with this, the correlation of incomes among siblings is quite a bit lower in Finland (and in the other Nordic countries) than, for example, in the U.S. (Björklund et al., 2002, Black and Devereux, 2011).

2.3 Descriptive statistics on inventors and social origins

The outcome variables are (see Appendix B, Table B1 for precise variable definitions): indicator variables first, for obtaining at least one patent (Inventor), being a medical doctor (MD), being a lawyer (Lawyer), number of patents obtained by the individual (Patent count), the number of forward citations obtained by the most cited patent of the individual (Citations), and an indicator for having invented a highly cited patent (High quality inventor).

The control variables we use are: age, region of residence (21 dummies), type of region (urban being the base, and indicator variables for semi-urban and rural), mother tongue (Finnish, Swedish and any other language) and for parental birth-of-decade (separate vectors of indicator variables for father and mother). Our variables (vectors) of interest are measures of parental wage, parental socioeconomic status, parental education, and the individual's own IQ.

We capture parental income by indicator variables for income quintiles, with separate indicators for fathers and mothers. To allow for non-linearitiese at the right tail of the income distribution, the highest income quintiles are divided into separate indicators for the 81st – 90th percentiles, the 91st – 95th percentiles, and the 96th – 100th percentiles. We use the 1975 (deflated) income from the census for parents of individuals born by 1975, and the 1985 census income information for parents of individuals born later than 1975. Income percentiles are based on the residuals of a log (wage) regression on year of birth dummies.[8]

We divide parents into four socioeconomic groups: blue-collar, junior and senior white-collar status, and others. We measure parental education by indicators for different education levels. The levels are: base (= 9 years of schooling), secondary, college, master's degree and PhD. We also include indicators for a STEM education, separately for both parents.

We include IQ using decile dummies. Just like with parental income, the highest IQ decile is divided into separate indicators for the 91st – 95th and the 96th – 100th percentiles.

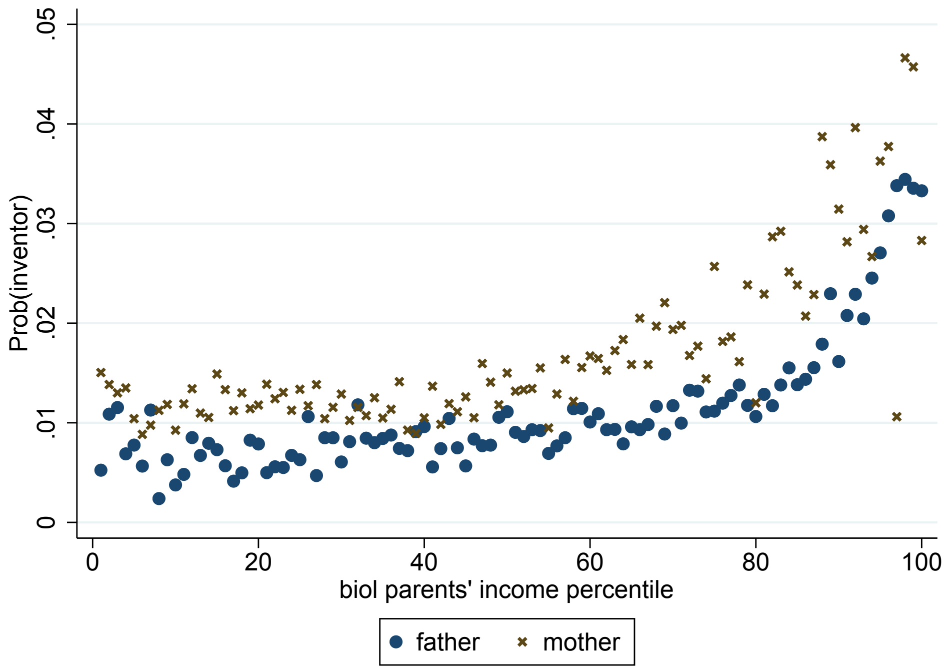
Inventors are on average slightly older than the overall population in our sample (44 vs 41) and are: 1) more likely to have a father (but not a mother) in the 5 percent of the income distribution (19 vs 8 percent); 2) less likely to have a blue-collar parent (29 vs 45 percent for fathers and 19 vs 31 for mothers) and more likely to have a white-collar father; 3) more likely to have highly educated parents (19 (9) percent probability of father (mother) having at least an MSc vs 6 (3) per cent) and more likely to have a mother (but not father) with a STEM education (14 vs 5 percent); and 4) are more likely to be in the top 5 per cent of the IQ distribution (19 vs 5 percent). All these differences are significant at the 1 percent level or better. We display the descriptive statistics in Appendix B, Table B2, conditioning on the inventor status of the individual.

Figure 3 reproduces Figure 1 for Finnish data, adding the relationship between mother's income percentile and the probability to invent. We observe that fixing the income percentile, the effect of mother's income is larger than that of father's. Further, starting from roughly the 60th percentile, the effect of mother's income starts to increase

[8] As a robustness test, we use an alternative income measure with no over-time variation.

faster than that of father's income.[9]

Figure 3: Mother and Father's Income and Becoming an Inventor



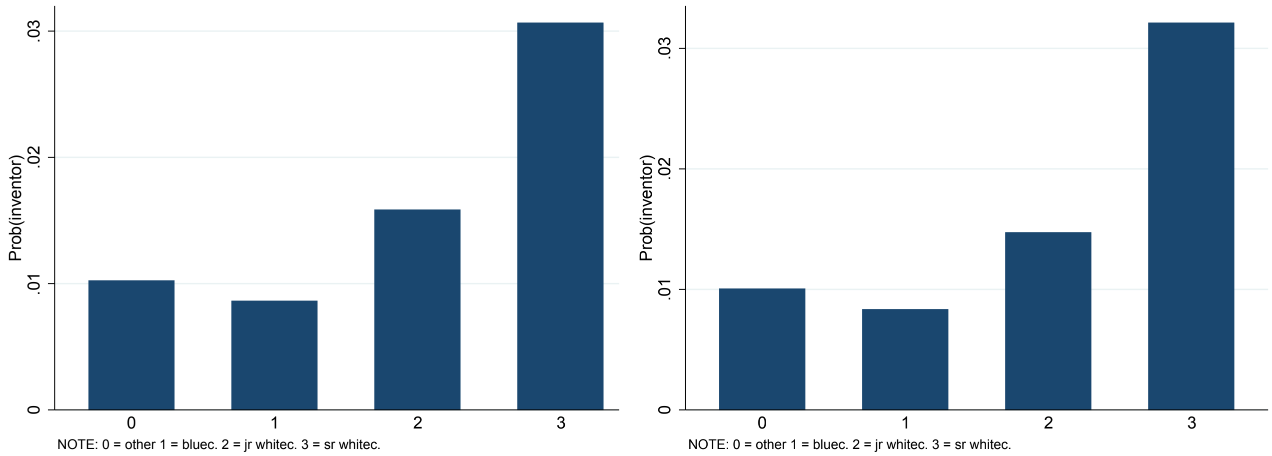
Notes: The figure displays the probability to invent as a function of father's and mother's income percentile. Parental income is measured in 1975 for individuals born 1961–1975, and in 1985 for individuals born in 1976–1985. We calculate the percentile ranks using residuals from a regression of the natural log of deflated income of fathers and mothers on year-of-birth dummies.

Figure 4 displays histograms where the probability to invent is conditioned on parental socioeconomic status, separately for fathers and mothers. The figure shows that those with senior white-collar parents are about three times as likely to invent as those with blue-collar parents. Parents are strongly positively assortatively matched along their socio-economic status (see Figure B5 in the Appendix), and income and socioeconomic status are positively correlated (Figure B6 in the Appendix). As an example, the average income percentile of blue-collar fathers is slightly above 60, but that of senior white collar fathers about 90. The association is weaker for mothers.

Figure 5 displays histograms where the probability to invent is conditioned on parental

[9]Notice that the distributions of mothers and fathers across the income percentiles are different, with a higher fraction of mothers in the low end of the income distribution; see Figure B3 in the Appendix. We observe positive assortative matching of parents regarding income. We display this association in Figure B4.

Figure 4: Parental Socioeconomic Status and Becoming an Inventor



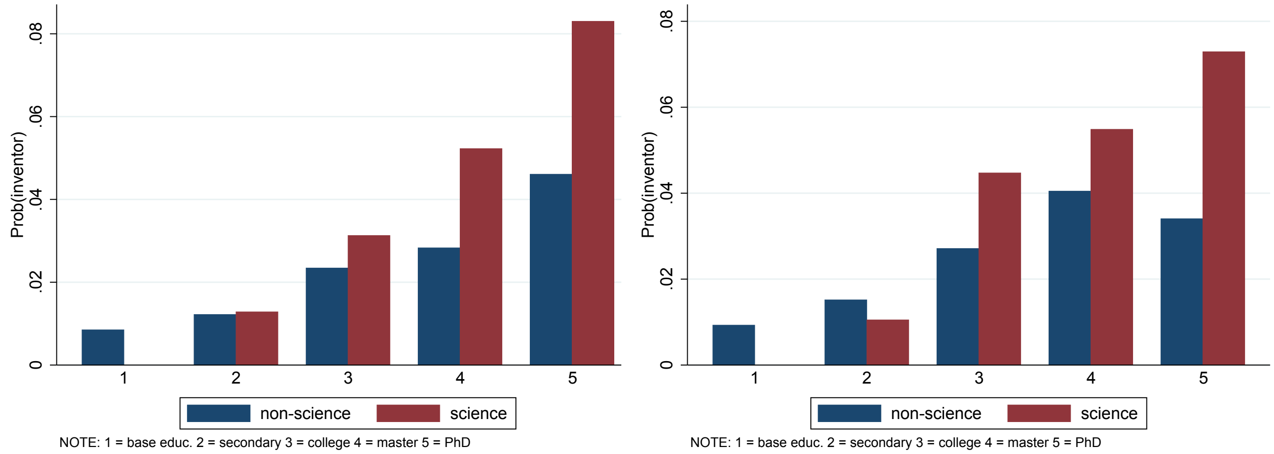
A. Father B. Mother

Notes: The figure displays the probability to invent conditional on the socio-economic status of the father (A) and mother (B). We divide parents into four groups by socioeconomic status: blue-collar, junior white collar, senior white collar, and others. Parental socioeconomic status is measured in 1975 for parents born before 1951, and in 1985 parents born in 1951 or thereafter.

education, separately for fathers and mothers and for STEM and non-STEM education. The figure shows clearly how the probability to invent is positively associated with the level of both parents' education, and conditional on the level of education, with the parent having a STEM education. Those with a father or mother with a STEM PhD are more than six times as likely to invent as those whose father or mother has only a base education. Having a father or a mother with a STEM instead of a non-STEM PhD increases the probability to invent by more than 50 percent.

Just like parental income and socio-economic status, parental education exhibits positive assortative matching. The probability that an individual whose father has a base education has a mother also with base education is over 60 percent; if the father has a PhD, the probability of the mother having at least an MSc is about 40 percent (Figure B7 in the Appendix). Education and income (Figure B8) and education and socioeconomic status (Figure B9) of parents are positively correlated. As an example, the mean income percentile of fathers with a base education is round 60, whereas the corresponding number for fathers with a PhD is 90. The strong positive correlations of these parental characteristics suggest that one should control for all of them in exploring the relation between parental background, income in particular, and the probability of becoming an inventor.

Figure 5: Parental Education Status and Becoming an Inventor



A. Father B. Mother

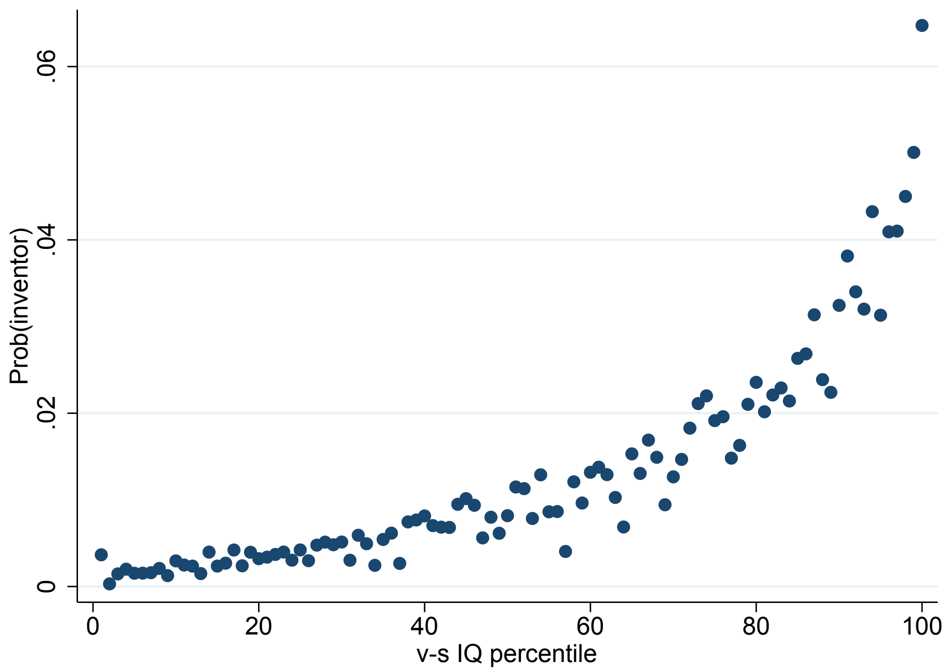
Notes: The figure displays the probability to invent conditional on the education of the father (A) and mother (B). We divide parents into five groups by level of education: base education (up to 9 years, depending on age of parent), secondary, tertiary, MSc, and PhD. We also condition all other levels of education but base education on a parent having a STEM education. A STEM base education does not exist. Parental education is measured in 1975 unless unavailable, in which case 1985 data used.

We then turn to the association between own ability and inventing. Figure 6 plots the probability to invent against IQ percentiles to allow for a comparison to Figures 1 and 3. 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.

Own IQ and parental income, socioeconomic status and education are all positively correlated. We display the correlation between father's and mother's income and IQ in Figure B10 in the Appendix. Both curves display an increasing, convex relationship. Individuals whose parents are white collar have above average IQ (Figure B11 , as do individuals with more educated parents (Figure B12).

Summarizing the descriptive statistics, Figures 3 - 6 suggest that parental income, socioeconomic status and education as well as own IQ all are strongly positively associated with the probability to invent. In particular, the probability is highly convex at the right end of the distribution for all parental characteristics and own IQ. These patterns suggest potential misallocation of innovative talent.

Figure 6: Own Visuo-Spatial IQ and Becoming an Inventor



Notes: The figure displays the probability to invent conditional on the visuo-spatial IQ percentile of the individual. IQ percentiles are calculated on the basis of the normalized IQ score, where normalization was done separately for each conscription cohort to avoid the Flynn effect.

3 Becoming an inventor

In this section, we study the determinants of becoming an inventor. In particular, we estimate a linear probability model where we regress the probability of becoming an inventor on base controls (see below), parental income, parental socioeconomic status, parental education, and own IQ.

3.1 Regression equation

The regression equation that will serve as the basis for the estimations in this section can be written as:

$$D_i = \alpha + \sum_f \beta_f fcontrols_{fi} + \sum_m \beta_m mcontrols_{mi} + \sum_k \theta_k IQ_{ki} + \sum_o \beta_o ocontrols_{oi} + \epsilon_i,$$

where: (i) D is a dummy for being an inventor / MD / lawyer (and in robustness tests, patent count, count of citations to the most cited patent of the individual, and a dummy for being a star inventor); (ii) the $fcontrols$ variables measure father characteristics; (iii) the $mcontrols$ variables measure mother characteristics; (iv) the $ocontrols$ variables measure other background characteristics; (v) the IQ variables measure the individual's own IQ; (vi) α, βs and θs are the parameters to be estimated; and (vii) ϵ is the error term.

Parental income is introduced via quintile indicators, with the highest quintile split as explained above (fa income, mo income). The excluded income group for both parents is the lowest quintile. The socioeconomic groups for both parents include "blue collar", "white collar junior", and "white collar senior" (fa sose, mo sose, sose =bluecollar, jr whitecollar, sr whitecollar). We take the "other" category as the base. For parental education the excluded group is base education, and we use indicators for secondary, college, masters, and PhD level education (fa educ, mo educ, educ =secondary, college, MSc, PhD). We also include separate dummies to indicate that a parent has a STEM education (fa STEM, mo STEM). For IQ, the base group is the 4th IQ decile; and we include dummies for 1st - 3rd and 5th - 9th IQ deciles (IQd); we split the highest decile into two separate dummies for 91st-95th and 96th-100th percentiles. Finally, we include in all specifications the following control variables: a 4th order polynomial in (log) age, 21 region dummies; dummies for suburban and urban areas; dummies for Swedish and other than Finnish language as mother tongue; and parental decade of birth dummies, separately for both parents.

3.2 Baseline results

The regression results are shown in Table 1. Since these are very lengthy regressions with too many independent variables, we report here only the coefficients of the most interesting variables.[10]

In column 1 of Table 1 we regress D on base controls and parental income. We see from column 1 in Table 1 that having either the father or the mother belong to the second highest 91-95 or the highest 96-100 income bracket has a positive and significant association with the probability of becoming an inventor. Comparing the coefficients of the order of more than 1 and 2 percentage points to the sample mean of 0.9 percent for the probability of becoming an inventor suggests that these are economically important effects. The estimated coefficients for father's and mother's income for a given income bracket are close to each other. We display the estimated relationship between father

[10]Tables displaying the coefficients of all but control variables can be found in Appendix Table C1.

Table 1: Who Becomes Inventor Regressions

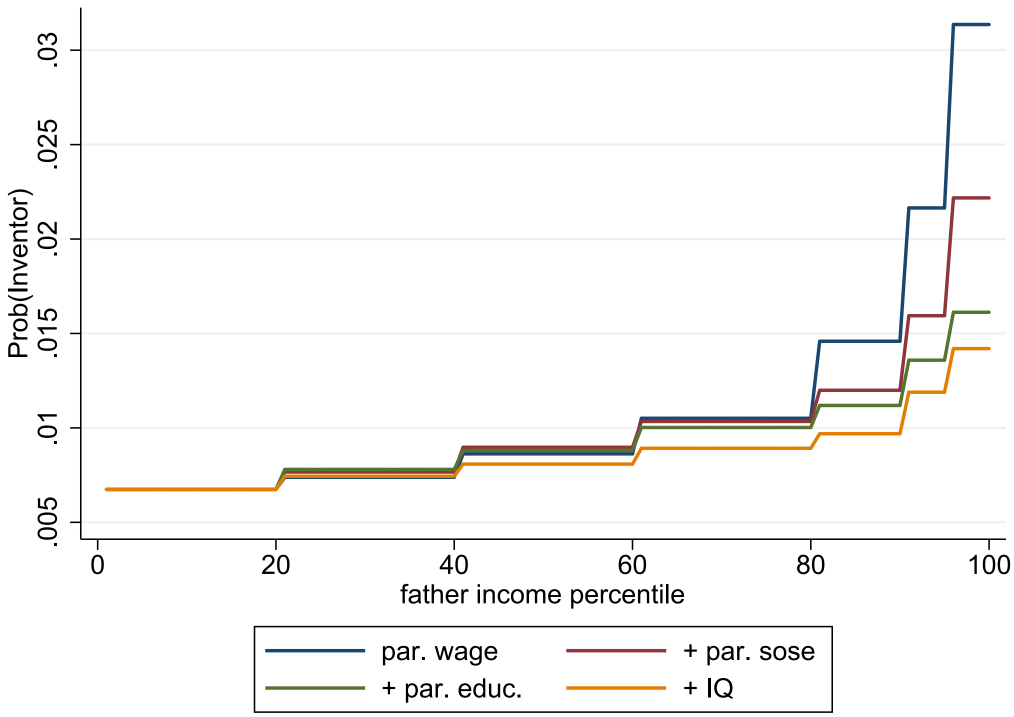
VARIABLES	(1)	(2)	(3)	(4)
fa income 91-95	0.0149***	0.00919***	0.00684***	0.00515***
	(0.00107)	(0.00109)	(0.00109)	(0.00108)
fa income 96-100	0.0246***	0.0154***	0.00938***	0.00745***
	(0.00126)	(0.00131)	(0.00130)	(0.00130)
mo income 91-95	0.0126***	0.00627**	-0.000846	-0.00186
	(0.00307)	(0.00311)	(0.00315)	(0.00314)
mo income 96-100	0.00260**	0.00216*	0.000139	-0.000410
	(0.00114)	(0.00115)	(0.00112)	(0.00112)
fa bluecollar		-0.00121**	-0.000999*	-0.000759
		(0.000543)	(0.000542)	(0.000540)
fa jr whitec.		0.00269***	0.00281***	0.00184**
		(0.000727)	(0.000738)	(0.000735)
fa sr whitec.		0.00883***	0.00402***	0.00270**
		(0.00102)	(0.00112)	(0.00112)
mo bluecollar		-0.00101*	-0.000263	4.32e-05
		(0.000551)	(0.000551)	(0.000550)
mo jr whitec.		0.00186***	0.00211***	0.00146**
		(0.000621)	(0.000621)	(0.000619)
mo sr whitec.		0.00884***	0.00431***	0.00333***
		(0.00119)	(0.00125)	(0.00125)
fa MSc			0.0119***	0.00876***
			(0.00175)	(0.00175)
fa PhD			0.0310***	0.0275***
			(0.00435)	(0.00434)
mo MSc			0.0152***	0.0119***
			(0.00242)	(0.00242)
mo PhD			0.0123	0.00826
			(0.00957)	(0.00957)
fa STEM			0.00889***	0.00861***
			(0.000801)	(0.000798)
mo STEM			-0.00112	-0.00116
			(0.000734)	(0.000732)
IQ 91-95				0.0236***
				(0.00144)
IQ 96-100				0.0351***
				(0.00165)
Nobs	352,668	352,668	352,668	352,668

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All specifications include a 4th order polynomial in log(age), 21 region dummies, dummies for suburban and urban regions, dummies for mother tongue, and dummies for parental decade of birth.

income and the individual's probability of becoming an inventor in Figure 7 (upper curve).[11] This curve mirrors the non-parametric Figure 1.

[11] We set the probability of becoming an inventor, conditional on having a father in the lowest income quintile, to the sample mean of those individuals (0.00675).

Figure 7: Decomposing the Impact of Father's Income



Notes: The figure displays the estimated probability to invent conditional on father's income quintile, based on the regression results reported in Table 1. The probability to invent conditional on the father being in the lowest quintile (base group in the regression) is the sample mean for individuals with a father in that income quintile.

The positive association of parental income on the probability of becoming an inventor can emerge through a number of channels. A first channel is that high-income parents typically enjoy a higher socioeconomic status (Figure B6 in the Appendix). Socioeconomic status broadly mirrors a family's economic and social resources, including the parents' work experience, position in the labor market and social networks. All these may influence a child's likelihood of becoming an inventor. Thus in Column 2 of Table 1 we add controls for the socioeconomic status of father and mother. We see that having the father or mother being a white collar has a positive and significant effect on the individual's probability of being an inventor, the effect being stronger if the parent is a senior rather than a junior white collar worker, and having a blue-collar parent has a negative effect. The impact of having a senior white collar father is about the same as having a father in the income percentile 91-95. Moreover, the coefficients of parental income are reduced by 40% or more compared to column 1, the exception being the coefficient of the mother being in the top-5% of the income distribution which is reduced by only 17%.

The overall relationship between father income and the individual's probability of inventing, based on results in column 2, is captured by the second highest curve in Figure 7: we see that this curve is somewhat less steep than the upper curve obtained by regressing the probability of inventing on father income only. The curve flattens at the higher income levels, and the estimated probability of becoming an inventor conditional on having a father in the top-5% of the income distribution is reduced from 3% to 2%.

A second channel is that higher-income parents tend to be more educated (Figure B8 in the Appendix). More educated parents may invest more money and effort to educate their kids, thereby inducing them to become inventors with a higher likelihood. Descriptive statistics support this explanation, particularly for parents with a science education. Thus in Column 3 of Table 1 we add variables capturing parental education. We see that having a father with a PhD has a direct and important impact on the probability of making an invention of 3 percentage points. The effects from having a parent with a master's degree are also sizeable, above 1 percentage point.[12] The impact of parental STEM education is almost 1 percentage point for fathers, but zero for mothers. Second, controlling for parental education reduces the effect of the father belonging to the highest 96-100 income bracket by a further 40%, and renders the mother income effects small in absolute value and statistically not significant. The fact that father's and mother's income coefficients diverge suggests that our income measures (after controlling for other parental characteristics) capture not only pure financial resources of the family, but also some other aspects. Finally, we note that the introduction of parental education reduces the impact of having a senior white collar parent by half.

As can be seen from Figure 7, the relationship between father's income and the probability of inventing becomes markedly less steep after adding variables capturing not only the socio-economic status of parents, but also their education. The estimated effect of having a father in the top-5% of the income distribution has halved from round 3% to round 1.5%.

A third potential channel for the positive relationship between parental income and the individual's probability of inventing, could be that higher income parents have higher IQ children and that higher IQ children are more likely to innovate.[13] Figure

[12]The estimated effect from having a mother with a PhD is not significant, most likely due to the small number (1.4 percent of observations) of invididuals with a PhD mother in our sample.

[13]The relationship between parental income and the individual's IQ (Figure B10) may in turn reflect both socioeconomic (Figures B11 and B12) and genetic considerations, see e.g. Pinker (2005).

6 strongly suggests that the individual's IQ is positively correlated with his probability of innovating.

To take invididual ability into account, we add measures of the individual's IQ in Column 4 of Table 1. IQ has a direct effect on the probability of becoming an inventor. Being in the 91st-95th or the 96th-100th percentile of the IQ distribution increases the probability of inventing by 2-3 percentage points. This is an economically significant impact, on par with the impact of having a father with a PhD. Second, controlling for IQ further reduces the effect of parental income on the probability of becoming an inventor again by a further 25% relative to that seen in Column 3. Overall, the estimated impact of having a father in the top 5% of the income distribution has been reduced to one third of the estimate in column 1 by the inclusion of parental socioeconomic status, parental education, and own IQ. The impact of mother's (high) income became insignificant already after including parental socioeconomic status and education.

Consequently, going again back to Figure 7, we see that the curve depicting the relationship between father income and the probability of becoming an inventor further shifts down when controlling for the individual's IQ.

The above results are robust to (see Tables C2-C5 in the Appendix): (i) using patent counts as the dependent variable in Table C2; (ii) using the number of citations to the highest cited patent of an individual in the first 5 years of patent life to account for patent quality in Table C3; iii) using an indicator variable that takes value one for star inventors, defined as having an invention the citations to which are in the top-5% of the citation distribution, and is zero otherwise, as the dependent variable in Table C4; (iv) using parental income measured as an average over several years as the basis for creating parental income percentile variables in Table C5;[14] and (v) introducing analytic and verbal IQ measures, modeling them in similar fashion to the visuospatial IQ in Table C6. Regarding the last robustness test, we find that coefficients of parental characteristics are further reduced, but not by much. The coefficients for visuospatial IQ go down as expected (e.g., the coefficient of being in the top-5% is reduced from 0.035 to 0.022), and the coefficients of verbal and analytic IQ are somewhat lower than those of visuospatial IQ (e.g., that of being in the top-5% are 0.022, 0.015 and 0.019 for visuospatial, verbal

[14]The alternative income measure is calculated as follows: for parents born before 1921, we use the 1975 deflated wage; for parents born 1921 - 1950, we calculate the wage as the average of their (deflated) wage in 1975 and 1985; for parents born between 1951 and 1955, we take their wage in 1985; for parents born 1955-1960, we take the average of their wage in 1988-1990 (1988 is the first year of FLEED); and for parents born thereafter, we take the average wage in years when they are age 28-30. We then regress (logs of) these wage measures on year-of-birth dummies and take the residuals. We then use these residuals to generate the percentile wage ranks.

and analytic IQ respectively).

Summarizing, we find evidence that although the effect of parental income is much smaller than what Figure 1 would suggest, it nonetheless exists even in as equitable an environment as Finland. To illustrate this, compare two individuals, one with a father in the lowest income quintile, one with a father in the top-5% of the income distribution. Ceteris paribus, to have (at least) the same probability to become an inventor, the former individual would have to be 3 deciles (30 percentiles) higher in the IQ distribution when the latter is in our base IQ category (4th IQ quintile. See Appendix Table C1 for the coefficients used in the calculation). Similarly, to compensate for the difference a senior whitecollar father makes compared to a bluecollar father, an individual would have to be 20 percentiles higher in the IQ distribution (comparing the 4th decile to the 6th). Finally, the compensating IQ differential for the difference that a PhD father makes compared to a father with a base education is 50 IQ percentiles.

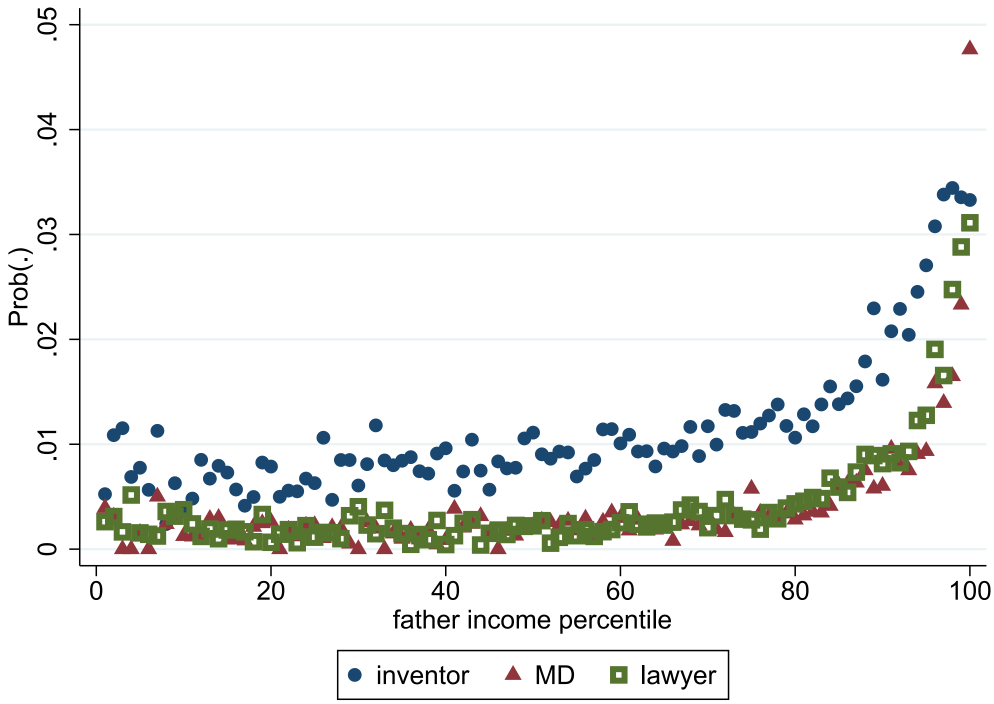
3.3 Becoming an inventor vs. becoming a lawyer or a medical doctor

To which extent what we said above regarding the determinants of becoming an inventor, should not equally apply to other high-ability professions such as lawyer or medical doctor? In this subsection we perform the same regression exercises as in the previous subsection, but replacing the indicator variable of becoming an inventor on the left-hand side of the regression equation by the indicator variable of becoming a medical doctor or a lawyer.

A first remark: in our cross-section data sample, 1.13% of individuals are inventors, whereas 0.48% are medical doctors and 0.49% are lawyers. This information will help us compare the magnitudes of the effects of parental income, parental education, and IQ on the probability of becoming a lawyer or a medical doctor with the magnitudes of the effects of the same variables on the probability of becoming an inventor. For example, if we find the same coefficient for parental education in the regression tables for becoming an inventor as in the regression tables for becoming a lawyer, that will mean that the actual effect of parental income is $1.13/.48 \approx 2.35$ higher on the probability of becoming an inventor than on the probability of becoming a medical doctor.

Figure 8 shows the three curves depicting respectively the probability of becoming an inventor, the probability of becoming a medical doctor and the probability of becoming a lawyer, as a function of father income, not controlling for any other characteristic. We see that all three curves have similar shapes, with the same non-linear effect which becomes steeper at the highest levels of father's income. However the probability of becoming an inventor lies significantly above the probabilities of becoming a lawyer or a medical doctor until we reach the highest father income percentiles. In other words, becoming an inventor is easier than becoming a lawyer or a medical doctor at all except the highest father income percentiles.

Figure 8: Father's Income and Becoming an Inventor, Doctor, or Lawyer



Notes: The figure displays the probability to invent, to become an MD, and to become a lawyer, all as functions of father's income percentile. Father's income is measured in 1975 for individuals born 1961-1975, and in 1985 for individuals born in 1976-1985. We calculate the percentile ranks using residuals from a regression of the natural log of deflated income of fathers and mothers on year-of-birth dummies.

Table 2 shows the R-squared decompositions for the probability of becoming an inventor, the probability of becoming a medical doctor, and the probability of becoming a lawyer, respectively.[15] In particular we see that IQ is by far the main characteristic for the probability of becoming an inventor in terms of the share of variation it explains, followed by parental education. These two groups of variables account for 66% and 16% of the overall variation captured by our model. In contrast, IQ plays a relatively speaking

[15] We report the regression results using dummies for becoming and MD or a lawyer as the dependent variables and the specifications used in Table 1, in the Appendix Table C7 and Table C8.

much more minor role for becoming a medical doctor or a lawyer. Parental education is the main explanatory variable for the probability of becoming a medical doctor or a lawyer (40% and 53%), with base controls and parental income also playing clearly more important roles than for inventors.

Table 2: Decomposing the Explained Impact on Becoming an Inventor

	– A. Partial R-squared –		
Explanatory variables	**Inventor**	**MD**	**Lawyer**
Base controls	0.002	0.004	0.002
Parental income	0.000	0.001	0.001
Parental socecon	0.000	0.000	0.000
Parental education	0.002	0.004	0.003
IQ	0.008	0.001	0.000
Sum of partial R-sq's	0.012	0.010	0.006
R-sq	0.017	0.018	0.013
	– B. Fraction of Partial R-squared –		
Explanatory variables	**Inventor**	**MD**	**Lawyer**
Base controls	0.148	0.418	0.263
Parental income	0.017	0.082	0.140
Parental socecon	0.017	0.020	0.018
Parental education	0.157	0.398	0.526
IQ	0.661	0.082	0.053

Notes: The upper panel displays the partial R-squared for a given dependent variable (column) and given vector or explanatory variables (row), their sum, and the R-squared of the estimation. The used specification is that in column 4 of Table 1. The lower panel displays the share of partial R-squared obtained for a given dependent variable (column) by a given vector of explanatory variables. For example, the 0.148 for Base controls for Inventor in the lower panel for Inventor is obtained by dividing 0.002 (upper panel, Base controls) by 0.012 (Sum of partial R-sq's). Base controls are: a 4th order polynomial in log(age), 21 region dummies, dummies for suburban and urban regions, dummies for mother tongue, and dummies for parental decade of birth. We follow Bound et al. (1995) in calculating the partial R-squared.

To further illustrate the economic significance of the different family characteristics and own ability, we perform a back-of-the-envelope calculation. We look at how much the overall probability of inventing would increase if: (i) all individuals had a father in the top income decile; (ii) all individuals had a father who is a senior white collar worker; (iii) all individuals had a father who obtained at least a master's degree ; (iv) all individuals belonged to the highest IQ decile.

The corresponding results are shown in Columns 1-4 of Table 3, where 100 is the base (pre-reallocation) value. If all individuals had a father in the top income decile,

Table 3: Counterfactuals with Father's Income, Status, Education, and Own IQ

Outcome		Data	Father income highest 10%	Father white collar mngr.	Father with at least MSc	IQ highest 10%
Inventor	mean	0.013	0.017	0.016	0.029	0.038
	% change	100	128	117	216	283
MD	mean	0.005	0.008	0.005	0.014	0.009
	% change	100	172	110	291	186
Lawyer	mean	0.005	0.010	0.006	0.016	0.005
	% change	100	181	112	288	100

Notes: In the "Data" column we display the mean predicted probability from our main specification (column 4 in Table 1). In the Father income column, we randomly allocated those whose fathers are not in the top decile to either the 91st-95th or the 96-100th percentiles. In the Father white collar mngr. column, we change all those fathers who are not white collar managers to being white collar managers. In the Father education columns, we change all those with a father with less than an MSc to the category of father having an MSc. In the IQ column, we randomly allocated those who are not in the top IQ decile to either the 91st-95th or the 96-100th percentiles. The row % change reports the change compared to the Data column.

the probability of becoming an inventor would increase by nearly a third, whereas the probabilities of becoming a medical doctor or a lawyer would increase by much more (respectively by 72% and 78%). If everybody had senior white collar fathers, the increases would be more modest at round 10-17%. Granting everyone fathers with a master's degree would have a large impact, increasing the probability of becoming an inventor by more than 100% and those of becoming a medical doctor or a lawyer much more, by almost 200%. In contrast, if all individuals belonged to the highest IQ decile, the probability of inventing would increase by an additional 183% whereas the probability of becoming a lawyer would remain unchanged and the probability of becoming a medical doctor would increase by (only) 86%. This exercise further underlines the result that parental income and parental education are more and own IQ less important in becoming a medical doctor or a lawyer than in becoming an inventor.

4 Endogeneity of IQ and close brothers

One might worry about the possible endogeneity of IQ in our regressions. For example, it could be that better educated and/or higher income parents provide a better environment for their kids. Such differences could not only have a direct impact on the probability of becoming an inventor as our results suggest, but they could also have a positive impact on IQ, rendering IQ endogenous (recall the evidence in subsection 2.3)

To ameliorate endogeneity concerns, we use visuospatial IQ instead of "regular" IQ. There is some evidence that visuospatial IQ reflects innate ability: for example Pekkala Kerr et al. (2013) find no effect of schooling on visuospatial IQ based on FDF data. Using Swedish Defense Forces (SDF) data, Carlsson et al. (2015) find no effect of schooling on visuospatial IQ, and Cesarini et al. (2016) find no causal impact of parental wealth on the cognitive skills of the children in the family, using Swedish contemporary lotteries.

Nonetheless, we take one further step in this section to address the issue of the potential endogeneity of IQ. We look at the effect of an IQ differential between the individual and close brother(s) born at most three years apart.[16] This allows us to include family fixed effects and thereby control for family-level time-invariant unobservables, such as genes shared by siblings, parenting style, and fixed family resources. Table 4 shows the results from the regression with family-fixed effects. The first column shows the baseline OLS results using the sample on brothers born at most three years apart. Notice that we include a dummy for the individual being the first born son in the family to account for birth-order effects. The second column shows the results from a regression where we introduce family fixed effects. We lose other parental characteristics than income due to their time-invariant nature.[17] The main finding in Table 4 is that the coefficients on "IQ 91-95" and "IQ 96-100" in Column 2 (i.e. when we perform the regression with family fixed effects) are the same as in the OLS Column 1. This suggests that these coefficients capture an effect of IQ on the probability of inventing which is largely independent of unobserved family background characteristics, as otherwise the OLS coefficients would be biased and different from the fixed effects estimates.[18]

[16]Ideally, we would have liked to restrict attention to twin brothers, but then we run into a small sample problem as there are very few innovators with twin brothers in Finland over our sample period. As robustness tests, we used samples with brothers at most zero and at most one year apart. The results in Table 4 survive the time-window of one year, but not that of zero years age difference. We prefer the time window of at most three year age difference as the OLS results remain similar to those obtained using our full sample. See Table D1 in the Appendix for the results of the robustness tests, as well as for the full set of coefficients for the regressions reported in Table 4.

[17]Admittedly, over-time variation is limited even for parental income and therefore a possible explanation for the insignificant coefficients.

[18]The reader may wonder whether this latter conclusion is consistent with the recent psychology literature pointing at both, a genetic and a social component of IQ (e.g. see Mc Gue et al., 1993; Ceci, 2001; Pinker, 2005; and Plomin and Spinath, 2004). Note first that here we are focusing at visuospatial IQ, which is supposed to more independent from family and social factors than verbal and quantitative IQ indicators. Second, here we are focusing on the effect of IQ on the probability of inventing rather than on the determinants of IQ per se.

Table 4: Comparing Close Brothers

	(1)	(2)
first born	-0.00209**	-0.000933
	(0.000869)	(0.00139)
fa income 91-95	0.00277	-0.0101
	(0.00231)	(0.0212)
fa income 96-100	0.0113***	-0.0272
	(0.00292)	(0.0276)
mo income 91-95	0.00375	-0.00512
	(0.00790)	(0.0481)
mo income 96-100	0.00393	0.00693
	(0.00269)	(0.00778)
fa bluecollar	0.000190	
	(0.00114)	
fa jr whitec.	0.00381**	
	(0.00168)	
fa sr whitec.	0.00631**	
	(0.00256)	
mo bluecollar	-0.00127	
	(0.00118)	
mo jr whitec.	0.00174	
	(0.00142)	
mo sr whitec.	-0.000173	
	(0.00279)	
fa MSc	0.000658	
	(0.00370)	
fa PhD	0.0281***	
	(0.00915)	
mo MSc	0.0139***	
	(0.00524)	
mo PhD	-0.0166	
	(0.0147)	
fa STEM	0.0101***	
	(0.00179)	
mo STEM	-0.000522	
	(0.00166)	
IQ 91-95	0.0216***	0.0202***
	(0.00309)	(0.00409)
IQ 96-100	0.0353***	0.0320***
	(0.00355)	(0.00457)
Observations	82,054	82,054
Number of families		41,605

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All specifications include a 4th order polynomial in log(age), 21 region dummies, dummies for suburban and urban regions, dummies for mother tongue, and dummies for parental decade of birth. The estimation sample consists of all brothers born at most 3 years apart.

5 The role of family structure

In the previous section we identified an effect of IQ on the probability of inventing which was largely independent from family characteristics. Adverse shocks to family conditions and structure, such as divorce or health problems, may results in a a less-than ideal environment for children to develop theri knowledge and skills. We therefore focus in this section attention on the role of family structure by comparing individuals who grow up with their biological parents, individuals that do not grow up with at least one biological parent, and individuals that grow up with a non-biological ("step") parent.[19]

Figure 9 shows scatter-plots of the probability of inventing as functions of the income of the biological father and the income of the step ("social") father respectively. We see that both curves have the same "J-form" shape, with the step-father curve lying slightly below the biological father curve.

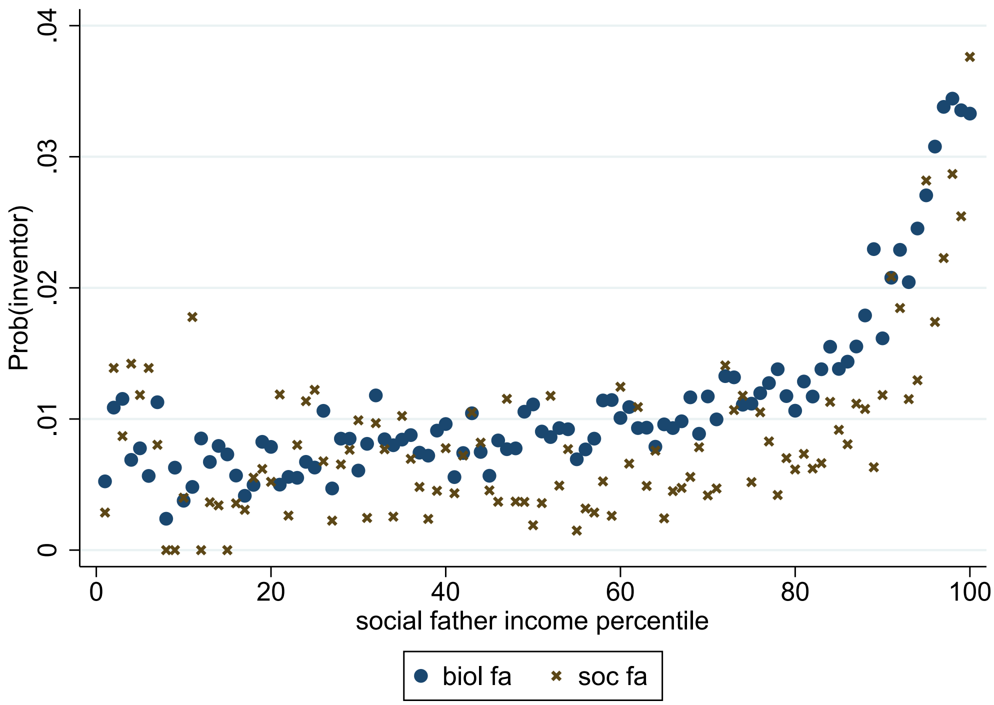
This is however not enough to conclude that the income of the step parent should matter, and to a similar extent as the income of the biological parent. We therefore augment our base regression by introducing: (i) indicator variables for different family structures (the base category is living with both biological parents); (ii) interactions between the income and education measures of biological parents and indicator variables for the individual not growing up with the biological parents; and (iii) income variables for step parents living with the child. We introduce the interactions only for the highest income and education dummies. The results from this extended regression are shown in Table 5.

Column 1 of Table 5 brings in family structure dummies into the specification. We see a negative and significant effect of not living with one or the other the biological. The estimated effects are not small when compared to the average probability of inventing of 0.013. Furthermore, the coefficients of parental income, social status and education as well as those for IQ are essentially unchanged compared to those reported earlier in Table 1 (the full results are reported in Table E1 in the Appendix).

Columns 2 to 4 show results from the same regression where we introduce the in-

[19]We utilize the 1975 and 1985 census to construct the family structure dummies. We use the 1975 census for the individuals born by that year, and the 1985 census for others (to maintain comparability). It is very likely that even those individuals that we categorize as not living with one or the other biological children have actually lived with both of them for some period in their lives. About 95% of individuals in our sample lived with their two biological parents. The rest lived at least for a while without their both biological parents at some point before their fifthteenth birthday. 3% lived with the biological mother but without the biological father; 1% lived with the biological father but not the biological mother; and less than 1% lived with neither of the biological parents.

Figure 9: Biological and Social Father's Income and Becoming an Inventor



Notes: The figure displays the probability to invent as a function both of the biological and the step father's income percentile. Fathers' incomes are measured in 1975 for individuals born 1961-1975, and in 1985 for individuals born in 1976-1985. We calculate the percentile ranks using residuals from a regression of the natural log of deflated income of fathers and mothers on year-of-birth dummies.

teractions of parental income and education with dummies for not growing up with the biological parent or for growing up with a step parent. In Column 2 the positive and significant coefficient of 0.008 on "biol fa income 96-100" captures the impact of growing up with a father who is in the top 5% in the income distribution. The coefficient of -0.012 for "biol fa income 96-100 x away", the interaction between the biological father being in the top-5% and the individual not growing with him, suggests that the positive direct impact of a high income father only materializes if the individual grows with the biological father. In Column 3 we introduce income variables for the step parents. These obtain negative coefficients throughout, suggesting that step parent income at best plays no role in leveling the road towards innovation. The coefficients on biological parents' income hardly budge after the introduction of step parent income variables. Finally, in Column 4 we bring interactions between biological parent education and family structure dum-

Table 5: Role of Family Structure and Resources

	(1)	(2)	(3)	(4)
biol fa away	-0.00399***	-0.00309***	-0.00311***	-0.00295***
biol mo away	-0.00343**	-0.00410**	-0.00398**	-0.00417**
biol fa & mo away	-1.27e-05	0.00116	0.00107	0.00126
91-95	0.00500***	0.00528***	0.00577***	0.00574***
biol fa income 96-100	0.00730***	0.00772***	0.00845***	0.00836***
biol mo income 91-95	-0.00137	-0.000708	-0.000442	-0.000549
biol mo income 96-100	0.000214	0.000258	0.000772	0.000756
biol fa income 91-95 x away		-0.00625*	-0.00669*	-0.00613*
biol fa income 96-100 x away		-0.0118**	-0.0125***	-0.00993**
biol mo income 91-95 x away		-0.0148***	-0.0152***	-0.0141***
biol mo income 96-100 x away		-0.000510	-0.00106	-0.000925
step fa income 91-95			-0.00327	-0.00329
step income 96-100			-0.00501*	-0.00504*
step mo income 91-95			-0.00381	-0.00344
step mo income 96-100			-0.0191**	-0.0190**
fa bluecollar	-0.000861	-0.000830	-0.000825	-0.000826
fa jr whitec.	0.00173**	0.00175**	0.00175**	0.00174**
fa sr whitec.	0.00261**	0.00257**	0.00258**	0.00255**
mo bluecollar	5.85e-05	3.07e-05	7.68e-05	8.20e-05
mo jr whitec.	0.00140**	0.00137**	0.00141**	0.00142**
mo sr whitec.	0.00326***	0.00314**	0.00316**	0.00315**
biol fa MSc	0.00874***	0.00874***	0.00880***	0.00884***
biol fa PhD	0.0275***	0.0275***	0.0275***	0.0278***
biol mo MSc	0.0117***	0.0117***	0.0121***	0.0125***
biol mo PhD	0.00794	0.00808	0.00908	0.0110
biol fa STEM	0.00860***	0.00859***	0.00855***	0.00854***
biol mo STEM	-0.00111	-0.00112	-0.00111	-0.00113
biol fa MSc x away				-0.000712
biol fa PhD x away				-0.0128
biol mo MSc x away				-0.00776
biol mo PhD x away				-0.0346***
IQ 91-95	0.0236***	0.0236***	0.0235***	0.0235***
IQ 96-100	0.0351***	0.0351***	0.0350***	0.0351***
Observations	352,668	352,668	352,668	352,668

Notes: Robust standard errors are reported in Appendix Table A1 to save space. *** p<0.01, ** p<0.05, * p<0.1. All specifications include a 4th order polynomial in log(age), 21 region dummies, dummies for suburban and urban regions, dummies for mother tongue, and dummies for parental decade of birth.

mies into the specification. We find that all of them, with the exception of the negative and significant mother's PhD interaction, carry insignificant coefficients.[20] Overall, these

[20] As discussed earlier, the very small share of mothers with a PhD renders the mother PhD coefficients

results suggest that the effects of father income on the probability of becoming an inventor are conditional on the father living with the individual, whereas this is not the case for the effects of parental education. It is also noticeable that the coefficients of the other variables are hardly affected by the introduction of the family structure dummies, nor by the introduction of the interactions.

The result that parental education matters even if the child does not grow up with her biological parents may appear at first sight in contradiction with our above finding that parental education partly underlies the positive correlation between parental income and the probability of inventing. However, the following considerations help reconcile this result with our overall analysis and findings. First, parental education may partly reflect parental ability which in turn may be genetically transmitted to the child. Second, even if he/she does not live with the child, the biological parent may still serve as a role model: for example, having a parent with a PhD in Science may encourage the child to pursue a similar curriculum. Third, even if she does not live with the child, the biological parent may still follow and monitor the child's studies, which helps the child complete higher education even though she is still losing, if only emotionally, from not living full time with her biological parent.

These results suggest that even children born to high income fathers but not living with them do not benefit from the increased probability of inventing that the father's high income could bring about. The results however also point that family breakdown does not affect innovation so much through the loss of the positive impact of parental education.

6 The interaction between IQ and parental characteristics

In the previous two sections we have shown that IQ has an impact on the probability of inventing independently from family characteristics (Section 4), and that family structure matters on top of IQ (Section 5). While there are many interactions we could look at, in this section we ask whether own IQ and parental income and parental education are complements or substitutes. High IQ children may benefit more from the environment that higher parental income and better education bring about, leading to complementarity. Alternatively, parents may use increased resources to even out differences in children's innate ability, leading to substitutability.

Table 6 shows the results from a regression where we introduce interactions between

somewhat suspicious.

the indicator for being in the top 5% of the IQ distribution, and the two highest parental income and education indicators (the full results are reported in Table F1 in the Appendix). We first observe that the direct effects of the interacted variables remain essentially unchanged compared to the results reported in Table 1, as do the coefficients of other variables. Turning to the interactions, we find only one significant coefficient, that between high IQ and father being in the top 5% of the income distribution. This effect survives at a weaker level of statistical significance, to the use of the close brothers sample and the introduction of family fixed effects (Columns 3- 6). Higher IQ individuals seem to benefit more from high father income.

Table 6: Potential Misallocation

	(1)	(2)	(3)	(4)	(5)	(6)
fa income 91-95	0.00527***	0.00515***	0.00527***	-0.00979	-0.0102	-0.00984
fa income 96-100	0.00617***	0.00745***	0.00615***	-0.0280	-0.0273	-0.0281
mo income 91-95	-0.00192	-0.00185	-0.00192	-0.00368	-0.00522	-0.00403
mo income 96-100	-0.000202	-0.000400	-0.000231	0.00561	0.00693	0.00562
fa bluecollar	-0.000793	-0.000761	-0.000794			
fa jr whitec.	0.00182**	0.00183**	0.00182**			
fa sr whitec.	0.00265**	0.00270**	0.00265**			
mo bluecollar	2.87e-05	4.40e-05	2.87e-05			
mo jr whitec.	0.00146**	0.00147**	0.00146**			
mo sr whitec.	0.00331***	0.00334***	0.00331***			
fa MSc	0.00862***	0.00877***	0.00862***			
fa PhD	0.0272***	0.0266***	0.0273***			
mo MSc	0.0119***	0.0119***	0.0118***			
mo PhD	0.00809	0.0112	0.0114			
fa STEM	0.00856***	0.00861***	0.00856***			
mo STEM	-0.00116	-0.00116	-0.00116			
IQ 91-95	0.0237***	0.0236***	0.0237***	0.0204***	0.0203***	0.0204***
IQ 96-100	0.0331***	0.0350***	0.0331***	0.0268***	0.0319***	0.0269***
fa inc 96-100 x IQ 96-100	0.0144***		0.0147***	0.0256*		0.0270*
mo inc 96-100 x IQ 96-100	-0.00358		-0.00275	0.0339		0.0336
fa PhD x IQ 96-100		0.00596	-0.000873		-0.00299	-0.0174
mo PhD x IQ 96-100		-0.0178	-0.0205		0.0481	0.0401
Sample	IQ	IQ	IQ	Brothers	Brothers	Brothers
Estimator	OLS	OLS	OLS	FE	FE	FE
Observations	352,668	352,668	352,668	82,054	82,054	82,054
Number of families				41,605	41,605	41,605

Notes: Robust standard errors are reported in Appendix Table A2 to save space. *** p<0.01, ** p<0.05, * p<0.1. All specifications include a 4th order polynomial in log(age), 21 region dummies, dummies for suburban and urban regions, dummies for mother tongue, and dummies for parental decade of birth. In columns (1)-(3) the sample is the IQ sample used in Table 1. In columns (4)-(6) the sample is the brothers sample used in Table 4.

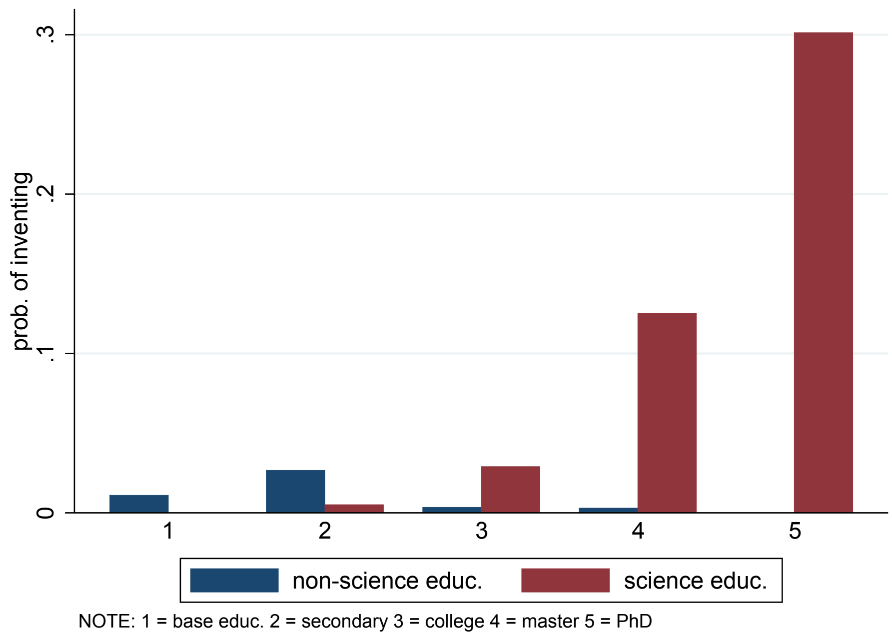
The positive interaction coefficient suggests a misallocation: namely, a positive fraction of potential inventors with very high IQ will not perform as well as they could, due to insufficient parental resources. To better illustrate this point, consider two individuals A and B whose fathers belong to the lowest income quintile. Individual A has average IQ, i.e. a visuospatial IQ in the 4th percentile. Individual B has a top IQ, i.e. a visuospatial IQ in the 96-100 IQ range. According to Table 6, reallocating individual A to a father with wage income in the 96-100 income range, will increase A's probability of inventing by 0.00617 (Column 1). In contrast, reallocating individual B to a father with wage income in the 96-100 income range, will increase B's probability of inventing by 0.00617 + 0.0144 = 0.0206. The ratio between these two probabilities is equal to 0.0206/0.00657 which is approximately equal to 3.3; this ratio, minus one, measures the extent of misallocation.This calculation suggests that high IQ individuals are particularly affected by misallocation. Inadequate parental resources may thus be disproportionately harmful for the highly talented, eroding the utilization of the innovative potential of the economy.

7 Own education

One particular channel whereby parental income may affect the individual's probability of inventing, and through which parental income and IQ may interact, is the individual's own education. Figure 10 shows that completing a STEM master's degree (and a fortiori a PhD) is associated with a much higher probability of inventing. Using our estimation sample, we find that the probability of inventing conditional on a STEM MSc is nearly four times as large as that of having a father in the top-5% of the income distribution or having a white collar father or a mother; 50% higher than having a PhD father or mother, and almost 100% higher than having an IQ at the top of the IQ distribution. The effect of having a STEM PhD in turn is more than twice as large as that of having a STEM MSc.

Education is not randomly distributed. The majority of individuals in our data (Figure G1 in the Appendix) have a secondary education as their highest education (measured in the year they turn 35); and some 10% have a master's degree or a PhD. The probability of obtaining a STEM MSc (Figure G2 in the Appendix) displays a similar convex increasing pattern as a function of parental income as the probability of inventing (or becoming a medical doctor or a lawyer). Education is also increasing in parent's socioeconomic status (Figure G3) and education (Figure G4). Finally, own IQ and education are also strongly positively correlated (Figure G5). To explore these relations further, we regress a dummy that takes value one if the individual has at least a mas-

Figure 10: Own Education and Probability to Invent



Notes: The figure displays the probability to invent conditional on the education of the individual. We divide education into five groups by level of education: base education (up to 9 years, depending on age of parent), secondary, tertiary, MSc, and PhD. We also condition all other levels of education but base education on a parent having a STEM education. A STEM base education does not exist. We measure education at age 35.

ter's degree on the same variables we've used hitherto in our analysis. We then execute the same R-squared decomposition exercise as above for becoming an inventor, a medical doctor, or a lawyer. The results, displayed in full in Tables G1 and G2 in Appendix G, show that parental income, parental socioeconomic status and parental education are all positively associated with obtaining an MSc, as is own IQ. In terms of magnitudes, the effect of having a father in the top-5% of the income distribution has, at 6.7 percentage points (Column 4 in Table G1), twice the effect of having a white collar senior father, but less than half the impact of having a father with a master's degree and only a third of the impact of a PhD father; the effects we estimate for mother's characteristics are smaller throughout.[21] The effect of being in the top-5% of the IQ distribution is at

[21] In Table G1 we also report estimates using the sample on brothers close in age. The first born son is about one percentage point less likely to obtain an MSc. Unsurprisingly, parental income variables lose significance once we employ family fixed effects. Both IQ and education results remain when we use the sample with brothers with at most one year age difference (Columns 3 and 4), and the education results

17 percentage points commensurate to having a father with a master's degree. When we look at the partial R-squared contributions (see Table G2 in Appendix), we find that own IQ dominates with a share of more than 40%, followed by parental education (27%) and base controls (26%), with parental income and parental socioeconomic status each contributing less than 5%.

We test the "own education" channel to invention in Table 7. Table 7 is identical to Table 1, except that we added indicator variables capturing the level and type (STEM or not) of the individual. In contrast to parental education, we make use of the full set of level and type of education dummies; as with parental education, base education is the base category. A first finding from this table is that having a Master or a PhD in Science has a much bigger effect on the probability of inventing than any other variable. Second, we see that once we control for the individual's own education, the effect of all other variables are reduced; for example, now neither the father's nor the mother's socioeconomic status has an impact on the probability of becoming an inventor. The effect of having a STEM MSc is 25 times as large as that of having a father in the top-5% of the income distribution, 10 times as large as having a PhD father, and six times as large as the impact of being in the top-5% of the IQ distribution. The effect of a STEM PhD is more than twice as large as that of a STEM MSc.

When we depict the regression on the probability of inventing on parental income (see Figure 11), we see that once we control for the individual's own education, the relationship between father income and the probability of inventing becomes almost flat, with only a small blip at the highest father income levels.

A last finding from Table 7, is that the impact of IQ remains positive and significant even when we control for the individual's own education. Overall, the above findings suggest a prominent role for own education and for IQ when explaining an individual's probability of becoming an inventor.

To further test this conjecture, we now compute partial R^2's in order to assess the relative explanatory powers of the observable background variables in our data sample. The findings, summarized in Table 8 below, indicate that out of the variation in the probability of becoming an inventor which we can explain using all our observed variables: (i) the individual's own education comes first, explaining almost 97% of that variation; (ii) second comes the individual's IQ (2.0%); and (iii) each of the remaining variables accounts for at most 0.6% of the total explained variation in the probability of becoming

remain even when we use the sample of brothers born in the same year (Columns 1 and 2). The OLS coefficients on IQ indicators for higher IQ deciles display an upward bias of round 30-40%. This suggests that IQ is correlated with the family-level unobservables affecting the probability of obtaining an MSc.

Table 7: Role of Own Education

fa income 91-95	0.00224**	fa MSc	0.000430
	(0.00105)		(0.00170)
fa income 96-100	0.00404***	fa PhD	0.00974**
	(0.00126)		(0.00420)
mo income 91-95	-0.00189	mo MSc	0.00129
	(0.00305)		(0.00234)
mo income 96-100	-0.000279	mo PhD	-0.00546
	(0.00108)		(0.00911)
fa bluecollar	-0.000736	fa STEM	0.00460***
	(0.000526)		(0.000771)
fa jr whitec.	-1.99e-05	mo STEM	-0.000634
	(0.000715)		(0.000711)
fa sr whitec.	0.000491	IQ 91-95	0.0103***
	(0.00109)		(0.00136)
mo bluecollar	0.000166	IQ 96-100	0.0157***
	(0.000537)		(0.00156)
mo jr whitec.	0.000315	STEM MSc	0.104***
	(0.000604)		(0.00227)
mo sr whitec.	0.000723	STEM PhD	0.225***
	(0.00121)		(0.00897)
Observations: 352,668			

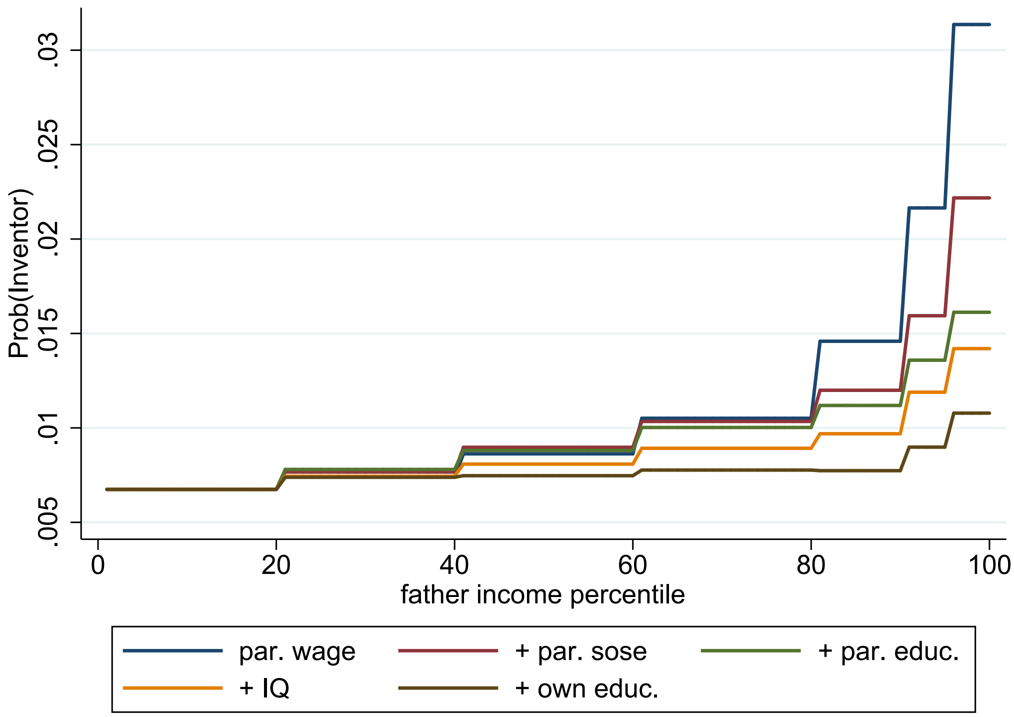
Notes: *** p<0.01, ** p<0.05, * p<0.1. All specifications include a 4th order polynomial in log(age), 21 region dummies, dummies for suburban and urban regions, dummies for mother tongue, and dummies for parental decade of birth.

an inventor, with the share of parental income and socioeconomic status reduced to zero. Coupling this with our earlier finding that IQ is the dominant source of explained variation for education provides further proof for the importance of IQ and own education. Consistent with this, Table 9 shows that if all individuals had achieved at least a STEM Master's degree, the aggregate probability of inventing would be multiplied by 8.11.

An issue with own education is endogeneity. Toivanen and Väänänen (2016) instrument university (engineering) education with the distance to the nearest technical university. They report OLS coefficients of the order of 0.02 - 0.05, and IV estimates of the order 0.09 - 0.16; i.e., they observe a substantial downward OLS bias. We can control for family background and own ability in particular in a much richer way than Toivanen and Väänänen (2016) and find results that are in line with their OLS estimates. This suggests that our results are unlikely to be upward biased.

This notwithstanding, we re-estimated our extended regression model which includes own education measures using the brothers with at most three years age differ-

Figure 11: Including Own Education in Decomposition



Notes: The figure displays the estimated probability to invent conditional on father's income quintile, based on the regression results reported in Tables 1 and 7. The probability to invent conditional on the father being in the lowest quintile (base group in the regression) is the sample mean for individuals with a father in that income quintile.

ence - subsample and family fixed effects. These results are reported in Table 10. They show that 1) our full sample OLS results (Table 7) are replicated (Column 5 in Table G3); 2) there is no evidence of a birth-order effect; and 3) that we reproduce the IQ and own education coefficients reported in Table 7 even after the introduction of family fixed effects (Column 6 in Table G3).[22] The last result suggests that, just like IQ, own education is not subject to endogeneity bias arising from unobserved family level effects.[23]

[22]This is true also when we include in unreported regressions the analytic and verbal IQ measures. With own education variables included, verbal IQ indicators no longer obtain significant coefficients even when using OLS, and the point estimates are further reduced with family fixed effects. Those of (high) visuospatial and analytic IQ remain significant even with family fixed effects and do not change appreciably from OLS.

[23]We report robustness tests using estimation samples with brothers with at most one and zero year age difference in Table G3 in Appendix G. These robustness tests show that while parental characteristics lose their significance even with OLS estimation, the IQ results are replicated for the one year age difference

Table 8: Decomposition with Education

Explanatory variables	Inventor
– A. Partial R-squared –	
Base controls	0.0004
Parental income	0.0000
Parental socecon	0.0000
Parental education	0.0003
IQ	0.0013
Education	0.0602
Sum of partial R-sq's	0.0622
R-sq	0.0766
– B. Fraction of Partial R-squared –	
Base controls	0.0064
Parental income	0.0000
Parental socecon	0.0000
Parental education	0.0048
IQ	0.0209
Education	0.9678

Notes: The upper panel displays the partial R-squared for a given dependent variable (column) and given vector or explanatory variables (row), their sum, and the R-squared of the estimation. The lower panel displays the share of partial R-squared obtained for a given dependent variable (column) by a given vector of explanatory variables. For example, the 0.0064 for Base controls in the lower panel is obtained by dividing 0.0004 (upper panel, Base controls) by 0.0622 (Sum of partial R-sq's). Base controls are: a 4th order polynomial in log(age), 21 region dummies, dummies for suburban and urban regions, dummies for mother tongue, and dummies for parental decade of birth. We follow Bound et al. (1995) in calculating the partial R-squared.

In additional robustness exercises, we have estimated the model with own education variables using i) the patent count; ii) the number of citations to the most cited patent of each inventor; and iii) an indicator for being a star inventor as the dependent variable. The results (see Table G4 in Appendix G) are similar to those reported in Table 7. We also re-estimated the model using our alternative measure for parental income, again replicating the results.

Finally, we have repeated the family structure and the IQ-interaction regressions in regression models that include own education variables (see Table G5 in Appendix G). The family structure variables and family structure - parental income and education - interactions on the one hand and the parental income and IQ interactions on the other hand obtain coefficients that are very similar to those reported. At the same time, the own education coefficients remain stable across all specifications and are also similar

sample, and the own education results for both the one and zero year age difference samples.

Table 9: Counterfactual Analysis with Education

Outcome		Data	Father income highest 10%	Father white collar mngr.	Father education at least MSc	IQ highest 10%	Education science at least MSc
Inventor	mean	0.013	0.015	0.014	0.018	0.024	0.110
	% change	100	114	96	136	181	813

Notes: In the "Data" column we display the mean predicted probability and its standard deviation from our main specification (controls for age, region, language, parental income, parental education, parental wealth, IQ and (for inventors and the last column only), own education), as well as the percentage change compared to this. In the Father income column, we randomly allocated those whose fathers are not in the top decile to either the 91st-95th or the 96-100th percentiles. In the Father white collar mngr. column, we change all those fathers who are not white collar managers to being white collar managers. In the Father education columns, we change all those with a father with less than an MSc to the category of father having an MSc. In the IQ column, we randomly allocated those who are not in the top IQ decile to either the 91st-95th or the 96-100th percentiles. In the education column, we change the education of everybody to have a science degree, and those with a degree less than an MSc to have an MSc.

to those reported Table 7. When we add the interaction of own high STEM education (MSc or PhD) and own high IQ to the specification, it obtains a positive and significant coefficient while at the same time the interaction with father's high income (top-5%) and high IQ loses significance. This last result may suggest that the high father income - IQ effect works through education.

One might take the results of this section to suggest that misallocation is reduced when own education is controlled for. The correct interpretation however is that while direct misallocation may be smaller than our earlier results suggest, the total degree of misallocation is as high as before once the indirect channel through own education is taken into account. Thus, even the free Finnish educational system does not remove misallocation through educational choices.

8 Conclusion

In this paper we have exploited the merger of three datasets -data on individual characteristics that include information on parents, patenting data, and IQ data- to analyze how family resources, social and educational background, own ability, and the interaction between background and ability impact an individual's probability of inventing.

Our main findings can be summarized as follows. First, while parental income matters even after controlling for other background variables and for IQ, the estimated im-

Table 10: Extended Model with Close Brothers

	(1)	(2)
first born	-0.00105	-0.000350
fa income 91-95	-0.000296	-0.0121
fa income 96-100	0.00742***	-0.0279
mo income 91-95	0.00156	-0.00346
mo income 96-100	0.00405	0.00664
fa bluecollar	0.000642	
fa jr whitec.	0.00246	
fa sr whitec.	0.00447*	
mo bluecollar	-0.000713	
mo jr whitec.	0.00103	
mo sr whitec.	-0.00223	
fa MSc	-0.00814**	
fa PhD	0.00908	
mo MSc	0.00236	
mo PhD	-0.0213	
fa STEM	0.00664***	
mo STEM	-0.000195	
IQ 91-95	0.00853***	0.0116***
IQ 96-100	0.0156***	0.0178***
STEM MSc	0.0966***	0.0871***
STEM PhD	0.229***	0.221***
Observations	82,054	82,054
Number of families		41,605

Notes: Robust standard errors are reported in Appendix Table A3 to save space. *** p<0.01, ** p<0.05, * p<0.1. All specifications include a 4th order polynomial in log(age), 21 region dummies, dummies for suburban and urban regions, dummies for mother tongue, and dummies for parental decade of birth. The sample is the brothers sample used in Table 4.

pact of parental income is greatly diminished once parental education and the individual's IQ are controlled for, and even more so when controlling for the individual's own education. Second, an R-squared decomposition shows that IQ matters more than all family background variables combined; moreover, IQ has both, a direct and an indirect impact through education on the probability of inventing, and finally the impact of IQ is larger and more convex for inventors than for medical doctors or lawyers. Third, to address the potential endogeneity of IQ, we focused on potential inventors with brothers close in age. This allowed us to control for family-specific time-invariant unobservables. We showed that the effect of visuospatial IQ on the probability of inventing is maintained when adding these controls. Fourth, to provide evidence on the importance of social family interactions, we compared individuals brought up by their biological parents with individuals with a missing biological parent and/or individuals with a step parent. We found that parental divorce decreases the probability of becoming an inventor and that the income of biological parents matters only when the child lives with them. Fifth, we discovered a positive and significant interaction between the individual's IQ and his father's income, which in turn points to a potential source of misallocation between individuals with heterogeneous IQ levels and parental resources.

Our analysis has interesting policy implications. In particular, it suggests that parental education and its impact on the child's education is a major source of potential misallocation, with parental income and socioeconomic status also affecting misallocation, but to a lesser extent. Such intergenerational spillovers in education affect an economy's ability to innovate. Indeed, we showed that achieving a higher education STEM degree increases an individual's probability of becoming an inventor significantly, while making it much less dependent upon parental income. We found that the impact of father's income on innovation is higher for high ability individuals, and that this extra impact also works mainly through education. This is evidence for misallocation being more severe for high ability individuals.

Taken together, our results suggest that by massively investing in education up to (STEM) PhD level, a country should significantly increase its aggregate innovation potential while making innovation more inclusive. Next, our results suggest that while IQ is of major importance in determining the level of education, family background in general and parental education in particular, play also an important role even in an equitable welfare state like Finland, where education is free up to and including university education.

We plan to extend our current analysis in several directions. A first extension is to replicate our analysis for other countries: do we get a pattern always similar to that in Figure 1[24] for the relationship between parental income and the probability of becoming an inventor, and do we explain it primarily by education and IQ (as we did here for Finland) or more by credit constraints? A second extension would be to look at the income mobility of inventors and how that mobility depends upon characteristics of the employing firm or sector (in particular firm size, firm age, the degree of competition in the firm's sector). A first step in this direction is by Aghion et al. (2017) which analyzes the returns to innovation and how these returns are shared between the inventor and the other employees and stakeholders in the same firm. These and other extensions of the analysis in this paper are left to future research.

[24]We already know that Bell et al. (2017) and Akcigit et al. (2017) obtain a similar pattern in the US, respectively using contemporaneous data and historical data.

References

Aghion, P., U. Akcigit, A. Bergeaud, R. Blundell, and D. Hémous: 2015, 'Innovation and Top Income Inequality'.

Aghion, P., U. Akcigit, A. Hyytinen, and O. Toivanen: 2017, 'On the Returns to Invention within Firms: Evidence from Finland'. Working Paper.

Aghion, P. and J. Tirole: 1994, 'The Management of Innovation'. *Quarterly Journal of Economics* **109**(4), 1185–1209.

Akcigit, U., S. Baslandze, and S. Stantcheva: 2016, 'Taxation and the International Mobility of Inventors'. *American Economic Review* **106**(10), 2930–2981.

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.

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

Balkin, D. B., G. D. Markman, and L. R. Gomez-Mejia: 2000, 'Is CEO Pay in High-technology Firms Related to Innovation?'. *Academy of Management Journal* pp. 1118–1129.

Bell, A., R. Chetty, X. Jaravel, N. Petkova, and J. Van Reenen: 2017, 'The Lifecycle of Inventors'. Stanford University mimeo.

Björklund, A., T. Eriksson, M. Jäntti, O. Raaum, and E. Österbacka: 2002, 'Brother Correlations in Earnings in Denmark, Finland, Norway and Sweden Compared to the United States'. *Journal of Population Economics* **15**(4), 757–772.

Black, S. E. and P. J. Devereux: 2011, 'Recent Developments in Intergenerational Mobility'. In: O. Ashenfelter and D. Card (eds.): *Handbook of Labor Economics*, Vol. 4. Elsevier, pp. 1487–1541.

Bound, J., D. A. Jaeger, and R. M. Baker: 1995, 'Problems with Instrumental Variables Estimation When the Correlation between the Instruments and the Endogenous Explanatory Variable Is Weak'. *Journal of the American statistical association* **90**(430), 443–450.

Carlsson, M., G. B. Dahl, B. Öckert, and D.-O. Rooth: 2015, 'The Effect of Schooling on Cognitive Skills'. *Review of Economics and Statistics* **97**(3), 533–547.

Ceci, S.: 2001, 'Intelligence: The Surprising Truth'. *Psychology Today* **34**(4), 46–53.

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

Cesarini, D., E. Lindqvist, R. Östling, and B. Wallace: 2016, 'Wealth, Health, and Child Development: Evidence from Administrative Data on Swedish Lottery Players'. *Quarterly Journal of Economics* **131**(2), 687–738.

Chetty, R., N. Hendren, P. Kline, and E. Saez: 2014, 'Where Is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States'. *Quarterly Journal of Economics* **129**(4), 1553–1623.

Corak, M.: 2004, *Generational Income Mobility in North America and Europe*. Cambridge University Press.

Dal Bó, E., F. Finan, O. Folke, T. Persson, and J. Rickne: 2017, 'Who Becomes a Politician?'. *Quarterly Journal of Economics* **132**(4), 1877–1914.

Depalo, D. and S. Di Addario: 2015, 'Inventors' Returns to Patents'. Bank of Italy mimeo.

Franco, A. M. and M. F. Mitchell: 2008, 'Covenants Not to Compete, Labor Mobility, and Industry Dynamics'. *Journal of Economics & Management Strategy* **17**(3), 581–606.

Giuri, P., M. Mariani, S. Brusoni, G. Crespi, D. Francoz, A. Gambardella, W. Garcia-Fontes, A. Geuna, R. Gonzales, D. Harhoff, et al.: 2007, 'Inventors and Invention Processes in Europe: Results from the PatVal-EU Survey'. *Research Policy* **36**(8), 1107–1127.

Hertz, T., T. Jayasundera, P. Piraino, S. Selcuk, N. Smith, and A. Verashchagina: 2007, 'The inheritance of educational inequality: International comparisons and fifty-year trends'. *The BE Journal of Economic Analysis & Policy* **7**(2).

Holmstrom, B.: 1989, 'Agency Costs and Innovation'. *Journal of Economic Behavior & Organization* **12**(3), 305–327.

Hsieh, C.-T., E. Hurst, P. Klenow, and C. Jones: 2013, 'The Allocation of Talent and U.S. Economic Growth'.

Jaravel, X., N. Petkova, and A. Bell: 2015, 'Team-specific Capital and Innovation'. Unpublished, Harvard University.

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.

Manso, G.: 2011, 'Motivating Innovation'. *The Journal of Finance* **66**(5), 1823–1860.

Mc Gue, M., T. J. Bouchard Jr, W. G. Iacono, and D. T. Lykken: 1993, 'Behavioral genetics of cognitive ability: A life-span perspective.'.

Pakes, A. and S. Nitzan: 1983, 'Optimum Contracts for Research Personnel, Research Employment, and the Establishment of "Rival" Enterprises'. *Journal of Labor Economics* **1**(4), 345–365.

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.

Pinker, S.: 2005, *The Blank Slate*. Southern Utah University.

Plomin, R. and F. M. Spinath: 2004, 'Intelligence: Genetics, Genes, and Genomics.'. *Journal of Personality and Social Psychology* **86**(1), 112.

Scotchmer, S.: 2004, *Innovation and Incentives*. MIT press.

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. and J. Vartiainen: 2009, 'Firm Factors in Wages and Wage Changes'. University of Chicago Press, pp. 149–178.

Appendix A: Additional Tables for the Main Text

Table A1: Standard Errors for Table 5

	(1)	(2)	(3)	(4)
biol fa away	(0.000808)	(0.000829)	(0.000829)	(0.000818)
biol mo away	(0.00163)	(0.00164)	(0.00164)	(0.00162)
biol fa & mo away	(0.00137)	(0.00157)	(0.00157)	(0.00155)
biol fa income 91-95	(0.00108)	(0.00111)	(0.00119)	(0.00119)
biol fa income 96-100	(0.00130)	(0.00132)	(0.00142)	(0.00142)
biol mo income 91-95	(0.00315)	(0.00327)	(0.00356)	(0.00356)
biol mo income 96-100	(0.00117)	(0.00132)	(0.00134)	(0.00134)
biol fa income 91-95 x away		(0.00341)	(0.00344)	(0.00352)
biol fa income 96-100 x away		(0.00477)	(0.00480)	(0.00467)
biol mo income 91-95 x away		(0.00376)	(0.00401)	(0.00407)
biol mo income 96-100 x away		(0.00248)	(0.00249)	(0.00250)
step fa income 91-95			(0.00223)	(0.00223)
step income 96-100			(0.00282)	(0.00282)
step mo income 91-95			(0.00859)	(0.00859)
step mo income 96-100			(0.00833)	(0.00833)
fa bluecollar	(0.000541)	(0.000541)	(0.000541)	(0.000541)
fa jr whitec.	(0.000736)	(0.000736)	(0.000736)	(0.000736)
fa sr whitec.	(0.00112)	(0.00112)	(0.00112)	(0.00112)
mo bluecollar	(0.000551)	(0.000551)	(0.000551)	(0.000551)
mo jr whitec.	(0.000621)	(0.000621)	(0.000621)	(0.000621)
mo sr whitec.	(0.00125)	(0.00125)	(0.00125)	(0.00125)
biol fa MSc	(0.00175)	(0.00175)	(0.00175)	(0.00178)
biol fa PhD	(0.00434)	(0.00434)	(0.00434)	(0.00442)
biol mo MSc	(0.00242)	(0.00242)	(0.00243)	(0.00251)
biol mo PhD	(0.00957)	(0.00957)	(0.00961)	(0.0101)
biol fa STEM	(0.000798)	(0.000798)	(0.000797)	(0.000797)
biol mo STEM	(0.000732)	(0.000732)	(0.000732)	(0.000732)
biol fa MSc x away				(0.00767)
biol fa PhD x away				(0.0202)
biol mo MSc x away				(0.00877)
biol mo PhD x away				(0.0122)
IQ 91-95	(0.00144)	(0.00144)	(0.00144)	(0.00144)
IQ 96-100	(0.00165)	(0.00165)	(0.00165)	(0.00165)

Table A2: Standard Errors for Table 6

	(1)	(2)	(3)	(4)	(5)	(6)
fa income 91-95	(0.00108)	(0.00108)	(0.00108)	(0.0213)	(0.0212)	(0.0213)
fa income 96-100	(0.00128)	(0.00129)	(0.00127)	(0.0276)	(0.0276)	(0.0276)
mo income 91-95	(0.00314)	(0.00314)	(0.00313)	(0.0485)	(0.0479)	(0.0480)
mo income 96-100	(0.00113)	(0.00118)	(0.00113)	(0.00779)	(0.00778)	(0.00779)
fa bluecollar	(0.000740)	(0.000741)	(0.000740)			
fa jr whitec.	(0.000873)	(0.000873)	(0.000873)			
fa sr whitec.	(0.000934)	(0.000934)	(0.000934)			
mo bluecollar	(0.000715)	(0.000715)	(0.000715)			
mo jr whitec.	(0.000756)	(0.000757)	(0.000756)			
mo sr whitec.	(0.000940)	(0.000940)	(0.000940)			
fa MSc	(0.00166)	(0.00166)	(0.00166)			
fa PhD	(0.00431)	(0.00447)	(0.00447)			
mo MSc	(0.00236)	(0.00236)	(0.00236)			
mo PhD	(0.00953)	(0.0102)	(0.0102)			
fa STEM	(0.000794)	(0.000794)	(0.000794)			
mo STEM	(0.000733)	(0.000733)	(0.000733)			
IQ 91-95	(0.00144)	(0.00144)	(0.00144)	(0.00408)	(0.00409)	(0.00408)
IQ 96-100	(0.00174)	(0.00165)	(0.00174)	(0.00467)	(0.00455)	(0.00466)
fa inc 96-100 x IQ 96-100	(0.00527)		(0.00534)	(0.0148)		(0.0150)
mo inc 96-100 x IQ 96-100	(0.00816)		(0.00811)	(0.0271)		(0.0269)
fa PhD x IQ 96-100		(0.0143)	(0.0145)		(0.0392)	(0.0399)
mo PhD x IQ 96-100		(0.0282)	(0.0279)		(0.0716)	(0.0673)

Table A3: Standard Errors for Table 10

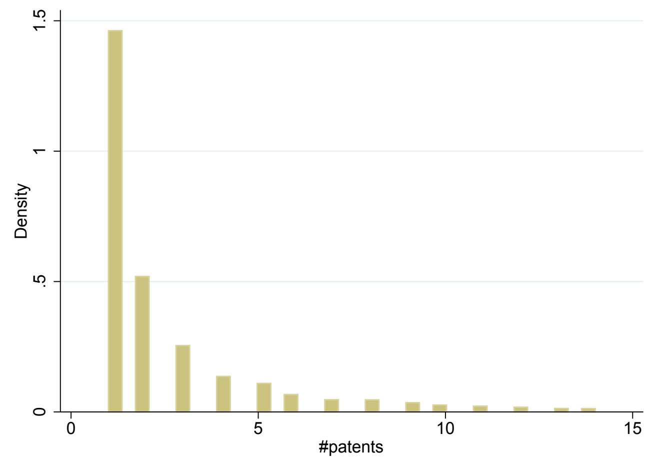
	(1)	(2)
first born	(0.000847)	(0.00135)
fa income 91-95	(0.00220)	(0.0201)
fa income 96-100	(0.00280)	(0.0275)
mo income 91-95	(0.00753)	(0.0481)
mo income 96-100	(0.00258)	(0.00727)
fa bluecollar	(0.00111)	
fa jr whitec.	(0.00162)	
fa sr whitec.	(0.00247)	
mo bluecollar	(0.00115)	
mo jr whitec.	(0.00138)	
mo sr whitec.	(0.00270)	
fa MSc	(0.00361)	
fa PhD	(0.00871)	
mo MSc	(0.00509)	
mo PhD	(0.0136)	
fa STEM	(0.00172)	
mo STEM	(0.00161)	
IQ 91-95	(0.00295)	(0.00397)
IQ 96-100	(0.00341)	(0.00444)
STEM MSc	(0.00454)	(0.00588)
STEM PhD	(0.0183)	(0.0206)

Online Appendices: B, C, D, E, F, and G

– Not intended for publication unless requested –

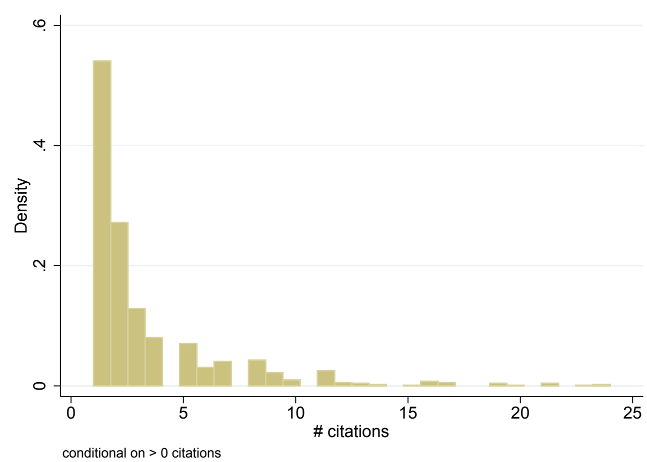
B Supplementary Materials for Section 2

Figure B1: The Distribution of Number of Patents by Inventor



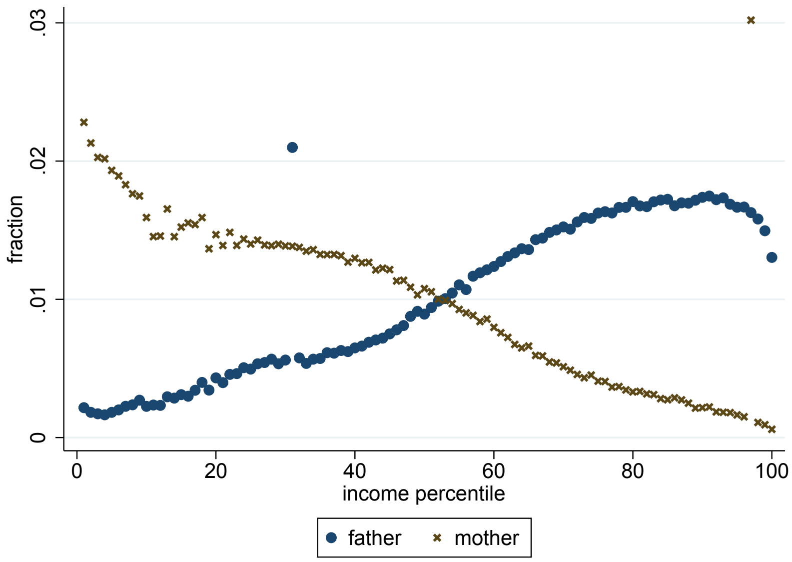
Notes: The number of EPO patents by inventor. Number of patents is calculated 1978-2012.

Figure B2: The Distribution of Number of Citations by Inventor



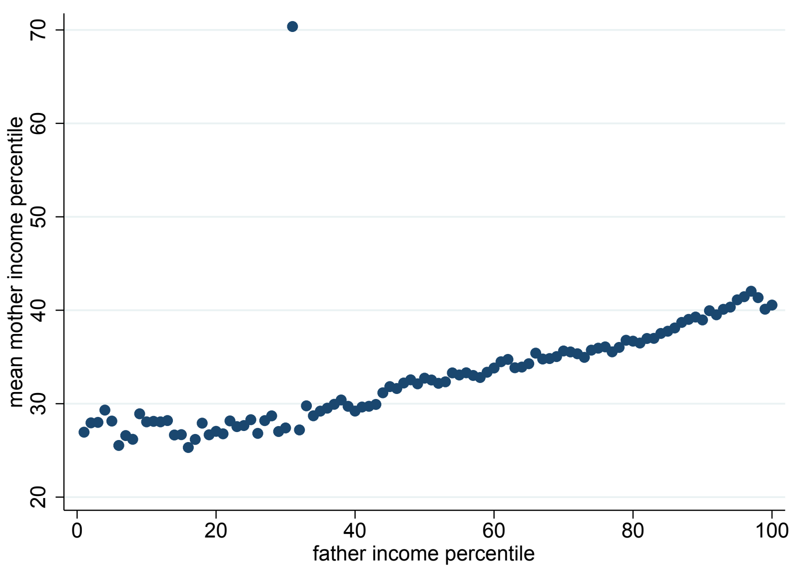
Notes: The number citations to all EPO patents of an inventor in the first 5 years of the patent.

Figure B3: Mothers and Fathers Distribution over Income Percentiles



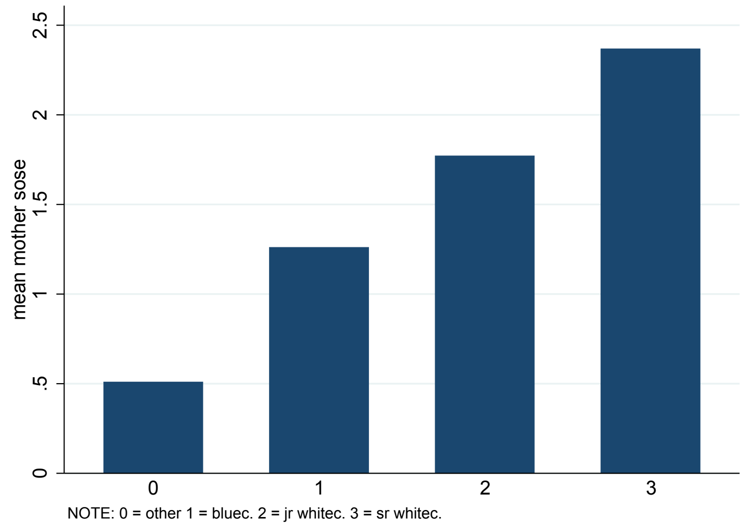
Notes: The figure displays the distribution of fathers and mothers across income percentiles. Parental incomes are measured in 1975 for individuals born 1961-1975, and in 1985 for individuals born in 1976-1985. We calculate the percentile ranks using residuals from a regression of the natural log of deflated income of fathers and mothers on year-of-birth dummies.

Figure B4: Wage Income of Fathers and Mothers



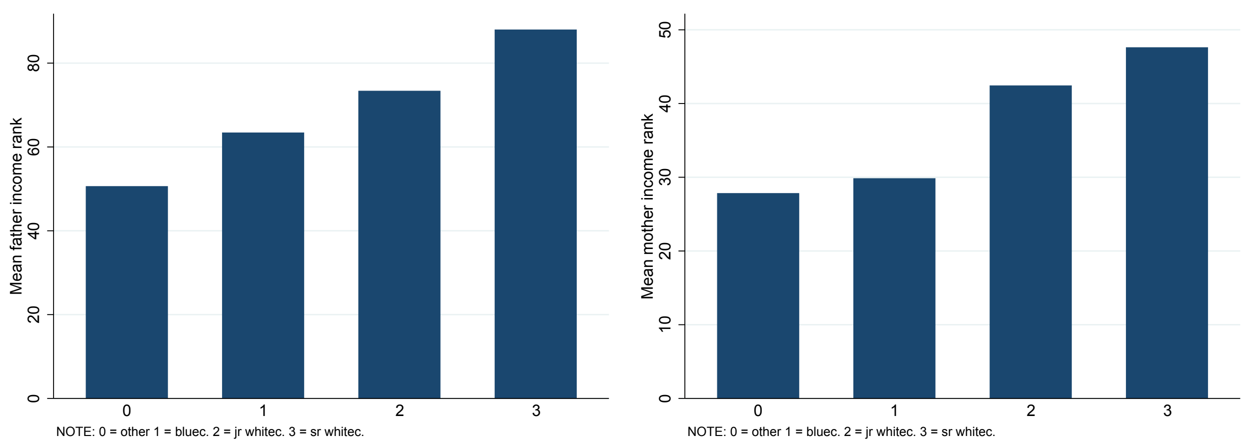
Notes: The figure displays average income percentile of mothers of an individual as a function of the income percentile of the father of that individual. Parental incomes are measured in 1975 for individuals born 1961-1975, and in 1985 for individuals born in 1976-1985. We calculate the percentile ranks using residuals from a regression of the natural log of deflated income of fathers and mothers on year-of-birth dummies.

Figure B5: Socioeconomic Status of Fathers and Mothers



Notes: The figure displays average socioeconomic class of mothers of an individual as a function of the socioeconomic class of the father of that individual. We code socioeconomic class as follows: 0 = other (base); 1 = blue-collar; 2 = junior white collar; 3 = senior white collar. Parental socioeconomic status is measured in 1975 for parents born before 1951, and in 1985 parents born in 1951 or thereafter.

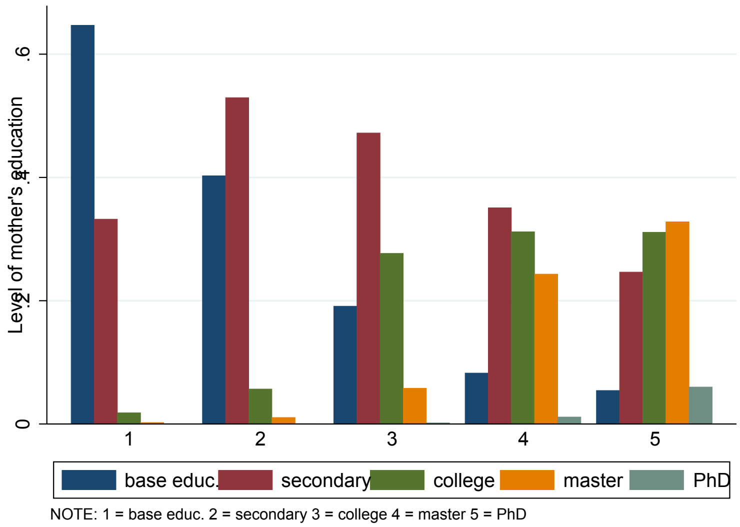
Figure B6: Socioeconomic Status and Income



A. Fathers B. Mothers

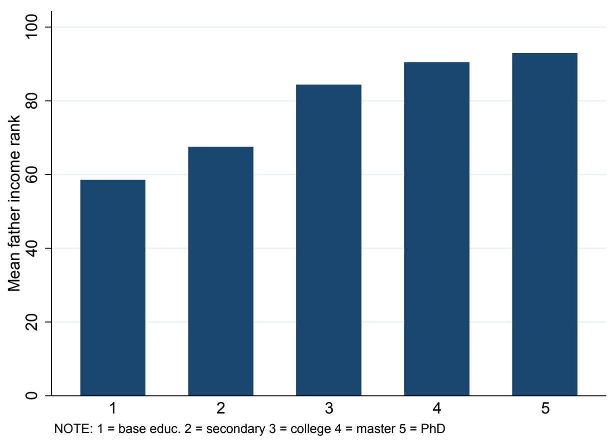
Notes: The figure displays average income percentile of a parent (A: fathers, B: mothers) as a function of the socioeconomic class of the parent. Parental incomes are measured in 1975 for individuals born 1961 âĂŞ 1975, and in 1985 for individuals born in 1976-1985. We calculate the percentile ranks using residuals from a regression of the natural log of deflated income of fathers and mothers on year-of-birth dummies. We code socioeconomic class as follows: 0 = other (base); 1 = blue-collar; 2 = junior white collar; 3 = senior white collar. Parental socioeconomic status is measured in 1975 for parents born before 1951, and in 1985 parents born in 1951 or thereafter.

Figure B7: Mother's Education vs Father's Education

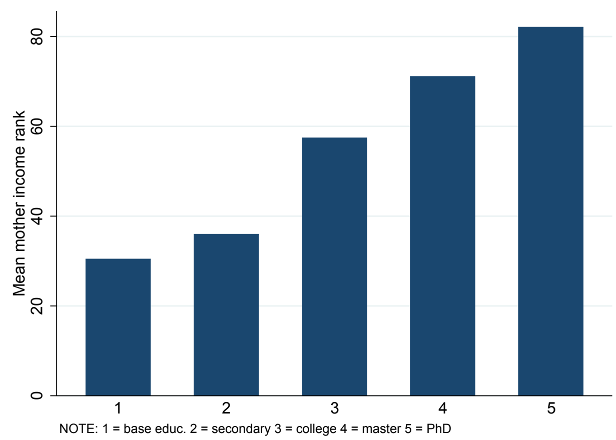


Notes: The education level of the mother of an individual conditional on the level of education of the father of the same individual. We divide education into five groups by level of education: base education (up to 9 years, depending on age of parent), secondary, tertiary, MSc, and PhD. Parental education is measured in 1975 unless unavailable, in which case 1985 data used.

Figure B8: Parental Education vs Parental Income



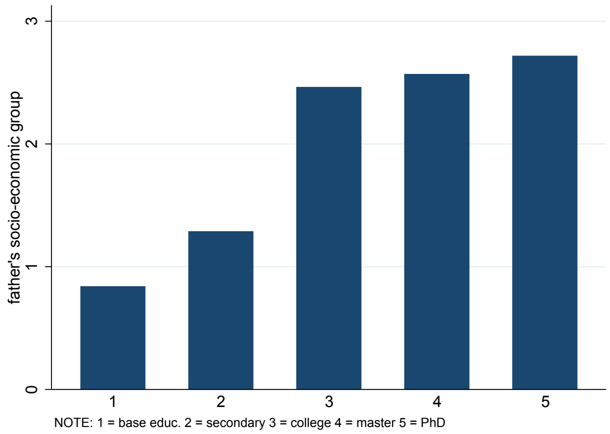
A. Fathers



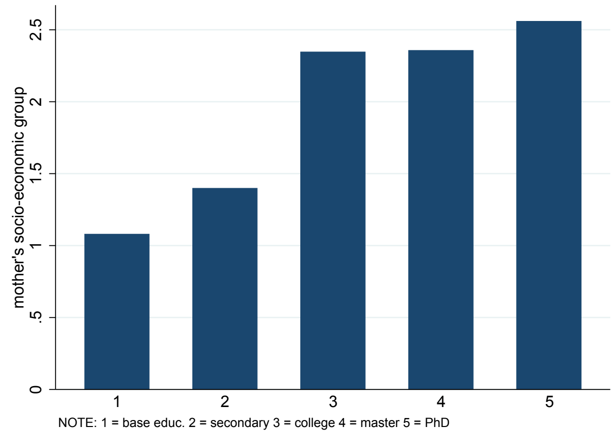
B. Mothers

Notes: The figure displays the mean income percentile of a parent as a function of the education of that parent. Parental incomes are measured in 1975 for individuals born 1961-1975, and in 1985 for individuals born in 1976-1985. We calculate the percentile ranks using residuals from a regression of the natural log of deflated income of fathers and mothers on year-of-birth dummies. We divide education into five groups by level of education: base education (up to 9 years, depending on age of parent), secondary, tertiary, MSc, and PhD. Parental education is measured in 1975 unless unavailable, in which case 1985 data used.

Figure B9: Parental Education vs Parental Socioeconomic Status



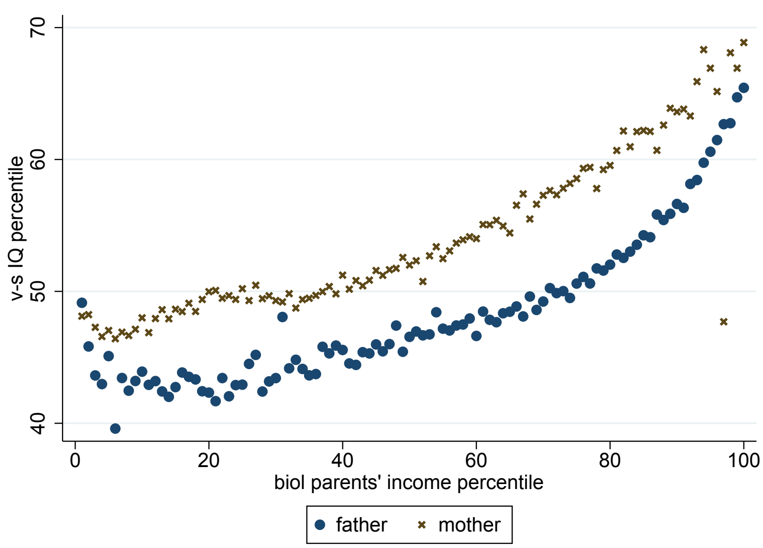
A. Fathers



B. Mothers

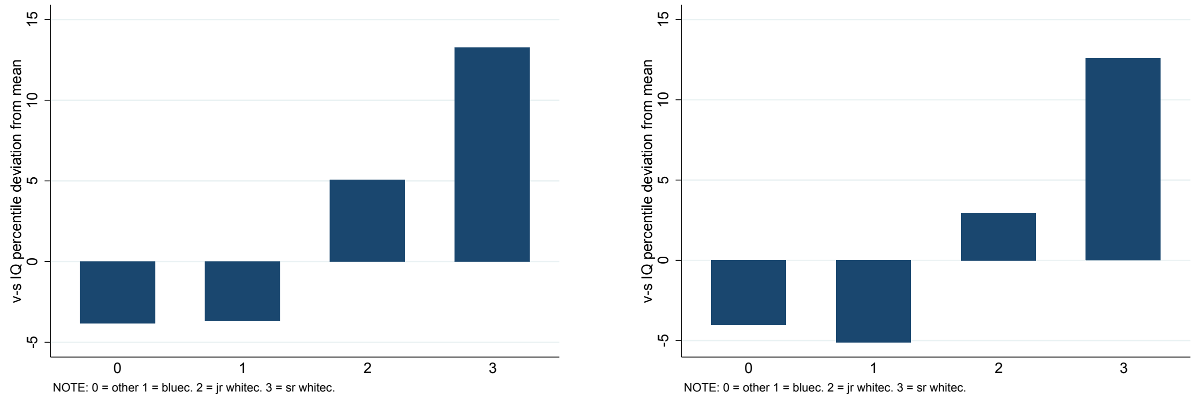
Notes: The figure displays average socioeconomic class of a parent (A: father, B: mother) as a function of the level of education of that parent. We code socioeconomic class as follows: 0 = other (base); 1 = blue-collar; 2 = junior white collar; 3 = senior white collar. Parental socioeconomic status is measured in 1975 for parents born before 1951, and in 1985 parents born thereafter. We divide education into five groups by level of education: base education (up to 9 years, depending on age of parent), secondary, tertiary, MSc, and PhD. Parental education is measured in 1975 unless unavailable, in which case 1985 data used.

Figure B10: Parental Income vs Own IQ



Notes: The figure displays the average IQ percentile of individuals with a parent in a given income percentile. IQ percentiles are calculated on the basis of the normalized IQ score, where normalization was done separately for each conscription cohort to avoid the Flynn effect. Parental incomes are measured in 1975 for individuals born 1961-1975, and in 1985 for individuals born in 1976-1985. We calculate the percentile ranks using residuals from a regression of the natural log of deflated income of fathers and mothers on year-of-birth dummies.

Figure B11: Own IQ vs Parents' Socioeconomic Status

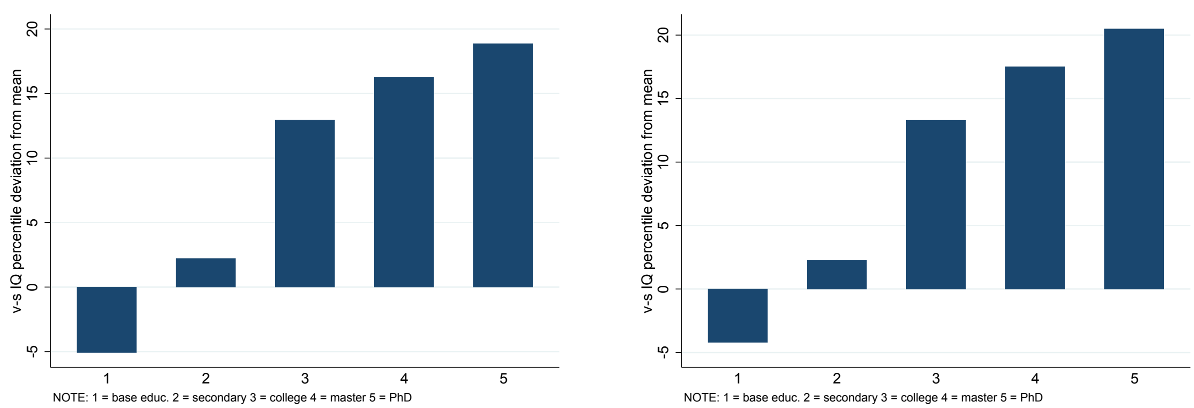


A. Father

B. Mother

Notes: The figure displays the average deviation from the mean IQ conditional on the socioeconomic status of the father. IQ percentiles are calculated on the basis of the normalized IQ score, where normalization was done separately for each conscription cohort to avoid the Flynn effect. We code socioeconomic class as follows: 0 = other (base); 1 = blue-collar; 2 = junior white collar; 3 = senior white collar. Parental socioeconomic status is measured in 1975 for individuals born 1961-1975, and in 1985 for individuals born in 1976-1985.

Figure B12: Own IQ vs Parents' Education



A. Father

B. Mother

Notes: The figure displays the average deviation from the mean IQ conditional on the level of education of the parents (A: father, B: mother). IQ percentiles are calculated on the basis of the normalized IQ score, where normalization was done separately for each conscription cohort to avoid the Flynn effect. We divide education into five groups by level of education: base education (up to 9 years, depending on age of parent), secondary, tertiary, MSc, and PhD. Parental education is measured in 1975 unless unavailable, in which case 1985 data used.

B.1 Data sources and data matching

Our data come from Statistics Finland (SF), European Patent Office (EPO) and Finnish Defence Forces (FDF). SF is our source of individual characteristics. They come from the Finnish Linked Employer-Employee Data (FLEED) for the period 1988-2012 and from the population census 1975 and 1985. EPO data allows us to identify Finnish inventors. FDF provided us with information on IQ test results for conscripts who served in 1982 or later.

FLEED and the population census data are standard administrative register-based data, collected and maintained by SF. Our EPO data are derived from OECD's REGPAT database, which includes patent applications to the EPO and PCT filings. FDF's IQ data are based on measured cognitive ability, based on subtests of verbal, arithmetic, and visuospatial reasoning that conscripts take at the beginning of their military service; see Pekkala Kerr et al. (2013) and Jokela et al. (2017) for a more detailed description.

We describe the type of data and variables drawn from each source in the main text (as well as later in this Appendix; see "B.3: Variable definitions and descriptive statistics" below). Here we provide a more detailed account of how the datasets have been matched.

SF's FLEED and population census contain unique but anonymized individual identifiers, which are based on unique social security numbers that everybody has. The FDF's IQ data also contains social security numbers. FDF's IQ data were transferred to SF, which then created individual identifiers for the purpose of data matching from the social security numbers. The three data sets were then linked using the individual identifiers.

EPO data, in contrast, does not contain linkable 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 1988-2012.

B.2 Institutional environment

B.2.1 Overall economic environment in 1988-2012

Finland has been a member of EU since 1995 and has a population of 5.5 million. It has been a member of the euro area since its introduction in 1999/2002.

During our observation period from 1988 to 2012, Finland's gross domestic product (GDP) grew on average 2.1% per year. The average masks a lot of variation (std = 3.6%), because the economy experienced boom periods in the late 1980s and late 1990s and two major economic slumps, one in the early 1990s and another in 2008/2009. In 1988/1989, unemployment rate was low, at around 3.1%. Unemployment peaked in the economic crisis of the early 1990s at around 16% (1993-1994), but decreased then to 7.7% by 2012.

At the beginning of our observation period, the employment rate among the population aged 15-74 was 67.3%. The employment rate has fluctuated somewhat, and decreased to 60.9% by 2012, mostly due to the ageing of the population. Commerce, hotel and restaurant services, education, social services and health services and transport employ the greatest number of people, with the public sector (municipalities, government) being a major employer in many of these sectors.

In 1988, 51% of population aged 15 or over had basic education, but the share dropped to 31% by 2012. The share of population having higher level tertiary (ISCED 7) or doctorate level (ISCED 8) education increased from 7% (1988) to nearly 18% (2012) over our observation period. Research and development expenditures also increased steadily during our observation period, reaching their peak in 2011 when the total R&D expenditure by business sector and public sector amounted to 3.8% of the GDP. Based on its Global Competitiveness Index, World Economic Forum has quite consistently ranked Finland to be one of the ten most competitive countries in the world.

B.2.2 Educational system

First versions of the national schooling system and public education in Finnish were launched in the 1860s. A key objective of the Finnish educational system has for along been to provide all citizens with equal opportunities. For example, according to the Finnish Ministry of Education and Culture, *"[o]ne of the basic principles of Finnish education is that all people must have equal access to high-quality education and training. The same educational opportunities should be available to all citizens irrespective of their ethnic origin, age, wealth or where they live."* (Ministry of Education and Culture, p. 6).

Basic education in Finland includes primary and lower secondary education. Compulsory education starts at age 7 and, with some minor exceptions, ends when nine years of comprehensive school has been completed. Upper secondary education lasts typically for three years and consists of upper secondary general education, which prepares for the matriculation examination, and vocational education.

Applicants to most field-specific degree programs of the Finnish polytechnics and universities are required to take an entrance examination. The applicants are then ranked on the basis of their prior performance (in e.g. the matriculation examination) and their performance in the entrance examination. Polytechnics give less weight on the entrance examination. Entry into degree programs in law and medicine, as well certain fields of science, technology and business is competitive. It takes from three to four years to complete a polytechnic degree or a lower university degree (bachelor). A higher university degree (master) can be acquired in five years, though in practice it often takes longer.

Post-graduate education leads to a licentiate, which is a lower level research degree that has become rare in most fields, or to a doctoral degree. In medicine, students directly pursue the licentiate degree.

Bar some compulsory but low fees associated with registration (e.g., student union membership fee in universities), education is offered free of charge at all levels of schooling and for the degrees officially recognized and endorsed by the Finnish schooling legislation.

There is a system of study grants and loans, which in an international comparison is rather well-developed and comprehensive. The system awards financial aid for full-time studies in upper secondary schools and vocational institutions, and also in institutions of higher education, including the universities.

The study grans are relatively generous and the study loans are backed by the government. At the end of our sample period, the average amount of financial aid (excluding the loan guarantee) was about 450 euros per month for students in institutes of higher education and polytechnics (source: Statistics of The Finnish Social Insurance Institution, KELA).

B.2.3 Wage setting

The Finnish labor market is characterized by widespread organization of employees (unionization) and employers, as well as by centralized wage-setting (bargaining and co-operation), which have resulted in various types of collective wage and labor agreements. A special feature of the Finnish labor market are national incomes policy settlements, which cover issues related to wage setting and salaries, taxation, pensions and unemployment benefits and which are agreements between the government and the central confederations of employees and employers (the tripartite system). About every three out of four Finnish employees are members of a trade union, and also those with higher education belong often to unions. In 2007, the system of centralized agreements largely ended when the private sector employers' association called for industry level negotiations. In 2011 there was a partial and temporary return to signing a national framework agreement, which was triggered by perceptions of deterioration of national price-cost competitiveness.

Despite these centralized features, wage setting is a mixture of collective and individual mechanisms. As Uusitalo and Vartiainen (2009) have emphasized, a key feature of the centralized agreements is that they coordinate the overall rate of wage increases. This does not prevent a firm from increasing its workers' wages by more than the coordinated overall increase. The collective agreements also restrict local bargaining by instituting agreed minimum wages for certain occupations and job levels. If a firm wants to employ somebody, the bargaining of his/her initial salary is subject to the minimum tariffs. However, as Uusitalo and Vartiainen (2009) stress, for most employees in the manufacturing sector the minimum wages rarely bind. These features of the Finnish labor market mean that relative wages have largely been set by market forces and that wage bargaining is to a significant extent local. Moreover, various firm-specific arrangements and performance-related pay components became more widespread in the 1990s.

B.2.4 Remuneration of inventors

A specific law governs innovations made by employees ("Act on the Right in Employee Inventions", originally given in 1967, augmented in 2000). The provisions of the act apply to inventions (potentially) patentable in Finland.

The employee inventions act says, in particular, that i) an employer may acquire the right in the invention (made by its employee) if the use of the invention falls within the field of activity of the employer's enterprise; that ii) an employee who makes an invention has to notify the employer of it without delay, and that the employer has to notify the employee, if the employer wishes to acquire the right in the invention; and, finally, that iii) if the employer acquires the right in the invention, the employee is entitled to a reasonable compensation from the employer.

When determining the amount of the compensation, particular attention is to be paid to the value of the invention, the scope of the right which the employer acquires, as well as to the terms and conditions of the employment contract of the employee and the contribution which other circumstances connected with the employment had to the conception of the invention.

In sum, the act assigns the right to ownership of an employee invention, but it does not directly determine the amount firms have to pay if they exercise the right. Rather, the determination of the amount of compensation is largely left to the market forces.

The Finnish act is by no means unique in an international comparison: For example, the Swedish "Act on the Right to Employee's Inventions" (introduced in 1949) shares many features with the corresponding Finnish act. Moreover, the German "Employee Invention Act" is in many ways similar: e.g. it states that when the employer claims the rights to an employee-made invention, the employer owes the employee an "adequate" remuneration. Things are a bit more complex in the UK, but when an employer owns his employee's invention, it is possible for the employee to claim compensation if his invention or the patent is of outstanding benefit to his employer and it is just to award such compensation.

B.2.5 Economic inequality

In an international comparison, income inequality is relatively low in Finland. Using disposable cash income (excl. capital gains) as the income measure, the Gini-coefficient has ranged from 20.7 in 1992 to 26.4 in 2007 (source: Statistics Finland). On average, it has been 23.8 during our sample period. Consistent with the relatively low income inequality, intergenerational mobility is in Finland - like in other Nordic countries - quite high, exceeding that of the UK and US (see, e.g., Black and Devereux, 2011). In line with this, the correlation of incomes among siblings is quite a bit lower in Finland (and in the other Nordic countries) than, for example, in the U.S. (Björklund et al., 2002, Black and Devereux, 2011).

The estimated intergenerational correlations of educational outcomes are relatively low in Finland (see, e.g., Hertz et al., 2007, Table 7). The available evidence suggests that schooling reforms have contributed to this, as they have reduced the intergenerational income elasticity in Finland. For example, Pekkarinen et al. (2009) find that the Finnish comprehensive school reform that took place in the 1970s reduced the intergenerational income elasticity by 23%.

B.3 Variable definitions and descriptive statistics

TABLE B1: VARIABLE DEFINITIONS

Variable	Acronym	Period	Definition
		– Dependent variables –	
Inventor	Inventor	1978-2012	= 1 if individual invents at least once by 2012, 0 otherwise
Medical doctor	MD	-2012	= 1 if individual has a medical degree by age 35, 0 otherwise
Lawyer	Lawyer	-2012	= 1 if individual has a law degree by age 35, 0 otherwise
Patent count	Patcount	1978-2012	= number of patents obtained by individual i by 2012
Star inventor	Highqual	1978-2012	= 1 if individual has at least one patent with at least 5 citations within first 5 years (top 5% of the citation distribution)
		– Explanatory variables: –	
Age	Age	1988- 2012	= biological age of the individual in year 2012 (cross-section analyses), year t (panel data)
Region of residence	Region	1988- 2012	= region of residence of the individual in year of entry to FLEED (cross-section analyses), year t (panel data). 20 regions
Urban	Urban	1988- 2012	= i) dummy for the region of residence being urban; ii) semi-urban
Language	Language	1988- 2012	= indicator variables for mother tongue being i) Swedish; ii) other non-Finnish language. Base language Finnish.

– Continued on the next page –

– Table B1 (Continued) –

Parental birth decade	DoB	1988- 2012	= indicator variables for parental decade of birth, separately for fathers and mothers.
Parental income quintiles	Fwage, Mwage	1975, 1985, 1988-	= residual from a wage regression (see text). Highest quintiles are divided into separate indicators for the 81st - 90th percentiles, the 91st - 95th percentiles, and the 96th - 100th percentiles.
Parental socioeconomic status	Fsose, Msose	1975, 1985	= socioeconomic status as reported in the 1975 and 1985 census. Categories: bluecollar, junior whitecollar, senior whitecollar, other. Base category "other".
Parental education	Feduc, Meduc	1975, 1985, 1988-	= Parental education at age 35 (or earliest observed age > 35). Five categories: Base, secondary, college, Master's, PhD. Separate indicator for STEM education.
IQ	IQ	1961 birth cohort-	= deciles of visuo-spatial IQ, with highest decile split into two (91-95; 96-100). For figures, we normalize the mean to 100 with s.d. 15. Both deciles and normalization are by year of entering military service.
Education	Educ	1988- 2012	= Own education. Cross-section: at age 35 (or earliest observed age > 35). Panel: in year t. Five categories: Base, secondary, college, Master's, PhD, separately for non-STEM and STEM degrees.
Patent count	Patcount	1988- 2012	= number of patents obtained by individual i in year t

Notes: The data sources are: Statistics Finland FLEED and 1975 and 1985 census individual information; EPO PATSTAT (OECD REGPAT) for patent data; and Finnish Defense Forces for IQ data.

Table B2: Summary Statistics

VARIABLE	*— Non-inventors —* mean	sd	p50	*— Inventors —* mean	sd	p50
inventor	0.000	0.000	0	1.000	0.000	1
patent count	0.000	0.000	0	1.275	2.651	1
5-year citation count	0.000	0.000	0	0.846	2.271	0
star inventor	0.000	0.000	0	0.106	0.308	0
MD	0.005	0.069	0	0.010	0.101	0
lawyer	0.005	0.074	0	0.001	0.038	0
MSc by age 35	0.108	0.310	0	0.623	0.485	1
age in 2014	43.303	5.860	44	44.751	4.773	45
fa income 21-40	0.126	0.332	0	0.073	0.260	0
fa income 41-60	0.189	0.391	0	0.124	0.329	0
fa income 61-80	0.303	0.459	0	0.246	0.431	0
fa income 81-90	0.170	0.375	0	0.195	0.397	0
fa income 91-95	0.085	0.278	0	0.147	0.354	0
fa income 96-100	0.075	0.264	0	0.189	0.391	0
mo income 21-40	0.275	0.446	0	0.239	0.426	0
mo income 41-60	0.209	0.406	0	0.203	0.403	0
mo income 61-80	0.103	0.303	0	0.138	0.345	0
mo income 81-90	0.027	0.163	0	0.056	0.230	0
mo income 91-95	0.009	0.095	0	0.022	0.147	0
mo income 96-100	0.034	0.182	0	0.036	0.187	0
wage percentile biol fathers	65.533	24.088	70	75.593	22.307	83
wage percentile biol mothers	35.629	25.523	32	40.189	27.790	36
wage percentile socialfathers	62.097	25.791	67	71.304	26.500	79
wage percentile social mothers	32.695	24.683	30	37.895	27.845	34
fa bluecollar	0.449	0.497	0	0.286	0.452	0
fa jr whitec.	0.169	0.375	0	0.200	0.400	0
fa sr whitec.	0.145	0.352	0	0.335	0.472	0
mo bluecollar	0.312	0.463	0	0.192	0.394	0
mo jr whitec.	0.355	0.478	0	0.388	0.487	0
mo sr whitec.	0.102	0.303	0	0.248	0.432	0
fa secondary	0.388	0.487	0	0.366	0.482	0
fa college	0.066	0.249	0	0.137	0.344	0
fa MSc	0.050	0.218	0	0.149	0.356	0
fa PhD	0.009	0.093	0	0.041	0.199	0
mo secondary	0.418	0.493	0	0.438	0.496	0
mo college	0.067	0.251	0	0.142	0.349	0
mo MSc	0.024	0.153	0	0.079	0.270	0
mo PhD	0.001	0.038	0	0.005	0.072	0
fa STEM	0.366	0.482	0	0.486	0.500	0
mo STEM	0.104	0.305	0	0.107	0.309	0
IQ 1-10	0.100	0.300	0	0.014	0.116	0
IQ 11-20	0.102	0.302	0	0.021	0.144	0
IQ 21-30	0.100	0.300	0	0.031	0.172	0
IQ 31-40	0.100	0.300	0	0.042	0.202	0
IQ 51-60	0.106	0.307	0	0.082	0.274	0
IQ 61-70	0.096	0.294	0	0.094	0.291	0
IQ 71-80	0.101	0.302	0	0.147	0.354	0

– Continued on the next page –

– Table B2 (Continued) –

	Non-inventors			Inventors		
VARIABLE	*mean*	*sd*	*p50*	*mean*	*sd*	*p50*
IQ 81-90	0.097	0.295	0	0.184	0.387	0
IQ 91-95	0.052	0.222	0	0.141	0.348	0
IQ 96-100	0.050	0.218	0	0.187	0.390	0
secondary	0.170	0.376	0	0.077	0.267	0
college	0.091	0.288	0	0.021	0.145	0
MSc	0.049	0.217	0	0.026	0.159	0
PhD	0.002	0.048	0	0.001	0.038	0
secondary STEM	0.278	0.448	0	0.036	0.187	0
college STEM	0.127	0.333	0	0.204	0.403	0
MSc STEM	0.051	0.221	0	0.484	0.500	0
PhD STEM	0.005	0.070	0	0.112	0.315	0
Region 1	0.221	0.415	0	0.245	0.430	0
Region 2	0.084	0.277	0	0.090	0.286	0
Region 4	0.053	0.224	0	0.050	0.218	0
Region 5	0.032	0.177	0	0.032	0.176	0
Region 6	0.083	0.276	0	0.111	0.314	0
Region 7	0.040	0.196	0	0.028	0.166	0
Region 8	0.039	0.193	0	0.032	0.176	0
Region 9	0.028	0.166	0	0.029	0.167	0
Region 10	0.036	0.186	0	0.031	0.174	0
Region 11	0.053	0.225	0	0.044	0.206	0
Region 12	0.036	0.186	0	0.025	0.155	0
Region 13	0.052	0.222	0	0.050	0.218	0
Region 14	0.046	0.210	0	0.042	0.201	0
Region 15	0.036	0.185	0	0.033	0.179	0
Region 16	0.016	0.127	0	0.014	0.118	0
Region 17	0.079	0.269	0	0.089	0.285	0
Region 18	0.021	0.142	0	0.019	0.136	0
Region 19	0.045	0.208	0	0.036	0.185	0
Region 21	0.000	0.014	0	0.000	0.015	0
Urban	0.613	0.487	1	0.679	0.467	1
Semi-Urban	0.187	0.390	0	0.171	0.377	0
Rural	0.200	0.400	0	0.149	0.356	0
Father birth 1901 - 1920	0.010	0.040	0	0.012	0.108	0
Father birth 1921 - 1930	0.093	0.290	0	0.098	0.297	0
Father birth 1931 - 1940	0.308	0.461	0	0.358	0.479	0
Father birth 1941 - 1950	0.442	0.497	0	0.474	0.499	0
Father birth 1951 - 1960	0.140	0.347	0	0.057	0.233	0
Father birth 1961 - 1970	0.008	0.090	0	0.001	0.032	0
Mother birth 1900 - 1920	0.002	0.050	0	0.003	0.050	0
Mother birth 1911 - 1920	0.055	0.228	0	0.063	0.244	0
Mother birth 1921 - 1930	0.245	0.430	0	0.293	0.455	0
Mother birth 1931 - 1940	0.489	0.500	0	0.550	0.498	1
Mother birth 1941 - 1950	0.188	0.391	0	0.090	0.286	0
Mother birth 1951 - 1960	0.021	0.145	0	0.002	0.046	0
Finnish	0.952	0.213	1	0.957	0.203	1
Swedish	0.047	0.211	0	0.042	0.201	0
Other mother tongue	0.001	0.029	0	0.001	0.029	0
Number of Observations		347,914			4,754	

C Supplementary Materials for Section 3

Table C1: Who Becomes Regressions with All Covariates Reported

VARIABLES	(1)	(2)	(3)	(4)
fa income 21-40	0.000640	0.000920	0.00106	0.000694
	(0.000741)	(0.000746)	(0.000746)	(0.000744)
fa income 41-60	0.00188***	0.00223***	0.00206***	0.00134*
	(0.000708)	(0.000732)	(0.000732)	(0.000730)
fa income 61-80	0.00377***	0.00360***	0.00328***	0.00218***
	(0.000695)	(0.000727)	(0.000727)	(0.000725)
fa income 81-90	0.00784***	0.00525***	0.00444***	0.00295***
	(0.000796)	(0.000827)	(0.000825)	(0.000822)
fa income 91-95	0.0149***	0.00919***	0.00684***	0.00515***
	(0.00107)	(0.00109)	(0.00109)	(0.00108)
fa income 96-100	0.0246***	0.0154***	0.00938***	0.00745***
	(0.00126)	(0.00131)	(0.00130)	(0.00130)
mo income 21-40	0.000346	0.000348	-0.000160	-0.000272
	(0.000478)	(0.000493)	(0.000492)	(0.000490)
mo income 41-60	0.00207***	0.00150***	0.000656	0.000340
	(0.000543)	(0.000569)	(0.000568)	(0.000566)
mo income 61-80	0.00513***	0.00267***	0.00105	0.000485
	(0.000783)	(0.000813)	(0.000819)	(0.000817)
mo income 81-90	0.0109***	0.00505***	0.00157	0.000761
	(0.00170)	(0.00174)	(0.00177)	(0.00176)
mo income 91-95	0.0126***	0.00627**	-0.000846	-0.00186
	(0.00307)	(0.00311)	(0.00315)	(0.00314)
mo income 96-100	0.00260**	0.00216*	0.000139	-0.000410
	(0.00114)	(0.00115)	(0.00112)	(0.00112)
fa bluecollar		-0.00121**	-0.000999*	-0.000759
		(0.000543)	(0.000542)	(0.000540)
fa jr whitec.		0.00269***	0.00281***	0.00184**
		(0.000727)	(0.000738)	(0.000735)
fa sr whitec.		0.00883***	0.00402***	0.00270**
		(0.00102)	(0.00112)	(0.00112)
mo bluecollar		-0.00101*	-0.000263	4.32e-05
		(0.000551)	(0.000551)	(0.000550)
mo jr whitec.		0.00186***	0.00211***	0.00146**
		(0.000621)	(0.000621)	(0.000619)
mo sr whitec.		0.00884***	0.00431***	0.00333***
		(0.00119)	(0.00125)	(0.00125)
fa secondary			-0.00420***	-0.00529***
			(0.000785)	(0.000782)
fa college			0.00391***	0.00127
			(0.00139)	(0.00138)

– Continued on the next page –

– Table C1 (Continued) –

fa MSc			0.0119***	0.00876***
			(0.00175)	(0.00175)
fa PhD			0.0310***	0.0275***
			(0.00435)	(0.00434)
mo secondary			0.00392***	0.00272***
			(0.000467)	(0.000466)
mo college			0.00645***	0.00397***
			(0.00128)	(0.00128)
mo MSc			0.0152***	0.0119***
			(0.00242)	(0.00242)
mo PhD			0.0123	0.00826
			(0.00957)	(0.00957)
fa STEM			0.00889***	0.00861***
			(0.000801)	(0.000798)
mo STEM			-0.00112	-0.00116
			(0.000734)	(0.000732)
IQ 1-10				-0.00409***
				(0.000535)
IQ 11-20				-0.00404***
				(0.000558)
IQ 21-30				-0.00320***
				(0.000591)
IQ 31-40				-0.00193***
				(0.000629)
IQ 51-60				0.00147**
				(0.000714)
IQ 61-70				0.00412***
				(0.000785)
IQ 71-80				0.00960***
				(0.000871)
IQ 81-90				0.0147***
				(0.000968)
IQ 91-95				0.0236***
				(0.00144)
IQ 96-100				0.0351***
				(0.00165)
Observations	352,668	352,668	352,668	352,668

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All specifications include a 4th order polynomial in log(age), 21 region dummies, dummies for suburban and urban regions, dummies for mother tongue, and dummies for parental decade of birth.

Table C2: Patent Count Regression

VARIABLES	(1)	(2)	(3)	(4)
fa income 21-40	-0.00300	-0.00269	-0.00252	-0.00301
	(0.00230)	(0.00228)	(0.00228)	(0.00228)
fa income 41-60	0.000430	0.000679	0.000403	-0.000587
	(0.00233)	(0.00231)	(0.00230)	(0.00230)
fa income 61-80	0.00198	0.00137	0.00114	-0.000380
	(0.00231)	(0.00226)	(0.00225)	(0.00226)
fa income 81-90	0.00753***	0.00324	0.00253	0.000445
	(0.00254)	(0.00248)	(0.00245)	(0.00245)
fa income 91-95	0.0196***	0.0108***	0.00759**	0.00521
	(0.00346)	(0.00331)	(0.00324)	(0.00325)
fa income 96-100	0.0363***	0.0225***	0.0107**	0.00802*
	(0.00459)	(0.00460)	(0.00444)	(0.00443)
mo income 21-40	0.00184	0.00157	0.000492	0.000350
	(0.00132)	(0.00132)	(0.00128)	(0.00128)
mo income 41-60	0.00563***	0.00444***	0.00272*	0.00230
	(0.00153)	(0.00160)	(0.00160)	(0.00159)
mo income 61-80	0.00708***	0.00330	-0.000316	-0.00111
	(0.00213)	(0.00224)	(0.00223)	(0.00223)
mo income 81-90	0.0273***	0.0192***	0.0114*	0.0102
	(0.00646)	(0.00641)	(0.00648)	(0.00647)
mo income 91-95	0.0189*	0.0102	-0.00444	-0.00596
	(0.0100)	(0.0101)	(0.0100)	(0.0100)
mo income 96-100	0.00773*	0.00745*	0.00327	0.00239
	(0.00407)	(0.00406)	(0.00367)	(0.00367)
fa bluecollar		-0.00179	-0.00145	-0.00113
		(0.00136)	(0.00136)	(0.00135)
fa jr whitec.		0.00507**	0.00572***	0.00427**
		(0.00197)	(0.00201)	(0.00201)
fa sr whitec.		0.0150***	0.00637*	0.00437
		(0.00346)	(0.00338)	(0.00336)
mo bluecollar		-0.000224	0.00132	0.00175
		(0.00138)	(0.00134)	(0.00134)
mo jr whitec.		0.00345**	0.00469***	0.00378**
		(0.00176)	(0.00171)	(0.00170)
mo sr whitec.		0.0110***	0.00117	-0.000253
		(0.00418)	(0.00470)	(0.00471)
fa secondary			-0.00762***	-0.00915***
			(0.00260)	(0.00260)

– Continued on the next page –

– Table C2 (Continued) –

fa college			0.00266	-0.00123
			(0.00380)	(0.00380)
fa MSc			0.0226***	0.0179***
			(0.00602)	(0.00600)
fa PhD			0.0937***	0.0883***
			(0.0226)	(0.0225)
mo secondary			0.00504***	0.00331***
			(0.00124)	(0.00123)
mo college			0.0165***	0.0129***
			(0.00449)	(0.00446)
mo MSc			0.0295***	0.0246***
			(0.00870)	(0.00869)
mo PhD			-0.0340**	-0.0403***
			(0.0154)	(0.0155)
fa STEM			0.0108***	0.0103***
			(0.00259)	(0.00259)
mo STEM			-0.000564	-0.000636
			(0.00199)	(0.00198)
IQ 1-10				-0.00383***
				(0.00111)
IQ 11-20				-0.00456***
				(0.00110)
IQ 21-30				-0.00406***
				(0.00119)
IQ 31-40				-0.00339***
				(0.00131)
IQ 51-60				0.00222
				(0.00178)
IQ 61-70				0.00641***
				(0.00223)
IQ 71-80				0.0126***
				(0.00224)
IQ 81-90				0.0198***
				(0.00248)
IQ 91-95				0.0340***
				(0.00438)
IQ 96-100				0.0589***
				(0.00576)
Observations	352,668	352,668	352,668	352,668

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All specifications include a 4th order polynomial in log(age), 21 region dummies, dummies for suburban and urban regions, dummies for mother tongue, and dummies for parental decade of birth.

Table C3: Citations Regression

VARIABLES	(1)	(2)	(3)	(4)
fa income 21-40	-0.00218	-0.00178	-0.00166	-0.00195
	(0.00157)	(0.00158)	(0.00158)	(0.00158)
fa income 41-60	0.00112	0.00176	0.00159	0.000998
	(0.00164)	(0.00169)	(0.00169)	(0.00169)
fa income 61-80	0.00203	0.00224	0.00200	0.00108
	(0.00157)	(0.00165)	(0.00165)	(0.00165)
fa income 81-90	0.00607***	0.00419**	0.00351*	0.00225
	(0.00186)	(0.00190)	(0.00190)	(0.00190)
fa income 91-95	0.0157***	0.0112***	0.00905***	0.00761**
	(0.00295)	(0.00302)	(0.00300)	(0.00298)
fa income 96-100	0.0204***	0.0132***	0.00712**	0.00546*
	(0.00292)	(0.00294)	(0.00298)	(0.00298)
mo income 21-40	0.000192	-0.000123	-0.000614	-0.000696
	(0.00111)	(0.00114)	(0.00113)	(0.00113)
mo income 41-60	0.00258*	0.00151	0.000714	0.000453
	(0.00132)	(0.00145)	(0.00143)	(0.00143)
mo income 61-80	0.00131	-0.00134	-0.00295	-0.00344*
	(0.00172)	(0.00180)	(0.00187)	(0.00187)
mo income 81-90	0.0146***	0.00971**	0.00629	0.00551
	(0.00412)	(0.00431)	(0.00440)	(0.00441)
mo income 91-95	0.0173*	0.0123	0.00565	0.00473
	(0.00974)	(0.00972)	(0.00946)	(0.00947)
mo income 96-100	0.00605**	0.00575*	0.00374	0.00321
	(0.00306)	(0.00306)	(0.00268)	(0.00268)
fa bluecollar		-0.00270**	-0.00241*	-0.00223*
		(0.00132)	(0.00131)	(0.00130)
fa jr whitec.		0.000841	0.00105	0.000167
		(0.00185)	(0.00187)	(0.00186)
fa sr whitec.		0.00619**	0.00105	-0.000179
		(0.00268)	(0.00296)	(0.00296)
mo bluecollar		0.000399	0.00109	0.00135
		(0.00114)	(0.00113)	(0.00113)
mo jr whitec.		0.00393***	0.00429***	0.00373**
		(0.00149)	(0.00147)	(0.00147)
mo sr whitec.		0.00767***	0.00309	0.00223
		(0.00296)	(0.00325)	(0.00324)
fa secondary			-0.00386*	-0.00480**
			(0.00205)	(0.00204)
fa college			0.00460	0.00223
			(0.00378)	(0.00378)

– Continued on the next page –

– Table C3 (Continued) –

fa MSc			0.0128***	0.00999**
			(0.00463)	(0.00462)
fa PhD			0.0376***	0.0343***
			(0.0117)	(0.0117)
mo secondary			0.00355***	0.00250**
			(0.00110)	(0.00110)
mo college			0.00687**	0.00463
			(0.00335)	(0.00334)
mo MSc			0.0120**	0.00902
			(0.00569)	(0.00568)
mo PhD			0.0193	0.0155
			(0.0287)	(0.0287)
fa STEM			0.00715***	0.00688***
			(0.00210)	(0.00210)
mo STEM			0.000253	0.000207
			(0.00183)	(0.00183)
IQ 1-10				-0.00109
				(0.00100)
IQ 11-20				-0.00169
				(0.00103)
IQ 21-30				-0.000966
				(0.00111)
IQ 31-40				-0.000110
				(0.00123)
IQ 51-60				0.000652
				(0.00141)
IQ 61-70				0.00538***
				(0.00179)
IQ 71-80				0.0109***
				(0.00209)
IQ 81-90				0.0128***
				(0.00214)
IQ 91-95				0.0214***
				(0.00352)
IQ 96-100				0.0371***
				(0.00431)
Observations	352,668	352,668	352,668	352,668

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All specifications include a 4th order polynomial in log(age), 21 region dummies, dummies for suburban and urban regions, dummies for mother tongue, and dummies for parental decade of birth. The dependent variable is the number of citations in the first 5 years of the patent's life.

Table C4: Star Inventor Regression

VARIABLES	(1)	(2)	(3)	(4)
fa income 21-40	0.000232	0.000272	0.000290	0.000253
	(0.000220)	(0.000221)	(0.000221)	(0.000221)
fa income 41-60	0.000497**	0.000560**	0.000544**	0.000467**
	(0.000215)	(0.000226)	(0.000226)	(0.000225)
fa income 61-80	0.000447**	0.000467**	0.000444**	0.000326
	(0.000205)	(0.000217)	(0.000217)	(0.000216)
fa income 81-90	0.000932***	0.000705***	0.000648***	0.000487*
	(0.000243)	(0.000252)	(0.000250)	(0.000249)
fa income 91-95	0.00172***	0.00117***	0.000951***	0.000768**
	(0.000339)	(0.000344)	(0.000341)	(0.000341)
fa income 96-100	0.00259***	0.00166***	0.000974**	0.000764*
	(0.000401)	(0.000415)	(0.000419)	(0.000420)
mo income 21-40	9.87e-05	0.000118	6.83e-05	5.71e-05
	(0.000158)	(0.000164)	(0.000163)	(0.000163)
mo income 41-60	0.000223	0.000203	0.000125	9.14e-05
	(0.000177)	(0.000187)	(0.000186)	(0.000186)
mo income 61-80	8.88e-05	-0.000126	-0.000272	-0.000336
	(0.000234)	(0.000246)	(0.000244)	(0.000245)
mo income 81-90	0.00198***	0.00135**	0.000990	0.000895
	(0.000625)	(0.000638)	(0.000648)	(0.000648)
mo income 91-95	0.00254**	0.00185	0.000997	0.000891
	(0.00118)	(0.00118)	(0.00120)	(0.00120)
mo income 96-100	0.000322	0.000268	-1.47e-05	-7.80e-05
	(0.000381)	(0.000385)	(0.000373)	(0.000373)
fa bluecollar		-0.000229	-0.000215	-0.000191
		(0.000190)	(0.000188)	(0.000188)
fa jr whitec.		0.000177	0.000202	9.59e-05
		(0.000247)	(0.000251)	(0.000251)
fa sr whitec.		0.000818**	0.000347	0.000200
		(0.000341)	(0.000394)	(0.000394)
mo bluecollar		6.09e-06	7.80e-05	0.000111
		(0.000187)	(0.000187)	(0.000187)
mo jr whitec.		0.000145	0.000187	0.000117
		(0.000212)	(0.000212)	(0.000211)
mo sr whitec.		0.00105***	0.000564	0.000458
		(0.000399)	(0.000424)	(0.000424)
fa secondary			-0.000415	-0.000534*
			(0.000272)	(0.000272)
fa college			0.000137	-0.000151
			(0.000463)	(0.000464)

– Continued on the next page –

– Table C4 (Continued) –

fa MSc			0.00157**	0.00122*
			(0.000627)	(0.000626)
fa PhD			0.00326**	0.00286*
			(0.00150)	(0.00149)
mo secondary			0.000321**	0.000192
			(0.000154)	(0.000154)
mo college			0.000580	0.000310
			(0.000430)	(0.000429)
mo MSc			0.00135*	0.000989
			(0.000812)	(0.000810)
mo PhD			0.00723	0.00678
			(0.00473)	(0.00474)
fa STEM			0.000813***	0.000783***
			(0.000271)	(0.000271)
mo STEM			-5.45e-05	-5.96e-05
			(0.000246)	(0.000246)
IQ 1-10				-0.000293*
				(0.000165)
IQ 11-20				-0.000274
				(0.000172)
IQ 21-30				-0.000196
				(0.000182)
IQ 31-40				-8.07e-06
				(0.000198)
IQ 51-60				9.60e-05
				(0.000211)
IQ 61-70				0.000636**
				(0.000252)
IQ 71-80				0.00164***
				(0.000300)
IQ 81-90				0.00132***
				(0.000291)
IQ 91-95				0.00260***
				(0.000467)
IQ 96-100				0.00423***
				(0.000562)
Observations	352,668	352,668	352,668	352,668

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All specifications include a 4th order polynomial in log(age), 21 region dummies, dummies for suburban and urban regions, dummies for mother tongue, and dummies for parental decade of birth.

Table C5: Who becomes regression with parental wages measured over several years

VARIABLES	(1)	(2)	(3)	(4)
fa income 21-40	-4.91e-05	0.000199	0.000132	2.82e-05
	(0.000726)	(0.000734)	(0.000734)	(0.000732)
fa income 41-60	0.00177**	0.00216***	0.00182**	0.00115
	(0.000698)	(0.000727)	(0.000727)	(0.000725)
fa income 61-80	0.00378***	0.00376***	0.00326***	0.00224***
	(0.000680)	(0.000721)	(0.000721)	(0.000719)
fa income 81-90	0.00723***	0.00512***	0.00420***	0.00275***
	(0.000777)	(0.000813)	(0.000812)	(0.000810)
fa income 91-95	0.0137***	0.00812***	0.00579***	0.00405***
	(0.00102)	(0.00105)	(0.00105)	(0.00105)
fa income 96-100	0.0240***	0.0148***	0.00906***	0.00709***
	(0.00116)	(0.00121)	(0.00121)	(0.00121)
mo income 21-40	-0.000218	4.07e-05	-0.000231	-0.000388
	(0.000472)	(0.000487)	(0.000486)	(0.000484)
mo income 41-60	0.00135**	0.00108*	0.000438	1.75e-05
	(0.000533)	(0.000567)	(0.000568)	(0.000566)
mo income 61-80	0.00561***	0.00348***	0.00207**	0.00133
	(0.000792)	(0.000831)	(0.000840)	(0.000837)
mo income 81-90	0.00613***	0.000266	-0.000820	-0.00181
	(0.00181)	(0.00185)	(0.00188)	(0.00187)
mo income 91-95	0.0102***	0.00608*	-0.000114	-0.000965
	(0.00324)	(0.00326)	(0.00333)	(0.00332)
mo income 96-100	0.0157***	0.0119***	0.00170	0.000573
	(0.00421)	(0.00422)	(0.00434)	(0.00432)
fa bluecollar		-0.00148***	-0.00118**	-0.000890
		(0.000547)	(0.000546)	(0.000544)
fa jr whitec.		0.00238***	0.00264***	0.00173**
		(0.000730)	(0.000740)	(0.000738)
fa sr whitec.		0.00845***	0.00385***	0.00258**
		(0.00102)	(0.00113)	(0.00112)
mo bluecollar		-0.00107*	-0.000280	8.01e-05
		(0.000553)	(0.000552)	(0.000551)
mo jr whitec.		0.00170***	0.00199***	0.00145**
		(0.000629)	(0.000628)	(0.000626)
mo sr whitec.		0.00891***	0.00443***	0.00354***
		(0.00120)	(0.00126)	(0.00125)
fa secondary			-0.00424***	-0.00530***
			(0.000786)	(0.000783)
fa college			0.00386***	0.00125
			(0.00139)	(0.00139)

– Continued on the next page –

– Table C5 (Continued) –

fa MSc			0.0118***	0.00863***
			(0.00176)	(0.00175)
fa PhD			0.0309***	0.0273***
			(0.00436)	(0.00435)
mo secondary			0.00387***	0.00269***
			(0.000467)	(0.000466)
mo college			0.00644***	0.00401***
			(0.00128)	(0.00127)
mo MSc			0.0151***	0.0120***
			(0.00250)	(0.00249)
mo PhD			0.0120	0.00809
			(0.00960)	(0.00960)
fa STEM			0.00892***	0.00863***
			(0.000801)	(0.000798)
mo STEM			-0.00106	-0.00112
			(0.000734)	(0.000732)
IQ 1-10				-0.00408***
				(0.000536)
IQ 11-20				-0.00404***
				(0.000559)
IQ 21-30				-0.00319***
				(0.000591)
IQ 31-40				-0.00193***
				(0.000630)
IQ 51-60				0.00146**
				(0.000715)
IQ 61-70				0.00412***
				(0.000786)
IQ 71-80				0.00960***
				(0.000872)
IQ 81-90				0.0147***
				(0.000969)
IQ 91-95				0.0236***
				(0.00144)
IQ 96-100				0.0351***
				(0.00165)
Observations	352,221	352,221	352,221	352,221

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All specifications include a 4th order polynomial in log(age), 21 region dummies, dummies for suburban and urban regions, dummies for mother tongue, and dummies for parental decade of birth.

Table C6: Who becomes inventor regression with all IQ measures

VARIABLES		VARIABLES		VARIABLES	
fa income 21-40	0.000634	fa college	-0.000369	verbal IQ 1-10	0.000396
	(0.000742)		(0.00138)		(0.000593)
fa income 41-60	0.00114	fa MSc	0.00620***	verbal IQ 11-20	-0.000254
	(0.000728)		(0.00175)		(0.000569)
fa income 61-80	0.00182**	fa PhD	0.0243***	verbal IQ 21-30	-0.00134**
	(0.000723)		(0.00433)		(0.000580)
fa income 81-90	0.00235***	mo secondary	0.00208***	verbal IQ 31-40	-0.000747
	(0.000821)		(0.000465)		(0.000630)
fa income 91-95	0.00435***	mo college	0.00209	verbal IQ 51-60	-0.000813
	(0.00108)		(0.00128)		(0.000713)
fa income 96-100	0.00647***	mo MSc	0.00873***	verbal IQ 61-70	0.00123
	(0.00130)		(0.00242)		(0.000799)
mo income 21-40	-0.000204	mo PhD	0.00495	verbal IQ 71-80	0.00250***
	(0.000489)		(0.00955)		(0.000871)
mo income 41-60	0.000321	fa STEM	0.00884***	verbal IQ 81-90	0.00493***
	(0.000565)		(0.000798)		(0.000981)
mo income 61-80	0.000222	mo STEM	-0.000752	verbal IQ 91-95	0.00815***
	(0.000816)		(0.000731)		(0.00146)
mo income 81-90	0.000483	visuospatial IQ 1-10	-0.000907	verbal IQ 96-100	0.0149***
	(0.00176)		(0.000553)		(0.00174)
mo income 91-95	-0.00238	visuospatial 11-20	-0.00131**	analytic IQ 1-10	-0.000785
	(0.00314)		(0.000560)		(0.000527)
mo income 96-100	-1.96e-05	visuospatial IQ 21-30	-0.00131**	analytic IQ 11-20	-0.000698
	(0.00111)		(0.000589)		(0.000516)
fa bluecollar	-0.000580	visuospatial IQ 31-40	-0.000927	analytic IQ 21-30	-0.000442
	(0.000539)		(0.000629)		(0.000539)
fa jr whitec.	0.00132*	visuospatial IQ 51-60	0.000239	analytic IQ 31-40	-0.000486
	(0.000734)		(0.000714)		(0.000576)
fa sr whitec.	0.00203*	visuospatial IQ 61-70	0.00147*	analytic IQ 51-60	0.00152**
	(0.00112)		(0.000788)		(0.000687)
mo bluecollar	0.000265	visuospatial IQ 71-80	0.00519***	analytic IQ 61-70	0.00349***
	(0.000549)		(0.000873)		(0.000770)
mo jr whitec.	0.00114*	visuospatial IQ 81-90	0.00787***	analytic IQ 71-80	0.00544***
	(0.000618)		(0.000974)		(0.000843)
mo sr whitec.	0.00286**	visuospatial IQ 91-95	0.0143***	analytic IQ 81-90	0.0105***
	(0.00124)		(0.00146)		(0.000984)
fa secondary	-0.00591***	visuospatial IQ 96-100	0.0219***	analytic IQ 91-95	0.0154***
	(0.000783)		(0.00169)		(0.00153)
				analytic IQ 96-100	0.0187***
					(0.00174)
Observations: 352,668					

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All specifications include a 4th order polynomial in log(age), 21 region dummies, dummies for suburban and urban regions, dummies for mother tongue, and dummies for parental decade of birth.

Table C7: Who becomes MD regression

VARIABLES	(1)	(2)	(3)	(4)
fa income 21-40	-0.00117***	-0.00112***	-0.00102***	-0.00111***
	(0.000384)	(0.000388)	(0.000386)	(0.000386)
fa income 41-60	0.000304	0.000133	1.04e-06	-0.000163
	(0.000381)	(0.000399)	(0.000399)	(0.000398)
fa income 61-80	0.000714*	0.000197	6.05e-05	-0.000188
	(0.000376)	(0.000397)	(0.000396)	(0.000396)
fa income 81-90	0.00300***	0.00105**	0.000653	0.000320
	(0.000451)	(0.000470)	(0.000466)	(0.000465)
fa income 91-95	0.00588***	0.00189***	0.000823	0.000453
	(0.000637)	(0.000648)	(0.000645)	(0.000645)
fa income 96-100	0.0192***	0.0127***	0.00796***	0.00754***
	(0.000970)	(0.000941)	(0.000886)	(0.000886)
mo income 21-40	0.000105	0.000112	-0.000425	-0.000454*
	(0.000261)	(0.000272)	(0.000271)	(0.000271)
mo income 41-60	0.00121***	0.000950***	-5.04e-06	-7.90e-05
	(0.000309)	(0.000325)	(0.000324)	(0.000324)
mo income 61-80	0.00438***	0.00286***	0.000879	0.000752
	(0.000520)	(0.000540)	(0.000545)	(0.000545)
mo income 81-90	0.00887***	0.00453***	0.000431	0.000278
	(0.00128)	(0.00130)	(0.00132)	(0.00132)
mo income 91-95	0.0174***	0.0126***	0.00497*	0.00477*
	(0.00280)	(0.00282)	(0.00283)	(0.00283)
mo income 96-100	0.00642***	0.00634***	0.00387***	0.00377***
	(0.000926)	(0.000934)	(0.000860)	(0.000861)
fa bluecollar		0.000343	0.000703**	0.000758**
		(0.000305)	(0.000304)	(0.000304)
fa jr whitec.		0.00172***	0.00120***	0.00100**
		(0.000426)	(0.000432)	(0.000431)
fa sr whitec.		0.00592***	0.00145*	0.00119
		(0.000685)	(0.000739)	(0.000740)
mo bluecollar		-0.000282	0.000427	0.000493
		(0.000303)	(0.000302)	(0.000301)
mo jr whitec.		0.00103***	0.00149***	0.00134***
		(0.000374)	(0.000371)	(0.000371)
mo sr whitec.		0.00770***	0.00343***	0.00323***
		(0.000871)	(0.000903)	(0.000903)
fa secondary			0.00317***	0.00293***
			(0.000513)	(0.000513)
fa college			0.00466***	0.00412***
			(0.000900)	(0.000899)

– Continued on the next page –

– Table C7 (Continued) –

fa MSc			0.0140***	0.0134***
			(0.00132)	(0.00132)
fa PhD			0.0365***	0.0358***
			(0.00396)	(0.00396)
mo secondary			0.00196***	0.00170***
			(0.000255)	(0.000254)
mo college			0.00674***	0.00624***
			(0.000935)	(0.000934)
mo MSc			0.0133***	0.0127***
			(0.00197)	(0.00197)
mo PhD			0.0134	0.0127
			(0.00875)	(0.00875)
fa STEM			-0.00282***	-0.00286***
			(0.000545)	(0.000545)
mo STEM			-0.00172***	-0.00172***
			(0.000403)	(0.000403)
IQ 1-10				-0.00131***
				(0.000340)
IQ 11-20				-0.00159***
				(0.000357)
IQ 21-30				-0.000917**
				(0.000398)
IQ 31-40				-0.000466
				(0.000430)
IQ 51-60				0.000711
				(0.000483)
IQ 61-70				9.83e-05
				(0.000497)
IQ 71-80				0.00254***
				(0.000554)
IQ 81-90				0.00317***
				(0.000599)
IQ 91-95				0.00436***
				(0.000835)
IQ 96-100				0.00559***
				(0.000909)
Observations	352,668	352,668	352,668	352,668

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All specifications include a 4th order polynomial in log(age), 21 region dummies, dummies for suburban and urban regions, dummies for mother tongue, and dummies for parental decade of birth.

Table C8: Who becomes lawyer regression

VARIABLES	(2)	(3)	(4)	(5)
fa income 21-40	-0.000616	-0.000536	-0.000469	-0.000533
	(0.000398)	(0.000401)	(0.000399)	(0.000399)
fa income 41-60	-0.000554	-0.000493	-0.000636	-0.000753*
	(0.000377)	(0.000396)	(0.000395)	(0.000395)
fa income 61-80	0.000340	8.98e-05	-8.42e-05	-0.000258
	(0.000383)	(0.000405)	(0.000405)	(0.000405)
fa income 81-90	0.00316***	0.00154***	0.00106**	0.000839*
	(0.000474)	(0.000490)	(0.000489)	(0.000489)
fa income 91-95	0.00599***	0.00262***	0.00161**	0.00138**
	(0.000674)	(0.000694)	(0.000697)	(0.000697)
fa income 96-100	0.0185***	0.0133***	0.00901***	0.00875***
	(0.000971)	(0.000997)	(0.000983)	(0.000983)
mo income 21-40	0.000976***	0.00103***	0.000517*	0.000491*
	(0.000280)	(0.000295)	(0.000294)	(0.000294)
mo income 41-60	0.00177***	0.00155***	0.000599*	0.000545
	(0.000330)	(0.000356)	(0.000354)	(0.000354)
mo income 61-80	0.00483***	0.00363***	0.00159***	0.00151***
	(0.000550)	(0.000572)	(0.000573)	(0.000573)
mo income 81-90	0.00906***	0.00612***	0.00187	0.00183
	(0.00133)	(0.00136)	(0.00137)	(0.00137)
mo income 91-95	0.0169***	0.0137***	0.00616**	0.00607**
	(0.00284)	(0.00286)	(0.00287)	(0.00287)
mo income 96-100	0.00241***	0.00217***	-0.000251	-0.000264
	(0.000778)	(0.000787)	(0.000751)	(0.000751)
fa bluecollar		-8.36e-05	0.000435	0.000481
		(0.000317)	(0.000315)	(0.000315)
fa jr whitec.		0.00225***	0.00145***	0.00137***
		(0.000468)	(0.000475)	(0.000475)
fa sr whitec.		0.00613***	0.00135	0.00127
		(0.000761)	(0.000848)	(0.000848)
mo bluecollar		-0.00106***	-0.000427	-0.000395
		(0.000330)	(0.000327)	(0.000327)
mo jr whitec.		0.000259	0.000665	0.000572
		(0.000411)	(0.000411)	(0.000411)
mo sr whitec.		0.00344***	-0.000531	-0.000631
		(0.000904)	(0.000961)	(0.000961)
fa secondary			0.00567***	0.00552***
			(0.000570)	(0.000571)
fa college			0.00699***	0.00676***
			(0.00100)	(0.00100)

– Continued on the next page –

– Table C8 (Continued) –

fa MSc			0.0177***	0.0175***
			(0.00154)	(0.00154)
fa PhD			0.0173***	0.0171***
			(0.00329)	(0.00329)
mo secondary			0.00154***	0.00140***
			(0.000275)	(0.000275)
mo college			0.00592***	0.00571***
			(0.000988)	(0.000989)
mo MSc			0.0145***	0.0143***
			(0.00204)	(0.00205)
mo PhD			0.00404	0.00387
			(0.00707)	(0.00708)
fa STEM			-0.00574***	-0.00573***
			(0.000583)	(0.000582)
mo STEM			-0.00149***	-0.00148***
			(0.000406)	(0.000406)
IQ 1-10				-0.00194***
				(0.000399)
IQ 11-20				-0.00184***
				(0.000432)
IQ 21-30				-0.00123***
				(0.000469)
IQ 31-40				-0.000318
				(0.000510)
IQ 51-60				0.000809
				(0.000557)
IQ 61-70				0.000621
				(0.000579)
IQ 71-80				0.00143**
				(0.000598)
IQ 81-90				0.00178***
				(0.000633)
IQ 91-95				0.000948
				(0.000808)
IQ 96-100				-0.00107
				(0.000794)
Observations	352,668	352,668	352,668	352,668

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All specifications include a 4th order polynomial in log(age), 21 region dummies, dummies for suburban and urban regions, dummies for mother tongue, and dummies for parental decade of birth.

D Supplementary Materials for Section 4

Table D1: Who becomes inventor - family FE estimations

	– window size 0 years –		*– window size 1 year –*		*– window size 3 years –*	
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
first born	-0.00289	-0.00364	-0.00215	-0.00325	-0.00209**	-0.000933
	(0.00335)	(0.00356)	(0.00148)	(0.00214)	(0.000869)	(0.00139)
fa income 21-40	0.000662	-0.00484	-0.00200	-0.0153	-0.00214	0.00892
	(0.00463)	(0.00391)	(0.00287)	(0.0560)	(0.00155)	(0.0172)
fa income 41-60	0.00459	-0.0123*	0.000256	-0.0175	0.000524	0.00620
	(0.00503)	(0.00669)	(0.00273)	(0.0555)	(0.00154)	(0.0155)
fa income 61-80	0.00457	-0.00320	0.00157	-0.0443	0.00147	-0.00227
	(0.00484)	(0.00489)	(0.00278)	(0.0579)	(0.00154)	(0.0162)
fa income 81-90	-0.000511	-0.00789	4.71e-05	-0.0267	0.00218	0.00212
	(0.00665)	(0.00874)	(0.00327)	(0.0569)	(0.00182)	(0.0186)
fa income 91-95	0.00133	-0.00893	0.00203	-0.00613	0.00277	-0.0101
	(0.00856)	(0.00792)	(0.00438)	(0.0626)	(0.00231)	(0.0212)
fa income 96-100	0.0182	-0.00299	0.0145**	-0.0989	0.0113***	-0.0272
	(0.0136)	(0.00958)	(0.00599)	(0.0748)	(0.00292)	(0.0276)
mo income 21-40	0.00402	0.00382	0.00182	0.0130	-0.00121	8.55e-05
	(0.00386)	(0.00599)	(0.00194)	(0.0286)	(0.00105)	(0.0103)
mo income 41-60	0.00236	-5.73e-05	0.000232	-0.000192	-0.00131	-0.00417
	(0.00407)	(0.00632)	(0.00225)	(0.0196)	(0.00127)	(0.00720)
mo income 61-80	0.00609	0.00619	0.000388	-0.00286	-0.00196	-0.00424
	(0.00658)	(0.00706)	(0.00333)	(0.0325)	(0.00184)	(0.00801)
mo income 81-90	0.0189		0.00655	-0.0676	0.000476	-0.0191
	(0.0266)		(0.0101)	(0.153)	(0.00449)	(0.0329)
mo income 91-95	-0.0349*		-0.0120	-0.211	0.00375	-0.00512
	(0.0183)		(0.0107)	(0.161)	(0.00790)	(0.0481)
mo income 96-100	-0.00429	0.00593	0.000810	0.0117	0.00393	0.00693
	(0.0101)	(0.00407)	(0.00506)	(0.0222)	(0.00269)	(0.00778)
fa bluecollar	-0.00423		0.000159		0.000190	
	(0.00274)		(0.00200)		(0.00114)	
fa jr whitec.	0.00427		0.00155		0.00381**	
	(0.00551)		(0.00301)		(0.00168)	
fa sr whitec.	0.0242**		0.00569		0.00631**	
	(0.0113)		(0.00497)		(0.00256)	
mo bluecollar	-0.00241		-0.00376*		-0.00127	
	(0.00265)		(0.00211)		(0.00118)	
mo jr whitec.	0.00189		-0.00167		0.00174	
	(0.00329)		(0.00270)		(0.00142)	
mo sr whitec.	-0.0110		0.000191		-0.000173	
	(0.0123)		(0.00530)		(0.00279)	

– Continued on the next page –

– Table D1: (Continued) –

fa secondary	-0.00370		-0.00398		-0.00700***	
	(0.00833)		(0.00345)		(0.00178)	
fa college	0.0129		0.00999		0.00210	
	(0.0152)		(0.00688)		(0.00314)	
fa MSc	-0.0335***		-0.00219		0.000658	
	(0.0119)		(0.00756)		(0.00370)	
fa PhD	-0.0216		0.0113		0.0281***	
	(0.0277)		(0.0170)		(0.00915)	
mo secondary	-0.00326		0.00238		0.00399***	
	(0.00438)		(0.00186)		(0.00103)	
mo college	0.00285		0.00692		0.00462*	
	(0.0151)		(0.00600)		(0.00278)	
mo MSc	0.00901		0.00820		0.0139***	
	(0.0235)		(0.00989)		(0.00524)	
mo PhD	0.0826		-0.00326		-0.0166	
	(0.0759)		(0.0306)		(0.0147)	
fa STEM	0.00475		0.00779**		0.0101***	
	(0.00853)		(0.00354)		(0.00179)	
mo STEM	0.00408		0.000653		-0.000522	
	(0.00643)		(0.00291)		(0.00166)	
IQ 1-10	-0.00184	-0.00616	-0.00170	-0.00406	-0.00644***	-0.00644***
	(0.00433)	(0.00919)	(0.00212)	(0.00318)	(0.00121)	(0.00185)
IQ 11-20	0.00109	0.00204	-0.00352*	-0.00429	-0.00586***	-0.00580***
	(0.00530)	(0.00817)	(0.00201)	(0.00301)	(0.00128)	(0.00188)
IQ 21-30	0.0117	0.0179*	0.000173	-0.00192	-0.00315**	-0.00296
	(0.00766)	(0.0106)	(0.00242)	(0.00346)	(0.00143)	(0.00216)
IQ 31-40	0.00238	0.00338	-0.000773	-0.00214	-0.00383***	-0.00262
	(0.00577)	(0.00803)	(0.00237)	(0.00315)	(0.00143)	(0.00191)
IQ 51-60	0.00252	-0.00734	0.000773	-0.00504	-0.000293	-0.00338
	(0.00577)	(0.00882)	(0.00262)	(0.00383)	(0.00160)	(0.00226)
IQ 61-70	0.00403	-0.000872	0.00745**	0.000838	0.00226	-0.00114
	(0.00646)	(0.00870)	(0.00326)	(0.00426)	(0.00174)	(0.00245)
IQ 71-80	0.00437	0.000564	0.0136***	0.0141***	0.0105***	0.00698**
	(0.00698)	(0.00917)	(0.00366)	(0.00477)	(0.00203)	(0.00273)
IQ 81-90	0.0144	-0.0137	0.0108***	-7.49e-05	0.0119***	0.00783***
	(0.00904)	(0.0127)	(0.00357)	(0.00478)	(0.00208)	(0.00287)
IQ 91-95	0.0240	-0.000841	0.0237***	0.0168**	0.0216***	0.0202***
	(0.0148)	(0.0193)	(0.00569)	(0.00754)	(0.00309)	(0.00409)
IQ 96-100	0.0203*	-0.00273	0.0272***	0.0265***	0.0353***	0.0320***
	(0.0120)	(0.0139)	(0.00604)	(0.00750)	(0.00355)	(0.00457)
Observations	4,117	4,117	23,141	23,141	82,054	82,054
Number of families		2,076		11,588		41,605

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All specifications include a 4th order polynomial in log(age), 21 region dummies, dummies for suburban and urban regions, dummies for mother tongue, and dummies for parental decade of birth.

E Supplementary Materials for Section 5

Table E1: Who becomes inventor regressions allowing for family structure

VARIABLES	(1)	(2)	(3)	(4)
biol fa away	-0.00399***	-0.00309***	-0.00311***	-0.00295***
	(0.000808)	(0.000829)	(0.000829)	(0.000818)
biol mo away	-0.00343**	-0.00410**	-0.00398**	-0.00417**
	(0.00163)	(0.00164)	(0.00164)	(0.00162)
biol fa& mo away	-1.27e-05	0.00116	0.00107	0.00126
	(0.00137)	(0.00157)	(0.00157)	(0.00155)
fa income 21-40	0.000655	0.000632	0.000599	0.000606
	(0.000744)	(0.000743)	(0.000744)	(0.000744)
fa income 41-60	0.00122*	0.00125*	0.00124*	0.00124*
	(0.000730)	(0.000730)	(0.000730)	(0.000730)
fa income 61-80	0.00201***	0.00206***	0.00206***	0.00206***
	(0.000726)	(0.000726)	(0.000726)	(0.000726)
fa income 81-90	0.00279***	0.00285***	0.00285***	0.00285***
	(0.000823)	(0.000823)	(0.000823)	(0.000823)
fa income 91-95	0.00500***	0.00528***	0.00577***	0.00574***
	(0.00108)	(0.00111)	(0.00119)	(0.00119)
fa income 96-100	0.00730***	0.00772***	0.00845***	0.00836***
	(0.00130)	(0.00132)	(0.00142)	(0.00142)
mo income 21-40	-0.000167	-0.000169	-0.000187	-0.000196
	(0.000492)	(0.000492)	(0.000492)	(0.000492)
mo income 41-60	0.000526	0.000527	0.000507	0.000488
	(0.000569)	(0.000569)	(0.000569)	(0.000569)
mo income 61-80	0.000806	0.000826	0.000794	0.000759
	(0.000827)	(0.000828)	(0.000828)	(0.000828)
mo income 81-90	0.00120	0.00127	0.00115	0.00108
	(0.00177)	(0.00177)	(0.00178)	(0.00178)
mo income 91-95	-0.00137	-0.000708	-0.000442	-0.000549
	(0.00315)	(0.00327)	(0.00356)	(0.00356)
mo income 96-100	0.000214	0.000258	0.000772	0.000756
	(0.00117)	(0.00132)	(0.00134)	(0.00134)
biol fa income 91-95 x away		-0.00625*	-0.00669*	-0.00613*
		(0.00341)	(0.00344)	(0.00352)
biol fa income 96-100 x away		-0.0118**	-0.0125***	-0.00993**
		(0.00477)	(0.00480)	(0.00467)
biol mo income 91-95 x away		-0.0148***	-0.0152***	-0.0141***
		(0.00376)	(0.00401)	(0.00407)
biol mo income 96-100 x away		-0.000510	-0.00106	-0.000925
		(0.00248)	(0.00249)	(0.00250)

– Continued on the next page –

– Table E1: (Continued) –

step fa income 91-95			-0.00327	-0.00329
			(0.00223)	(0.00223)
step income 96-100			-0.00501*	-0.00504*
			(0.00282)	(0.00282)
step mo income 91-95			-0.00381	-0.00344
			(0.00859)	(0.00859)
step mo income 96-100			-0.0191**	-0.0190**
			(0.00833)	(0.00833)
fa bluecollar	-0.000861	-0.000830	-0.000825	-0.000826
	(0.000541)	(0.000541)	(0.000541)	(0.000541)
fa jr whitec.	0.00173**	0.00175**	0.00175**	0.00174**
	(0.000736)	(0.000736)	(0.000736)	(0.000736)
fa sr whitec.	0.00261**	0.00257**	0.00258**	0.00255**
	(0.00112)	(0.00112)	(0.00112)	(0.00112)
mo bluecollar	5.85e-05	3.07e-05	7.68e-05	8.20e-05
	(0.000551)	(0.000551)	(0.000551)	(0.000551)
mo jr whitec.	0.00140**	0.00137**	0.00141**	0.00142**
	(0.000621)	(0.000621)	(0.000621)	(0.000621)
mo sr whitec.	0.00326***	0.00314**	0.00316**	0.00315**
	(0.00125)	(0.00125)	(0.00125)	(0.00125)
fa secondary	-0.00529***	-0.00527***	-0.00526***	-0.00525***
	(0.000782)	(0.000782)	(0.000782)	(0.000782)
fa college	0.00126	0.00128	0.00127	0.00132
	(0.00138)	(0.00138)	(0.00138)	(0.00138)
fa MSc	0.00874***	0.00874***	0.00880***	0.00884***
	(0.00175)	(0.00175)	(0.00175)	(0.00178)
fa PhD	0.0275***	0.0275***	0.0275***	0.0278***
	(0.00434)	(0.00434)	(0.00434)	(0.00442)
mo secondary	0.00267***	0.00268***	0.00266***	0.00267***
	(0.000466)	(0.000466)	(0.000466)	(0.000466)
mo college	0.00381***	0.00380***	0.00382***	0.00385***
	(0.00128)	(0.00128)	(0.00128)	(0.00128)
mo MSc	0.0117***	0.0117***	0.0121***	0.0125***
	(0.00242)	(0.00242)	(0.00243)	(0.00251)
mo PhD	0.00794	0.00808	0.00908	0.0110
	(0.00957)	(0.00957)	(0.00961)	(0.0101)
fa STEM	0.00860***	0.00859***	0.00855***	0.00854***
	(0.000798)	(0.000798)	(0.000797)	(0.000797)
mo STEM	-0.00111	-0.00112	-0.00111	-0.00113
	(0.000732)	(0.000732)	(0.000732)	(0.000732)
biol fa MSc x away				-0.000712
				(0.00767)
biol fa PhD x away				-0.0128
				(0.0202)
biol mo MSc x away				-0.00776
				(0.00877)
biol mo PhD x away				-0.0346***
				(0.0122)

– Continued on the next page –

– Table E1: (Continued) –

IQ 1-10	-0.00407***	-0.00407***	-0.00407***	-0.00406***
	(0.000535)	(0.000535)	(0.000535)	(0.000535)
IQ 11-20	-0.00403***	-0.00403***	-0.00401***	-0.00402***
	(0.000558)	(0.000558)	(0.000558)	(0.000558)
IQ 21-30	-0.00319***	-0.00318***	-0.00317***	-0.00317***
	(0.000591)	(0.000591)	(0.000591)	(0.000591)
IQ 31-40	-0.00193***	-0.00192***	-0.00192***	-0.00192***
	(0.000629)	(0.000629)	(0.000629)	(0.000629)
IQ 51-60	0.00146**	0.00146**	0.00145**	0.00146**
	(0.000714)	(0.000714)	(0.000714)	(0.000714)
IQ 61-70	0.00412***	0.00412***	0.00412***	0.00412***
	(0.000785)	(0.000785)	(0.000785)	(0.000785)
IQ 71-80	0.00959***	0.00959***	0.00960***	0.00960***
	(0.000870)	(0.000871)	(0.000871)	(0.000870)
IQ 81-90	0.0147***	0.0147***	0.0147***	0.0147***
	(0.000968)	(0.000968)	(0.000968)	(0.000968)
IQ 91-95	0.0236***	0.0236***	0.0235***	0.0235***
	(0.00144)	(0.00144)	(0.00144)	(0.00144)
IQ 96-100	0.0351***	0.0351***	0.0350***	0.0351***
	(0.00165)	(0.00165)	(0.00165)	(0.00165)
Observations	352,668	352,668	352,668	352,668

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All specifications include a 4th order polynomial in log(age), 21 region dummies, dummies for suburban and urban regions, dummies for mother tongue, and dummies for parental decade of birth.

F Supplementary Materials for Section 6

Table F1: Who becomes inventor regression with IQ interactions

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
fa income 21-40	0.000694	0.000702	0.000694	0.00882	0.00888	0.00881
	(0.000744)	(0.000744)	(0.000744)	(0.0171)	(0.0172)	(0.0171)
fa income 41-60	0.00136*	0.00134*	0.00136*	0.00641	0.00618	0.00637
	(0.000729)	(0.000730)	(0.000729)	(0.0155)	(0.0155)	(0.0155)
fa income 61-80	0.00221***	0.00218***	0.00221***	-0.00221	-0.00233	-0.00219
	(0.000725)	(0.000725)	(0.000725)	(0.0162)	(0.0162)	(0.0162)
fa income 81-90	0.00301***	0.00295***	0.00302***	0.00195	0.00205	0.00195
	(0.000821)	(0.000822)	(0.000821)	(0.0185)	(0.0186)	(0.0185)
fa income 91-95	0.00527***	0.00515***	0.00527***	-0.00979	-0.0102	-0.00984
	(0.00108)	(0.00108)	(0.00108)	(0.0213)	(0.0212)	(0.0213)
fa income 96-100	0.00617***	0.00745***	0.00615***	-0.0280	-0.0273	-0.0281
	(0.00129)	(0.00130)	(0.00129)	(0.0276)	(0.0276)	(0.0276)
mo income 21-40	-0.000269	-0.000272	-0.000269	1.78e-05	0.000155	-2.20e-07
	(0.000490)	(0.000490)	(0.000490)	(0.0103)	(0.0103)	(0.0103)
mo income 41-60	0.000340	0.000340	0.000340	-0.00424	-0.00411	-0.00420
	(0.000566)	(0.000566)	(0.000566)	(0.00721)	(0.00720)	(0.00720)
mo income 61-80	0.000494	0.000485	0.000492	-0.00441	-0.00411	-0.00436
	(0.000817)	(0.000817)	(0.000817)	(0.00803)	(0.00800)	(0.00800)
mo income 81-90	0.000776	0.000757	0.000771	-0.0191	-0.0192	-0.0185
	(0.00176)	(0.00176)	(0.00176)	(0.0330)	(0.0324)	(0.0322)
mo income 91-95	-0.00192	-0.00185	-0.00192	-0.00368	-0.00522	-0.00403
	(0.00314)	(0.00314)	(0.00314)	(0.0485)	(0.0479)	(0.0480)
mo income 96-100	-0.000202	-0.000400	-0.000231	0.00561	0.00693	0.00562
	(0.00106)	(0.00111)	(0.00106)	(0.00779)	(0.00778)	(0.00779)
fa bluecollar	-0.000793	-0.000761	-0.000794			
	(0.000539)	(0.000540)	(0.000539)			
fa jr whitec.	0.00182**	0.00183**	0.00182**			
	(0.000735)	(0.000735)	(0.000735)			
fa sr whitec.	0.00265**	0.00270**	0.00265**			
	(0.00112)	(0.00112)	(0.00112)			
mo bluecollar	2.87e-05	4.40e-05	2.87e-05			
	(0.000549)	(0.000550)	(0.000549)			
mo jr whitec.	0.00146**	0.00147**	0.00146**			
	(0.000618)	(0.000619)	(0.000618)			
mo sr whitec.	0.00331***	0.00334***	0.00331***			
	(0.00125)	(0.00125)	(0.00125)			
fa secondary	-0.00524***	-0.00529***	-0.00524***			
	(0.000782)	(0.000782)	(0.000782)			
fa college	0.00137	0.00128	0.00136			
	(0.00138)	(0.00138)	(0.00138)			
fa MSc	0.00862***	0.00877***	0.00862***			
	(0.00175)	(0.00175)	(0.00175)			
fa PhD	0.0272***	0.0266***	0.0273***			
	(0.00434)	(0.00450)	(0.00450)			

– Continued on the next page –

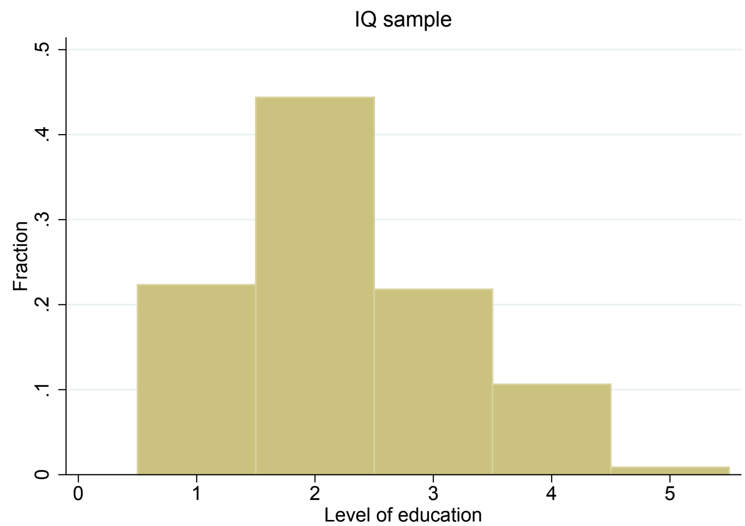
– Table F1: (Continued) –

mo secondary	0.00272***	0.00272***	0.00272***			
	(0.000466)	(0.000466)	(0.000466)			
mo college	0.00394***	0.00397***	0.00394***			
	(0.00128)	(0.00128)	(0.00128)			
mo MSc	0.0119***	0.0119***	0.0118***			
	(0.00242)	(0.00241)	(0.00241)			
mo PhD	0.00809	0.0112	0.0114			
	(0.00956)	(0.0102)	(0.0102)			
fa STEM	0.00856***	0.00861***	0.00856***			
	(0.000797)	(0.000798)	(0.000797)			
mo STEM	-0.00116	-0.00116	-0.00116			
	(0.000732)	(0.000732)	(0.000732)			
IQ 1-10	-0.00413***	-0.00409***	-0.00413***	-0.00648***	-0.00646***	-0.00647***
	(0.000535)	(0.000535)	(0.000535)	(0.00185)	(0.00185)	(0.00185)
IQ 11-20	-0.00408***	-0.00405***	-0.00408***	-0.00579***	-0.00579***	-0.00578***
	(0.000558)	(0.000558)	(0.000558)	(0.00187)	(0.00187)	(0.00187)
IQ 21-30	-0.00322***	-0.00320***	-0.00321***	-0.00296	-0.00296	-0.00294
	(0.000591)	(0.000591)	(0.000591)	(0.00216)	(0.00216)	(0.00216)
IQ 31-40	-0.00195***	-0.00193***	-0.00194***	-0.00263	-0.00262	-0.00261
	(0.000629)	(0.000629)	(0.000629)	(0.00191)	(0.00191)	(0.00191)
IQ 51-60	0.00148**	0.00147**	0.00148**	-0.00338	-0.00338	-0.00336
	(0.000714)	(0.000714)	(0.000714)	(0.00226)	(0.00226)	(0.00225)
IQ 61-70	0.00415***	0.00412***	0.00415***	-0.00113	-0.00113	-0.00112
	(0.000785)	(0.000785)	(0.000785)	(0.00245)	(0.00245)	(0.00245)
IQ 71-80	0.00965***	0.00960***	0.00964***	0.00700**	0.00697**	0.00700**
	(0.000870)	(0.000870)	(0.000870)	(0.00273)	(0.00273)	(0.00273)
IQ 81-90	0.0147***	0.0147***	0.0147***	0.00799***	0.00785***	0.00801***
	(0.000968)	(0.000968)	(0.000967)	(0.00287)	(0.00287)	(0.00287)
IQ 91-95	0.0237***	0.0236***	0.0237***	0.0204***	0.0203***	0.0204***
	(0.00144)	(0.00144)	(0.00144)	(0.00408)	(0.00409)	(0.00408)
IQ 96-100	0.0331***	0.0350***	0.0331***	0.0268***	0.0319***	0.0269***
	(0.00174)	(0.00165)	(0.00174)	(0.00467)	(0.00455)	(0.00466)
fa income 96-100 x IQ 96-100	0.0144***		0.0147***	0.0256*		0.0270*
	(0.00527)		(0.00534)	(0.0148)		(0.0150)
mo income 96-100 x IQ 96-100	-0.00358		-0.00275	0.0339		0.0336
	(0.00816)		(0.00811)	(0.0271)		(0.0269)
fa PhD x IQ 96-100		0.00596	-0.000873		-0.00299	-0.0174
		(0.0143)	(0.0145)		(0.0392)	(0.0399)
mo PhD x IQ 96-100		-0.0178	-0.0205		0.0481	0.0401
		(0.0282)	(0.0279)		(0.0716)	(0.0673)
Sample	IQ	IQ	IQ	Brothers	Brothers	Brothers
Estimator	OLS	OLS	OLS	FE	FE	FE
Observations	352,668	352,668	352,668	82,054	82,054	82,054
Number of families				41,605	41,605	41,605

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All specifications include a 4th order polynomial in log(age), 21 region dummies, dummies for suburban and urban regions, dummies for mother tongue, and dummies for parental decade of birth.

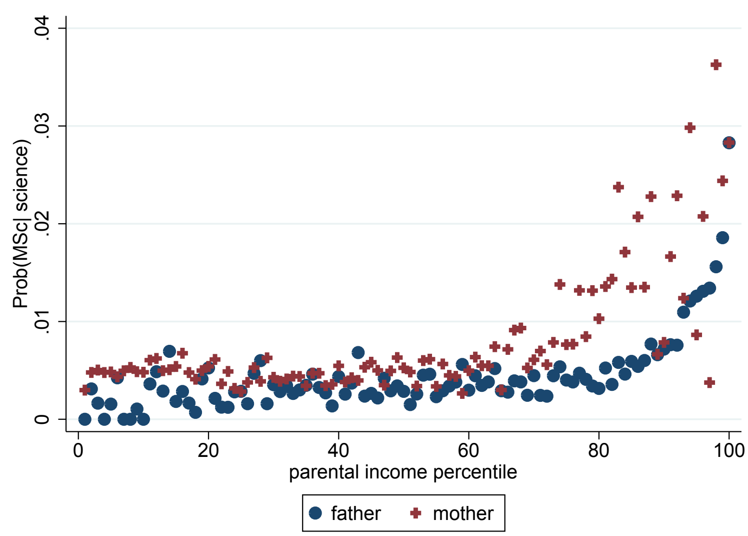
G Supplementary Materials for Section 7

Figure G1: Distribution of Educational Attainment



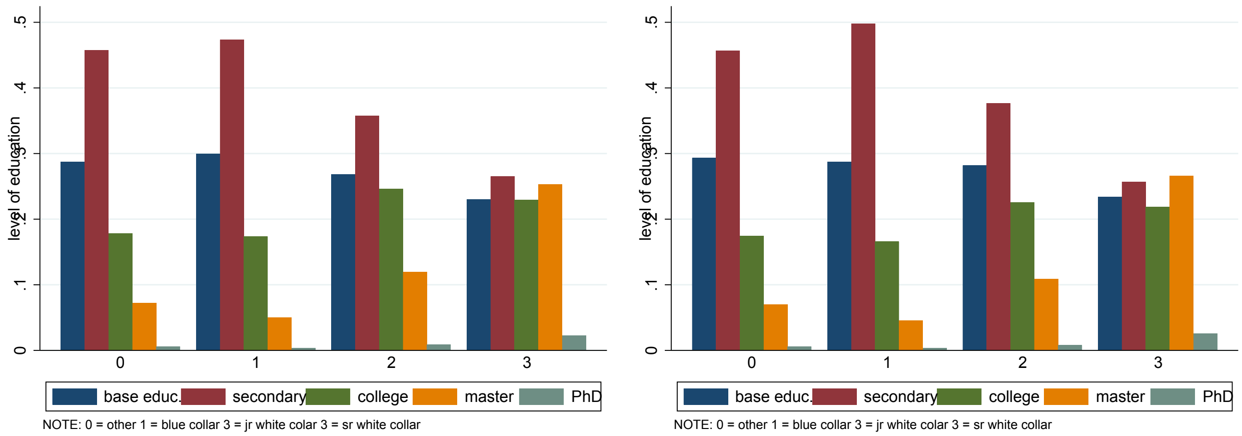
Notes: The figure displays the educational attainment of the individuals in our sample. We divide education into five groups by level of education: base education (up to 9 years, depending on age of parent), secondary, tertiary, MSc, and PhD. We measure education at age 35.

Figure G2: Probability of Having MSc vs Parental Income



Notes: The figure displays the probability of obtaining an MSc as a function of parental income. Parental incomes are measured in 1975 for individuals born 1961-1975, and in 1985 for individuals born in 1976-1985. We calculate the percentile ranks using residuals from a regression of the natural log of deflated income of fathers and mothers on year-of-birth dummies.

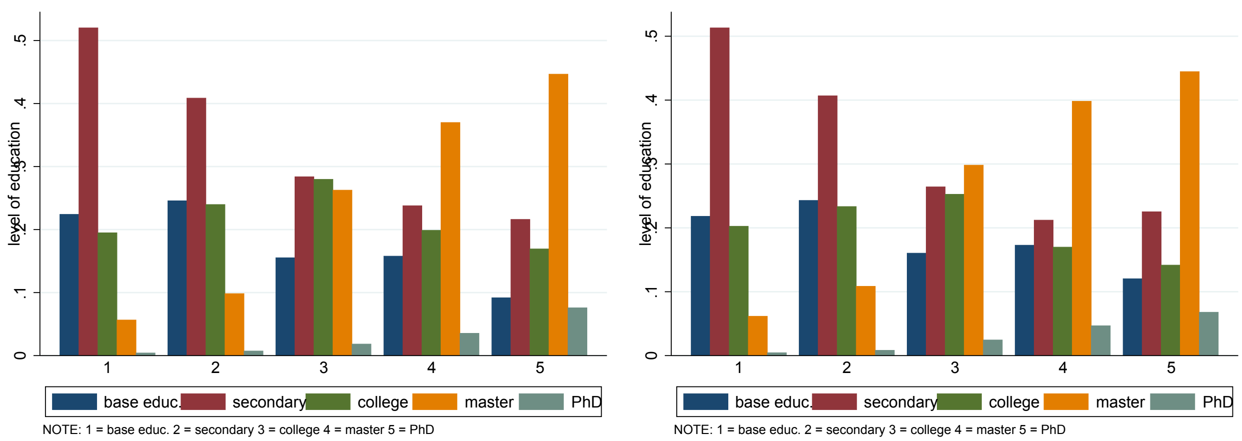
Figure G3: Child's Education vs Parental Socioeconomic Background



A. Father B. Mother

Notes: The figure displays the education of an individual as a function of parental socioeconomic class (A: father, B: mother). We code socioeconomic class as follows: 0 = other (base); 1 = blue-collar; 2 = junior white collar; 3 = senior white collar. Parental socioeconomic status is measured in 1975 for parents born before 1951, and in 1985 parents born in 1951 or thereafter.

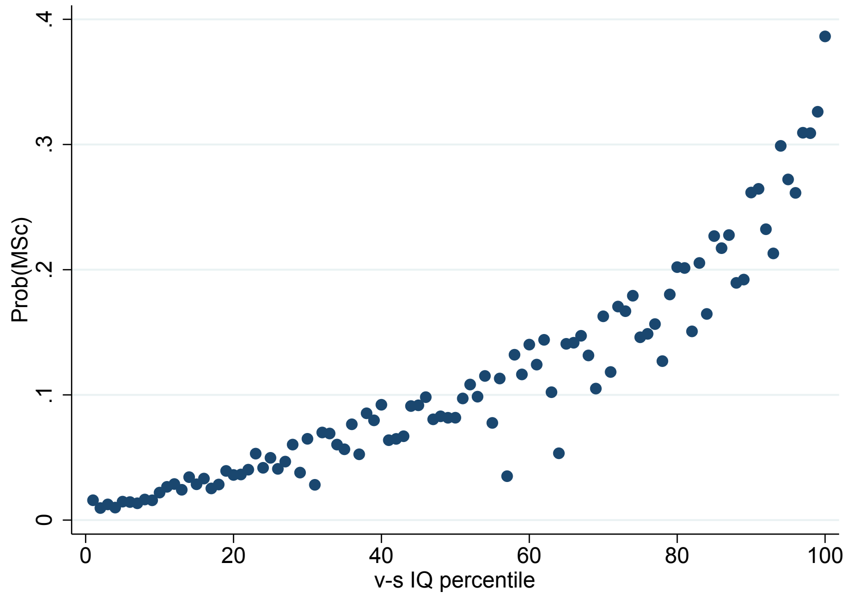
Figure G4: Child's Education vs Parents' Education



A. Father B. Mother

Notes: The figure displays the education of an individual as a function of parental education (A: father, B: mother). We divide education into five groups by level of education: base education (up to 9 years, depending on age of parent), secondary, tertiary, MSc, and PhD. We measure education of an individual at age 35. Parental education is measured in 1975 unless unavailable, in which case 1985 data used.

Figure G5: Education vs IQ



Notes: The figure displays the probability of obtaining an MSc by age 35 as a function of the visuospatial IQ percentile. IQ percentiles are calculated on the basis of the normalized IQ score, where normalization was done separately for each conscription cohort to avoid the Flynn effect.

Table G1: Education Regression

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
first born					-0.0103***	-0.0116***
					(0.00209)	(0.00318)
fa income 21-40	-0.00146	0.00128	0.00257	0.000121	-0.00228	-0.00505
	(0.00201)	(0.00200)	(0.00198)	(0.00194)	(0.00403)	(0.0217)
fa income 41-60	0.00886***	0.0121***	0.0102***	0.00543***	0.00485	0.0107
	(0.00193)	(0.00198)	(0.00196)	(0.00192)	(0.00404)	(0.0234)
fa income 61-80	0.0247***	0.0219***	0.0182***	0.0110***	0.0108***	0.0136
	(0.00190)	(0.00197)	(0.00194)	(0.00191)	(0.00404)	(0.0258)
fa income 81-90	0.0726***	0.0419***	0.0319***	0.0223***	0.0216***	0.00643
	(0.00220)	(0.00224)	(0.00221)	(0.00217)	(0.00478)	(0.0343)
fa income 91-95	0.137***	0.0707***	0.0488***	0.0382***	0.0393***	-0.0182
	(0.00288)	(0.00290)	(0.00286)	(0.00281)	(0.00638)	(0.0489)
fa income 96-100	0.244***	0.137***	0.0791***	0.0671***	0.0651***	0.132*
	(0.00332)	(0.00348)	(0.00346)	(0.00340)	(0.00764)	(0.0702)
mo income 21-40	0.00701***	0.00773***	0.000114	-0.000692	0.00463*	0.0192
	(0.00125)	(0.00127)	(0.00126)	(0.00123)	(0.00276)	(0.0206)
mo income 41-60	0.0267***	0.0213***	0.00730***	0.00521***	0.00442	-0.00848
	(0.00144)	(0.00150)	(0.00148)	(0.00146)	(0.00339)	(0.0227)
mo income 61-80	0.0714***	0.0445***	0.0137***	0.0101***	0.0176***	-0.00525
	(0.00214)	(0.00218)	(0.00217)	(0.00214)	(0.00537)	(0.0320)
mo income 81-90	0.150***	0.0838***	0.0242***	0.0198***	0.0311***	0.0340
	(0.00455)	(0.00457)	(0.00455)	(0.00449)	(0.0111)	(0.0771)
mo income 91-95	0.197***	0.125***	0.0365***	0.0305***	0.0471**	0.0576
	(0.00811)	(0.00806)	(0.00804)	(0.00794)	(0.0192)	(0.0836)
mo income 96-100	0.0363***	0.0310***	0.00425	0.00138	-0.00179	-0.0120
	(0.00303)	(0.00298)	(0.00283)	(0.00278)	(0.00614)	(0.0160)
fa bluecollar		-0.00938***	-0.00309**	-0.00143	-0.00749**	
		(0.00147)	(0.00145)	(0.00143)	(0.00323)	
fa jr whitec.		0.0358***	0.0277***	0.0222***	0.0204***	
		(0.00200)	(0.00199)	(0.00196)	(0.00463)	
fa sr whitec.		0.108***	0.0436***	0.0362***	0.0403***	
		(0.00285)	(0.00307)	(0.00302)	(0.00721)	
mo bluecollar		-0.0165***	-0.00724***	-0.00544***	-0.00850***	
		(0.00148)	(0.00146)	(0.00144)	(0.00325)	
mo jr whitec.		0.0147***	0.0180***	0.0139***	0.00630	
		(0.00170)	(0.00167)	(0.00165)	(0.00384)	
mo sr whitec.		0.0932***	0.0339***	0.0280***	0.0147*	
		(0.00326)	(0.00336)	(0.00331)	(0.00762)	
fa secondary			0.0304***	0.0235***	0.0237***	
			(0.00212)	(0.00209)	(0.00492)	
fa college			0.101***	0.0859***	0.0958***	
			(0.00380)	(0.00374)	(0.00866)	

– Continued on the next page –

– Table G1: (Continued) –

fa MSc			0.174***	0.156***	0.175***	
			(0.00475)	(0.00468)	(0.0108)	
fa PhD			0.253***	0.233***	0.249***	
			(0.00945)	(0.00931)	(0.0204)	
mo secondary			0.0376***	0.0302***	0.0359***	
			(0.00122)	(0.00119)	(0.00275)	
mo college			0.116***	0.102***	0.107***	
			(0.00349)	(0.00344)	(0.00783)	
mo MSc			0.176***	0.158***	0.168***	
			(0.00603)	(0.00596)	(0.0137)	
mo PhD			0.186***	0.164***	0.161***	
			(0.0219)	(0.0218)	(0.0484)	
fa STEM			-0.00184	-0.00316	-0.00666	
			(0.00212)	(0.00209)	(0.00493)	
mo STEM			-0.0172***	-0.0173***	-0.0195***	
			(0.00182)	(0.00179)	(0.00407)	
IQ 1-10				-0.0410***	-0.0421***	-0.0338***
				(0.00156)	(0.00333)	(0.00459)
IQ 11-20				-0.0353***	-0.0357***	-0.0249***
				(0.00167)	(0.00356)	(0.00482)
IQ 21-30				-0.0248***	-0.0311***	-0.0289***
				(0.00178)	(0.00374)	(0.00497)
IQ 31-40				-0.00972***	-0.0145***	-0.00868*
				(0.00190)	(0.00404)	(0.00519)
IQ 51-60				0.0161***	0.0116***	0.00532
				(0.00207)	(0.00435)	(0.00549)
IQ 61-70				0.0317***	0.0317***	0.0285***
				(0.00221)	(0.00478)	(0.00594)
IQ 71-80				0.0574***	0.0600***	0.0423***
				(0.00230)	(0.00496)	(0.00618)
IQ 81-90				0.0899***	0.0910***	0.0693***
				(0.00249)	(0.00533)	(0.00663)
IQ 91-95				0.122***	0.127***	0.0935***
				(0.00332)	(0.00708)	(0.00880)
IQ 96-100				0.169***	0.174***	0.125***
				(0.00357)	(0.00753)	(0.00946)
Sample	All	All	All	All	3 year window	3 year window
Estimators	OLS	OLS	OLS	OLS	OLS	FE
Observations	352,668	352,668	352,668	352,668	82,054	82,054
Number of families						41,605

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All specifications include a 4th order polynomial in log(age), 21 region dummies, dummies for suburban and urban regions, dummies for mother tongue, and dummies for parental decade of birth.

Table G2: Decomposition of the explained impact on own education

A. Partial R-squared	
Explanatory variables	MSc
Base controls	0.023
Parental income	0.003
Parental socecon	0.003
Parental education	0.023
IQ	0.035
Sum of partial R-sq's	0.086
R-sq	0.168
B. Fraction of partial R-squared	
	MSc
Base controls	0.264
Parental income	0.029
Parental socecon	0.034
Parental education	0.267
IQ	0.406

Notes: The upper panel displays the partial R-squared for a given dependent variable (column) and given vector or explanatory variables (row), their sum, and the R-squared of the estimation. The lower panel displays the share of partial R-squared obtained for a given dependent variable (column) by a given vector of explanatory variables. For example, the 0.0064 for Base controls in the lower panel for Inventor is obtained by dividing 0.0004 (upper panel, Base controls) by 0.0622 (Sum of partial R-sq's). Base controls are: a 4th order polynomial in log(age), 21 region dummies, dummies for suburban and urban regions, dummies for mother tongue, and dummies for parental decade of birth.

Table G3: Who becomes regression with own education - family FE estimations

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
first born	-0.00325	-0.00311	-0.00170	-0.00262	-0.00105	-0.000350
	(0.00328)	(0.00345)	(0.00145)	(0.00204)	(0.000847)	(0.00135)
fa income 21-40	0.00106	-0.0291	-0.00218	-0.0126	-0.00210	0.00985
	(0.00415)	(0.0234)	(0.00272)	(0.0469)	(0.00147)	(0.0153)
fa income 41-60	0.00487	-0.0256	-0.000391	-0.0152	-9.00e-06	0.00734
	(0.00445)	(0.0240)	(0.00262)	(0.0484)	(0.00146)	(0.0142)
fa income 61-80	0.00241	-0.0381	-0.000153	-0.0426	3.56e-05	-0.00199
	(0.00440)	(0.0246)	(0.00268)	(0.0506)	(0.00147)	(0.0149)
fa income 81-90	-0.00518	0.0154	-0.00241	-0.0239	-1.91e-05	0.00186
	(0.00610)	(0.0453)	(0.00313)	(0.0489)	(0.00173)	(0.0175)
fa income 91-95	-0.00420	-0.0232	-0.00208	0.0103	-0.000296	-0.0121
	(0.00827)	(0.0323)	(0.00417)	(0.0545)	(0.00220)	(0.0201)
fa income 96-100	0.0151	0.0744	0.00915	-0.0796	0.00742***	-0.0279
	(0.0125)	(0.0947)	(0.00570)	(0.0696)	(0.00280)	(0.0275)
mo income 21-40	0.00362	-0.00807	0.00133	0.00663	-0.00112	-0.000968
	(0.00381)	(0.0126)	(0.00184)	(0.0232)	(0.00101)	(0.00910)
mo income 41-60	0.00147	-0.00207	-0.000623	0.00177	-0.00120	-0.00550
	(0.00408)	(0.0118)	(0.00220)	(0.0171)	(0.00123)	(0.00675)
mo income 61-80	0.00195	0.00974	-0.000717	0.00644	-0.00237	-0.00167
	(0.00631)	(0.0130)	(0.00326)	(0.0269)	(0.00179)	(0.00748)
mo income 81-90	0.0145		0.00449	-0.0160	-0.00111	-0.0192
	(0.0213)		(0.00920)	(0.122)	(0.00432)	(0.0294)
mo income 91-95	-0.0365*		-0.0151	-0.205	0.00156	-0.00346
	(0.0192)		(0.0112)	(0.149)	(0.00753)	(0.0481)
mo income 96-100	-0.00171	-0.000882	0.00286	0.00924	0.00405	0.00664
	(0.00995)	(0.00845)	(0.00479)	(0.0205)	(0.00258)	(0.00727)
fa bluecollar	-0.00435		0.00112		0.000642	
	(0.00306)		(0.00197)		(0.00111)	
fa jr whitec.	0.00197		0.00143		0.00246	
	(0.00540)		(0.00295)		(0.00162)	
fa sr whitec.	0.0227**		0.00557		0.00447*	
	(0.0107)		(0.00481)		(0.00247)	
mo bluecollar	-0.000497		-0.00298		-0.000713	
	(0.00289)		(0.00208)		(0.00115)	
mo jr whitec.	0.00386		-0.00156		0.00103	
	(0.00355)		(0.00262)		(0.00138)	
mo sr whitec.	-0.0134		-0.00274		-0.00223	
	(0.0118)		(0.00509)		(0.00270)	
fa secondary	-0.000425		-0.00329		-0.00569***	
	(0.00752)		(0.00326)		(0.00171)	
fa college	0.00879		0.00429		-0.000938	
	(0.0146)		(0.00662)		(0.00303)	
fa MSc	-0.0311***		-0.0107		-0.00814**	
	(0.0111)		(0.00742)		(0.00361)	
fa PhD	-0.0286		-0.00205		0.00908	
	(0.0261)		(0.0151)		(0.00871)	
mo secondary	-0.00438		0.000406		0.00184*	
	(0.00413)		(0.00177)		(0.000990)	
mo college	0.000912		0.00185		-0.000723	
	(0.0135)		(0.00575)		(0.00271)	

– *Continued on the next page* –

– Table G3 Continued –

mo MSc	0.00975		-0.000530		0.00236	
	(0.0218)		(0.00963)		(0.00509)	
mo PhD	0.0819		-0.0118		-0.0213	
	(0.0712)		(0.0312)		(0.0136)	
fa STEM	0.00192		0.00534		0.00664***	
	(0.00776)		(0.00333)		(0.00172)	
mo STEM	0.00278		0.00159		-0.000195	
	(0.00632)		(0.00283)		(0.00161)	
IQ 1-10	0.00277	-0.00516	0.00238	-0.00135	-0.00265**	-0.00378**
	(0.00410)	(0.00842)	(0.00203)	(0.00307)	(0.00115)	(0.00181)
IQ 11-20	0.00487	0.00241	0.000702	-0.00186	-0.00234*	-0.00355*
	(0.00530)	(0.00808)	(0.00192)	(0.00295)	(0.00123)	(0.00185)
IQ 21-30	0.0150**	0.0186*	0.00342	0.000305	-0.000267	-0.000686
	(0.00743)	(0.0101)	(0.00234)	(0.00334)	(0.00139)	(0.00211)
IQ 31-40	0.00572	0.00352	0.00110	-0.00105	-0.00229*	-0.00126
	(0.00556)	(0.00760)	(0.00223)	(0.00305)	(0.00137)	(0.00188)
IQ 51-60	0.00407	-0.00689	0.00143	-0.00428	-0.00138	-0.00364*
	(0.00548)	(0.00852)	(0.00250)	(0.00371)	(0.00154)	(0.00220)
IQ 61-70	0.00112	-0.00368	0.00516	-0.000310	-0.000400	-0.00285
	(0.00634)	(0.00888)	(0.00316)	(0.00421)	(0.00170)	(0.00241)
IQ 71-80	-0.00286	-0.00477	0.00846**	0.0107**	0.00470**	0.00374
	(0.00677)	(0.00919)	(0.00344)	(0.00460)	(0.00195)	(0.00267)
IQ 81-90	0.00206	-0.0194	0.00231	-0.00466	0.00246	0.00148
	(0.00827)	(0.0126)	(0.00341)	(0.00468)	(0.00202)	(0.00281)
IQ 91-95	0.0147	-0.00450	0.0139***	0.0109	0.00853***	0.0116***
	(0.0144)	(0.0196)	(0.00538)	(0.00726)	(0.00295)	(0.00397)
IQ 96-100	-0.000183	-0.0127	0.00977*	0.0141**	0.0156***	0.0178***
	(0.0112)	(0.0141)	(0.00578)	(0.00718)	(0.00341)	(0.00444)
secondary	-0.00302	-0.00938	0.000224	-0.00104	0.000657	-0.000184
	(0.00397)	(0.00825)	(0.00156)	(0.00312)	(0.000859)	(0.00163)
college	-0.000633	-0.0211*	-0.00138	0.00269	-0.00265***	-0.00148
	(0.00543)	(0.0119)	(0.00195)	(0.00376)	(0.000985)	(0.00202)
MSc	-0.00968*	-0.0241	-0.00115	0.00544	0.000663	0.00482
	(0.00498)	(0.0198)	(0.00335)	(0.00690)	(0.00183)	(0.00352)
PhD	-0.0240	0.00334	0.0152	0.0518**	-0.00371	-0.0188
	(0.0174)	(0.0172)	(0.0220)	(0.0247)	(0.00585)	(0.0173)
secondary STEM	-0.00327	-0.00797	-0.00155	-0.00229	-0.00242***	-0.00233*
	(0.00327)	(0.00625)	(0.00113)	(0.00224)	(0.000587)	(0.00120)
college STEM	0.0135*	0.00203	0.0141***	0.00839*	0.0164***	0.0107***
	(0.00700)	(0.0106)	(0.00286)	(0.00430)	(0.00156)	(0.00244)
MSc STEM	0.0862***	0.0711**	0.0973***	0.0905***	0.0966***	0.0871***
	(0.0184)	(0.0306)	(0.00949)	(0.0120)	(0.00454)	(0.00588)
PhD STEM	0.226**	0.249**	0.250***	0.254***	0.229***	0.221***
	(0.0940)	(0.105)	(0.0389)	(0.0418)	(0.0183)	(0.0206)
Observations	4,117	4,117	23,141	23,141	82,054	82,054
Number of families		2,076		11,588		41,605

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All specifications include a 4th order polynomial in log(age), 21 region dummies, dummies for suburban and urban regions, dummies for mother tongue, and dummies for parental decade of birth.

Table G4: Who becomes inventor regressions

	Inventor	Patent count	Citations	Star inventor	Alternative income measure
fa income 21-40	0.00064	-0.00295	-0.00193	0.000248	-0.000121
	(0.000720)	(0.00226)	(0.00158)	(0.000221)	(0.000709)
fa income 41-60	0.000726	-0.00131	0.000577	0.000407*	0.000435
	(0.000707)	(0.00229)	(0.00168)	(0.000224)	(0.000703)
fa income 61-80	0.00102	-0.00180	0.000225	0.000210	0.000958
	(0.000702)	(0.00225)	(0.00164)	(0.000215)	(0.000697)
fa income 81-90	0.000991	-0.00215	0.000677	0.000289	0.000663
	(0.000797)	(0.00245)	(0.00188)	(0.000248)	(0.000785)
fa income 91-95	0.00224**	0.00104	0.00509*	0.000472	0.00106
	(0.00105)	(0.00323)	(0.00295)	(0.000339)	(0.00102)
fa income 96-100	0.00404***	0.00276	0.00230	0.000418	0.00353***
	(0.00126)	(0.00440)	(0.00297)	(0.000418)	(0.00117)
mo income 21-40	0.000199	0.00114	-0.000202	0.000108	-1.29e-05
	(0.000476)	(0.00128)	(0.00113)	(0.000162)	(0.000470)
mo income 41-60	0.000231	0.00225	0.000404	7.92e-05	-6.67e-05
	(0.000549)	(0.00158)	(0.00142)	(0.000185)	(0.000550)
mo income 61-80	0.000306	-0.00134	-0.00359*	-0.000355	0.000766
	(0.000792)	(0.00221)	(0.00186)	(0.000244)	(0.000812)
mo income 81-90	0.000186	0.00917	0.00486	0.000832	-0.00152
	(0.00171)	(0.00640)	(0.00437)	(0.000646)	(0.00183)
mo income 91-95	-0.00189	-0.00620	0.00461	0.000891	-0.00205
	(0.00305)	(0.00994)	(0.00942)	(0.00119)	(0.00323)
mo income 96-100	-0.000279	0.00239	0.00316	-7.18e-05	-0.00238
	(0.00108)	(0.00365)	(0.00267)	(0.000372)	(0.00415)
fa bluecollar	-0.000736	-0.00112	-0.00222*	-0.000197	-0.000824
	(0.000526)	(0.00134)	(0.00130)	(0.000188)	(0.000530)
fa jr whitec.	-1.99e-05	0.00144	-0.00157	-0.000102	-6.51e-05
	(0.000715)	(0.00199)	(0.00186)	(0.000251)	(0.000718)
fa sr whitec.	0.000491	0.000833	-0.00234	-3.17e-05	0.000457
	(0.00109)	(0.00333)	(0.00295)	(0.000394)	(0.00109)
mo bluecollar	0.000166	0.00181	0.00139	0.000115	0.000215
	(0.000537)	(0.00134)	(0.00113)	(0.000186)	(0.000539)
mo jr whitec.	0.000315	0.00211	0.00273*	-3.69e-06	0.000346
	(0.000604)	(0.00168)	(0.00146)	(0.000211)	(0.000611)
mo sr whitec.	0.000723	-0.00436	-0.000282	0.000182	0.000908
	(0.00121)	(0.00470)	(0.00322)	(0.000423)	(0.00122)
fa secondary	-0.00365***	-0.00647**	-0.00322	-0.000361	-0.00363***
	(0.000759)	(0.00258)	(0.00203)	(0.000271)	(0.000760)
fa college	-0.00142	-0.00558	-0.000445	-0.000432	-0.00138
	(0.00134)	(0.00378)	(0.00376)	(0.000463)	(0.00134)

– Continued on the next page –

– Table G4 Continued –

fa income 21-40	0.00064	-0.00295	-0.00193	0.000248	-0.000121
	(0.000720)	(0.00226)	(0.00158)	(0.000221)	(0.000709)
fa income 41-60	0.000726	-0.00131	0.000577	0.000407*	0.000435
	(0.000707)	(0.00229)	(0.00168)	(0.000224)	(0.000703)
fa income 61-80	0.00102	-0.00180	0.000225	0.000210	0.000958
	(0.000702)	(0.00225)	(0.00164)	(0.000215)	(0.000697)
fa income 81-90	0.000991	-0.00215	0.000677	0.000289	0.000663
	(0.000797)	(0.00245)	(0.00188)	(0.000248)	(0.000785)
fa income 91-95	0.00224**	0.00104	0.00509*	0.000472	0.00106
	(0.00105)	(0.00323)	(0.00295)	(0.000339)	(0.00102)
fa income 96-100	0.00404***	0.00276	0.00230	0.000418	0.00353***
	(0.00126)	(0.00440)	(0.00297)	(0.000418)	(0.00117)
mo income 21-40	0.000199	0.00114	-0.000202	0.000108	-1.29e-05
	(0.000476)	(0.00128)	(0.00113)	(0.000162)	(0.000470)
mo income 41-60	0.000231	0.00225	0.000404	7.92e-05	-6.67e-05
	(0.000549)	(0.00158)	(0.00142)	(0.000185)	(0.000550)
mo income 61-80	0.000306	-0.00134	-0.00359*	-0.000355	0.000766
	(0.000792)	(0.00221)	(0.00186)	(0.000244)	(0.000812)
mo income 81-90	0.000186	0.00917	0.00486	0.000832	-0.00152
	(0.00171)	(0.00640)	(0.00437)	(0.000646)	(0.00183)
mo income 91-95	-0.00189	-0.00620	0.00461	0.000891	-0.00205
	(0.00305)	(0.00994)	(0.00942)	(0.00119)	(0.00323)
mo income 96-100	-0.000279	0.00239	0.00316	-7.18e-05	-0.00238
	(0.00108)	(0.00365)	(0.00267)	(0.000372)	(0.00415)
fa bluecollar	-0.000736	-0.00112	-0.00222*	-0.000197	-0.000824
	(0.000526)	(0.00134)	(0.00130)	(0.000188)	(0.000530)
fa jr whitec.	-1.99e-05	0.00144	-0.00157	-0.000102	-6.51e-05
	(0.000715)	(0.00199)	(0.00186)	(0.000251)	(0.000718)
fa sr whitec.	0.000491	0.000833	-0.00234	-3.17e-05	0.000457
	(0.00109)	(0.00333)	(0.00295)	(0.000394)	(0.00109)
mo bluecollar	0.000166	0.00181	0.00139	0.000115	0.000215
	(0.000537)	(0.00134)	(0.00113)	(0.000186)	(0.000539)
mo jr whitec.	0.000315	0.00211	0.00273*	-3.69e-06	0.000346
	(0.000604)	(0.00168)	(0.00146)	(0.000211)	(0.000611)
mo sr whitec.	0.000723	-0.00436	-0.000282	0.000182	0.000908
	(0.00121)	(0.00470)	(0.00322)	(0.000423)	(0.00122)
fa secondary	-0.00365***	-0.00647**	-0.00322	-0.000361	-0.00363***
	(0.000759)	(0.00258)	(0.00203)	(0.000271)	(0.000760)
fa college	-0.00142	-0.00558	-0.000445	-0.000432	-0.00138
	(0.00134)	(0.00378)	(0.00376)	(0.000463)	(0.00134)

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All specifications include a 4th order polynomial in log(age), 21 region dummies, dummies for suburban and urban regions, dummies for mother tongue, and dummies for parental decade of birth.

Table G5: Who becomes regression with own education - family FE estimations

	Family structure	Family structure	Family structure	Family structure	Interactions	Interactions	Interactions	Interactions
biol fa away	-0.00125	-0.000568	-0.000590	-0.000512				
	(0.000783)	(0.000805)	(0.000805)	(0.000792)				
biol mo away	-0.00214	-0.00265*	-0.00262	-0.00265*				
	(0.00159)	(0.00160)	(0.00160)	(0.00158)				
biol fa & mo away	0.00172	0.00273*	0.00271*	0.00275*				
	(0.00135)	(0.00156)	(0.00155)	(0.00153)				
fa income 21-40	0.000614	0.000596	0.000552	0.000555	0.000643	0.000641	0.000645	0.000655
	(0.000719)	(0.000719)	(0.000720)	(0.000720)	(0.000720)	(0.000720)	(0.000720)	(0.000720)
fa income 41-60	0.000696	0.000722	0.000714	0.000719	0.000737	0.000727	0.000740	0.000755
	(0.000708)	(0.000708)	(0.000708)	(0.000708)	(0.000707)	(0.000707)	(0.000707)	(0.000707)
fa income 61-80	0.000984	0.00102	0.00102	0.00102	0.00104	0.00102	0.00105	0.00105
	(0.000704)	(0.000704)	(0.000704)	(0.000704)	(0.000702)	(0.000702)	(0.000702)	(0.000702)
fa income 81-90	0.000953	0.00100	0.001000	0.00100	0.00103	0.000991	0.00103	0.00105
	(0.000797)	(0.000797)	(0.000797)	(0.000798)	(0.000796)	(0.000797)	(0.000796)	(0.000796)
fa income 91-95	0.00220**	0.00240**	0.00244**	0.00244**	0.00231**	0.00223**	0.00231**	0.00232**
	(0.00105)	(0.00107)	(0.00114)	(0.00114)	(0.00105)	(0.00105)	(0.00105)	(0.00105)
fa income 96-100	0.00401***	0.00434***	0.00448***	0.00446***	0.00325***	0.00405***	0.00316**	0.00335***
	(0.00126)	(0.00128)	(0.00137)	(0.00138)	(0.00125)	(0.00126)	(0.00125)	(0.00125)
mo income 21-40	0.000213	0.000210	0.000197	0.000189	0.000198	0.000198	0.000199	0.000199
	(0.000478)	(0.000478)	(0.000478)	(0.000478)	(0.000476)	(0.000476)	(0.000476)	(0.000476)
mo income 41-60	0.000262	0.000261	0.000238	0.000226	0.000229	0.000229	0.000227	0.000220
	(0.000552)	(0.000553)	(0.000553)	(0.000552)	(0.000549)	(0.000549)	(0.000549)	(0.000549)
mo income 61-80	0.000361	0.000375	0.000328	0.000302	0.000305	0.000298	0.000299	0.000312
	(0.000802)	(0.000803)	(0.000803)	(0.000803)	(0.000792)	(0.000792)	(0.000792)	(0.000792)
mo income 81-90	0.000259	0.000309	0.000169	0.000113	0.000178	0.000171	0.000170	0.000196
	(0.00172)	(0.00172)	(0.00172)	(0.00172)	(0.00171)	(0.00171)	(0.00171)	(0.00171)
mo income 91-95	-0.00182	-0.00118	-0.00163	-0.00172	-0.00196	-0.00192	-0.00199	-0.00204
	(0.00306)	(0.00317)	(0.00345)	(0.00345)	(0.00305)	(0.00305)	(0.00304)	(0.00305)
mo income 96-100	-0.000339	-0.000272	0.000322	0.000317	-1.63e-05	-0.000255	-6.21e-05	-5.65e-05
	(0.00113)	(0.00128)	(0.00129)	(0.00129)	(0.00103)	(0.00108)	(0.00103)	(0.00103)
biol fa inc 91-95 x away		-0.00415	-0.00415	-0.00401				
		(0.00336)	(0.00336)	(0.00345)				
biol fa inc 96-100 x away		-0.00981**	-0.00981**	-0.00845*				
		(0.00482)	(0.00482)	(0.00462)				
biol mo inc 91-95 x away		-0.0141***	-0.0141***	-0.0137***				
		(0.00506)	(0.00506)	(0.00508)				
biol mo inc 96-100 x away		-0.00123	-0.00123	-0.00119				
		(0.00240)	(0.00240)	(0.00240)				
step fa income 91-95			-0.000384	-0.000385				
			(0.00216)	(0.00216)				
step income 96-100			-0.000920	-0.000937				
			(0.00274)	(0.00274)				
step mo income 91-95			0.00137	0.00157				
			(0.00834)	(0.00834)				
step mo income 96-100			-0.0222**	-0.0223**				
			(0.00874)	(0.00874)				
fa bluecollar	-0.000763	-0.000740	-0.000739	-0.000741	-0.000756	-0.000735	-0.000757	-0.000758
	(0.000528)	(0.000528)	(0.000528)	(0.000528)	(0.000526)	(0.000526)	(0.000526)	(0.000526)
fa jr whitec.	-4.70e-05	-3.22e-05	-4.01e-05	-4.66e-05	-3.18e-05	-2.01e-05	-3.24e-05	-1.75e-05
	(0.000715)	(0.000716)	(0.000716)	(0.000716)	(0.000715)	(0.000715)	(0.000715)	(0.000715)
fa sr whitec.	0.000469	0.000442	0.000417	0.000399	0.000457	0.000487	0.000450	0.000465
	(0.00109)	(0.00109)	(0.00109)	(0.00109)	(0.00109)	(0.00109)	(0.00109)	(0.00109)
mo bluecollar	0.000195	0.000173	0.000213	0.000219	0.000160	0.000167	0.000159	0.000161
	(0.000538)	(0.000538)	(0.000538)	(0.000538)	(0.000537)	(0.000537)	(0.000537)	(0.000537)
mo jr whitec.	0.000332	0.000306	0.000356	0.000364	0.000314	0.000316	0.000314	0.000318
	(0.000606)	(0.000606)	(0.000606)	(0.000606)	(0.000603)	(0.000604)	(0.000603)	(0.000603)
mo sr whitec.	0.000737	0.000639	0.000728	0.000740	0.000709	0.000723	0.000705	0.000692
	(0.00121)	(0.00121)	(0.00121)	(0.00121)	(0.00121)	(0.00121)	(0.00121)	(0.00121)
fa secondary	-0.00365***	-0.00364***	-0.00362***	-0.00361***	-0.00362***	-0.00366***	-0.00363***	-0.00362***
	(0.000759)	(0.000759)	(0.000759)	(0.000759)	(0.000759)	(0.000759)	(0.000759)	(0.000758)
fa college	-0.00143	-0.00140	-0.00139	-0.00137	-0.00136	-0.00143	-0.00137	-0.00136
	(0.00134)	(0.00134)	(0.00134)	(0.00134)	(0.00134)	(0.00134)	(0.00134)	(0.00134)
fa MSc	0.000414	0.000417	0.000454	0.000294	0.000353	0.000415	0.000311	0.000359
	(0.00170)	(0.00170)	(0.00170)	(0.00172)	(0.00170)	(0.00169)	(0.00169)	(0.00169)

– Continued on the next page –

– Table G5 Continued –

fa PhD	0.00973**	0.00975**	0.00972**	0.0101**	0.00961**	0.0107**	0.0112**	0.0116***
	(0.00420)	(0.00420)	(0.00420)	(0.00428)	(0.00421)	(0.00436)	(0.00436)	(0.00436)
mo secondary	0.000617	0.000619	0.000628	0.000637	0.000623	0.000616	0.000620	0.000633
	(0.000451)	(0.000451)	(0.000451)	(0.000451)	(0.000451)	(0.000451)	(0.000451)	(0.000451)
mo college	-0.00136	-0.00137	-0.00134	-0.00131	-0.00135	-0.00134	-0.00135	-0.00135
	(0.00125)	(0.00125)	(0.00125)	(0.00125)	(0.00125)	(0.00125)	(0.00125)	(0.00125)
mo MSc	0.00128	0.00124	0.00167	0.00201	0.00128	0.00131	0.00129	0.00131
	(0.00234)	(0.00235)	(0.00236)	(0.00243)	(0.00234)	(0.00234)	(0.00234)	(0.00234)
mo PhD	-0.00547	-0.00536	-0.00431	-0.00253	-0.00544	-0.00257	-0.00238	-0.00203
	(0.00911)	(0.00911)	(0.00915)	(0.00965)	(0.00911)	(0.00966)	(0.00966)	(0.00966)
fa STEM	0.00460***	0.00459***	0.00457***	0.00456***	0.00457***	0.00461***	0.00457***	0.00458***
	(0.000771)	(0.000771)	(0.000770)	(0.000770)	(0.000771)	(0.000771)	(0.000770)	(0.000770)
mo STEM	-0.000626	-0.000632	-0.000649	-0.000663	-0.000638	-0.000631	-0.000633	-0.000637
	(0.000711)	(0.000711)	(0.000710)	(0.000711)	(0.000711)	(0.000711)	(0.000711)	(0.000711)
biol fa MSc x away				0.00527				
				(0.00761)				
biol fa PhD x away				-0.0166				
				(0.0206)				
biol mo MSc x away				-0.00728				
				(0.00845)				
biol mo PhD x away				-0.0317**				
				(0.0143)				
IQ 1-10	-0.000639	-0.000640	-0.000643	-0.000642	-0.000658	-0.000630	-0.000655	-0.000713
	(0.000517)	(0.000517)	(0.000517)	(0.000517)	(0.000517)	(0.000517)	(0.000517)	(0.000516)
IQ 11-20	-0.000850	-0.000848	-0.000848	-0.000850	-0.000865	-0.000845	-0.000862	-0.000916*
	(0.000540)	(0.000540)	(0.000540)	(0.000540)	(0.000539)	(0.000540)	(0.000539)	(0.000539)
IQ 21-30	-0.000835	-0.000827	-0.000826	-0.000828	-0.000844	-0.000832	-0.000842	-0.000884
	(0.000572)	(0.000572)	(0.000572)	(0.000572)	(0.000572)	(0.000572)	(0.000572)	(0.000572)
IQ 31-40	-0.000602	-0.000599	-0.000601	-0.000602	-0.000607	-0.000600	-0.000606	-0.000628
	(0.000611)	(0.000611)	(0.000611)	(0.000611)	(0.000611)	(0.000611)	(0.000611)	(0.000611)
IQ 51-60	9.72e-05	0.000102	9.55e-05	9.60e-05	0.000106	9.56e-05	0.000104	0.000141
	(0.000692)	(0.000692)	(0.000692)	(0.000692)	(0.000692)	(0.000692)	(0.000692)	(0.000691)
IQ 61-70	0.00131*	0.00132*	0.00131*	0.00131*	0.00133*	0.00131*	0.00132*	0.00139*
	(0.000762)	(0.000762)	(0.000762)	(0.000762)	(0.000762)	(0.000762)	(0.000762)	(0.000762)
IQ 71-80	0.00380***	0.00380***	0.00380***	0.00380***	0.00383***	0.00379***	0.00382***	0.00396***
	(0.000839)	(0.000840)	(0.000839)	(0.000839)	(0.000839)	(0.000839)	(0.000839)	(0.000838)
IQ 81-90	0.00549***	0.00549***	0.00549***	0.00550***	0.00553***	0.00548***	0.00552***	0.00575***
	(0.000928)	(0.000928)	(0.000928)	(0.000928)	(0.000928)	(0.000928)	(0.000928)	(0.000926)
IQ 91-95	0.0103***	0.0102***	0.0103***	0.0103***	0.0103***	0.0102***	0.0103***	0.0106***
	(0.00136)	(0.00136)	(0.00136)	(0.00136)	(0.00136)	(0.00136)	(0.00136)	(0.00136)
IQ 96-100	0.0157***	0.0157***	0.0157***	0.0157***	0.0146***	0.0159***	0.0147***	0.0116***
	(0.00156)	(0.00156)	(0.00156)	(0.00156)	(0.00166)	(0.00157)	(0.00166)	(0.00142)
fa inc 96-100 x IQ 96-100					0.00881*		0.0100**	0.00781
					(0.00503)		(0.00509)	(0.00504)
mo inc 96-100 x IQ 96-100					-0.00446		-0.00318	-0.00323
					(0.00782)		(0.00780)	(0.00781)
fa PhD x IQ 96-100						-0.00693	-0.0115	-0.0147
						(0.0138)	(0.0140)	(0.0140)
mo PhD x IQ 96-100						-0.0173	-0.0187	-0.0201
						(0.0271)	(0.0270)	(0.0271)
secondary	0.000956**	0.000961**	0.000964**	0.000966**	0.000976**	0.000965**	0.000974**	0.00107***
	(0.000397)	(0.000397)	(0.000397)	(0.000397)	(0.000396)	(0.000396)	(0.000396)	(0.000395)
college	-0.00224***	-0.00223***	-0.00223***	-0.00223***	-0.00218***	-0.00222***	-0.00219***	-0.00215***
	(0.000440)	(0.000440)	(0.000441)	(0.000440)	(0.000439)	(0.000439)	(0.000439)	(0.000439)
MSc	-0.000539	-0.000559	-0.000585	-0.000587	-0.000475	-0.000532	-0.000495	-0.000361
	(0.000775)	(0.000775)	(0.000777)	(0.000777)	(0.000774)	(0.000775)	(0.000774)	(0.000773)
PhD	0.000274	0.000276	0.000343	0.000268	0.000408	0.000323	0.000436	0.000578
	(0.00333)	(0.00333)	(0.00333)	(0.00333)	(0.00333)	(0.00333)	(0.00333)	(0.00333)
secondary STEM	-0.00164***	-0.00163***	-0.00163***	-0.00163***	-0.00164***	-0.00163***	-0.00164***	-0.00165***
	(0.000274)	(0.000274)	(0.000274)	(0.000274)	(0.000273)	(0.000273)	(0.000273)	(0.000273)
college STEM	0.0152***	0.0152***	0.0152***	0.0152***	0.0152***	0.0152***	0.0152***	0.0153***
	(0.000730)	(0.000730)	(0.000730)	(0.000730)	(0.000729)	(0.000729)	(0.000729)	(0.000729)
MSc STEM	0.103***	0.103***	0.103***	0.103***	0.103***	0.104***	0.103***	0.101***
	(0.00227)	(0.00227)	(0.00227)	(0.00227)	(0.00227)	(0.00227)	(0.00227)	(0.00245)
PhD STEM	0.225***	0.225***	0.225***	0.225***	0.225***	0.225***	0.225***	0.222***
	(0.00897)	(0.00897)	(0.00897)	(0.00897)	(0.00897)	(0.00897)	(0.00897)	(0.00902)
PhD x IQ 96 - 100								0.0178***
								(0.00607)
Observations	352,668	352,668	352,668	352,668	352,668	352,668	352,668	352,668