

NBER WORKING PAPER SERIES

WHO BECOMES AN INVENTOR IN AMERICA? THE IMPORTANCE OF EXPOSURE TO
INNOVATION

Alexander M. Bell
Raj Chetty
Xavier Jaravel
Neviana Petkova
John Van Reenen

Working Paper 24062
<http://www.nber.org/papers/w24062>

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

A preliminary draft of this paper was previously circulated under the title "The Lifecycle of Inventors." The opinions expressed in this paper are those of the authors alone and do not necessarily reflect the views of the Internal Revenue Service, U.S. Department of the Treasury, the National Institutes of Health, or the National Bureau of Economic Research. We would particularly like to thank Philippe Aghion, with whom we started thinking about these issues, for inspiration and many insightful comments. We would like to also thank Daron Acemoglu, Ufuk Akcigit, Olivier Blanchard, Erik Hurst, Danny Kahnemann, Pete Klenow, Henrik Kleven, Richard Layard, Eddie Lazear, Josh Lerner, Alex Olssen, Jim Poterba, Scott Stern, Otto Toivanen, Heidi Williams and numerous seminar participants for helpful comments and discussions. Trevor Bakker, Augustin Bergeron, Mike Droste, Jamie Fogel, Nikolaus Hidenbrand, Alexandre Jenni, Benjamin Scuderi, and other members of the Equality of Opportunity Project research team provided outstanding research assistance. This research was funded by the National Science Foundation, the National Institute on Aging Grant T32AG000186, the Lab for Economic Applications and Policy at Harvard, the European Research Council, the Economic and Social Research Council at CEP, the Kauffman Foundation, the Bill and Melinda Gates Foundation, and the Robert Wood Johnson Foundation.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2017 by Alexander M. Bell, Raj Chetty, Xavier L. Jaravel, Neviana Petkova, and John Van Reenen. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Who Becomes an Inventor in America? The Importance of Exposure to Innovation
Alexander M. Bell, Raj Chetty, Xavier Jaravel, Neviana Petkova, and John Van Reenen
NBER Working Paper No. 24062
November 2017
JEL No. E0,H0,J0

ABSTRACT

We characterize the factors that determine who becomes an inventor in America by using de-identified data on 1.2 million inventors from patent records linked to tax records. We establish three sets of results. First, children from high-income (top 1%) families are ten times as likely to become inventors as those from below-median income families. There are similarly large gaps by race and gender. Differences in innate ability, as measured by test scores in early childhood, explain relatively little of these gaps. Second, exposure to innovation during childhood has significant causal effects on children's propensities to become inventors. Growing up in a neighborhood or family with a high innovation rate in a specific technology class leads to a higher probability of patenting in exactly the same technology class. These exposure effects are gender-specific: girls are more likely to become inventors in a particular technology class if they grow up in an area with more female inventors in that technology class. Third, the financial returns to inventions are extremely skewed and highly correlated with their scientific impact, as measured by citations. Consistent with the importance of exposure effects and contrary to standard models of career selection, women and disadvantaged youth are as under-represented among high-impact inventors as they are among inventors as a whole. We develop a simple model of inventors' careers that matches these empirical results. The model implies that increasing exposure to innovation in childhood may have larger impacts on innovation than increasing the financial incentives to innovate, for instance by cutting tax rates. In particular, there are many “lost Einsteins” — individuals who would have had highly impactful inventions had they been exposed to innovation.

Alexander M. Bell
Harvard University
1805 Cambridge Street
Cambridge, MA 02138
ambell@gmail.com

Neviana Petkova
US Treasury
1500 Pennsylvania Ave
Washington, DC 20220
neviana.petkova@treasury.gov

Raj Chetty
Department of Economics
Stanford University
579 Serra Mall
Stanford, CA 94305
and NBER
chetty@stanford.edu

John Van Reenen
Department of Economics, E62-518
MIT
77 Massachusetts Avenue
Cambridge, MA 02139
and NBER
vanreene@mit.edu

Xavier Jaravel
London School of Economics
Houghton Street
London
WC2A 2AE
UK
x.jaravel@lse.ac.uk

I Introduction

Innovation is widely viewed as a central driver of economic growth (e.g., Romer 1990, Aghion and Howitt 1992). As a result, many countries use a variety of policies to spur innovation, ranging from tax incentives to technical education. One way to understand the effectiveness of such policies is to study the determinants of who becomes an inventor. What types of people become successful inventors today? What do their experiences teach us about the factors that affect rates of innovation?

Relatively little is known about the characteristics of inventors because most sources of data on innovation (e.g., patent records) do not record even basic demographic information, such as an inventor’s age or gender. In this paper, we present the first comprehensive portrait of inventors in the United States. Following standard practice in prior work on innovation, we define an “inventor” as an individual who holds a patent.¹ We link data on the universe of patent applications and grants in the U.S. between 1996 and 2014 to federal income tax returns to construct a panel dataset covering 1.2 million inventors (patent applicants or recipients). Using this new dataset, we track inventors’ lives chronologically from birth to adulthood to identify factors that determine who becomes an inventor and the types of policies that may be most effective in increasing innovation.

In the first part of our empirical analysis, we show that children’s characteristics at birth – their socioeconomic class, race, and gender – are highly predictive of their propensity to become inventors. Children born to parents in the top 1% of the income distribution are ten times as likely to become inventors as those born to families with below-median income.² Whites are more than three times as likely to become inventors as blacks. And 82% of 40-year-old inventors today are men. This gender gap in innovation is shrinking gradually over time, but at the current rate of convergence, it will take another 118 years to reach gender parity.

Why do rates of innovation vary so sharply based on characteristics at birth? In economic models, any choice can be traced to three exogenous factors: endowments (e.g., differences in genetic ability across subgroups), preferences (e.g., a greater taste for pursuing science or a career

¹The use of patents as a proxy for innovation has well-known limitations (e.g. Griliches 1990, OECD 2009). In particular, not all innovations are patented and not all patents are meaningful innovations. We address these measurement issues by showing that (a) our results hold if we focus on highly-cited (i.e., high-impact) patents and (b) the mechanisms that lead to the differences in rates of patenting across subgroups that we document are unlikely to be affected by these concerns.

²This pattern is not unique to innovation: children from high-income families are also substantially more likely to enter other high-skilled professional occupations and, more generally, reach the upper-tail of the income distribution. We focus on innovation here because it is thought to have particularly large social spillovers and because focusing on innovation has methodological advantages in understanding the mechanisms underlying career choice, as we discuss below.

with risky returns), or constraints (e.g., a lack of liquidity or opportunities to build human capital). Since each of these explanations has very different implications for policies that can be used to increase innovation, we structure most of our analysis around assessing the relative importance of these three mechanisms.

As a first step, we evaluate whether differences in ability explain these gaps in innovation using test scores in early childhood as a proxy for ability. We obtain data on test scores from 3rd to 8th grade by linking school district records for 2.5 million children who attended New York City public schools to the patent and tax records. Math test scores in 3rd grade are highly predictive of patent rates, but they account for less than one-third of the gap in innovation between children from high- vs. low-income families.³ This is because children from lower income families are much less likely to become inventors even conditional on having test scores at the top of their 3rd grade class. Differences in 3rd grade math scores also explain a small share of the gap in innovation by race, and virtually none of the gap in innovation by gender.

The gap in innovation explained by test scores grows in later grades, consistent with prior evidence that test score gaps widen as children progress through school (e.g., Fryer and Levitt 2004, Fryer 2011). Half of the gap in innovation by parent income can be explained by differences in math test scores in 8th grade. These results suggest that low-income children start out on relatively even footing with their higher-income peers in terms of innovation ability, but fall behind over time, perhaps because of differences in their childhood environment. However, they do not provide conclusive evidence about the role of environment because test scores are an imperfect measure of ability. If a child’s ability to innovate is poorly captured by standardized tests, particularly at early ages, ability could still account for a substantial share of gaps in innovation.⁴

In the second part of our empirical analysis, we address this issue by studying the impacts of childhood environment directly. We show that exposure to innovation during childhood through one’s family or neighborhood has a significant *causal* effect on a child’s propensity to become an inventor.⁵ We establish this result – which we view as the central empirical result of the paper – in a series of steps. We first show that children who grow up in commuting zones (CZs) with higher

³Test scores in English have no predictive power conditional on test scores in math, suggesting that tests in early childhood are diagnostic of specific skills that matter for innovation.

⁴On the other hand, since children from different socioeconomic backgrounds are exposed to different environments even before they enter school, these calculations could overstate the portion of the gap in innovation that is due to differences in ability.

⁵We use the term “exposure to innovation” to mean having contact with someone in the innovation sector, e.g. through one’s family or neighbors. We do not distinguish between the mechanisms through which such exposure matters, which could range from specific human capital accumulation to changes in aspirations.

patent rates are significantly more likely to become inventors, even conditional on the CZ in which they work in adulthood. We then show this pattern holds not just for whether a child innovates but also in the technology category in which he or she innovates. For example, among people living in Boston, those who grew up in Silicon Valley are especially likely to patent in computers, while those who grew up in Minneapolis – which has many medical device manufacturers – are especially likely to patent in medical devices. We find similar patterns at the family level: children whose parents or parents’ colleagues hold patents in a technology class are more likely to patent in exactly that field themselves.

These patterns of transmission hold even across the 445 narrowly defined technology subclasses into which patents can be classified. For example, a child whose parents hold a patent in amplifiers is much more likely to patent in amplifiers himself than in antennas. Moreover, the patterns are gender-specific: women are much more likely to patent in a specific technology class if female workers in their childhood CZ were especially likely to patent in that class. Conditional on women’s patent rates, men’s patent rates have no predictive power for women’s innovation. Conversely, men’s innovation rates are influenced by male rather than female inventors in their area.

Under the assumption that differences in genetic ability do not generate differences in propensities to innovate across narrow technology classes in a gender-specific manner, this set of results on patenting by technology class implies that exposure to innovation during childhood has a causal effect on innovation. Intuitively, as long as genetics do not govern one’s ability to invent an amplifier rather than an antenna in a gender-specific manner, the close alignment between the subfield in which children innovate and the type of innovation they were exposed to in their families or neighborhoods must be driven by causal exposure effects. Formally, the sharp variation in rates of innovation across technology classes and gender subgroups provides a set of overidentifying restrictions that allow us to distinguish exposure effects from plausible models of selection in observational data.

We estimate that moving a child from a CZ that is at the 25th percentile of the distribution in terms of innovation (e.g., New Orleans, LA) to the 75th percentile (e.g., Austin, TX) would increase his or her probability of becoming an inventor by at least 17% and potentially as much as 50%. These exposure effects are consistent with recent evidence documenting neighborhood exposure effects on earnings, college attendance, and other outcomes (Chetty et al. 2016, Chetty and Hendren 2017). Neighborhood effects have typically been attributed to factors that affect general human capital accumulation, such as the quality of local schools or residential segregation.

Our findings show that, at least in the context of innovation, such mechanisms are unlikely to be the sole reason that childhood environment matters, as it is implausible that some neighborhoods prepare children to innovate in one particular technology class such as amplifiers. Rather, they point to mechanisms such as transmission of specific human capital, mentoring, or networks (e.g., through internships) that lead children to pursue certain career paths. Children from low-income families, minorities, and women are less likely to have such exposure through their families and neighborhoods, which helps explain why they have significantly lower rates of innovation overall. For example, our estimates imply that if girls were as exposed to female inventors as boys are to male inventors in their childhood CZs, the gender gap in innovation would be half as large as it currently is.

Stepping forward chronologically in studying children’s environments, we next examine how rates of innovation vary across colleges. Students from low-income and high-income families at the colleges with the most inventors (e.g., MIT) go on to patent at relatively similar rates, supporting the view that factors that affect children *before* they enter the labor market are a key determinant of who becomes an inventor. For example, liquidity constraints in financing innovation or differences in risk preferences are unlikely to explain why low-income children innovate at lower rates, as such explanations would generate differences in rates of innovation even conditional on the college a child attends.

In the third part of our empirical analysis, we examine inventors’ careers after entering the labor market, with the aim of understanding how financial incentives affect individuals’ decisions to become inventors. We find that the financial returns to innovation are highly skewed and highly correlated with their scientific impact – two key facts which we later show (using a standard model of career choice) imply that small changes in financial incentives will not affect aggregate innovation significantly. In particular, the top 1% of inventors obtain more than 22% of total inventors’ income, implying that the distribution of income *within* patent-holders is even more skewed than in the population as a whole. Individuals with highly cited patents have much higher incomes, showing that the private benefits of innovation are correlated with their scientific impacts.⁶ The highest-impact patents are most commonly obtained when individuals are in their mid-forties – well after individuals make career choices – consistent with prior work (e.g., Jones et al. 2014). Much of this income is earned in the years before patents are granted, implying that they are not just returns

⁶We follow prior work (e.g., Jaffe et al. 1993) in using patent citations as a proxy for a patent’s scientific impact. Although citations are an imperfect proxy for impact, they are well correlated with other measures of value, such as firm’s profits and market valuations (Schere et al. 2000, Hall et al. 2005, Abrams et al. 2013, Kogan et al. 2017).

from the patent itself but from associated businesses or salaries.⁷

Inventors from under-represented groups (women, minorities, and those from low-income families) have very similar earnings and citations to other inventors on average. This result challenges standard models that explain differences in occupational choice by differences in barriers to entry across subgroups (e.g., Hsieh et al. 2016). Under the assumption that the ability to innovate does not vary across groups, such “rational sorting” models predict that the individuals from disadvantaged groups who become inventors will have *higher* productivity than inventors from more advantaged backgrounds, since the marginal inventors who are screened out are those with lower potential. In fact, we find the opposite: inventors from disadvantaged groups do not have higher-impact innovations on average. Put differently, women and disadvantaged youth are just as under-represented among star inventors as they are among inventors as a whole. This finding is consistent with the view that exposure is a central determinant of innovation. In particular, a lack of exposure may prevent some individuals from pursuing a career in innovation even though they would have had highly impactful innovations had they done so (“lost Einsteins,” as in Celik (2014)).⁸

We characterize the implications of our empirical findings for policies to increase innovation using a simple model of career choice in which three factors determine whether an individual pursues innovation: financial incentives (Roy 1951), barriers to entry (Hsieh et al. 2016), and exposure to innovation, which is the new element we introduce given our empirical results. We model exposure as a binary variable: individuals who do not receive exposure to innovation never pursue innovation, while those who receive exposure decide whether to pursue innovation by maximizing expected lifetime utility. Using this model, we contrast three types of policies to increase innovation: increasing exposure (e.g., through internships), reducing barriers to entry (e.g., by providing subsidies for certain subgroups), and increasing private financial returns (e.g., by cutting top income tax rates on inventors).

The model implies that the potential to increase innovation by reducing barriers to entry or increasing financial returns is limited, for three reasons. First, such policies only affect the subset of individuals who have exposure. Second, if the returns to innovation are forecastable at the point of career choice, such policies would only induce inventors of marginal quality to enter the field rather

⁷Aghion et al. (2017) and Depalo and Di Addario (2015) document analogous patterns in Finland and Italy. We focus exclusively on the impacts of patents on inventors’ own earnings; see Van Reenen (1996) and Kline et al. (2017) for analyses of how innovation affects co-workers wages in the inventor’s firm.

⁸These results could be explained by a model in which the hurdles that generate barriers to entry (e.g., discrimination) also reduce an individual’s productivity after entering innovation. Regardless of the underlying explanation, however, the under-representation of certain groups among star inventors implies that there are “lost Einsteins.”

than star inventors (Jaimovich and Rebelo 2017). The mean annual income of those with patents in the top 1% of the citation distribution is more than \$1 million between ages 40-50. The decisions of these star inventors are unlikely to be affected by small changes in financial incentives, making aggregate quality-weighted innovation relatively insensitive to tax rates. Third, if the returns to innovation are uncertain at the point of career choice, the elasticity of innovation with respect to top income tax rates is likely to be small in a standard expected utility model because tax changes only affect payouts when inventors have very high incomes and low marginal utility.

In contrast, the model implies that increasing exposure can have substantial impacts on quality-weighted innovation by drawing individuals who produce high-impact inventions into the innovation pipeline. In particular, increasing exposure to innovation specifically among children who (a) excel in math and science at early ages and (b) are from under-represented groups can have large returns, since the gaps in innovation rates by parental income, race, and gender are largest among children who have very high test scores in early grades.

It is important to keep two caveats in mind when interpreting the conclusions we draw from our model. First, our analysis focuses exclusively on the decisions of individual inventors. Taxes and financial incentives could potentially affect innovation through many other channels, for instance by changing the behavior of firms, other salaried workers who contribute to the innovation process, or through general equilibrium effects (e.g., Lerner and Wulf 2007, Akcigit et al. 2017). Taxes may also influence inventors' behavior on other margins, such as how much effort to supply or where to locate (Akcigit et al. 2016, Moretti and Wilson 2017), which are distinct from the extensive margin career choice decisions we focus on here. Second, our analysis does not provide guidance on the impacts of specific policies to increase exposure to innovation. To facilitate future work evaluating such policies, we construct a set of publicly available data tables that provide statistics on patent rates and citations by technology category, parent income group, gender, age, commuting zone, and college. In addition, we report statistics on inventors' income distributions by year and citations. These tables can be used to study a variety of issues, ranging from the impacts of local economic conditions on rates and types of innovation to how the returns to innovation have changed over time.

Related Literature. Our results build on and contribute to several literatures. First, our results relate to the literature on career choice (e.g., Topel and Ward 1992, Hall 2002). Some studies in this literature have used data on specific occupations – such as medicine and law – to show that children are particularly likely to pursue their parents' occupations (e.g., Laband and Lentz 1983,

Lentz and Laband 1989), but they have not separated causal exposure effects from selection effects as we do here. While the mechanisms we document may apply to other careers as well, we focus on innovation because of its importance for economic growth (e.g., Jones and Williams 1999, Bloom et al. 2013) and because inventors’ earnings exhibit much greater cross-sectional variance than most other highly skilled professions, leading to differences in the predicted impacts of policies such as tax changes. In addition, from a methodological perspective, patents have the advantage of being classified into narrow technological classes, allowing us to identify the causal effects of exposure and show that exposure matters not just for broad career choices but in a granular, subject-specific manner.

Second, our results relate to the literature on the misallocation of talent across occupations (e.g., Murphy et al. 1991, Hsieh et al. 2016). Our analysis does not offer any evidence that talent is misallocated, but our finding that the allocation of talent to innovation is driven partly by differences in exposure rather than ability is consistent with the premise of this literature.⁹ Indeed, our results raise the possibility that the welfare costs of distortions in the allocation of talent may be even greater than predicted by models such as Hsieh et al. (2016), since some of the individuals who fail to pursue innovation due to a lack of exposure are superstars rather than marginal entrants. More broadly, our findings suggest that improving opportunities for children from low-income or minority backgrounds (e.g., Heckman 2006, Card and Giulano 2014) could increase not just their own earnings but also economic growth by improving the allocation of talent.

Third, in the literature on innovation itself, most existing research focuses on the “demand side” of innovation, such as tax credits for research and development (e.g., Becker 2015). Some authors have called for greater focus on “supply side” policies that increase the number of inventors instead (e.g., Goolsbee 1998, Romer 2000). Our study takes a step toward understanding the supply side of innovation by characterizing the behavior of individual inventors. In doing so, it joins a nascent body of work studying the origins of inventors that draws primarily upon data from Scandinavian registries. For example, Aghion et al.’s (2017) recent study of inventors in Finland documents gaps in innovation by parental background consistent with our results and characterizes the predictive power of other factors that we do not observe in our data, such as IQ and parental education.¹⁰ Our analysis complements the work of Aghion et al. (2017) and other related studies by (a) identifying

⁹More precisely, we study the determinants of the allocation of talent across sectors, but do not present any normative evidence that welfare would be higher if individuals were to choose different occupations.

¹⁰Other studies include Giuri et al. (2007), Azoulay et al. (2011), Toivanen and Vaananen (2012), Dorner et al. (2014), Jung and Ejermo (2014), Lindquist et al. (2015), and Akcigit et al. (2017). A forerunner of this recent work was a classic study by Schmookler (1957) of 57 inventors.

different factors that affect innovation, most importantly the causal effect of exposure and (b) presenting comprehensive data and publicly available statistics on inventors’ origins and careers in the United States.

The paper is organized as follows. Section II describes the data. Sections III, IV, and V present our empirical results on inventors’ characteristics at birth, childhood environments, and career trajectories, respectively. In Section VI, we present the model of inventors’ career choices and discuss its implications for policies to increase innovation. Section VII concludes. Data tables on patent rates by subgroup can be downloaded from the Equality of Opportunity Project [website](#).

II Data

In this section, we describe our data sources, define the samples and key variables we use in our analysis, and present summary statistics.

II.A Data Sources

Patent Records. We obtain information on patents from two sources. First, we use information on patent grants from a database hosted by [Google](#), which contains the full text of all patents granted in the U.S. from 1976 to present. We focus on the 1.7 million patents that were granted between 1996 and 2014 to U.S. residents. Second, we use data on 1.6 million patent applications between 2001 and 2012 provided by Strumsky (2014).¹¹

We define an individual as an inventor if he or she is listed as an inventor on a patent application between 2001-2012 or grant between 1996-2014; for simplicity, we refer to this outcome as “inventing by 2014” below. Importantly, we include all individuals listed as inventors, not just those assigned intellectual property rights. In particular, inventors employed by companies are listed as inventors, while their company is typically listed as the assignee. In addition to inventors’ names, we also extract information on inventors’ geographic location (city and state) when they filed the patent and the 3-digit technology class to which the patent belongs, as assigned by the United States Patent and Trademark Office (USPTO). We classify patents into technology categories using the classification developed in the NBER Patent Data Project by Hall et al. (2001). We assign each inventor in our data a single technology class based on the class in which he or she has the most

¹¹In 2001, the U.S. began publishing patent *applications* (and not just patent grants) 18 months after filing. For a fee, applicants can choose to have their filing kept secret; 15% of applicants choose to do so. To ensure that this missing data problem does not generate selection bias, we verify that the results we report below are all robust to defining inventors purely using patent grants rather than applications.

patents, breaking ties randomly. We obtain data on the number of times each granted patent was cited from its issuance date until 2014 from the USPTO’s full-text issuance files.

Tax Records. We use federal income tax records spanning 1996-2012 to obtain information such as an individual’s gender and age, geographic location, and own and parental income. The tax records cover all individuals who appear in the Death Master file produced by the Social Security Administration, which includes all persons in the U.S. with a Social Security Number or Individual Taxpayer Identification Number (ITIN). The data include both income tax returns (1040 forms) and third-party information returns (e.g., W-2 forms), which give us information on the earnings of those who do not file tax returns.

The patent data were linked to the tax data using an inventor’s name, city, and state. In the tax data, these fields were obtained from the Death Master file, 1040 forms, and third-party information returns (see Online Appendix A for a complete description of the matching procedure). 88% of individuals who applied for or were granted a patent were successfully linked, with higher match rates in more recent years since information returns are unavailable prior to 1999.

We evaluate the quality of our matching algorithm by using external data on ages for a subset of inventors from Jones (2010). The age of the inventor recorded in the Death Master file matches the age reported in Jones’s dataset in virtually all cases, confirming that our algorithm generates virtually no false matches. The 12% of inventors who are not matched are individuals with common names that are difficult to link to unique records (e.g., “John Smith”), individuals with spelling errors in their names or addresses, or individuals who listed different addresses on their patent applications and tax forms. The observable characteristics (in the patent data) of unmatched inventors are very similar to those of those of matched inventors, suggesting that the individuals we match are representative of inventors in the U.S.

New York City School District Records. We use data from the New York City (NYC) school district to obtain information on test scores in childhood for the subset of individuals who attended New York City public schools. These data span the school years 1988-1989 through 2008-2009 and cover roughly 2.5 million children in grades 3-8. Test scores are available for English language arts and math for students in grades 3-8 in every year from the spring of 1989 to 2009, with the exception of 7th grade English scores in 2002. These data were linked to the tax data by Chetty et al. (2014a) with an 89% match rate, and we use their linked data directly in our analysis.

After these three databases were linked, the data were de-identified (i.e., individual identifiers were removed) and the analysis was conducted using the de-identified dataset.

II.B Sample Definitions

We use three different samples in our empirical analysis: full inventors, intergenerational, and New York City schools.

Full Inventors Sample. Our first analysis sample consists of all inventors (individuals with patent grants or applications) who were successfully linked to the tax data. There are approximately 1.2 million individuals in this sample. This sample is structured as a panel from 1996 to 2012, with data in each year on individual’s incomes, patents, and other variables. We use this sample to analyze inventors’ labor market careers in Section V.

Intergenerational Sample. Much of our empirical analysis compares inventors to non-inventors in terms of characteristics at birth (Section III) and childhood environment (Section IV). To measure conditions at birth and childhood location, we must link individuals to their parents. To do so, we use the sample constructed by Chetty et al. (2014b) to study intergenerational mobility, focusing on all children in the tax data who (1) were born in the 1980-84 birth cohorts, (2) can be linked to parents, and (3) were U.S. citizens as of 2013. Chetty et al. (2014b, Appendix A) describe how this intergenerational sample is constructed starting from the raw tax data; here, we briefly summarize its key features.

We define a child’s parents as the first tax filers between 1996 and 2012 to claim the child as a dependent and were between the ages of 15 and 40 when the child was born. Since children begin to leave the household after age 16, the earliest birth cohort that we can reliably link to parents is the 1980 birth cohort (who are 16 in 1996, when our data begin). Children are assigned parent(s) based on the first tax return on which they are claimed, regardless of subsequent changes in the parents’ marital status or dependent claiming. Although parents who never file a tax return cannot be linked to children, we still identify parents for more than 90% of children, as the vast majority of children are claimed at some point because of the tax benefits of claiming children. We restrict the sample to children who are citizens in 2013 to exclude individuals who are likely to have immigrated to the U.S. as adults, for whom we cannot measure parent income. We cannot directly restrict the sample to individuals born in the U.S. because the database only records current citizenship status.¹²

Since few individuals patent in or before their early twenties, we focus on individuals in the 1980-84 birth cohorts, who are between the ages of 28-32 in 2012, the last year of our data. There

¹²In addition, we limit the sample to parents with positive income (excluding 1.5% of children) because parents who file a tax return – as is required to link them to a child – yet have zero income are unlikely to be representative of individuals with zero income while those with negative income typically have large capital losses, which are a proxy for having significant wealth.

are 16.4 million individuals in our primary intergenerational analysis sample, of whom 34,973 are inventors. To assess whether our results are biased by focusing on innovation at relatively early ages (by age 32), we also examine a set of older cohorts using data from Statistics of Income (SOI) cross-sections, which provide 0.1% stratified random samples of tax returns prior to 1996. The SOI cross-sections provide identifiers for dependents claimed on tax forms starting in 1987, allowing us to link parents to children back to the 1971 birth cohort (Chetty et al. 2014b, Appendix A). There are approximately 11,000 individuals, of whom 131 are inventors, in the 1971-72 birth cohorts in the SOI sample that we use to study innovation rates up to age 40.

New York City Schools Sample. When analyzing whether test scores explain differences in rates of innovation (Section III), we focus on the sample of children in the NYC public schools data linked to the tax data. We also use this sample when analyzing differences in innovation rates by race and ethnicity, as race and ethnicity are only observed in the school district data. We focus on children in the 1979-1985 birth cohorts for the test score analysis because the earliest birth cohort observed in the NYC data is 1979. As in Chetty et al. (2014a), we exclude students who are in classrooms where more than 25% of students are receiving special education services and students receiving instruction at home or in a hospital. There are approximately 430,000 children in our NYC schools analysis sample, of whom 452 are inventors.

II.C Variable Definitions and Summary Statistics

In this subsection, we define the key variables we use in our analysis and present summary statistics. We measure all monetary variables in 2012 dollars, adjusting for inflation using the consumer price index (CPI-U).

Income. We use two concepts to measure individuals' incomes: wage earnings and total income. Wage earnings are total earnings reported on an individual's W-2 forms. wage earnings as well as self-employment income and capital income. Total income is defined for tax filers as Adjusted Gross Income (as reported on the 1040 tax return) plus tax-exempt interest income and the non-taxable portion of Social Security and Disability benefits minus the spouse's W-2 wage earnings (for married filers). Total income includes For non-filers, total income is defined as wage earnings. Individuals who do not file a tax return and who have no W-2 forms are assigned an income of zero.¹³ Because the database does not record W-2's and other information returns prior to 1999, we cannot reliably measure individual earnings prior to that year, and therefore measure individuals'

¹³Importantly, these observations are true zeros rather than missing data. Because the database covers all tax records, we know that these individuals have no taxable income.

incomes only starting in 1999. Income is measured prior to the deduction of individual income taxes and employee-level payroll taxes.

Parents' Incomes. Following Chetty et al. (2014b), we measure parent income as total pre-tax income at the household level. In years where a parent files a tax return, we define family income as Adjusted Gross Income (as reported on the 1040 tax return) plus tax-exempt interest income and the non-taxable portion of Social Security and Disability benefits. In years where a parent does not file a tax return, we define family income as the sum of wage earnings (reported on form W-2), unemployment benefits (reported on form 1099-G), and gross social security and disability benefits (reported on form SSA-1099) for both parents.¹⁴ In years where parents have no tax return and no information returns, family income is coded as zero. As in Chetty et al. (2014b), we average parents' family income over the five years from 1996 to 2000 to obtain a proxy for parent lifetime income that is less affected by transitory fluctuations. We use the earliest years in our sample to best reflect the economic resources of parents while the children in our sample are growing up.

Geographic Location. In each year, individuals are assigned ZIP codes of residence based on the ZIP code from which they filed their tax return. If an individual does not file in a given year, we search W-2 forms for a payee ZIP code in that year. Non-filers with no information returns are assigned missing ZIP codes. We map ZIP codes to counties and CZs using the crosswalks and methods described in Chetty et al. (2014b, Appendix A). For children whose parents were married when they were first claimed as dependents, we always track the mother's location if marital status changes.

College Attendance. Chetty et al. (2017) construct a roster of attendance at all colleges in the U.S. from 1999-2013 by combining information from IRS Form 1098-T, an information return filed by colleges on behalf of each of their students to report tuition payments, with Pell Grant records from the Department of Education.¹⁵ We assign each child in the intergenerational sample to the college he or she attends (if any) for the most years between ages 19-22. See Chetty et al. (2017, Appendix B) for further details on how colleges are identified.

Test Scores. We obtain data on standardized test scores directly from the New York City school district database. The tests were administered at the New York City school district level during

¹⁴Since we do not have W-2's prior to 1999, parent income is coded as 0 prior to 1999 for non-filers. Assigning non-filing parents 0 income has little impact on our estimates because only 3.1% of parents in the full analysis sample do not file in each year prior to 1999 and most non-filers have very low W-2 income (Chetty et al. 2014b). For instance, in 2000, the median W-2 income among non-filers in our baseline analysis sample was \$0.

¹⁵All institutions qualifying for federal financial aid under Title IV of the Higher Education Act of 1965 must file a 1098-T form in each calendar year for any student that pays tuition. The Pell Grant records are used to identify students who pay no tuition.

the period we study. Following Chetty et al. (2014a), we normalize the official scale scores from each exam (math and English) to have mean zero and standard deviation one by year and grade to account for changes in the tests across school years.

Summary Statistics. Table I presents descriptive statistics for the three analysis samples described above. Column 1 presents statistics for the full inventors sample; columns 2 and 3 consider inventors and non-inventors in the intergenerational sample; and columns 4 and 5 consider inventors and non-inventors in the NYC schools sample.

In the full inventors sample, the median number of patent applications between 1996-2012 is 1 and the median number of citations per inventor is also only 1. But these distributions are very skewed: the standard deviations of the number of patent applications and citations are 11.1 and 118.1, respectively. Inventors have median annual wage earnings of \$83,000 and total income of \$100,000. Again, these distributions are very skewed, with large standard deviations and mean incomes well above the medians. The mean age of inventors is 44 and 13% of inventors in the sample are women.

The intergenerational and NYC school samples have younger individuals because they are restricted to more recent birth cohorts. As a result, inventors in these subsamples have lower median incomes, patent applications, and citations than in the full sample.

III Inventors' Characteristics at Birth

In this section, we study how rates of innovation differ along three key dimensions determined at birth: parental income, race, and gender. We first document gaps in rates of innovation and then use test score data to assess the extent to which these gaps can be explained by differences in ability.

III.A Gaps in Innovation by Characteristics at Birth

Parental Income. Figure Ia plots the fraction of children who invent by 2014 vs. their parents' income percentile using our intergenerational analysis sample (children in the 1980-84 birth cohorts). We assign parents percentile ranks by ranking them based on their mean household income from 1996 to 2000 relative to other parents with children in the same birth cohort. Children from higher-income families are significantly more likely to become inventors. 8 out of 1,000 children born to parents in the top 1% of the income distribution become inventors, 10 times higher than the rate among those with below-median-income parents. The relationship is steeply upward sloping even

among high-income families: rates of innovation rise by 22% between the 95th and 99th percentile of the parental income distribution. This pattern suggests that liquidity constraints or differences in resources are unlikely to fully explain why parent income matters, as liquidity constraints are less likely to bind at higher income levels and resources presumably have diminishing marginal returns.

Figure Ib shows that the probability a child has highly-cited patents – defined as having total citations in the top 5% of his or her cohort’s distribution – has a very similar relationship to parental income. Hence, the relationship between patenting and parent income is not simply driven by children from high-income families filing low-value or defensive patents at higher rates. The pattern in Figure I also remains robust at older ages, allaying the concern that children from higher-income families may simply patent earlier than those from low-income families. In particular, using the Statistics of Income 0.1% sample, we find that the relationship between rates of innovation between ages 30 and 40 and parental income remains qualitatively similar (Online Appendix Figure Ia). Defining inventors purely on the basis of patent grants or patent applications also yields similar results (Online Appendix Figure Ib).

The relationship between innovation and parental income is representative of the relationship between achieving professional success and parental income more generally. Children’s propensities to reach the upper tail of the income distribution have a similarly convex and sharply increasing relationship with parental income (Online Appendix Figure II). For instance, children with parents in the top 1% of the parent income distribution are 27 times more likely to reach the top 1% of their birth cohort’s income distribution and 10.6 times more likely to reach the top 5% of their cohort’s income distribution than those born to parents below the median. As discussed in the introduction, we focus on innovation here (rather than professional success in general) because of innovation’s relevance for economic growth, its unique risk profile, and its advantages in characterizing mechanisms more precisely. However, the results and mechanisms we establish here may apply to other careers beyond innovation.

Race and Ethnicity. Next, we turn to gaps in innovation by race and ethnicity. Since we do not observe race or ethnicity in the tax data, we use the New York City school district sample for this analysis. The first set of bars in Figure II shows the fraction of children who patent by 2014 among white non-Hispanic, Black non-Hispanic, Hispanic, and Asian children. 1.6 per 1,000 white children and 3.3 per 1,000 Asian children who attend NYC public schools between grades 3-8 become inventors. These rates are considerably higher than those of Black children (0.5) and

Hispanics (0.2), consistent with evidence from Cook and Kongcharoen (2010).¹⁶

Since there are significant differences in parental income by race and ethnicity, the raw gaps across race and ethnicity partly reflect the income gradient shown in Figure I. To separate these two margins, we control for differences in income by non-parametrically reweighting the parental income distributions of Blacks, Hispanics, and Asians to match that of whites in the NYC sample, following the methodology of DiNardo et al. (1996). We divide the parental income distribution of children in the NYC sample into ventiles (20 bins) and compute mean patent rates across the 20 bins for each racial/ethnic group, weighting each bin by the fraction of white children whose parents fall in that income bin (i.e., integrating over the income distribution for whites).

The second set of bars in Figure II plot the resulting innovation rates, which can be interpreted as the innovation rates that would prevail for each group if it had the same income distribution as whites. Adjusting for income differences does not eliminate the racial and ethnic gaps, but changes their magnitudes. The Black-white gap falls by a factor of 2 (from 1.1/1000 to 0.6/1000). The white-Asian gap widens from 1.7/1000 to 2.6/1000 when we reweight by income, as Asian parents in NYC public schools have lower incomes on average than white parents. The Hispanic-white gap remains essentially unchanged.

Gender. Finally, we examine gaps in innovation by gender. Since gender is recorded in the tax data for all individuals in the population, we use the full inventors sample for this analysis. The advantage of doing so is that we can examine gender differences in rates of innovation not just for those born in the 1980s as in our intergenerational sample, but for older cohorts as well.

Figure III plots the fraction of female inventors – individuals who applied for or were granted a patent between 1996 and 2014 – by birth cohort.¹⁷ Consistent with prior work (Thursby and Thursby 2005, Ding et al. 2006, Hunt 2009, Kahn and Ginther 2017), we find substantial gender differences in innovation for those in the prime of their careers today; for instance, 18% of inventors born in 1980 are female. What is less well known from prior work is the rate at which this gap

¹⁶The innovation rates are lower than those in Figure Ia because NYC public schools have predominantly low-income students, with more than 75% of students from families with incomes below the national median. NYC public schools also have a much larger share of minorities than the U.S. population: 19.5% of the children in our NYC sample are white, 9.6% are Asian, 33.7% are Hispanic, and 36.0% are Black. Although we cannot be sure that the racial patterns within the NYC schools hold nationally, we do find that the relationship between parental income and innovation in the NYC sample is very similar to the national pattern in Figure Ia, suggesting that it provides representative evidence at least on the socioeconomic dimension (Online Appendix Figure Ic).

¹⁷Because we examine patenting in a fixed time window, we measure patent rates at different ages for different cohorts, ranging from ages 56-72 for the 1940 cohort to ages 16-32 for the 1980 cohort. This approach yields consistent estimates of the gender gap across cohorts if gender differences in patenting do not vary by age. While we cannot evaluate the validity of this assumption across all cohorts, examining patent rates at a fixed age (e.g., age 40) over the 17 cohorts we can analyze yields similar results (not reported).

is changing over time. Figure III shows that the fraction of female inventors was only 7% in the 1940 cohort and has risen monotonically and linearly over time. However, the rate of convergence is slow: a 0.27 percentage point (pp) increase in the fraction of female inventors per cohort on average, based on a linear regression. At this rate, it will take another 118 years to reach gender parity in innovation.

III.B Do Differences in Ability Explain the Gaps in Innovation?

Why do rates of innovation vary so widely across individuals with different characteristics at birth? In economic models, any choice can be traced to three exogenous factors: endowments (e.g., differences in ability), preferences (e.g., tastes for risk), or constraints (e.g., a lack of liquidity). In this subsection, we take a step toward evaluating the first of these factors – innate ability – by using data on childhood test scores for children in our New York City schools sample. Although students who attend New York City public schools are a selected subgroup, differences in innovation rates by parental income (Online Appendix Figure Ic) and gender (Table I) are very similar in the NYC school district sample as in the full intergenerational sample. We consider whether test scores explain the gap in innovation within the NYC sample by income, race, and gender in turn.

Parental Income. In Table IIa, we estimate the fraction of the gap in innovation by parental income that can be explained by math test scores in 3rd grade (the first grade we observe in the NYC data).¹⁸ We define “high-income” children as children with parents in the top income quintile within the NYC sample, placing all others in the “lower-income” category; using other thresholds to divide the two groups yields similar results. We focus on math test scores because scores in English do not predict innovation rates conditional on math scores (Online Appendix Table I).¹⁹

The first row of Table II shows that 1.93 out of 1,000 children from top-quintile families born between 1979-85 invent by 2014, as compared with 0.52 out of 1,000 children from lower-income families. The raw gap in innovation across these income groups is thus 1.41 inventors per 1,000 children. In the second row, we reweight the test scores of the lower-income students to match those of children from high income families, following the methodology of DiNardo et al. (1996) as in our analysis of income and race above. We divide the 3rd grade math test score distribution of

¹⁸Of course, 3rd grade test scores are not a pure measure of intrinsic ability, as children from different socioeconomic backgrounds are exposed to different environments even before 3rd grade. Nevertheless, we show that 3rd grade test scores can provide informative bounds on the extent to which innovation gaps are driven by ability, particularly when compared with test scores in later grades.

¹⁹The same is not true for success on other dimensions: for instance, both math and English scores are predictive of the probability that a child reaches the top 1% of the income distribution (Online Appendix Table I).

children in the NYC sample into ventiles (20 bins) and compute mean patent rates across the 20 bins for the lower-income group, weighting each bin by the fraction of high-income children with test scores in that bin. The second row of Table II shows that children from lower-income families would have a patent rate of 0.96 per 1,000 (rather than 0.52) if they had the same test scores as children from high-income families. The patent rate rises because children from high-income families have higher test scores in 3rd grade; for instance, children from the top income quintile score 0.65 SD higher on average than children from lower quintiles (Online Appendix Figure IIIa). However, these differences in test scores explain less than 1/3 of the raw gap in innovation, as the gap remains at 0.97 per 1,000 even after adjusting for differences in test scores, as shown in column 3 of Table II.

Figure IVa illustrates why test scores fail to fully explain the gap in innovation by plotting innovation rates vs. test scores for children with parents in the top quintile (circles) and those with lower-income parents (triangles). Each point in this figure shows the fraction of inventors within a ventile of the test score distribution. In high-income families, children who score highly on 3rd grade math tests are much more likely to become inventors than those with lower test scores. By contrast, in lower-income families, children with higher test scores do not have much higher innovation rates. As a result, among students with test scores in the top 5% of the distribution, those from high-income families are more than twice as likely to become inventors as those from lower-income families. This result suggests that becoming an inventor in America relies on two traits: having high ability (as proxied for by test scores early in childhood) *and* being born into a high-income family.

To obtain further insight into the role of ability, we repeat the preceding analysis using test scores in later grades. Figure V plots the fraction of the raw gap in innovation that is explained by math test scores in each grade from grades 3-8. As children get older, test scores account for more of the gap in innovation by parental income. By 8th grade, 48% of the gap can be explained by differences in test scores, significantly higher than the 31% in 3rd grade. Based on a linear regression across the six grades in which we observe scores, we estimate that on average an additional 3.2 percentage points of the gap is explained by test scores each year ($p < 0.01$).²⁰

Extrapolating linearly back to birth, our estimates imply that only 5.7% of the gap in innovation would be explained by test scores (ability) at birth. Conversely, test scores at the end of high school

²⁰This result is robust to a variety of alternative specification choices, including using a balanced panel of children across the grades, splitting the sample at the median instead of the 80th percentile of parental income, and controlling for test scores using regressions instead of reweighting (not reported).

would explain 60.1% of the gap. These results suggest that low-income children start out on even footing with their higher-income peers in terms of ability, but fall steadily behind as they progress through school. Indeed, we show below that conditional on the college that children attend, the gap in innovation by parent income is one-tenth as large as the raw gap shown in Figure I.

Race and Ethnicity. We use an analogous reweighting approach to estimate how much of the racial gaps in innovation can be accounted for by test scores. The third set of bars in Figure II show the innovation rates that would prevail if all children had 3rd grade math test scores comparable to those of whites. The gaps shrink modestly, showing that test scores explain very little of the racial gaps in innovation. For example, the Black-white gap shrinks from 1.1 to 1.0, a change of less than 10%, while the Asian-white gap falls by 9%. Figure IVb illustrates why this is the case by plotting patent rates vs. test scores by race and ethnicity. Even conditional on test scores, whites and Asians are substantially more likely to become inventors than Blacks and Hispanics. Very few of even the highest-scoring Black and Hispanic children pursue innovation.

Replicating the reweighting analysis by grade, we find that test scores in later grades explain more of the racial gaps in innovation, consistent with the patterns for income. For instance, 51% of the gap in patent rates between Asians and other racial and ethnic groups can be explained by 8th grade test scores.

Gender. Finally, we conduct an analogous exercise for gender, reweighting girls' test scores to match that of boys. Math test scores in 3rd grade account for only 2.4% of the difference in innovation rates between males and females (Online Appendix Table II). This is because the distribution of math test scores for boys and girls is extremely similar in 3rd grade (Online Appendix Figure IIIb). Similar to the patterns by race and parental income, high-scoring girls are much less likely to become inventors than high-scoring boys (Figure IVc).

Even in 8th grade, test scores account for only 8.5% of the gender gap in innovation. One explanation for why the gender gap in test scores expands less across grades than racial and class gaps is that boys and girls attend similar schools and grow up in similar neighborhoods, whereas children with different parental income and racial backgrounds do not.

Overall, the results in this section are consistent with evidence from other domains that disparities in ability are small at birth and expand gradually over time (e.g., Fryer and Levitt 2006, Fryer 2011). One explanation for these patterns is that differences in childhood environment – e.g., in the quality of schools or the degree of exposure to science and innovation – affect the amount students learn or the amount of time they study. However, as noted in prior work, one must be

cautious in attributing these results to environmental differences. If tests at later ages are more effective at capturing intrinsic ability, one may find the patterns across grades documented above even in the absence of differences in childhood environment. In light of this limitation, we directly examine the causal effects of childhood environment in the next section.

IV Childhood Environment and Exposure to Innovation

In this section, we study how differences in a child’s environment prior to entering the labor market affect rates of innovation. We first characterize the effects of a child’s family and neighborhood, showing in particular that exposure to innovation has a significant causal effect on a child’s propensity to innovate. We then examine how rates of innovation vary within and across colleges, which sheds further light on the importance of pre-labor-market factors.

IV.A Parents

To characterize the role that children’s parents play in shaping their decision to pursue innovation, we begin by asking whether children whose fathers are inventors are more likely to become inventors themselves.²¹ In our intergenerational analysis sample (children in the 1980-84 birth cohorts), 2.0 out of 1,000 children whose parents were not inventors become inventors by 2014. In contrast, 18.0 per 1,000 children of inventors become inventors themselves – a nine-fold difference.²² This pattern holds even conditional on parental income, across the parent income distribution (not reported).

The intergenerational persistence of innovation could be driven by the genetic transmission of ability to innovate across generations or by an *exposure effect* – the environmental effect of growing up in a family of innovators, holding one’s intrinsic ability fixed. These exposure effects could reflect the accumulation of specific human capital, changes in preferences, or simply increased awareness about innovation as a career pathway.

We distinguish between intrinsic ability and exposure effects by exploiting variation in the specific technology class in which a child innovates. Based on the USPTO’s classification system, patents can be grouped into seven broad categories (chemicals, computers and communications,

²¹We focus on fathers here because the vast majority of inventors, particularly in older generations, are male (Figure III). We examine the role of female inventors in the context of neighborhood differences, where we have greater power, in section IV.B below. We define a father as an inventor if he applied for a patent between 2001-2012 or was granted a patent between 1996-2014, analogous to the definition for children.

²²Part of this association reflects the fact that children and their fathers sometimes are co-inventors on the same patent. However, this is relatively rare: 13.7 out of 1,000 children of inventors file patents on which their parent is not a co-inventor, still far higher than the rate for non-inventors. Additionally, our measure of parental inventor status suffers from measurement error because we do not observe parents’ patents prior to 1996 in our data, likely attenuating our estimate of the difference.

drugs and medical, electrical and electronic, mechanical, design and plant, and other). Within these categories, patents are further classified into 37 sub-categories and 445 specific technology classes. These technology classes are very narrow: for instance, within the communications category, there are separate classes for modulators, demodulators, and oscillators; within the resins subcategory, there are separate classes for synthetic and natural resins.

We isolate the causal effects of exposure by analyzing whether children are particularly likely to patent in the same technology classes as their parents. The idea underlying our research design is that genetic differences in ability are unlikely to lead to differences in propensities to innovate across similar, narrowly-defined technology classes. For instance, a child is unlikely to have a gene that codes specifically for ability to invent in modulators rather than oscillators. Under this assumption, the degree of alignment between the specific technology classes in which children and their parents innovate can be used to estimate causal exposure effects.

Implementing this research design requires a metric for the degree of similarity between technology classes. We define the distance between two technology classes A and B based on the share of inventors in class A who also invent in class B; the higher the share of common inventors, the lower the distance between A and B. Online Appendix Table III gives an example that illustrates this distance metric by showing the technology classes that are closest to technology class 375, “pulse or digital communications.” Pulse or digital communications has a distance of zero with itself by definition. Inventors who had a patent in pulse or digital communications were most likely to have another patent in demodulators, which is therefore assigned an ordinal distance of $d = 1$ from the pulse and digital communications class. The next closest class is modulators ($d = 2$), and so on.

Figure VIa plots the fraction of children who patent in a technology class d units away from their father’s technology class, among children of inventors in our intergenerational sample.²³ Nearly 1 in 1,000 children patent in exactly the same technology class as their father ($d = 0$). In contrast, the probability of inventing in the next closest technology class (with distance $d = 1$) is less than 0.2 per 1,000, an estimate that is significantly different from the value at $d = 0$ with $p < 0.01$. The child’s probability of inventing in a given class then falls gradually as d rises, although the gradient is relatively flat compared to the jump between $d = 0$ and $d = 1$.

The jump in innovation rates at $d = 0$ suggests that part of the reason that children of inventors are more likely to become inventors themselves is due to exposure to innovation rather than

²³Children or fathers who patent in multiple technology classes are assigned the technology class in which they patent most frequently. We omit observations where a child and his or her father are co-inventors on the same patent to eliminate mechanical effects on the rate of patenting in the same class.

differences in innate ability. To formalize the identification assumption underlying this conclusion, let $e_{ic} \in \{0, 1\}$ represent an indicator for whether child i 's father has a patent in technology class c (i.e., if child i is “exposed” to innovation in class c) and α_{ic} represent the child’s intrinsic ability to innovate in class c . Suppose that child i patents in technology class c if $\alpha_{ic} + \beta e_{ic} > 0$. Our identification assumption is that:

$$\lim_{d \rightarrow 0} Cov(\alpha_{i,c} - \alpha_{i,c+d}, e_{i,c} - e_{i,c+d}) = 0. \quad (1)$$

Equation (1) requires that an individual’s intrinsic ability to innovate in a technology class does not covary with whether his father innovates in that particular technology class among technology classes that are very similar. Under this assumption, we can identify the causal effects of exposure (β) even though ability is correlated with exposure ($Cov(e_{ic}, \alpha_{ic}) > 0$) by analyzing how a child’s propensity to innovate in a given technology class varies with the distance between that class and the class in which his parents patented. In particular, the jump in rates of innovation at $d = 0$ in Figure VI cannot be generated by differences in ability under the assumption in (1) and must therefore be driven by the causal effect of exposure.²⁴

Interpreting the difference in innovation rates between technology class $d = 0$ and $d = 1$ as purely driven by exposure, we infer that having a parent who is an inventor increases a child’s probability of being an inventor by at least 0.078 pp through exposure effects.²⁵ Given the mean invention rate of 0.21 pp in the intergenerational sample, this result implies that exposure to innovation through one’s parents increases innovation by at least 35%. This calculation is conservative because it only attributes innovation within the same class to exposure effects. If exposure has more diffuse impacts across technology classes – as is likely the case – the total impact of exposure to a parental inventor could be considerably larger. For instance, replicating the analysis in Figure VIa at the patent sub-category level instead of the class level, we find that children are 0.2 pp more likely to invent in the same sub-category as their father than the next closest sub-category. If this increase is entirely due to exposure (i.e., if intrinsic ability to innovate does not vary sharply across similar

²⁴Equation (1) is a convenient way to conceptualize our research design, but we cannot literally take the limit as $d \rightarrow 0$ because of the discreteness of technology classes. In practice, we effectively assume that $Cov(\alpha_{i,c} - \alpha_{i,c+1}, e_{i,c} - e_{i,c+1}) = 0$, i.e. that a child’s ability to invent in a technology class does not covary with parental exposure across two adjacent classes.

²⁵If patent frequencies vary substantially across technology classes, patent rates could vary between the $d = 0$ and $d = 1$ groups for mechanical reasons unrelated to exposure, effectively because (1) would be violated by differences in the size of technology classes. To gauge the extent of such mechanical biases, we verify that there is essentially no change in patent rates from $d = 0$ to $d = 1$ in placebo tests where children were randomly reassigned fathers (not reported).

sub-categories), then exposure to a parental inventor doubles the probability that a child becomes an inventor.

In sum, exposure to innovation in one’s family substantially increases the likelihood that a child pursues innovation.²⁶ Although this result is useful in establishing that exposure matters, replicating the level of exposure one obtains through one’s parents is likely to be challenging from a policy perspective. Moreover, parents are only one of many potential sources through which children may acquire knowledge about careers in innovation. We therefore turn to two broader sources of exposure outside one’s immediate family: parents’ colleagues and residential neighbors.

IV.B Parents’ Colleagues

In this subsection, we examine how exposure to innovation through parents’ colleagues affects a child’s propensity to become an inventor. To do so, we first assign each father in our intergenerational sample an industry based on the six-digit NAICS code of his most frequent employer between 1999-2012.²⁷ We then measure the patent rate among workers in the father’s industry – whom we term the father’s “colleagues” – as the average number of patents issued to individuals in that industry per year (between 1996-2012) divided by the average number of workers in that industry per year between 1999-2012 based on counts of W-2 forms in the tax data. To ensure that we do not capture the effects of parental exposure itself, we drop children whose own parents were inventors throughout the remainder of this section.

In column 1 of Table III, we regress the fraction of children who become inventors among those with fathers in a given industry on patent rates for workers in that industry. This regression has one observation for each of the 345 industries and is weighted by the number of fathers in each industry.²⁸ The estimate of 0.250 (s.e. = 0.028) implies that a 1 percentage point increase in the patent rate among a father’s colleagues is associated with a 0.25 percentage point increase in the probability that a child becomes an inventor. This estimate implies that a one standard deviation (0.24 pp) increase in the fraction of inventors in the father’s industry is associated with a 25.3% increase in children’s innovation rates.

²⁶This result is consistent with Aghion et al.’s (2017) finding that parental education is highly predictive of a child’s propensity to innovate even conditional on parental income: children with more educated parents may be more exposed to science and innovation.

²⁷For individuals receiving W-2s from multiple firms in a given year, we define the employer in that year to be the firm that issued the W-2 with the highest salary. We exclude fathers working in industries with fewer than 50,000 individuals (5% of fathers), as patent rates are measured imprecisely for these industries.

²⁸This regression is equivalent to regressing an indicator for whether a child is an inventor on the rate of innovation in his or her father’s industry in a dataset with one observation per child, clustering standard errors by industry, because the innovation rate (the right hand side variable) does not vary within industries.

The association in column 1 of Table III could reflect either the causal effect of exposure to innovation through a parents' colleagues or a correlation with other unobservables, such as a child's own intrinsic ability to innovate. As above, we isolate exposure effects by testing whether children are more likely to innovate in exactly the same technology classes as their parents' colleagues. Using the same measure of distance d between technology classes defined in Section IV.A, we estimate OLS regressions of the form:

$$y_{cj} = \kappa_c + b_d P_{c+d,j} + \varepsilon_{cj}, \quad (2)$$

where y_{cj} denotes the patent rate in technology class c of children with fathers who work in industry j , $P_{c+d,j}$ denotes the patent rate in the class $c + d$ among workers in industry j , and κ_c represents a class-specific intercept. We estimate these regressions at the industry by technology class level, weighting by the number of children with fathers in each industry. We include class fixed effects (κ_c) to account for the variation in size across classes and identify b_d from variation across industries in class-specific patent rates.

Figure VIb plots estimates from regressions analogous to (2). Each bar plots estimates of b_d from a separate regression, varying the distance d used to define workers' patent rates $P_{c+d,j}$ in (2). The first bar plots b_0 , the relationship between children's patent rates in a given class and their fathers' colleagues patent rates in the same class ($d = 0$). In the second bar, we define $P_{c+d,j}$ as the mean patent rate in the father's industry in the next 10 closest classes ($d = 1$ to 10). The third bar uses the average patent rate in classes with $d = 11$ to 20, and so on. The coefficient b_j on parents' colleagues' patent rates drops by 85% from the same class ($d = 0$) to the next closest classes ($p < 0.01$). That is, children are much more likely to patent in exactly the same class as their parents' colleagues than in very similar classes. This result implies that an increase in parents' colleagues patent rates *causes* an increase in a child's propensity to innovate under the following identification assumption:

$$\lim_{d \rightarrow 0} Cov(\varepsilon_{c,j} - \varepsilon_{c+d,j}, P_{c,j} - P_{c+d,j}) = 0. \quad (3)$$

This assumption, which is analogous to (1), requires that as the distance d between technology classes grows small, differences in unobservable determinants of children's innovation rates in class c vs. $c + d$ are orthogonal to differences in parents' colleagues' innovation rates in those classes. Intuitively, we require that children whose fathers work in an industry where many workers patent in amplifiers rather than antennas do not have greater intrinsic ability to invent in amplifiers relative to antennas themselves. Under this assumption, we can infer from Figure VIb that a 1 pp increase

in patent rates among parental colleagues increases a child’s probability of becoming an inventor by at least $b_0 - b_{1-10} = 0.065$ pp due to a causal exposure effect.

Our measure of distance between technology classes based on co-patenting rates is one of many potential approaches to identifying “similar” patent classes. To assess the sensitivity of our results to this choice, we use the USPTO’s hierarchical classification system, which groups patents into similar fields (categories, sub-categories, and classes), as an alternative way to identify similar patent classes. In columns 2-5 of Table III, we estimate a series of regressions to assess whether children patent in the same fields as workers in their father’s industry using the USPTO’s classification system. In column 2, we test whether children are more likely to invent in the same categories as their father’s colleagues using a regression specification analogous to (2) estimated at the category by industry level with $d = 0$. Columns 3 and 4 replicate the specification in column 2 at the sub-category and technology class levels. Finally, in column 5 of Table III, we replicate column 4 with three additional controls: patent rates in (i) the same sub-category but in a different patent class, (ii) the same category but a different sub-category, and (iii) other categories.

At all levels of the hierarchy, we find a strong, statistically significant association between children’s patent rates and their parents’ colleagues patent rates. Moreover, column 5 shows that innovation among parents’ colleagues leads to a 10 times larger increase in innovation in exactly the same technology class (e.g., synthetic resins) as it does in other classes even within the same sub-category (e.g., natural resins). The coefficient on the own-class patent rate is not statistically different from the specification in column 4, while the coefficients on the other-class and category patent rates are very close to zero. Under our identification assumption in (3), the much smaller estimates in other classes imply that children’s higher propensity to invent in the same class as other workers in their parents’ industry is driven by the causal effect of exposure.

The class-specificity of the exposure effects also sheds light on the mechanism through which exposure matters. Transmission of general human capital or an interest in science would be unlikely to have impacts that vary so sharply by technology class. Instead, the data point to mechanisms such as transmission of specific human capital, access to networks that help children pursue a certain subfield, acquisition of information about a specific career pathway, or role model effects.

IV.C Neighborhoods

In this subsection, we study how rates of innovation in the neighborhood in which a child grows up affect his or her propensity to innovate. Following Chetty et al. (2014b), we assign children in our

intergenerational sample to commuting zones based on where they were first claimed as dependents by their parents.

Figure VIIa maps rates of innovation across the commuting zones where children grew up, with darker colors representing areas where more children become inventors. Figure VIIb lists the ten CZs where children are the most or least likely to grow up to become inventors (among the 100 most populated CZs). Children who grow up in the Northeast, coastal California, and the rural Midwest have the highest probability of becoming inventors, while those in the Southeast have the lowest probability. The areas where children grow up to become inventors generally have higher mean incomes (population-weighted correlation $\rho = 0.63$), fewer single parents ($\rho = -0.39$), and higher levels of absolute upward intergenerational mobility ($\rho = 0.32$), based on the CZ-level measures defined in Chetty et al. (2014b). However, there are some stark exceptions to these patterns, such as Detroit, MI, where children have among the highest likelihood of becoming inventors but where income mobility and mean incomes are relatively low.²⁹

The spatial analysis in Figure VII differs from previous analyses of “innovation clusters” and agglomeration (e.g., Porter and Stern 2001, Kim and Marschke 2005), as it reflects the locations where inventors grow up, which may differ from where they work and patent as adults. Nevertheless, children who grow up in the areas that have the most innovation tend to be most likely to go into innovation themselves. For instance, children who grow up in the San Jose commuting zone (which includes Silicon Valley) top the list in terms of the probability of becoming inventors themselves.

To examine this relationship more systematically, we define the patent rate for workers living in each CZ as the average number of patents issued per year (in the full USPTO data) to individuals from a given CZ between 1980 and 1990 divided by the CZ’s population between the ages of 15-64 in the 1990 Census. Figure VIII presents a scatter plot of the fraction of children who go on to become inventors vs. the patent rate for workers in their childhood CZ (among the 100 most populated CZs). There is a clear positive relationship between these variables, with a correlation of $\rho = 0.75$ (unweighted).

The correlation in Figure VIII is consistent with the hypothesis that exposure to innovation during childhood through one’s neighbors increases a child’s propensity to innovate, but could also reflect geographical sorting by ability to innovate. We distinguish between these explanations by examining variation in patent rates across technology classes, as in our analysis of industry-

²⁹We defer a more comprehensive examination of the spatial variation in innovation – which can be conducted using the CZ-level data that is publicly available in Online Data Table Ia – to future research, focusing here on the role of exposure to innovation during childhood.

level differences. Figure VIc plots regression coefficients of children’s innovation rates in a given technology class c on class-level patent rates among workers in their childhood CZs, using the same methodology as in Figure VIb. Once again, there is a drop of 85% in the coefficient on patent rates between the same class and the next closest classes ($p < 0.01$), showing that neighborhoods have causal exposure effects under an identification assumption analogous to (3).

In Table IV, we evaluate the robustness of this result and the mechanisms underlying it using a set of fixed effects regressions specifications. As a reference, in column 1, we regress the fraction of children who grow up to be inventors in each CZ on the patent rate of workers in their childhood CZ, replicating the analysis in Figure VIII including all 741 CZs rather than just the 100 largest ones. The coefficient of 2.9 implies that a 1 SD (0.02 pp) increase in the annual CZ-level patent rate is associated with a 0.058 pp (28.5%) increase in the fraction of children who file a patent application by 2012. One potential explanation for the results in column 1 (and Figure VIc) is that children tend to stay near the areas where they grew up, and may mechanically end up being more likely to patent if they live in an area like Silicon Valley simply because the jobs that are available in such areas tend to be in the innovation sector. To distinguish this mechanism from childhood exposure effects that change children’s behavior, we focus on the subset of children who move to a different CZ in adulthood from where they grew up. In column 2, we estimate a regression analogous to that in column 1 at the childhood CZ by current CZ level, limiting the sample to children whose current (2012) CZ differs from their childhood CZ. We regress the fraction of children who grow up to be inventors in each of these cells on the patent rate of the CZ in which they grew up, including fixed effects for the child’s 2012 CZ so that the coefficient of interest is identified purely from comparisons across individuals who grew up in different areas but currently live in the same area. The coefficient on the patent rate in the childhood CZ remains similar in this specification, showing that most of the relationship in column 1 is not mechanically driven by the types of jobs available in an area.

Next, we analyze whether this association continues to hold at the category level: do children go on to patent in the same categories as their neighbors did while they were growing up? We consider three different specifications. In column 3, we replicate the specification in column 1 at the CZ by patent category level, using the same specification as in (2) but letting j index CZs instead of industries. In column 4, we replicate the specification in column 2 at the category level. We restrict attention to movers and regress the share patenting in a given category (with one observation per childhood CZ, current CZ, and category) on the childhood CZ patent rate in that

category. We include current CZ fixed effects interacted with category in this specification. In column 5, we include all children and replace the CZ by category fixed effects with fixed effects for the father’s industry by category, estimating the model at the childhood CZ by father’s industry by category level. This specification isolates variation from one’s neighbors that is orthogonal to the variation from parents’ colleagues examined above. In all of these specifications, we find robust and significant positive coefficients on CZ-level patent rates. These specifications effectively show that children who grow up in Silicon Valley are especially likely to patent in computers, while children who grow up in Minneapolis (which has many medical device manufacturers) are especially likely to patent in medical devices. This is true even among children who live in the same place in adulthood and who have parents who work in the same industry.

In columns 6 and 7, we replicate the specification in column 3 at the sub-category and technology class level, respectively. We continue to find substantial positive coefficients in these specifications, showing that children go on to innovate in the same specific technology classes that those around them did during their childhood.³⁰ Column 8 replicates the specification in column 7 including controls for patent rates in other classes, sub-categories, and categories, as in column 5 of Table III. The coefficient on the own-class coefficient is not statistically different from the specification in column 7, while the coefficients on the other-class and category patent rates are close to zero. That is, rates of innovation in other classes and categories have very little predictive power for children’s propensity to innovate in a given class. As discussed above, this result implies that neighbors have causal exposure effects on children’s propensity to innovate in a given technology class under an identification assumption analogous to (1).

Interpreting the class-level coefficient of 1.02 in column 8 as a pure exposure effect, we infer that at least $1.02/2.9 = 35\%$ of the raw cross-sectional association between CZ-level patent rates and children’s propensities to innovate in column 1 is due to a causal exposure effect. It follows that a 1 SD increase in patent rates in a CZ increases children’s propensities to innovate by at least 10% through exposure. This effect is about 1/3 as large as the impact of direct parental exposure, showing that exposure to innovation through one’s peers and neighbors – which is potentially much more scalable through well-designed internship or mentorship programs – can have meaningful impacts on children’s propensities to innovate.

Gender-Specific Exposure Effects. Next, we examine the heterogeneity of exposure effects by

³⁰There are insufficient individuals patenting in the same sub-category or class within the same CZ to implement these specifications with current CZ fixed effects.

gender, focusing specifically on whether girls are more likely to go into innovation if they are exposed to female inventors as children. As a first step, Figure IX shows how gender gaps in innovation vary across the areas in which children grow up using our intergenerational analysis sample. Panel A maps the fraction of female inventors by the state in which inventors grew up, while Panel B shows this statistic for the top 10 and bottom 10 CZs among the 100 largest CZs.³¹ Although no state comes close to gender parity, there is significant variation in the magnitude of the gender gap: 28.7% of children who grow up to become inventors in Rhode Island are female, as compared with 11.3% in Idaho.³²

To examine the extent to which the heterogeneity in gender gaps in Figure IX can be explained by gender-specific differences in exposure to innovation, we first estimate gender-specific patent rates for workers in each CZ. We do so using our linked patent-tax sample instead of all patents in the USPTO data as above because gender is not observed in the USPTO data.³³ As a benchmark, Column 1 of Table V replicates the specification in column 1 of Table IV using this alternative measure of the CZ-level innovation rate. The raw magnitude of the coefficient differs because the mean and SD of the tax-data-based innovation rate differs from the USPTO-based measure. However, a 1 SD increase in the CZ innovation rate is associated with a 30.8% increase in children’s propensities to innovate, very similar to the 28.5% estimate obtained above.

In column 2 of Table V, we regress the fraction of females who go on to patent in each CZ on the innovation rates for women and men in that CZ. The coefficient on female innovation rate is significant and positive, while the coefficient on the male innovation rate is small and statistically insignificant. Symmetrically, column 3 shows that male innovation rates are more predictive of boys’ propensities to innovate than female innovation rates.³⁴

³¹We present the map at the state level because gender-specific patent rates are noisy in small CZs since there are so few female inventors in general.

³²The gender gap is generally smaller in states that score higher on Pope and Sydnor (2010)’s gender stereotype adherence index on standardized tests in 8th grade, which measures the extent to which children in a state adhere to the stereotype that boys are better at math and science while girls are better at English (population-weighted correlation = 0.21; Online Appendix Figure IV).

³³Specifically, we define the innovation rate for gender g in CZ j as the total number of patent applications filed by individuals of gender g born before 1980 in our full inventors sample divided by the number of individuals between ages 15 and 64 of gender g in CZ j in the 1990 Census. We convert this measure to an annual rate by dividing by 17, as we observe patent applications between 1996-2012. We restrict attention to inventors born before 1980 to avoid overlap with the intergenerational analysis sample that we use to study outcomes. Pooling genders, the population-weighted correlation across CZs between this measure of innovation rates and the baseline measure above is 0.65.

³⁴We find similar patterns at the individual level – daughters are more likely to become inventors if their mothers were inventors while sons are more likely to become inventors if their fathers are inventors – but the coefficients are imprecisely estimated because there are so few female inventors in the cohorts who are parents in our intergenerational sample.

One potential concern with the analysis in columns 2 and 3 is that women may have particularly strong tastes or abilities to innovate in certain categories (e.g., biology). This could lead to the gender-specific associations in columns 2 and 3 even in the absence of exposure effects if children live in the same areas as their parents and the supply of jobs varies across places. Columns 4 and 5 evaluate this possibility by examining variation in innovation rates across patent categories, using a specification with one observation per CZ by patent category and including category fixed effects, as in column 3 of Table IV.³⁵ We find very similar patterns in these specifications: women are more likely to innovate in a particular category if there were more women innovating in that category in the area where they grew up, showing that our baseline results are not driven by selection across categories. We reject the null hypothesis that the coefficients are the same for both genders with $p < 0.02$ in both of these specifications.

Table V further supports the hypothesis that exposure to innovation in childhood through one’s neighbors has a causal effect on children’s propensities to pursue innovation by providing an additional overidentification test of that hypothesis. In particular, the results in Table V imply that any confounding variable (such as ability) would have to vary not just across technology classes, but also in a gender-specific manner. Moreover, these findings suggest that the differences in rates of innovation across areas where children grow up are unlikely to be driven purely by the general factors such as schools or segregation emphasized in prior work on neighborhood effects, as such factors would be unlikely to generate impacts that vary so sharply by gender and technology class.

Implications for Gaps in Innovation. The estimates in column 2 of Table V imply that if girls were as exposed to female inventors in their childhood CZs as boys are to male inventors, female innovation rates would rise by 164% and the gender gap in innovation would fall by 55%. It is more difficult to gauge the extent to which differences in exposure account for the racial and income gaps in innovation for two reasons. First, we lack measures of exposure by race and parental income because we only observe race in the New York City school district sample and there are very few inventors who have low incomes. Second, because there is substantial segregation by race and income across neighborhoods (unlike by gender), broad measures of exposure (e.g., at the CZ or industry level) are unlikely to capture differences in exposure across racial and socioeconomic groups. The prevalence of segregation suggests that minorities and low-income children are much less exposed to inventors than their peers, but we defer quantification of the extent to which

³⁵Unfortunately, there is insufficient power to test this hypothesis by focusing on movers because there are relatively few female inventors.

exposure explains gaps in innovation along these dimensions to future work.

Our findings imply that increasing exposure to innovation among children who excel in math and science early in childhood – e.g., as measured by early childhood test scores – has the greatest potential to increase innovation in currently under-represented groups (low-income, minority, female). In particular, Figure IV shows that children who score below the 90th percentile on 3rd grade math tests are unlikely to become inventors even if they are from advantaged (high-income, non-minority, male) backgrounds. Since greater exposure is unlikely to raise the innovation rates of children from the under-represented groups beyond those observed for children with comparable test scores from advantaged backgrounds, there may be limited scope to increase innovation rates among children who score below the 90th percentile. In contrast, the large gaps in innovation rates by parental income, race, and gender among the highest-scoring children imply that there is substantial potential to increase innovation rates among children from under-represented groups who exhibit strong quantitative skills early in childhood.

IV.D Colleges

Stepping forward chronologically in studying children’s environments, we next examine how rates of innovation vary within and across colleges. Following Chetty et al. (2017), we assign each child in the intergenerational sample to the college he or she attends for the most years between ages 19-22 (if any). Figure Xa presents a list of the ten colleges (among colleges with at least 500 students per cohort) whose students have the highest rate of becoming inventors, as measured by having a patent application or grant between 1996-2014.³⁶ These colleges include highly selective institutions, such as the Massachusetts Institute of Technology – where more than 12% of students file patent applications by their early 30s – as well as smaller specialized colleges such as the Rochester Institute of Technology. These institutions account for a large share of future inventors in the U.S.: students from these ten colleges account for 7.9% of inventors, but only 0.5% of college enrollment in the U.S.

Figure Xb examines how innovation rates vary across students with different levels of parent income *within* these ten colleges. We construct this figure by dividing the parent income distribution at these colleges into 20 equal-sized bins (ventiles) and plotting the mean invention rate of children vs. mean parental income rank within each bin. There are fewer points on the left because selective

³⁶Innovation rates for every college in the U.S. that has at least 10 inventors in our sample are provided in Online Data Table III. The college-level estimates are blurred to protect confidentiality using the procedure in Chetty et al. (2017, Appendix C). The degree of error due to the blurring procedure is smaller than the degree of sampling error in the estimates.

colleges have relatively few low-income students (Chetty et al. 2017). Figure Xb shows that 7.1% of children with parents in the top 1% of the national income distribution become inventors at these colleges, compared with 4.0% of children from below-median-income families. This gap is an order of magnitude smaller than the 10 to 1 gap shown in Figure Ia for the nation as a whole.

Figure X suggests that liquidity constraints in financing innovation are unlikely to explain why low-income children innovate at lower rates, as such constraints would generate differences in rates of innovation even conditional on the college a child attends. Liquidity may not be critical because most inventors are not “garage” inventors; 70% of patents are filed by individuals working at companies with more than 100 employees. Similarly, differences in intrinsic preferences (such as tastes for risk) are unlikely to be responsible for the differences in rates of innovation across socioeconomic groups, as that explanation would also predict that low-income students within a given college would be less likely to become inventors than their higher-income peers.³⁷

More broadly, the preceding analysis supports the view that factors that affect children *before* they enter the labor market – such as childhood exposure to innovation – are key determinants of who becomes an inventor. In the next section, we assess the role of labor market conditions more directly by studying the career trajectories of inventors.

V Career Trajectories

We now examine inventors’ career paths and outcomes after they enter labor market. We first document a set of facts on the income distributions of inventors that shed light on the effects of financial incentives on innovation. We then show how financial returns and citations to patents vary with inventors’ characteristics at birth, which provides further information on the factors that drive differences in innovation rates across subgroups.

V.A *Financial Returns to Innovation*

Cross-Sectional Distribution of Permanent Income. We measure inventors’ permanent incomes by computing their average annual incomes between the ages of 40 and 50. Since our data on individual incomes begin in 1999 and end in 2012, we focus on individuals in our full inventors sample who are born between 1959 and 1962, for whom we see income at all ages between 40 and 50.³⁸ These

³⁷Of course, exposure to innovation might itself affect people’s preferences. In saying that “preferences” do not explain gaps in innovation, we refer specifically to determinants of preferences over careers that are orthogonal to a child’s environment and ability, such as a taste to pursue a less risky career.

³⁸We focus here on total individual income, which includes all sources of income (see Section II). Wage earnings comprise 95% of total income for the average inventor (conditional on having total income above \$1,000).

individuals applied for or were granted patents between ages 34 and 53, as our patent data span 1996-2012.

The income distribution of inventors, plotted in Online Appendix Figure V, is extremely skewed. The median annual income between ages 40-50 (in 2012 dollars) is \$114,000, the mean is \$192,000, and the 99th percentile is \$1.6 million. 22% of total income earned by inventors accrues to individuals in the top 1% of the inventors' income distribution, a top income share that is similar to the 23% top income share in the mid-2000s in the population as a whole (Atkinson et al. 2011). The degree of inequality *among* inventors is similar to that in the general population. In contrast, most other high-skilled professions, such as medicine or law, have much more homogeneous income distributions. One of the only other professions with comparable heterogeneity in income is the financial sector (Lockwood et al. 2017). Innovation thus differs from many other high-skilled occupations in that a small group of individuals obtain most of the returns.

Inventor's incomes reflect the private returns to innovation, which may differ from the social returns to innovation. Prior work has used the future citations to a patent as a measure of its scientific impact (e.g., Jaffe et al. 1993). Figure XI shows that the private returns to patents are highly correlated with their scientific impact, as measured by citations. It presents a binned scatter plot of average annual income between ages 40-50 vs. the total number of citations an inventor obtains. We restrict the sample to patent applications in 1996 in this figure to maximize the time horizon over which we can measure future citations. This plot is constructed by dividing citations into 21 bins and plotting mean income vs. mean citations within each bin. The first 19 bins include inventors in the first 19 ventiles (5% bins) of the citations distribution, while the last two bins plot the same relation for the 95th to 98th percentiles and the 99th percentile of the citation distribution. There is a strong positive relationship between citations and income. Notably, inventors who have patents in the top 1% of the citation distribution earn more than \$1 million *per year* between ages 40 and 50, showing that individuals with highly impactful innovations from a scientific perspective obtain large private returns over their lifetimes.

Lifecycle Profiles of Innovation and Income. Next, we turn to the dynamics of inventors' careers. Figure XIIa plots the cross-sectional age distribution of individuals who filed a patent application in 2000 that was subsequently granted. The modal age of patenting is 38, with symmetric declines at younger and older ages, consistent with Jones et al. (2014). This pattern is partly driven by the fact that the fraction of people who work falls at older ages. Figure XIIb plots the fraction of workers in the population (individuals with positive W-2 earnings) who patent in 2000 by age. Innovation

per worker still peaks around 40, but falls more gradually at older ages, with a 33% decline from age 40 to 60. Figure XIIc plots the fraction of workers whose patents went on to become highly-cited (in the top 5% of patents filed in 2000) by age. The rate of high-impact innovation falls by 66% from age 40 to 60. This result is broadly consistent with Acemoglu et al.’s (2014) hypothesis that the “young and restless” have higher impact discoveries, although individuals’ most impactful innovations tend to come in the middle rather than beginning of their professional careers.

With this background on ages of innovation, we now turn to the dynamics of income over inventors’ lifecycles. Figure XIIIa plots the median income of individuals who apply for a patent at age 30, 40, or 50. In each case, we see a steep increase in income in the years immediately preceding the patent application, following by a leveling off or decline. Figure XIIIb generalizes this analysis using an event study framework, defining year 0 as the year in which an individual files a patent application and other years relative to this reference year (e.g., +1 is the year after the application).³⁹ Consistent with the findings in Figure XIIIa, median and mean incomes rise sharply and peak at the point of patent application. A similar pattern is observed in the upper tail: the 99th percentile of the distribution peaks at \$1.8 million shortly after the year of application and falls slightly thereafter.

Figure XIIIc presents a breakdown of median income for three groups: unsuccessful applications (patents that were not granted before 2014), all granted patents, and highly-cited patents (those in the top 5% of the citation distribution among all patents granted in the same year). As noted above, individuals with highly-cited patents have higher incomes, and much of that higher income again comes from a much steeper earnings trajectory in the years prior to the point of patent application. The result in Figure XIII show that a patent application marks the peak of a successful career in innovation rather than an event that itself produces high returns, perhaps because the patent event itself is not news to the firm or the market. Indeed, patent royalties account for less than 3% of income even for inventors with highly cited patents five years after a patent is granted. Together, Figures XII and XIII suggest that the returns to innovation may be uncertain when individuals are deciding whether to pursue a career in innovation (e.g., at the point of college graduation), as the most impactful inventions tend to occur around age 40, and incomes tend to rise rapidly shortly before that point.

In summary, the private returns to innovation are highly skewed, correlated with their scientific

³⁹We limit the sample to individuals who file patent applications between ages 35 and 50. For individuals who file multiple patents in this age window, we choose one of the patents at random.

impact, and potentially uncertain at the time of career choice. In Section VI below, we show that these facts imply that changes in financial incentives are unlikely to have large effects on rates of innovation using a simple model of career choice.

V.B Earnings and Citations by Characteristics at Birth

In this subsection, we study how the returns to innovation, as measured by individuals’ incomes and patent citations, vary with inventors’ characteristics at birth. This analysis sheds light on the types of constraints that generate the differences in rates of innovation across subgroups documented in Section III above. In particular, prior work has modeled such constraints as differences in fixed costs of entry across subgroups. Most notably, Hsieh et al. (2016) argue that reductions in such entry costs for certain groups such as women and minorities could explain macroeconomic trends in the allocation of talent across sectors and economic growth.

If differences in innovation rates are due to barriers to entry, inventors from groups that face higher barriers to entry will have *higher* levels of productivity on average than inventors from more advantaged backgrounds, assuming that the distribution of ability does not vary across groups. Intuitively, higher barriers to entry screen out low-productivity inventors, so only “superstar” inventors make it through the pipeline.

We test this prediction in Figure XIVa by comparing the mean incomes of inventors with different characteristics at birth, who appear to have roughly similar distributions of ability at birth based on the analysis in Section III.B. The first pair of bars in Figure XIVa compares individuals from families with incomes above vs. below the 80th percentile of the parental income distribution using inventors in our intergenerational analysis sample. The second pair compares minorities (Blacks and Hispanics) and non-minorities using inventors in the New York City schools sample. The third pair compares males and females using the full inventors sample. In all cases, we reject the prediction of a simple barriers to entry model: inventors from the under-represented groups have similar or lower incomes on average. Figure XIVb replicates this analysis using the probability of having a highly-cited patent (in the top 5% of the distribution of citations among inventors in a given birth cohort) as the outcome. The patterns are analogous: inventors from under-represented groups also do not have higher-impact inventions.

The results in Figure XIV imply the probability that an individual becomes a superstar inventor is just as sensitive to his or her conditions at birth as the probability that he innovates at all, as shown in Figure 1b in the context of income. Hence, differences in the costs of entering the

innovation sector are unlikely to explain the gaps in innovation across subgroups. Instead, these results are consistent with the view that the key constraints at play are differences in exposure to innovation. In particular, a lack of exposure (e.g., awareness of innovation as a potential career) could reduce the probability that individuals pursue innovation at all levels of productivity.⁴⁰

Irrespective of the source of the gaps in innovation, the most important implication of Figure XIV is that there are many “lost Einsteins” – individuals who do not pursue a career in innovation even though they would have had highly impactful innovations had they done so. These results raise the possibility that the productivity costs of potential distortions in the allocation of talent may be even greater than predicted by models such as Hsieh et al. (2016), since some of the individuals who fail to pursue innovation could have been superstars. If the social planner’s goal is to increase innovation, the key question is what types of policy changes can bring these “lost Einsteins” into the innovation pipeline. In the next section, we use a simple model of career choice motivated by our empirical findings to explore this question.

VI A Model of Inventors’ Careers

In this section, we characterize the implications of our findings for policies to increase innovation by developing a static two-sector model of career choice that matches our empirical results. Using this model, we compare the effects of three types of policies to increase innovation: increasing private financial returns, reducing barriers to entry, and increasing exposure. Derivations and proofs of the propositions that follow are given in Online Appendix B.

VI.A Model Setup

A continuum of agents, indexed by i (with total mass one), choose to enter one of two sectors: the innovation sector (I) or another sector (\bar{I}). There are three factors that govern each agent’s choice of occupation: financial payoffs, barriers to entry, and exposure.

Financial Payoffs. Outside the innovation sector, agents receive a fixed wage $w_{\bar{I}}$. In the innovation sector, agents’ payoffs are determined by their innovation-specific abilities $\alpha_i \geq 1$, which follow a Pareto distribution $F_\alpha(\alpha) = 1 - (\frac{1}{\alpha})^{\beta_\alpha}$, and a stochastic shock $\pi_i \geq 1$ that is independently drawn from a Pareto distribution $F_\pi(\pi) = 1 - (\frac{1}{\pi})^{\beta_\pi}$. Agents know their ability α_i when deciding

⁴⁰While the results in Figure XIV are consistent with exposure effects, they could also be explained by alternative models, such as a model in which the same barriers that affect entry rates (e.g., discrimination) also affect individuals’ productivity after entering innovation. We discuss these issues further and compare the effects of exposure to innovation and barriers to entry formally using a model of career choice in Section VI below.

whether or not to enter the innovation sector but do not know π_i . Agent i 's realized payoff from entering the innovation sector is given by the product of ability and the stochastic shock:

$$r_i \equiv \alpha_i \cdot \frac{\beta_\alpha - 1}{\beta_\alpha} \cdot \pi_i \frac{\beta_\pi - 1}{\beta_\pi}.$$

With this specification, changes in the shape parameters β_π and β_α affect the skewness of payoffs while leaving the mean return $E[r_i]$ unchanged. We assume that wages and returns to innovation are fixed, and in particular, do not respond to the number of individuals who enter each sector.

Taxes and Barriers to Entry. Individuals must pay a tax τ on their incomes in the innovation sector, resulting in a net-of-tax payoff to innovation of $(1 - \tau)r_i$.⁴¹ This tax τ can alternatively be interpreted as a way to model barriers to entering innovation, as in Hsieh et al. (2016).⁴² Note that this model of barriers to entry rules out factors that may affect individuals' productivity even after they enter the innovation sector (e.g., discrimination). The results below on the effects of barriers to entry therefore pertain only to barriers that are isomorphic to monetary costs.

Exposure. Finally, individuals' decisions are influenced by whether they are exposed to innovation. We model exposure as a binary variable λ_i that follows a Bernoulli distribution $\lambda_i \sim B(\lambda)$. Individuals who do not receive exposure to innovation ($\lambda_i = 0$) never pursue innovation regardless of their ability, while those who receive exposure ($\lambda_i = 1$) choose their sector by maximizing expected lifetime utility.

Preferences and Agent Behavior. To obtain closed-form solutions, we assume that agents have constant relative risk aversion (CRRA) utility functions, although all of the results that follow hold with any smooth and concave utility function. Let $u(c_i) = \frac{c_i^{1-\theta}}{1-\theta}$ denote agent i 's utility as a function of his consumption c_i , with $\theta \geq 0$. Working in the innovation sector yields expected utility $V_i^I = E_\pi [u(r_i \cdot (1 - \tau_i))]$. Agent i therefore enters the innovation sector if $\lambda_i V_i^I > V_i^J = u(w_J)$. It is straightforward to show that agents follow a threshold rule when deciding to enter innovation: agents with innovation-specific ability $\alpha_i > \bar{\alpha}$ enter the innovation sector. Taking into account exposure effects, the share of agents who become inventors is therefore $\phi = \lambda \cdot (1 - F_a(\bar{\alpha}))$.

Given the evidence in Figure XI above that inventors' salaries are linearly related to their patent citations, we assume that the social value of innovation is proportional to the private return:

⁴¹We assume that the tax applies only to the innovation sector as a simple way to capture the fact that top income tax rates may affect the payoffs to innovation (which can sometimes be very high) more than payoffs to other careers that have lower (fixed) salaries. Insofar as taxes also affect payoffs in other sectors, career choices will be less sensitive to tax rates, reinforcing our results below.

⁴²In Hsieh et al.'s model, the barriers to entry τ vary across subgroups (e.g., women and minorities effectively face higher tax rates). Our model can be interpreted as applying to one such subgroup; the comparative static results below show how differences in τ affect innovation rates across subgroups.

$s_i = \nu \cdot r_i$, where $\nu \geq 0$. We define aggregate quality-weighted innovation as

$$\Phi = \phi E[\nu \cdot r_i | \alpha_i > \bar{\alpha}].$$

Intuitively, aggregate innovation depends upon the number of inventors (ϕ) and the average quality of their innovations. In the next two subsections, we characterize how changes in tax rates or barriers to entry (τ_i) and exposure (λ) affect ϕ and Φ .

VI.B Effects of Changes in Financial Returns or Barriers to Entry

In this subsection, we characterize the impact of changes in the tax rate τ , which we assume to be constant across agents for simplicity, on innovation. Note that a reduction in τ can be interpreted as an increase in the financial return to innovation or as a reduction in barriers to entry.

Proposition 1. *Reducing the tax rate (τ) increases the fraction of inventors (ϕ) and aggregate innovation (Φ). The magnitude of the response is characterized by three properties:*

1. *[Exposure dampening] The absolute impacts of changes in τ on ϕ and Φ are proportional to exposure λ .*

2. *[Forecastable returns] When returns to innovation vary purely because of heterogeneity in ability ($\beta_\pi \rightarrow \infty$), the elasticities of ϕ and Φ with respect to $(1 - \tau)$ converge to $\epsilon_{\phi, (1-\tau)} \rightarrow \beta_\alpha$ and $\epsilon_{\Phi, (1-\tau)} \rightarrow \beta_\alpha - 1$.*

3. *[Stochastic returns] As the skewness of stochastic returns to innovation increases ($\beta_\pi \rightarrow 1$), at a given initial level of innovation ϕ_0 , the elasticities of ϕ and Φ with respect to $(1 - \tau)$ both converge to zero if $\theta > 0$: $\epsilon_{\phi, (1-\tau)} \rightarrow 0$ and $\epsilon_{\Phi, (1-\tau)} \rightarrow 0$.*

The first result in Proposition 1 (exposure dampening) implies that the response of the number of inventors (and in turn of aggregate innovation) to changes in financial incentives is muted when exposure to innovation is low. Naturally, only the agents who are exposed to innovation respond to a change in τ . Given our empirical finding that rates of innovation are low in many subgroups of the population because of a lack of exposure, this result implies that tax cuts or changes in barriers to entry (e.g., changes in hiring practices) are unlikely to increase innovation among such groups.⁴³

The second result in Proposition 1 (forecastable returns) focuses on the case where heterogeneity in inventors' incomes is driven entirely by forecastable differences in abilities. In this case, the elasticity of aggregate innovation Φ with respect to changes in financial returns is determined purely

⁴³Although the absolute impacts of tax changes ($\frac{d\phi}{d\tau}$ and $\frac{d\Phi}{d\tau}$) are proportional to exposure λ , the elasticities of ϕ and Φ with respect to τ are invariant to λ . A lower value of λ reduces the rate of initial innovation at the same rate as the derivatives, leaving the elasticity (percentage impacts) unchanged.

by the skewness of the distribution of innovation abilities. The elasticity falls as the skewness of the ability distribution rises ($\beta_\alpha \rightarrow 1$) because there are fewer individuals who are on the margin of entering the innovation sector, whose decisions would be influenced by small tax changes. Moreover, aggregate quality-weighted innovation (Φ) is less responsive to changes in the tax rate τ than the number of agents entering innovation (ϕ), as shown by Jaimovich and Rebelo (2017).⁴⁴ Intuitively, the marginal entrants who are induced to enter the innovation sector because of a reduction in the tax rate τ must have lower ability than the average inventor already in the sector, thereby increasing quality-weighted innovation by less than the total number of inventors. As the ability distribution becomes more skewed ($\beta_\alpha \rightarrow 1$), the elasticity of quality-weighted innovation with respect to the tax rate converges to zero. In the limiting case, aggregate innovation is driven by a small fraction of superstar inventors whose behavior is insensitive to taxes because they have very high earnings in the innovation sector relative to the outside option.

Figure XVa illustrates this result by plotting the number of inventors ϕ and total quality-weighted innovation Φ as a function of the tax rate on inventors' earnings. In this simulation, we set $\beta_\pi = \infty$ (no stochastic shocks) and set the skewness of the ability distribution $\beta_\alpha = 1.26$ to match the Pareto shape parameter of 1.26 estimated using the inventors' empirical earnings distribution shown in Online Appendix Figure V. Both ϕ and Φ are normalized to 100% at a tax rate of $\tau = 0$. As predicted by the proposition, the number of inventors declines much faster than total innovation as the tax rate on inventors' earnings increases. For example, at a tax rate of $\tau = 40\%$ on inventors' earnings, the total number of inventors ϕ is 48% smaller than it would be in the absence of taxes ($\tau = 0$), but aggregate quality-weighted innovation Φ is only depressed by 12.5%. While the exact numbers naturally depend upon model specification, these calculations suggest that aggregate quality-weighted innovation may not be very sensitive to small changes in tax rates under parameters that match the empirical distribution of income.

The third result in Proposition 1 (stochastic returns) focuses on the case where heterogeneity in inventors' incomes is driven primarily by unforecastable shocks rather than ability heterogeneity, i.e. where $\beta_\pi \rightarrow 1$. The level of innovation ϕ converges to 0 as $\beta_\pi \rightarrow 1$ when $\theta > 0$ because the expected value of innovation V_I falls to 0 as the variance of payoffs grows large, holding the mean payoff fixed, when individuals are risk averse. To obtain comparative statistics at the same level of

⁴⁴Building on Jaimovich and Rebelo (2017), our contribution in the second part of Proposition 1 is to derive a formula that can be directly calibrated using the parameters and relationships we estimate empirically, namely the skewness of the income distribution of inventors and the linear relationship between the social returns to innovation (as measured by citations) and inventors' incomes.

innovation ϕ_0 as in the case with ability heterogeneity, we reduce the wage in the non-innovation sector $w_{\bar{I}}$ as $\beta_{\pi} \rightarrow 1$ to keep the fraction of inventors fixed at ϕ_0 .⁴⁵

In this setting, we find that as the skewness of stochastic shocks rises, both the elasticities of the number of inventors and quality-weighted innovation with respect to tax rates converge to zero if agents are risk averse ($\theta > 0$). The logic underlying this result is easiest to understand in the context of a limiting example with two states of the world: a bad state in which innovation has zero return and a good state in which innovation has a large payoff, say \$10 million. In the bad state, taxes have no impact on utility. In the good state, a slightly smaller payout (e.g., \$9 million instead of \$10 million) does not reduce an agent's incentive to become an inventor by very much because the marginal utility of money is already low. Intuitively, when returns are very skewed, taxes only affect inventor's payoffs when they are very deep in the money and are not sensitive to financial incentives, resulting in small behavioral responses. Put differently, when innovation has very risky payoffs, inventors must enter innovation partly because of its non-monetary benefits, making their behavior less sensitive to financial incentives.

Figure XVb illustrates this result by plotting innovation rates vs. taxes when the heterogeneity in inventors' incomes is driven primarily by stochastic shocks rather than differences in ability. We calibrate the model so that stochastic returns account for 90% of the skewness in inventors' earnings and the income distribution has a Pareto shape parameter of 1.26 as above.⁴⁶ We consider two cases: $\theta = 0$ (risk neutral agents, linear utility) and $\theta = 1$ (risk averse agents, log utility). With linear utility, taxes have very large effects: a tax rate of $\tau = 40\%$ reduces quality-weighted innovation Φ by 70.5% relative to the benchmark with no taxes. But when agents are risk averse, taxes have modest effects: Φ falls by only 9.4% from the no-tax benchmark when $\tau = 40\%$. Once again, these calculations suggest that tax changes have modest effects on aggregate innovation even when the returns to innovation are uncertain, under the standard assumption that individuals have diminishing marginal utility of consumption.

In sum, Proposition 1 implies that the decisions of individuals who contribute most to aggregate innovation are unlikely to be sensitive to small changes in financial incentives in standard models

⁴⁵Formally, we change $w_{\bar{I}}$ to $\kappa(\beta_{\pi}) \cdot w_{\bar{I}}$ as we vary β_{π} , choosing the scaling factor κ to keep the fraction of inventors at ϕ_0 , which one can interpret as a fixed (empirically observed) level of innovation. See Online Appendix B for further details.

⁴⁶Formally, we set β_{π} and β_{α} such that $s \equiv \left(\frac{\beta_{\pi}}{\beta_{\pi}-1} - 1\right) / \left(\frac{\beta_{\alpha}}{\beta_{\alpha}-1} - 1 + \frac{\beta_{\pi}}{\beta_{\pi}-1} - 1\right) = 0.9$ and the equilibrium income distribution has a shape parameter of 1.26. We retain 10% of skewness from variation in ability because the model is degenerate if we only allow for heterogeneity from stochastic shocks, since there is no source of heterogeneity other than ability in our model.

of career choice that match our empirical findings regarding the skewness of inventors' incomes and the correlation between private and social returns. We caution that this result relates only to the decisions of individual inventors, whose behavior is our focus in this paper. Taxes could potentially affect innovation and economic growth through many other channels, for instance by changing the behavior of firms, other salaried workers who contribute to the innovation process, or through general equilibrium effects.

VI.C Effects of Changes in Exposure

We now consider the impact of changes in exposure (λ) on rates of innovation.

Proposition 2. *The elasticities of the number of inventors and aggregate innovation with respect to exposure λ are both equal to one: $\epsilon_{\Phi,\lambda} = \epsilon_{\phi,\lambda} = 1$.*

Proposition 2 shows that the impacts of changing exposure are invariant to the skewness of the distribution of inventors' earnings or other parameters of the model. Intuitively, increasing exposure simply scales up the fraction of individuals who are already entering innovation proportionally. For instance, moving λ from 10% to 20% mechanically doubles the (randomly selected) set of individuals who are exposed to and enter the innovation sector, thereby doubling aggregate quality-weighted innovation as well.

Proposition 2 implies that there may be great potential to increase aggregate innovation by increasing exposure in a subgroup g that currently has few inventors, *if* the low rate of innovation ϕ_g in that group is due to a lack of exposure (in which case there is scope to increase λ_g) rather than high barriers to entry τ_g . As discussed in Section V.B, one way to determine whether the low innovation rate is driven by exposure effects or barriers to entry is by examining the average quality of inventions for inventors in that subgroup. The following corollary formalizes this result.

Corollary 1. *If the distribution of innovation abilities is constant across groups, differences in the average quality of inventions reveal whether differences in innovation rates arise from barriers to entry or exposure effects.*

1. *[Barriers to entry] Groups that face higher barriers to entry τ have higher-quality inventions conditional on inventing: ϕ declines with τ , while $E[\nu \cdot r_i | \alpha_i > \bar{\alpha}]$ is increasing in τ .*
2. *[Exposure] Groups that have less exposure to innovation λ have the same quality of inventions conditional on inventing: ϕ declines with λ , but $E[\nu \cdot r_i | \alpha_i > \bar{\alpha}]$ does not vary with λ .*

The first result in this corollary follows from the logic in the second part of Proposition 1. The marginal inventor who is screened out in groups with higher barriers to entry tends to be of lower-quality on average, implying that the remaining inventors have higher quality patents on average. The second result follows from the logic underlying Proposition 2. Since an increase in exposure simply draws in randomly selected inventors, groups that have less exposure do not have inventors of different quality on average, provided that the ability to innovate does not vary across groups.

Empirically, individuals from under-represented groups who become inventors do not have higher-quality patents (Figure XIVb) despite having relatively similar innate abilities (Section III.B). Based on Corollary 1, these findings are consistent with the view that the low rates of innovation among women and disadvantaged youth are driven primarily by a lack of exposure rather than fixed costs of entry that screen out inventors of marginal quality, as in Hsieh et al. (2016).

In sum, the results above imply that there is substantial scope to increase aggregate innovation by increasing exposure among under-represented subgroups. While quantifying the effects of specific policies that increase exposure is outside the scope of this study, the estimates in Section IV suggest that changes in exposure could have substantial effects. For instance, we estimate that a one standard deviation change in exposure to innovation at the neighborhood (CZ) level would increase the number of inventors by 10-30% (Table IV). The fact that some neighborhoods in America induce many more children to become inventors demonstrates that it is feasible to design childhood environments that could significantly increase aggregate innovation. How exactly one can replicate the impacts of such environments in a cost-effective manner is a key question that we leave to future work.

VII Conclusion

This paper has presented new evidence on the factors that determine who becomes an inventor by tracking the lives of inventors in America from birth to adulthood. Most previous work on innovation has focused on factors such as financial incentives, barriers to entry, and STEM education. Our results point to a different channel – exposure to innovation during childhood – as a critical factor that determines who becomes an inventor. A lack of exposure to innovation can help explain why high-ability children in low-income families, minorities, and women are significantly less likely to become inventors. Importantly, such lack of exposure screens out not just marginal inventors but the “Einsteins” who produce innovations that have the greatest impacts on society. Policies that increase exposure therefore have the capacity to greatly increase quality-weighted aggregate

innovation. In contrast, changes in financial incentives (e.g., via tax cuts) are less likely to spur additional star inventors to enter the field because the private financial returns to high-impact innovations are already quite large.

Policies to increase exposure to innovation could range from mentoring by current inventors to internship programs and changes in peer networks. Our analysis does not provide guidance on which specific programs are most effective, but it does provide some guidance on how they should be targeted. In particular, targeting exposure programs to children who (a) excel in math and science at early ages and (b) are from under-represented groups may maximize their impacts on innovation, since the gaps in innovation rates by parental income, race, and gender are largest among children who have very high (top 5%) test scores in early grades. Furthermore, tailoring programs to participants' backgrounds may increase their impact; for example, our findings suggest that women are more influenced by female inventors rather than male inventors.

Beyond the literature on innovation, our findings contribute to the growing literature on how children's prospects for success are shaped by their environments. Prior studies have focused primarily on general human capital accumulation as the mechanism through which neighborhoods and schools affect outcomes. Our analysis suggests that environment matters through much narrower channels as well, for instance by influencing the specific career pathways that children choose to pursue, either via transmission of specific human capital or through changes in aspirations. Such mechanisms call for a different class of interventions than traditional investments in schools or neighborhoods, such as programs or networks that provide children exposure to specific careers that may be a good match for their talents.

More broadly, our findings suggest that there is not necessarily a strict trade-off between reducing inequality and increasing economic growth. Drawing more low-income and minority children into science and innovation could increase their incomes – thereby reducing the persistence of inequality across generations – while stimulating growth by harnessing currently under-utilized talent. Developing and testing methods to increase exposure to innovation among under-represented subgroups is therefore a particularly promising direction for research and policy.

References

- Abrams, D., U. Akcigit, and J. Popadak (2013). Patent Value and Citations: Creative Destruction or Strategic Disruption? *Faculty Scholarship*.
- Acemoglu, D., U. Akcigit, and M. Celik (2014). Young, Restless and Creative: Openness to Disruption and Creative Innovations. *NBER Working Paper No. 19894*.
- Aghion, P., U. Akcigit, A. Hyttinen, and O. Toivanen (2017). The Social Origins of Inventors. *mimeo*.
- Aghion, P. and P. Howitt (1992). A Model of Growth Through Creative Destruction. *Econometrica* 60(2), 323–351.
- Akcigit, U., S. Baslandze, and S. Stantcheva (2016). Taxation and the International Mobility of Inventors. *The American Economic Review* 106(10), 2930–2981.
- Akcigit, U., J. Grigsby, and T. Nicholas (2017). The Birth of American Ingenuity: Innovation and Invention in the Golden Age. *NBER Working Paper No. 23047*.
- Akcigit, U., D. Hanley, and S. Stantcheva (2017). Optimal Taxation and R&D Policy. *NBER Working Paper No. 22908*.
- Atkinson, A. B., T. Piketty, and E. Saez (2011). Top Incomes in the Long Run of History. *Journal of Economic Literature* 49(1), 3–71.
- Azoulay, P., J. Graff Zivin, and G. Manso (2011). Incentives and Creativity: Evidence from the Academic Life Sciences. *RAND Journal of Economics* 42(3), 527–554.
- Becker, B. (2015). Public R&D Policies and Private R&D Investment: A Survey of the Empirical Evidence. *Journal of Economic Surveys* 29(5), 917–942.
- Bloom, N., M. Schankerman, and J. Van Reenen (2013). Technology Spillovers and Product Market Rivalry. *Econometrica* 81(4), 1347–1393.
- Card, D. and L. Giulano (2014). Does Gifted Education Work? For Which Students? *NBER Working Paper No. 20453*.
- Celik, M. (2014). Does the Cream Always Rise to the Top? The Missallocation of Talent and Innovation. *Working Paper*.
- Chetty, R., J. Friedman, E. Saez, N. Turner, and D. Yagan (2017). Mobility Report Cards: The Role of Colleges in Intergenerational Mobility. *NBER Working Paper No. 23618*.
- Chetty, R., J. N. Friedman, N. Hilger, E. Saez, D. W. Schanzenbach, and D. Yagan (2011). How Does your Kindergarten Classroom Affect your Earnings? Evidence from Project STAR. *The Quarterly Journal of Economics* 126(4), 1593–1660.
- Chetty, R., J. N. Friedman, and J. E. Rockoff (2014a). Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates. *American Economic Review* 104(9), 2593–2632.
- Chetty, R. and N. Hendren (2017). The Effects of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects. *NBER Working Paper No. 23001*.

- Chetty, R., N. Hendren, and L. Katz (2016). The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment. *American Economic Review* 106(4), 855–902.
- Chetty, R., N. Hendren, P. Kline, and E. Saez (2014b). Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States. *The Quarterly Journal of Economics* 129(4), 1553–1623.
- Cook, L. and C. Kongcharoen (2010). The Idea Gap in Pink and Black. *NBER Working Paper No. 16331*.
- Depalo, D. and S. Di Addario (2015). Inventors’ Returns to Patents. *Bank of Italy mimeo*.
- DiNardo, J., N. M. Fortin, and T. Lemieux (1996). Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach. *Econometrica* 64(5), 1001–1044.
- Ding, W., F. Murray, and T. Stuart (2006). Gender Differences in Patenting in the Academic Life Sciences. *Science* 313(5787), 665–667.
- Dorner, M., D. Harhoff, and K. Hoisl (2014). Patterns and Determinants of Inventor Mobility – Evidence from the Employment Biographies of Inventors in Germany. *mimeo*.
- for Economic Co-operation, O. and Development (2009). *OECD Patent Statistics Manual 2009*. OECD.
- Fryer, R. and S. Levitt (2004). Understanding the Black-White Test Score Gap in the First Two Years of School. *The Review of Economics and Statistics* 86(2), 447–464.
- Fryer, R. and S. Levitt (2006). The Black-White Test Score Gap through Third Grade. *American Law and Economics Review* 8(2), 249–281.
- Fryer, R. G. (2011). Racial inequality in the 21st century: The declining significance of discrimination. In *Handbook of labor economics*, Volume 4, pp. 855–971. Elsevier.
- Giuri, P., M. Mariani, S. Brusoni, G. Crespi, D. Francoz, A. Gambardella, W. Garcia-Fontes, and A. Geuna (2007). Inventors and Invention Processes in Europe: Results from the PatVal-EU Survey. *Research Policy* 36(8), 1107–1127.
- Goolsbee, A. (1998). Does Government R&D Policy Mainly Benefit Scientists and Engineers? *American Economic Review* 88(2), 298–302.
- Griliches, Z. (1990). Patent Statistics as Economic Indicators: A Survey. *Journal of Economic Literature* 28(4), 1661–1707.
- Hall, B., A. Jaffe, and M. Trajtenberg (2001). The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools. *NBER Working Paper No. 8498*.
- Hall, B. H., A. Jaffe, and M. Trajtenberg (2005). Market Value and Patent Citations. *RAND Journal of Economics*, 16–38.
- Hall, D. T. (2002). *Careers In and Out of Organizations*. Sage Publications.
- Heckman, J. (2006). Skill Formation and the Economics of Investing in Disadvantaged Children. *Science* 312(5782), 1900–1902.

- Hsieh, C., E. Hurst, C. Jones, and P. Klenow (2016). The Allocation of Talent and U.S Economic Growth. *NBER Working Paper No.18693*.
- Hunt, J. (2009). Why Don't Women Patent? *NBER Working Paper No. 17888*.
- Jaffe, A. B., M. Trajtenberg, and R. Henderson (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics* 108(3), 577–598.
- Jaimovich, N. and S. Rebelo (2017). Non-linear Effects of Taxation on Growth. *Journal of Political Economy* 125(1), 265–291.
- Jones, B. (2010). Age and Great Invention. *The Review of Economics and Statistics* 92(1), 1–14.
- Jones, B., E. J. Reedy, and B. Weinberg (2014). Age and Scientific Genius. *NBER Working Paper No. 19866*.
- Jones, C. and J. Williams (1999). Too Much of a Good Thing? *Journal of Economic Growth* 5(1), 65–85.
- Jung, T. and O. Ejeremo (2014). Demographic Patterns and Trends in Patenting: Gender, Age and Education of Inventors. *Technological Forecasting and Social Change* 86, 110–124.
- Kahn, S. and D. Ginther (2017). Women and STEM. *NBER Working Paper No. 19894*.
- Kim, J. and G. Marschke (2005). Labor Mobility of Scientists, Technological Diffusion, and the Firm's Patenting Decision. *RAND Journal of Economics* 36(2), 298–317.
- Kline, P., N. Petkova, H. Williams, and O. Zidar (2017). Who Profits from Patents? Rent-sharing at Innovative Firms. *IRLE Working Paper*.
- Kogan, L., D. Papanikolaou, A. Seru, and N. Stoffman (2017). Technological innovation, resource allocation, and growth. *The Quarterly Journal of Economics* 132(2), 665–712.
- Laband, D. and B. Lentz (1983). Like Father, like Son: Toward an Economic Theory of Occupational Following. *Southern Economic Journal* 50(2), 474–493.
- Lentz, B. and D. Laband (1989). Why So Many Children of Doctors Become Doctors: Nepotism vs. Human Capital Transfers. *Journal of Human Resources* 24(3), 396–413.
- Lerner, J. and J. Wulf (2007). Innovation and Incentives: Evidence from Corporate R&D. *The Review of Economics and Statistics* 89(4), 634–644.
- Lindquist, M., J. Sol, and M. Van Praag (2015). Why Do Entrepreneurial Parents Have Entrepreneurial Children? *Journal of Labor Economics* 33(2), 269–296.
- Lockwood, B. B., C. G. Nathanson, and E. G. Weyl (2017). Taxation and the Allocation of Talent. *Journal of Political Economy* 125(5), 1635–1682.
- Moretti, E. and D. J. Wilson (2017). The Effect of State Taxes on the Geographical Location of Top Earners: Evidence from Star Scientists. *American Economic Review* 107(7), 1858–1903.
- Murphy, K., A. Shleifer, and R. Vishny (1991). The Allocation of Talent: Implications for Growth. *The Quarterly Journal of Economics* 106(2), 503–530.

- Pope, D. and J. Sydnor (2010). Geographic Variation in the Gender Differences in Test Scores. *Journal of Economic Perspectives* 24(2), 95–108.
- Porter, M. E. and S. Stern (2001). Innovation: Location Matters. *MIT Sloan Management Review* 42(4), 28–36.
- Romer, P. (2000). Should the Government Subsidize Supply or Demand in the Market for Scientists and Engineers? *NBER Working Paper No. 7723*.
- Romer, P. M. (1990). Endogenous Technological Change. *Journal of political Economy* 98(5, Part 2), S71–S102.
- Roy, A. (1951). Some Thoughts on the Distribution of Earnings. *Oxford Economic Papers* 3(2), 135–146.
- Schere, F., D. Harhoff, and J. Kukies (2000). Uncertainty and the Size Distribution of Rewards from Innovation. *Journal of Evolutionary Economics* 10(1), 175–200.
- Schmookler, J. (1957). Inventors Past and Present. *The Review of Economics and Statistics* 39(3), 321–333.
- Strumsky, D. (2014). Strumsky Patent Database (Available upon Request). Downloaded in December 2014.
- Thursby, M. and J. Thursby (2005). Gender Patterns of Research and Licensing Activity of Science and Engineering Faculty. *Journal of Technology Transfer* 30(4), 343–353.
- Toivanen, O. and L. Vaananen (2012). Returns to Inventors. *The Review of Economics and Statistics* 94(4), 1173–1190.
- Topel, R. and M. Ward (1992). Job Mobility and the Careers of Young Men. *Quarterly Journal of Economics* 107(2), 439–479.
- Van Reenen, J. (1996). The Creation and Capture of Rents: Wages and Innovation in a Panel of U.K. Companies. *Quarterly Journal of Economics* 111(1), 195–226.

ONLINE APPENDICES

Appendix A. Matching Algorithm

The patent data were linked to the tax records using a variant of the matching algorithm developed in Chetty et al. (2014a) to link the New York City school district data to the tax records. Chetty et al. (2011) show that the match algorithm outlined below yields accurate matches for approximately 99% of cases in a school district sample that can be exactly matched on social security number. Note that identifiers were used solely for the matching procedure. After the match was completed, the data were de-identified (i.e., individual identifiers such as names were stripped) and the statistical analysis was conducted using the de-identified dataset.

Before beginning the match process, the names were standardized as follows. First, suffixes sometimes appear at the end of taxpayers' first, middle, or last name fields. If these fields end with a space followed by "JR", "SR", or a numeral I-IV, the suffix is stripped out and stored separately from the name. Second, the USPTO database separates inventor names into "first" and "last," but the tax data often separates names into first, middle, and last. In practice, many inventors include a middle initial or name in the first name field. Whenever there is a single space in the inventor's first name field, for the purposes of matching, we allow the first string to be an imputed first name, and the second string to be an imputed middle name or initial. The use of these imputed names is described below.

The matching algorithm proceeds in seven steps. Inventors enter a match round only if they have not already been matched to a taxpayer in an earlier round. Each round consists of a name criterion and a location criterion. The share of data matched in each round is documented below.

- **Stage 1:** Exact match on name and location.
 - Name match: The inventor's last name exactly matches the taxpayer's last name. Either the inventor's first name field exactly matches the concatenation of the IRS first and middle name fields or the IRS middle name field is missing, but the first name fields match. If an imputed middle name is available for the inventor, candidate matches are removed if they have ever filed at the IRS with a middle name or initial that conflicts with the inventor's.
 - Location match: The inventor's city and state must match some city and state reported by that taxpayer exactly.
 - 49% of patents are uniquely matched in this stage.
- **Stage 2:** Exact match on imputed name data and location.
 - Name match: The inventor's last name exactly matches the taxpayer's last name and the taxpayer's last name is the same as the inventor's imputed first name. Either the inventor's imputed middle name/initial matches one of the taxpayer's middle/initial name fields, or one of the two is missing. For inventors with non-missing imputed middle names, priority is given to matches to correct taxpayer middle names rather than to taxpayers with missing middle names. As above, candidate matches are removed if they have ever filed at the IRS with a conflicting middle name or initial.
 - Location match: As above, the inventor's city and state must exactly match some city and state reported by that taxpayer.
 - 12% of patents are uniquely matched in this stage.
- **Stage 3:** Exact match on actual or imputed name data and 1040 zip crosswalked.

- Name match: The inventor’s last name exactly matches the taxpayer’s last name. The inventor’s first name matches the taxpayer’s first name in one of the following situations, in order of priority: (1) inventor’s first name is the same as the taxpayer’s combined first and middle name; (2) inventor’s imputed first name matches taxpayer’s and middle names match on initials; (3) inventor has no middle name data, but inventor’s first name is the same as the taxpayer’s middle name.
 - Taxpayers are removed if they are ever observed filing with middle names in conflict with the inventor’s.
 - Location match: The inventor’s city and state match one of the city/state fields associated with one of the taxpayer’s 1040 zip codes.
 - Location match: The inventor’s city and state match one of the city/state fields associated with one of the taxpayer’s 1040 zip codes.
 - 3% of patents are uniquely matched in this stage.
- **Stage 4:** Same as previous stage, but using names from 1040 forms instead of names from W-2 forms.
 - Name match: The inventor’s name matches the name of a 1040 (or matches without inventor’s middle initial/name and no taxpayer middle initials/names conflict with inventor’s).
 - Location match: The inventor’s city and state must match some city and state reported by that taxpayer exactly.
 - 6% of patents are uniquely matched in this stage.
- **Stage 5:** Match using W-2 full name field.
 - Name match: The inventor’s FULL name exactly matches the FULL name of a taxpayer on a W2.
 - Location match: The inventor’s city and state match one of the city/state fields associated with one of the taxpayer’s 1040 zip codes.
 - 8% of patents are uniquely matched in this stage.
- **Stage 6:** Fuzzy match using W-2 full name field.
 - Name match: The inventor’s full name (minus the imputed middle name) exactly matches the full name of a taxpayer on a W2.
 - Location match: The inventor’s city and state match one of the city/state fields associated with one of the taxpayer’s 1040 zip codes.
 - 1% of patents are uniquely matched in this stage.
- **Stage 7:** Match to all information returns.
 - Name match: The inventor’s full name exactly matches the full name of a taxpayer on any type of information return form.
 - Location match: The inventor’s city and state match one of the city/state fields associated with one of the taxpayer’s information return forms.
 - 6% of patents are uniquely matched in this stage.

Appendix B. Model of Inventors' Careers: Proofs

In this appendix, we first present analytical formulas for the key expressions in our model, then describe the comparative statics of interest, and finally present proofs of the propositions in Section VI.

Analytical Formulas. Expected utility in the innovation sector for agent i is given by:

$$V_i^I = \int_1^\infty \frac{((1-\tau) \cdot \frac{\beta_\pi-1}{\beta_\pi} x \cdot \frac{\beta_\alpha-1}{\beta_\alpha} \alpha_i)^{1-\theta}}{1-\theta} dF_\pi(x) = \frac{\beta_\pi}{\beta_\pi + \theta - 1} \left(\frac{\beta_\pi - 1}{\beta_\pi} \right)^{1-\theta} \frac{((1-\tau) \cdot \frac{\beta_\alpha-1}{\beta_\alpha} \alpha_i)^{1-\theta}}{1-\theta}$$

Since $\frac{\partial V_i^I}{\partial \alpha_i} > 0$ and the outside wage is fixed, there is an ability cutoff beyond which all agents enter the innovation sector. This cutoff is characterized by:

$$\begin{aligned} V^{\bar{I}} &= V_{\bar{\alpha}}^I \\ \Rightarrow \frac{(w_{\bar{I}})^{1-\theta}}{1-\theta} &= \frac{\beta_\pi}{\beta_\pi + \theta - 1} \left(\frac{\beta_\pi - 1}{\beta_\pi} \right)^{1-\theta} \frac{((1-\tau) \cdot \frac{\beta_\alpha-1}{\beta_\alpha} \bar{\alpha})^{1-\theta}}{1-\theta} \\ \Rightarrow \bar{\alpha} &= \frac{w_{\bar{I}}}{1-\tau} \frac{\beta_\pi}{\beta_\pi - 1} \frac{\beta_\alpha}{\beta_\alpha - 1} \left(\frac{\beta_\pi + \theta - 1}{\beta_\pi} \right)^{\frac{1}{1-\theta}} \end{aligned}$$

It follows that the fraction of agents entering the innovation sector is:

$$\phi = \lambda \cdot (1 - F_\alpha(\bar{\alpha})) = \lambda \cdot \left(\frac{1-\tau}{w_{\bar{I}}} \right)^{\beta_\alpha} \left(\frac{\beta_\pi - 1}{\beta_\pi} \frac{\beta_\alpha - 1}{\beta_\alpha} \right)^{\beta_\alpha} \left(\frac{\beta_\pi}{\beta_\pi + \theta - 1} \right)^{\frac{\beta_\alpha}{1-\theta}}. \quad (4)$$

Aggregate innovation is given by

$$\begin{aligned} \Phi &= \nu \frac{\beta_\pi - 1}{\beta_\pi} \frac{\beta_\alpha - 1}{\beta_\alpha} \lambda \int_{\bar{\alpha}}^\infty x dF_\alpha(x) \int_1^\infty y dF_\pi(y) \\ &= \lambda \cdot \nu \left(\frac{\beta_\pi - 1}{\beta_\pi} \frac{\beta_\alpha - 1}{\beta_\alpha} \right)^{\beta_\alpha - 1} \left(\frac{1-\tau}{w_{\bar{I}}} \right)^{\beta_\alpha - 1} \left(\frac{\beta_\pi}{\beta_\pi + \theta - 1} \right)^{\frac{\beta_\alpha - 1}{1-\theta}} \end{aligned} \quad (5)$$

The expected quality of innovations conditional on inventing is:

$$E[\nu \cdot r_i | \alpha_i > \bar{\alpha}] = \frac{\Phi}{\phi} = \nu \frac{\beta_\alpha}{\beta_\alpha - 1} \frac{\beta_\pi}{\beta_\pi - 1} \frac{w_{\bar{I}}}{1-\tau} \left(\frac{\beta_\pi + \theta - 1}{\beta_\pi} \right)^{\frac{1}{1-\theta}} \quad (6)$$

Comparative Statics. Our goal is to compute elasticities of innovation with respect to tax rates in two scenarios, holding fixed the fraction of inventors at a given (empirically observed) level: (a) the case where $\beta_\pi \rightarrow \infty$ (i.e., there are no stochastic shocks) and (b) the case where $\beta_\pi \rightarrow 1$ (i.e., the skewness of stochastic shocks grows arbitrarily large).

In the first case, we simply compute the elasticities of ϕ and Φ with respect to the net-of-tax rate $1 - \tau$ around the level of innovation ϕ_0 that prevails when $\beta_\pi \rightarrow \infty$. Using equations (4) and (5), these elasticities are:

$$\epsilon_{\phi, (1-\tau)} = \frac{d\phi}{d(1-\tau)} \frac{1-\tau}{\phi} = \beta_\alpha, \quad (7)$$

$$\epsilon_{\Phi,(1-\tau)} = \frac{d\Phi}{d(1-\tau)} \frac{1-\tau}{\Phi} = \beta_\alpha - 1. \quad (8)$$

In the second case, the level of innovation ϕ converges to 0 as $\beta_\pi \rightarrow 1$ when $\theta > 0$ because the expected value of innovation V_I falls to 0 as the variance of payoffs grows large holding the mean payoff fixed. To obtain comparative statistics that are comparable to the first case, we hold the fraction of inventors fixed at ϕ_0 (the same level as in the first case) by varying the wage in the non-innovation sector $w_{\bar{I}}$ as $\beta_\pi \rightarrow 1$. In particular, we change $w_{\bar{I}}$ to $\kappa(\beta_\pi) \cdot w_{\bar{I}}$ as we vary β_π , choosing the scaling factor κ to keep the fraction of inventors at ϕ_0 as a function of β_π . Formally, for a given change in skewness from a reference level β_π^B to the level of interest β_π , $\kappa(\beta_\pi)$ is chosen such that the threshold to enter innovation $\bar{\alpha}(\beta_\pi, \kappa) = \bar{\alpha}(\beta_\pi^B)$, i.e.

$$\kappa(\beta_\pi) = \frac{\beta_\pi - 1}{\beta_\pi} \frac{\beta_\pi^B}{\beta_\pi^B - 1} \left(\frac{\beta_\pi}{\beta_\pi + \theta - 1} \frac{\beta_\pi^B + \theta - 1}{\beta_\pi^B} \right)^{\frac{1}{1-\theta}}. \quad (9)$$

At any given level of β_π , the elasticity of innovation with respect to the net-of-tax rate around the original fraction of inventors ϕ_0 is:

$$\epsilon_{\phi,(1-\tau)} = \frac{d\phi}{d(1-\tau)} \frac{1-\tau}{\phi \cdot \left(\frac{1}{\kappa}\right)^{\beta_\alpha}} = \beta_\alpha \cdot \kappa(\beta_\pi)^{\beta_\alpha}.$$

When the reference level of skewness $\beta_\pi^B \rightarrow \infty$ (i.e., when $w_{\bar{I}}$ is adjusted to hold the fraction of inventors fixed at ϕ_0), the elasticity of innovation w.r.t. $1-\tau$ is:

$$\epsilon_{\phi,(1-\tau)} = \beta_\alpha \left(\frac{\beta_\pi - 1}{\beta_\pi} \right)^{\beta_\alpha} \left(\frac{\beta_\pi}{\beta_\pi + \theta - 1} \right)^{\frac{\beta_\alpha}{1-\theta}} \quad (10)$$

Likewise, the elasticity of aggregate innovation (Φ) w.r.t. the net of tax rate ($1-\tau$) is:

$$\epsilon_{\Phi,(1-\tau)} = \frac{d\Phi}{d(1-\tau)} \frac{1-\tau}{\Phi \cdot (\kappa)^{1-\beta_\alpha}} = (\beta_\alpha - 1) \left(\frac{\beta_\pi - 1}{\beta_\pi} \right)^{\beta_\alpha - 1} \left(\frac{\beta_\pi}{\beta_\pi + \theta - 1} \right)^{\frac{\beta_\alpha - 1}{1-\theta}} \quad (11)$$

Propositions. With these expressions in hand, it is straightforward to establish the propositions and corollaries in Section VI.

Proof of Proposition 1. (4) and (9) imply $\frac{d\phi}{d(1-\tau)} = \lambda \cdot \frac{\beta_\alpha(1-F_\alpha(\bar{\alpha}_0)) \left(\frac{\beta_\pi}{\beta_\pi-1}\right)^{\beta_\alpha} \left(\frac{\beta_\pi+\theta-1}{\beta_\pi}\right)^{\frac{\beta_\alpha}{1-\theta}}}{1-\tau}$; (5)

and (9) imply $\frac{d\Phi}{d(1-\tau)} = \lambda \cdot \frac{(\beta_\alpha-1)(1-F_\alpha(\bar{\alpha}_0)) \frac{\beta_\alpha-1}{\beta_\alpha} \left(\frac{\beta_\pi}{\beta_\pi-1}\right)^{\beta_\alpha-1} \left(\frac{\beta_\pi+\theta-1}{\beta_\pi}\right)^{\frac{\beta_\alpha-1}{1-\theta}}}{1-\tau}$; as $\beta_\pi \rightarrow \infty$, (7) and (8) establish that $\epsilon_{\phi,(1-\tau)} \rightarrow \beta_\alpha$ and $\epsilon_{\Phi,(1-\tau)} \rightarrow \beta_\alpha - 1$; as $\beta_\pi \rightarrow 1$ with $\theta > 0$, (10) and (11) imply that $\epsilon_{\phi,(1-\tau)} \rightarrow 0$ and $\epsilon_{\Phi,(1-\tau)} \rightarrow 0$.

Proof of Proposition 2. (4) implies $\epsilon_{\phi,\lambda} = 1$ and (5) implies $\epsilon_{\Phi,\lambda} = 1$.

Proof of Corollary 1. (6) implies that $E[\nu \cdot r_i | \alpha_i > \bar{\alpha}]$ is increasing in τ and (4) implies that ϕ is declining with τ ; (6) implies that $E[\nu \cdot r_i | \alpha_i > \bar{\alpha}]$ does not vary with λ and (4) implies that ϕ is declining with λ .

TABLE I
Summary Statistics

Sample:		Full	Intergenerational		New York City School District	
		Inventors (1)	Inventors (2)	Non-inventors (3)	Inventors (4)	Non-inventors (5)
<i>Patenting Outcomes</i>						
Patent Grants	Mean	3.0	1.4		1.3	
	Median	1.0	1.0		1.0	
	Std. Dev.	6.5	2.7		2.0	
Patent Applications	Mean	3.2	2.2		2.1	
	Median	1.0	1.0		1.0	
	Std. Dev.	11.1	4.3		3.4	
Patent Citations	Mean	26.2	1.2		1.3	
	Median	1.0	0.0		0.0	
	Std. Dev.	118.1	12.3		8.7	
Age at Application	Mean	43.7	27.5		27.7	
	Std. Dev.	11.5	2.3		2.7	
<i>Income in 2012</i>						
Individual Wage Earnings (\$)	Mean	111,457	82,902		94,622	
	Median	83,000	72,000		74,000	
	Std. Dev.	140,463	91,909		127,712	
Total Individual Income (\$)	Mean	147,360	92,665		104,136	
	Median	100,000	74,000		75,000	
	Std. Dev.	185,171	114,327		147,924	
Parent Household Income (\$)	Mean		183,303	85,992	108,049	47,509
	Median		109,000	59,000	66,000	33,000
	Std. Dev.		662,669	336,387	208,251	81,607
Attended College at Age 20			86.0%	47.7%		
<i>Test Scores</i>						
3rd Grade Mean Math Score					1.0	0.1
3rd Grade Mean English Score					0.8	1.0
8th Grade Mean Math Score					1.3	0.2
8th Grade Mean English Score					1.0	0.9
<i>Demographics</i>						
Female Share		13.1%	18.5%	49.8%	21.9%	48.8%
White Non-Hispanic Share					44.9%	19.5%
Black Non-Hispanic Share					17.3%	36.0%
Hispanic Share					8.4%	33.7%
Asian Share					27.4%	9.6%
Sample Size		1,200,689	34,973	16,360,910	452	433,863

Notes: This table presents summary statistics for the three samples of inventors and corresponding samples of non-inventors used in the empirical analysis. We define individuals as inventors if they were listed as an inventor on a patent application between 2001-2012 or grant between 1996-2014. The full inventors sample (Column 1) includes all inventors who were linked to the tax data using the procedure described in Online Appendix A. The intergenerational sample consists of U.S. citizens born in 1980-1984 matched to their parents in the tax data (Columns 2 and 3). The New York City School District sample includes children in the 1979-1985 birth cohorts who attended New York City public schools at some point between grades 3-8 and were linked to the tax data (Columns 4 and 5). Citations are measured as total patent citations between 1996-2014. For individuals with more than one patent application, age at application is the age at a randomly selected patent application filing. Incomes are measured in 2012. Individual wage earnings is defined as total earnings reported on an individual's W-2 forms. Individual total income is defined for tax filers as Adjusted Gross Income (as reported on the 1040 tax return) minus the spouse's W-2 wage earnings (for married filers). For non-filers, total income is defined as wage earnings. In this table only, wage earnings are top-coded at \$1 million and total income is top-coded at \$10 million. Parent income is measured as mean household income (AGI) between 1996-2000. Median income variables are rounded to the nearest thousand dollars. College attendance at age 20 is measured using 1098-T forms filed by colleges, as in Chetty et al. (2017). Test scores, which are based on standardized tests administered at the district level, are normalized to have mean zero and standard deviation one by year and grade. See Section II for further details on sample and variable definitions.

TABLE II
 Fraction of Gap in Innovation by Parental Income Explained by Differences in 3rd Grade Test Scores

	Patent Rates for Children with Parents Below 80th Percentile	Patent Rates for Children with Parents Above 80th Percentile	High vs. Low Income Innovation Gap
	(1)	(2)	(3)
Raw Estimates	0.52 (0.05)	1.93 (0.20)	1.41 (0.21)
Reweightd to Match 3rd Grade Scores of High-Income Children	0.96 (0.07)	1.93 (0.20)	0.97 (0.21)
Gap in Innovation Explained by 3rd Grade Test Scores:			31.2%

Notes: This table shows how much of the gap in patent rates by parental income can be explained by 3rd grade math test scores. The statistics in this table are based on the children in the New York City public schools sample, which consists of children in the 1979-1985 birth cohorts who attended New York City public schools and were linked to the tax data. We divide children into two groups: those with parents in the top quintile of the income distribution within the New York City sample ("high-income children") and all other children in the sample ("low-income children"). We define a child as an inventor if he or she is listed as an inventor on a patent application between 2001-2012 or grant between 1996-2014 (see Section II.B). The first row of the table lists the fraction of children who become inventors among low-income (Column 1) and high-income children (Column 2) along with the differences between these two values (Column 3). In the second row of the table, Column 1 shows the patent rate that low-income children would have if they had the same math test scores as the high-income children. We calculate this counterfactual rate by dividing the math test score distribution into ventiles (twenty bins) and then calculating the patent rate for low-income children weighting by the number of high-income children in each of the twenty bins. Column 2 repeats the patent rates for high-income children, and Column 3 shows the gap between the high-income patent rate and the counterfactual low-income patent rate in Column 1. This adjusted gap can be interpreted as the difference in patent rates that would remain if test scores were identical across low- and high-income children. The percentage of the raw gap in innovation explained by 3rd grade test score is the percentage reduction in the gap from the raw to the reweighted estimates. Standard errors are reported in parentheses.

TABLE III
Exposure to Innovation from Parents' Colleagues: Children's Innovation Rates vs. Patent Rates in Father's Industry

Dependent Variable:	(1)	(2)	(3)	(4)	(5)
	Fraction Inventing	Fraction Inventing in Patent Category	Fraction Inventing in Patent Sub-Category	Fraction Inventing in Patent Class	Fraction Inventing in Patent Class
Patent Rate in Father's Industry	0.250 (0.028)				
Patent Rate in Father's Industry in Same Category		0.163 (0.018)			
Patent Rate in Father's Industry in Same Sub-Category			0.155 (0.017)		
Patent Rate in Father's Industry in Same Class				0.078 (0.013)	0.0598 (0.0125)
Patent Rate in Father's Industry in Same Sub-Category but Other Class					0.0044 (0.0008)
Patent Rate in Father's Industry in Same Category but Other Sub- Category					0.0001 (0.0004)
Patent Rate in Father's Industry in Other Category					0.0002 (0.0000)
Fixed Effects	None	Patent Category	Patent Sub-Category	Patent Class	Patent Class
Unit of Observation	Father's Industry	Father's Industry by Patent Category	Father's Industry by Patent Sub- Category	Father's Industry by Patent Class	Father's Industry by Patent Class
Number of Cells	345	2,415	12,765	153,525	153,525

Notes: This table analyzes how a child's propensity to invent is related to patent rates in his or her father's industry. The sample consists of children in the intergenerational sample (1980-84 birth cohorts) whose parents are not inventors. Each column presents estimates from a separate OLS regression, with standard errors clustered by industry in parentheses. In Column 1, we regress the share of children who become inventors among those with fathers in industry j on the patent rate among workers in industry j , with one observation per industry (six digit NAICS code). We measure the patent rate among workers in each industry as the average number of patents issued to individuals in that industry per year between 1996-2012 divided by the average number of workers per year (based on W-2 counts) in each industry between 1999-2012. Column 2 is run at the industry by patent category level. Here, we regress the share of children with fathers in industry j who invent in patent category c on the share of workers in industry j who have patents in category c . We include patent category fixed effects in this regression to account for differences in patent rates across categories. Columns 3 and 4 are analogous to Column 2, but use more narrowly defined categorizations of patent types: patent sub-categories and patent classes. Column 5 replicates Column 4 with three additional controls: the fraction of inventors in (i) the same sub-category but in a different patent class, (ii) the same category but a different sub-category, and (iii) other categories. All regressions are weighted by the number of children in each cell. There are 10,213,731 children underlying these regressions, the set of children in the intergenerational sample whose fathers have a non-missing NAICS code.

TABLE IV
Neighborhood Exposure Effects: Children's Innovation Rates vs. Patent Rates in Childhood Commuting Zone

Dependent Variable:	(1) Fraction Inventing	(2) Fraction Inventing	(3) Fraction Inventing in Patent Category	(4) Fraction Inventing in Patent Category	(5) Fraction Inventing in Patent Category	(6) Fraction Inventing in Patent Sub-Cat.	(7) Fraction Inventing in Patent Class	(8) Fraction Inventing in Patent Class
Patent Rate in Childhood CZ	2.932 (0.417)	2.578 (0.531)						
Patent Rate in Same Category in Childhood CZ			1.759 (0.404)	1.114 (0.341)	1.722 (0.406)			
Patent Rate in Same Sub-Category in Childhood CZ						1.526 (0.375)		
Patent Rate in Same Technology Class in Childhood CZ							1.108 (0.181)	1.017 (0.162)
Patent Rate in Same Sub-Category, but Different Technology Class in Childhood CZ								0.0003 (0.0063)
Patent Rate in Same Category, but Different Sub-Category in Childhood CZ								-0.0015 (0.0028)
Patent Rate in Different Category of Childhood CZ								0.0054 (0.0006)
Fixed Effects	None	Current CZ	Category	Current CZ by Category	Father's NAICS by Category	Sub-Category	Class	Class
Unit of observation	Childhood CZ	Childhood CZ by Current CZ	Childhood CZ by Category	Childhood CZ by Current CZ by Category	Childhood CZ by Father's NAICS by Category	Childhood CZ by Sub-Category	Childhood CZ by Patent Class	Childhood CZ by Patent Class
Number of Cells	741	221,621	5,187	1,551,347	1,637,706	27,417	329,745	329,745

Notes: This table analyzes how a child's propensity to invent is related to patent rates in his or her childhood commuting zone. The sample consists of children in the intergenerational sample (1980-84 birth cohorts) whose parents are not inventors. Each child is assigned a childhood CZ based on the ZIP code from which their parents first claimed them as dependents. Each column presents estimates from a separate OLS regression, with standard errors clustered by CZ in parentheses. In Column 1, we regress the share of children who become inventors among those who grow up in CZ j on the patent rate among workers in CZ j , with one observation per CZ. We measure the patent rate among workers in each CZ as the average number of patents issued per year (in the full USPTO data) to individuals in a given CZ between 1980 and 1990 divided by the CZ's population between the ages of 15-64 in the 1990 Census. Column 2 is run at the childhood CZ by current CZ level, limiting the sample to children whose current (2012) CZ differs from their childhood CZ. Here, we regress the share of inventors in each cell on the patent rate in the childhood CZ and on fixed effects for the 2012 CZ, so that the coefficient on childhood CZ patent rates is identified from comparisons across individuals currently living in the same CZ. Column 3 is run at the childhood CZ by patent category level. Here, we regress the share of children from CZ j who invent in patent category c on the share of workers in CZ j who have patents in category c . We include patent category fixed effects in this regression to account for differences in patent rates across categories. Column 4 replicates Column 2 at the category level, limiting the sample to children who move and estimating the model at the childhood CZ by current CZ by category level, with current CZ by category fixed effects. In Column 5, we include all children and replace the CZ by category fixed effects with fixed effects for the father's industry by category, estimating the model at the childhood CZ by father's industry by category level. This specification isolates variation from one's neighbors that is orthogonal to the variation from parents' colleagues. Columns 6 and 7 are analogous to Column 3 but use more narrowly defined categorizations of patent types: patent sub-categories and patent classes. Column 8 replicates Column 7 with three additional controls: the fraction of inventors in (i) the same sub-category but in a different patent class, (ii) the same category but a different sub-category, and (iii) other categories. All regressions are weighted by the number of children in each cell. There are approximately 15.5 million children underlying the regressions in Columns 1, 3, 6, 7 and 8. Columns 2 and 4 are based on the subset of 5.4 million individuals who moved across CZs. Column 5 includes the 10.2 million children whose fathers have non-missing NAICS codes.

TABLE V
Gender-Specific Exposure Effects: Children's Innovation Rates vs. Innovation Rates by Gender in Childhood CZ

Dependent variable:	(1)	(2)	(3)	(4)	(5)
	Fraction Inventing in CZ	Fraction of Women Inventing	Fraction of Men Inventing	Fraction of Women Inventing in Patent Category	Fraction of Men Inventing in Patent Category
Innovation Rate in Childhood CZ	0.986 (0.145)				
Innovation Rate of Women in Childhood CZ		2.408 (1.265)	-0.356 (4.398)	2.232 (0.607)	-2.157 (1.300)
Innovation Rate of Men in Childhood CZ		0.174 (0.154)	1.784 (0.625)	0.102 (0.062)	1.693 (0.295)
Fixed Effects	None	None	None	Category	Category
Unit of Observation	Childhood CZ	Childhood CZ	Childhood CZ	Childhood CZ by Category	Childhood CZ by Category
Number of Cells	741	741	741	5,188	5,188
p-value from F-test for Equality of Coefficients		0.113	0.667	0.001	0.015

Notes: This table analyzes how a child's propensity to invent is related to the innovation rates of adults of the same gender in his or her childhood commuting zone (CZ). The sample consists of children in the intergenerational sample (1980-84 birth cohorts) whose parents are not inventors. Each column presents estimates from a separate OLS regression, with standard errors clustered by CZ in parentheses. Column 1 replicates the specification in Column 1 of Table IV, except that here we define the independent variable using the linked patent-tax data rather than just the patent data, since we do not observe gender in the patent data itself. Specifically, we define the innovation rate for workers in CZ j as the total number of patent applications filed by individuals born before 1980 in our full inventors sample divided by the number of individuals between ages 15 and 64 in CZ j in the 1990 Census. We convert this measure to an annual rate by dividing by 17, as we observe patents between 1996-2012. In Column 2, we regress the fraction of girls from CZ j who become inventors on the patent rates of female and male workers in CZ j . Column 3 replicates Column 2 using the share of boys who become inventors as the dependent variable. The regression in column 4 is run at the childhood CZ by patent category level. Here, we regress the share of girls from CZ j who invent in patent category c on the share of male and female workers in CZ j who have patents in category c . We include patent category fixed effects in this regression to account for differences in patent rates across categories. Column 5 replicates Column 4 using the share of boys who become inventors as the dependent variable. All regressions are weighted by the number of children in each cell. The last row of the table reports p-values from F-tests for equality of the coefficients on male and female innovation rates in each regression. There are 15,499,290 individuals underlying each of the regressions.

APPENDIX TABLE I
Association Between Patent Rates and Upper-Tail Incomes with 3rd grade Math vs. English Test Scores

Dependent Variable:	Inventor (per 1,000 Individuals)					In Top 1% of Income Distribn. (per 1,000 Individuals)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
3rd Grade Math Score (SD)	0.85*** (0.09)		0.85*** (0.13)	0.91*** (0.13)		11.07*** (0.36)		8.08*** (0.46)
3rd Grade English Score (SD)		0.68*** (0.09)	0.05 (0.11)		0.08 (0.12)		10.19*** (0.34)	4.15*** (0.42)
Fixed Effects	None	None	None	English Ventile	Math Ventile	None	None	None
Mean of Dependent Variable	0.83	0.86	0.86	0.86	0.86	9.68	9.72	9.84
Observations	222,369	214,265	209,722	209,722	209,722	165,422	161,275	158,016

Notes: This table examines the extent to which third grade test scores are predictive of patent rates and upper-tail earnings outcomes. The sample in columns 1-5 consists of children in the 1979-1985 birth cohorts who attended New York City public schools in 3rd grade and were linked to the tax data. The sample in columns 6-8 consists of children who appear in both the NYC school district and intergenerational samples (1980-84 birth cohorts). Each column shows the coefficients and robust standard errors (in parentheses) from a separate OLS regression run at the student level; *** denotes $p < 0.001$. In Columns 1-5, the dependent variable is an indicator for being an inventor, defined as applying for a patent between 2001-2012 or being granted a patent between 1996-2014. In columns 6-8, it is an indicator for being in the top 1% of the individual income distribution in 2012 when compared to other individuals in the same birth cohort in the NYC school district sample. The dependent variables in each column are math and English test scores in 3rd grade. Test scores, which are based on standardized tests administered at the district level, are normalized to have mean zero and standard deviation one by year and grade. In Columns 4 and 5, we control for English and math scores non-parametrically using ventile fixed effects (20 bins) rather than a linear control. In all columns, coefficients are scaled so that they can be interpreted as the effect of a 1 SD change in test scores on the number of individuals per 1,000 who have the relevant outcome.

APPENDIX TABLE II
 Fraction of Gender Gap in Innovation Explained by Differences in Test Scores

A. Percent of Innovation Gap Explained by 3rd Grade Math Test Scores

	Patent Rates for Women	Patent Rates for Men	Gender Gap
	(1)	(2)	(3)
Raw Estimates	0.43 (0.06)	1.13 (0.09)	0.70 (0.11)
Reweightd to Match 3rd Grade Scores of Men	0.45 (0.06)	1.13 (0.09)	0.68 (0.11)
Gap in Innovation Explained by 3rd Grade Test Scores:			2.4%

B. Percent of Gap Explained by Test Scores Grades 3-8

Grade	Percent of Innovation Gap Explained by Math Test Scores in Grade <i>g</i>
3	2.4% (0.4)
4	2.4% (0.3)
5	3.4% (0.4)
6	4.6% (0.5)
7	6.8% (0.8)
8	8.5% (1.0)
Slope:	1.3 (0.2)

Notes: This table shows how much of the gender gap in patent rates can be explained by test scores using the New York City school district sample. Panel A is constructed in exactly the same way as Table II, comparing girls with boys instead of low-income children with high-income children. Panel B presents estimates of the gender gap in innovation that can be explained by test scores in grades 3-8, analogous to the estimates in Figure V. The slope estimate reported at the bottom is estimated using an OLS regression of the six estimates on grade.

APPENDIX TABLE III
Distance Between Technology Classes: Illustrative Example

Category: Computers + Communications

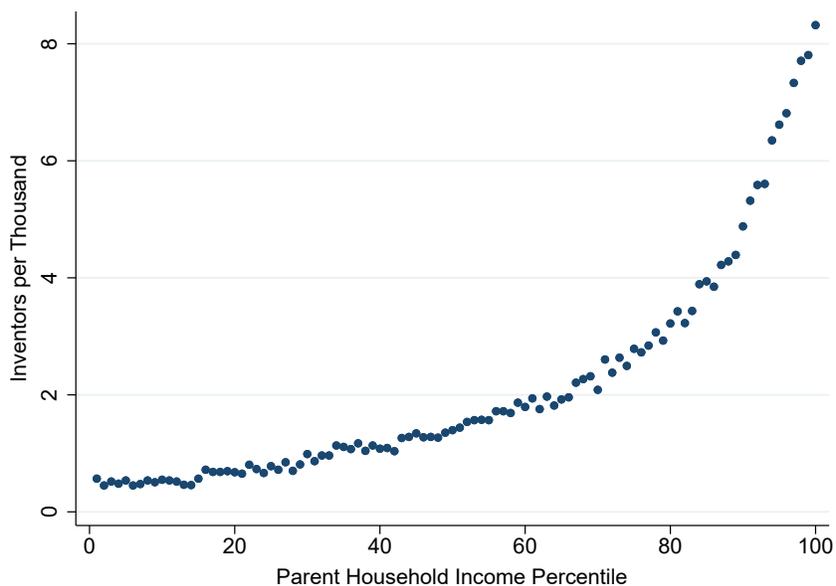
Sub-category: *Communications*

<u>Technology Class (= 375)</u>	<u>Distance Rank (<i>d</i>)</u>
<i>Pulse or Digital Communications</i>	0
Demodulators	1
Modulators	2
Coded Data Generation or Conversion	3
Electrical Computers: Arithmetic Processing and Calculating	4
Oscillators	5
Multiplex Communications	6
Telecommunications	7
Amplifiers	8
Motion Video Signal Processing for Recording or Reproducing	9
Directive Radio Wave Systems and Devices (e.g., Radar, Radio Navigation)	10

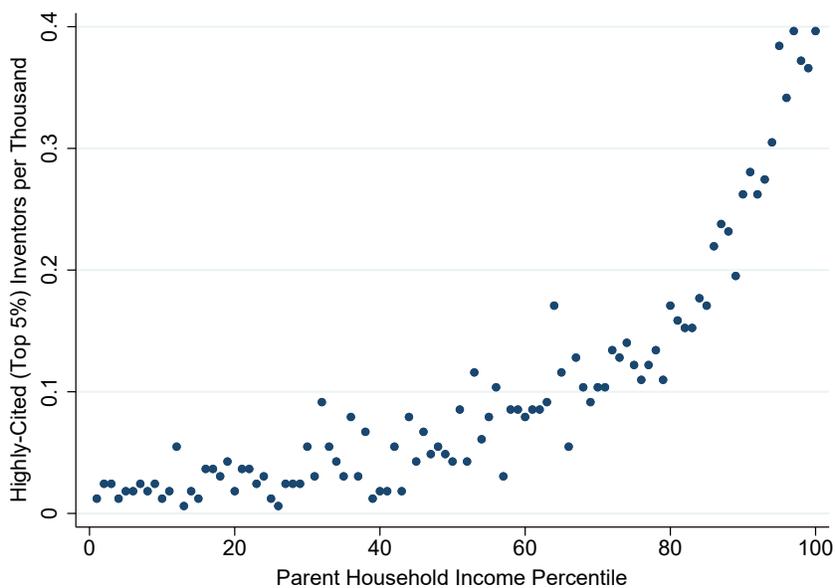
Notes: This table provides an example of our measures of distance between technology classes. We define the distance between two technology classes A and B by computing the share of inventors in class A who also invent in class B; the higher the share of common inventors, the lower the distance between A and B. We convert this distance metric to an ordinal measure, defining $d=0$ for the own class, $d=1$ for the next nearest class, etc. The table lists the 10 closest classes to the "Pulse or Digital Communications" class, which falls within the Communications subcategory of the Computers + Communications category.

FIGURE I: Patent Rates vs. Parent Income

A. All Inventors by 2014

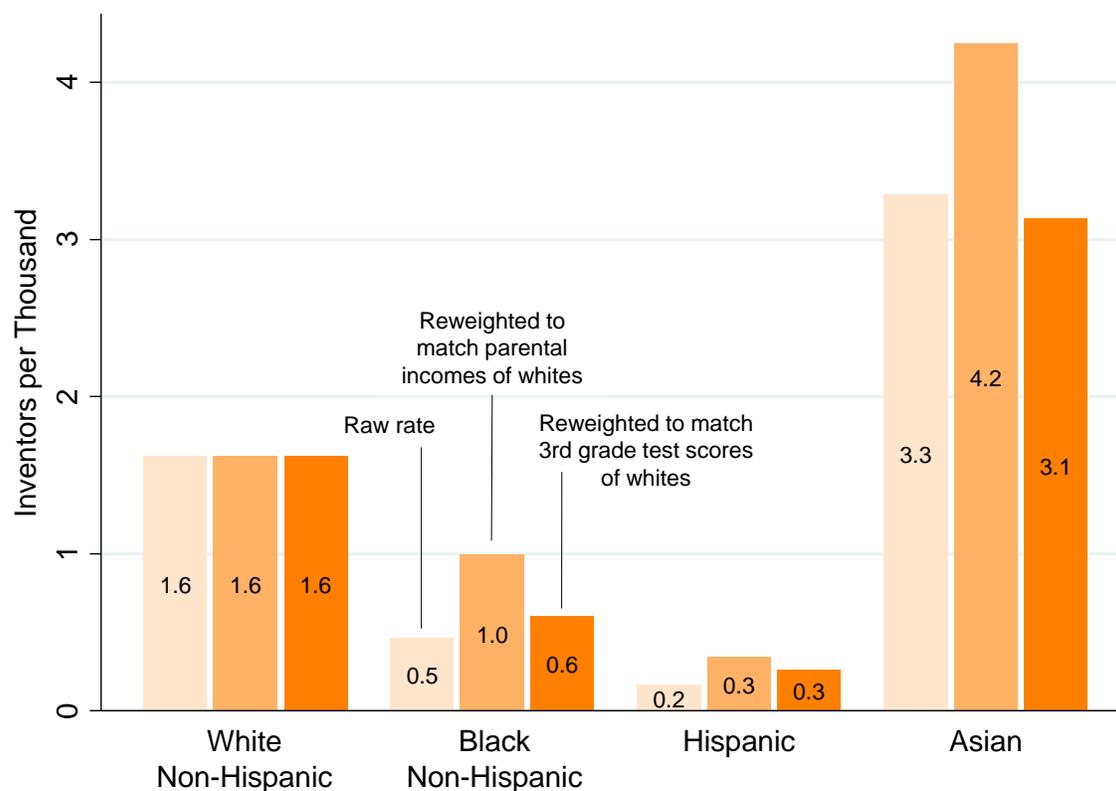


B. Highly-Cited Inventors



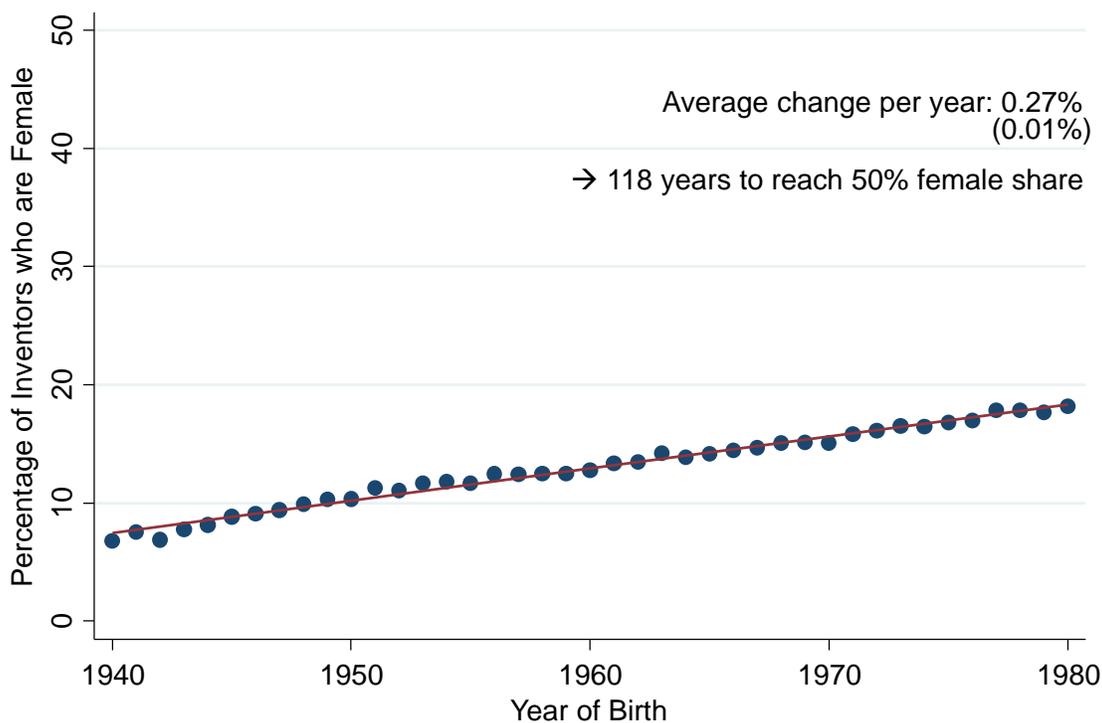
Notes: This figure characterizes the relationship between patent rates and parental income using our intergenerational analysis sample, which consists of U.S. citizens in the 1980-84 birth cohorts (see Section II.B for details). Panel A plots the number of children (per 1,000 individuals) who invent by 2014 vs. their parents' income percentile. Parents are assigned percentile ranks by ranking them based on their mean household income from 1996 to 2000 relative to other parents with children in the same birth cohort. Inventing by 2014 is defined as being listed as an inventor on a patent application between 2001-2012 or grant between 1996-2014 (see Section II.B). Panel B replicates Panel A, defining the outcome as being a highly-cited inventor, defined as having total citations in the top 5% of the distribution among inventors in the same birth cohort.

FIGURE II: Patent Rates by Race and Ethnicity



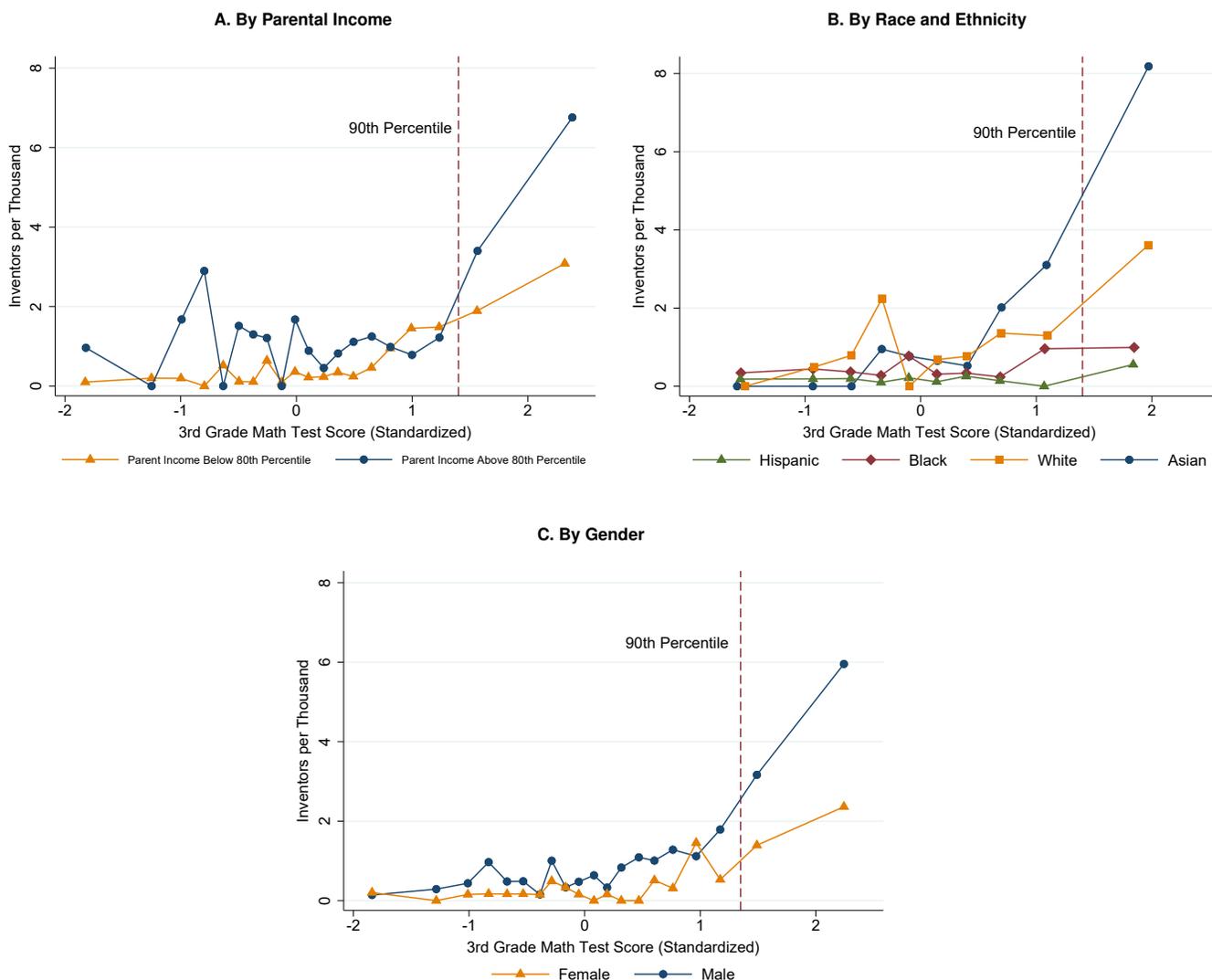
Notes: This figure presents patent rates by race and ethnicity using our New York City public schools sample, which consists of children in the 1979-1985 birth cohorts who attend NYC public schools at some point between grades 3-8. Each bar plots the number of children (per 1,000 individuals) who invent by 2014, as defined in the notes to Figure I. In each triplet, the first bar shows the raw patent rate for the relevant subgroup. The second bar plots the patent rate that would prevail if children in the relevant subgroup had the same distribution of parental income as white children. To construct these estimates, we divide children into 20 bins based on their parental incomes and compute mean patent rates across the 20 bins, weighting each bin by the fraction of white children with incomes in that bin. The third bar in each triplet shows the patent rate that would prevail if children in the relevant subgroup had the same distribution of 3rd grade math test scores as white children. These estimates are constructed by dividing children into 20 bins based on their test scores and computing mean patent rates across the 20 bins, weighting each bin by the fraction of white children with test scores in that bin.

FIGURE III: Percentage of Female Inventors by Birth Cohort



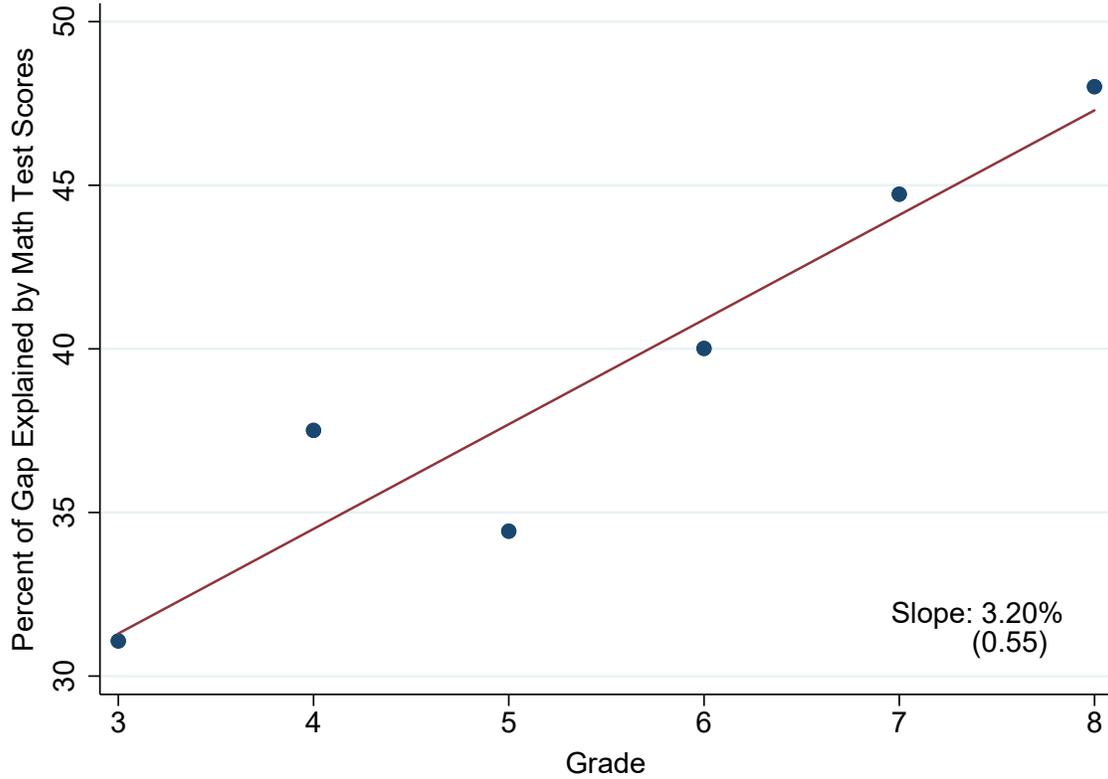
Notes: This figure plots the percentage of inventors who are female by year of birth using our full inventors sample, which consists of all 1.2 million individuals in the linked patent-tax data. Inventing is defined as being listed as an inventor on a patent application between 2001-2012 or grant between 1996-2014 (see Section II.B for details). The change per year is estimated using an unweighted OLS regression of the percentage of female inventors on birth year, depicted by the solid line. The standard error from this regression is shown in parentheses.

FIGURE IV: Patent Rates vs. 3rd Grade Math Test Scores



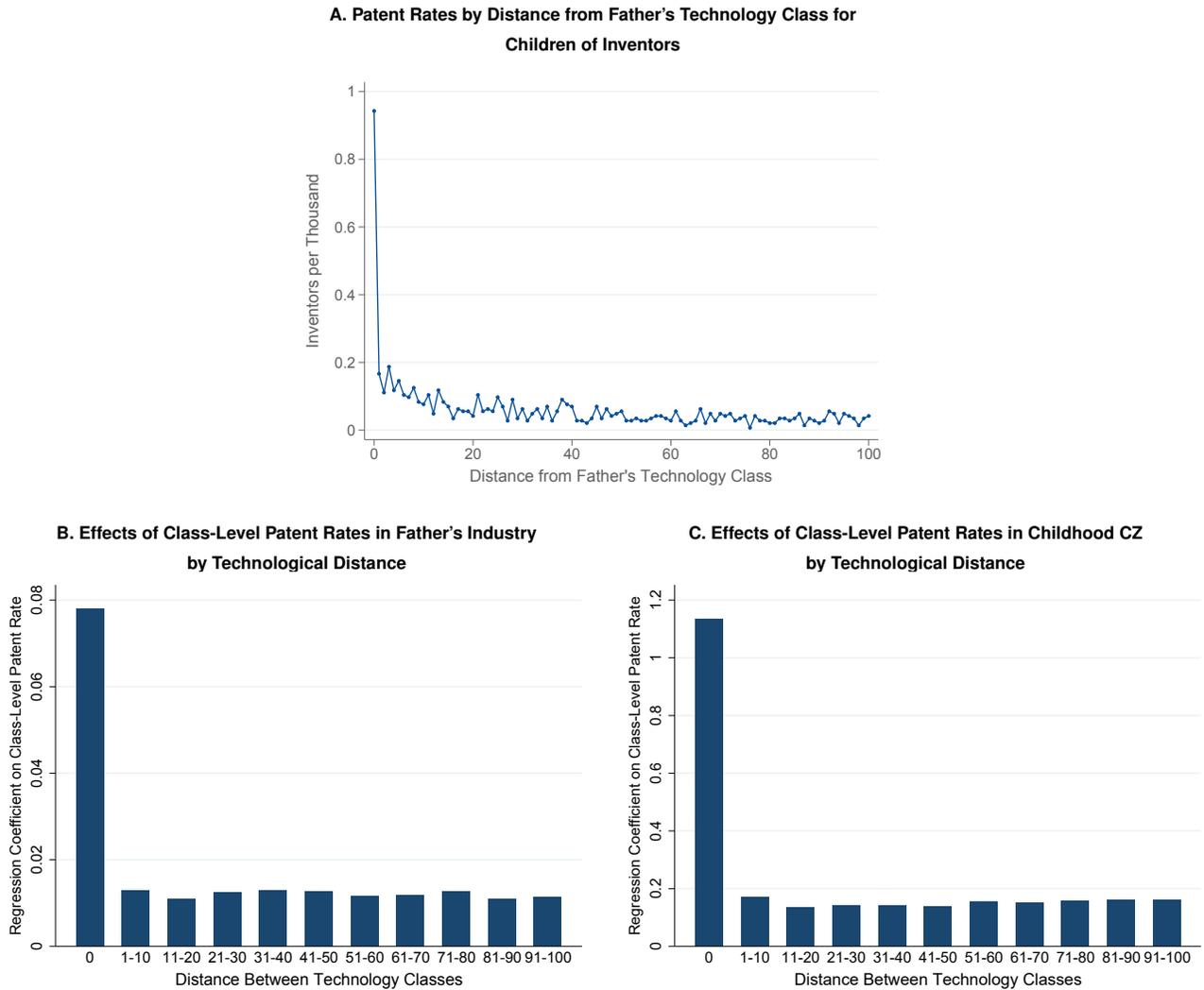
Notes: This figure shows the relationship between patent rates and math test scores in 3rd grade for various subgroups. The sample consists of children in the 1979-1985 birth cohorts who attended New York City public schools in 3rd grade. Test scores, which are based on standardized tests administered at the district level, are normalized to have mean zero and standard deviation one by year and grade. In Panel A, we divide children into two groups based on whether their parents' incomes fall below the 80th percentile of the income distribution of parents' income in the New York City schools sample. The figure presents a binned scatter plot of patent rates vs. test scores for these two subgroups. To construct the figure, we first divide children into 20 equal sized bins (ventiles) based on their test scores. We then plot the share of inventors (per 1000 individuals) vs. the mean test score within each bin for each of the two subgroups. Panel B and C replicate Panel A, dividing children by their race and ethnicity (Panel B) and gender (Panel C) instead of parental income. We use 10 bins rather than 20 bins of test scores in Panel B because of smaller sample sizes for some racial and ethnic groups. The vertical dashed lines depict the 90th percentile of the test-score distribution.

FIGURE V: Gap in Patent Rates by Parental Income Explained by Test Scores in Grades 3-8



Notes: This figure shows how much of the gap in patent rates by parental income can be explained by math test scores in grades 3-8. The sample consists of children in our New York City public schools sample (birth cohorts 1979-1985), whom we divide into two groups: those with parents in the top quintile of the income distribution within the New York City sample (“high-income children”) and all other children in the sample (“low-income children”). The gap in innovation explained by math test scores in grade g is the percentage reduction in the gap in innovation when we reweight low-income students’ grade g test score distribution to match that of high-income students. Table II illustrates how we construct this estimate using 3rd grade test scores (31.2%); estimates for later grades use the same methodology. The slope and best-fit line are estimated using an unweighted OLS regression on the six points, with standard error reported in parentheses.

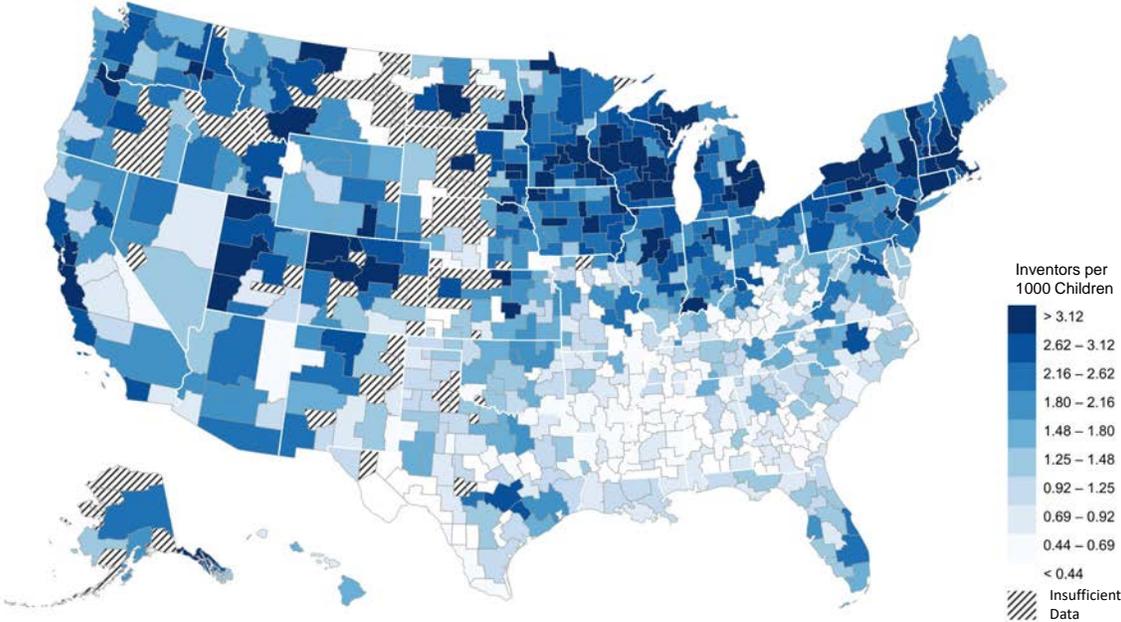
FIGURE VI: Children’s Patent Rates vs. Class-Level Patent Rates in Childhood Environment



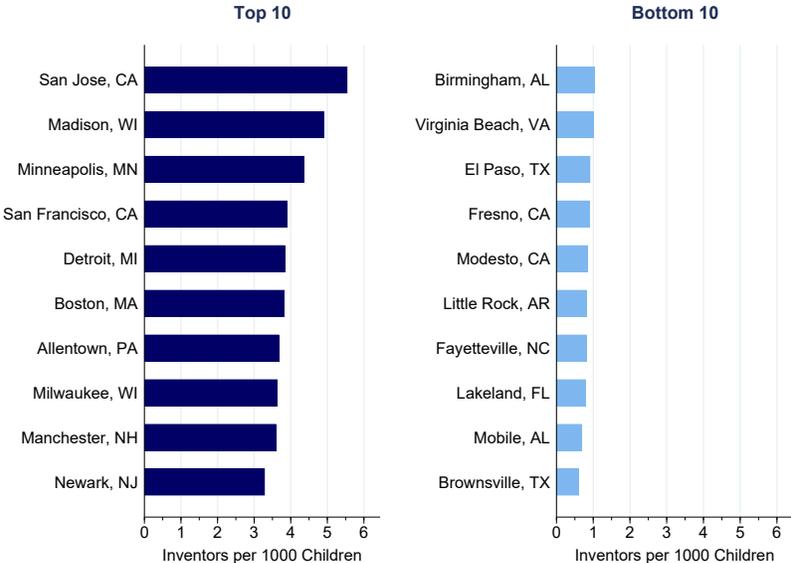
Notes: This figure shows how children’s propensities to patent in a technology class vary with the class in which their father (Panel A), father’s colleagues (Panel B), or childhood neighbors (Panel C) patented. In Panel A, the sample consists of all children in our intergenerational sample whose fathers are inventors (those who applied for a patent between 2001-2012 or were granted a patent between 1996-2014) and who were not listed as co-inventors on a patent with their fathers. To construct Panel A, we first assign fathers and children a technology class based on the class in which they have the most patents and patent applications. We then define the distance between two technology classes A and B based on the share of inventors in class A who also invent in class B. Using this distance metric, for each child, we define $d = 0$ as the class in which his or her father patents, $d = 1$ as the next closest class, etc. We then plot the share of children (per 1,000 individuals) who invent in a technology class that is d units away from their father’s class. Classes in which fewer than 100 inventors have a patent grant or application between 1996-2014 are omitted. In Panels B and C, the sample consists of all children in our intergenerational sample whose parents are *not* inventors. Each bar in Panel B plots estimates from a separate regression, with one observation per father’s industry (six digit NAICS code) and patent technology class. In the first bar, we regress the fraction of children who patent in technology class c among those with fathers in industry j on the patent rate among workers in industry j in the same technology class c . We measure the class-level patent rate among workers in each industry as the average number of patents in class c issued to individuals in that industry per year (between 1996-2012) divided by the average number of workers per year in each industry between 1999-2012. In the second bar, we regress the same dependent variable on the mean patent rate in the father’s industry in the 10 closest classes ($d = 1$ to 10). The third bar uses the average patent rate in classes with $d = 11$ to 20, etc. All regressions are weighted by the number of children in each cell and include class level fixed effects for class c . Panel C replicates Panel B, replacing patent rates in the father’s industry with patent rates of workers in the CZ where the child grew up. CZ-level patent rates are defined as the average number of patents issued in class c per year to individuals from a given CZ between 1980-1990 divided by the CZ’s population between ages 15-64 in the 1990 Census.

FIGURE VII: The Origins of Inventors: Patent Rates by Childhood Commuting Zone

A. Patent Rates by Childhood CZ

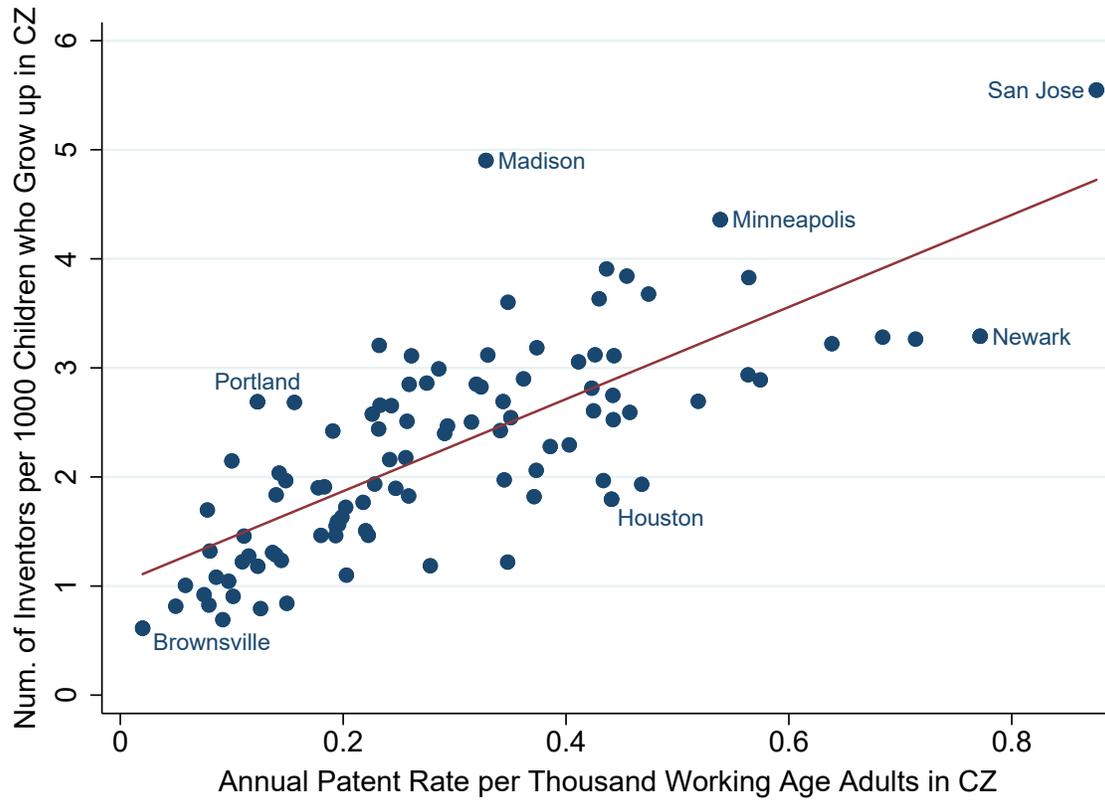


B. CZs with the Highest and Lowest Patent Rates among the 100 Largest CZs



Notes: Panel A maps the share of children who become inventors by the commuting zone (CZ) in which they grew up using our intergenerational sample (U.S. citizens in the 1980-84 birth cohorts). Each child is assigned a CZ based on ZIP code from which their parents filed their 1040 tax return in the year they were first claimed as dependents (which is typically 1996, as our data begin in 1996). The map is constructed by dividing the CZ into unweighted deciles based on patent rates, with darker shades representing areas where more children grow up to become inventors. Data for CZs with fewer than 1,000 children, which account for 0.3% of the children in the sample, are omitted. Panel B lists the CZs with the ten highest and lowest shares of inventors per thousand children among the 100 CZs with the largest populations in the 2000 Census.

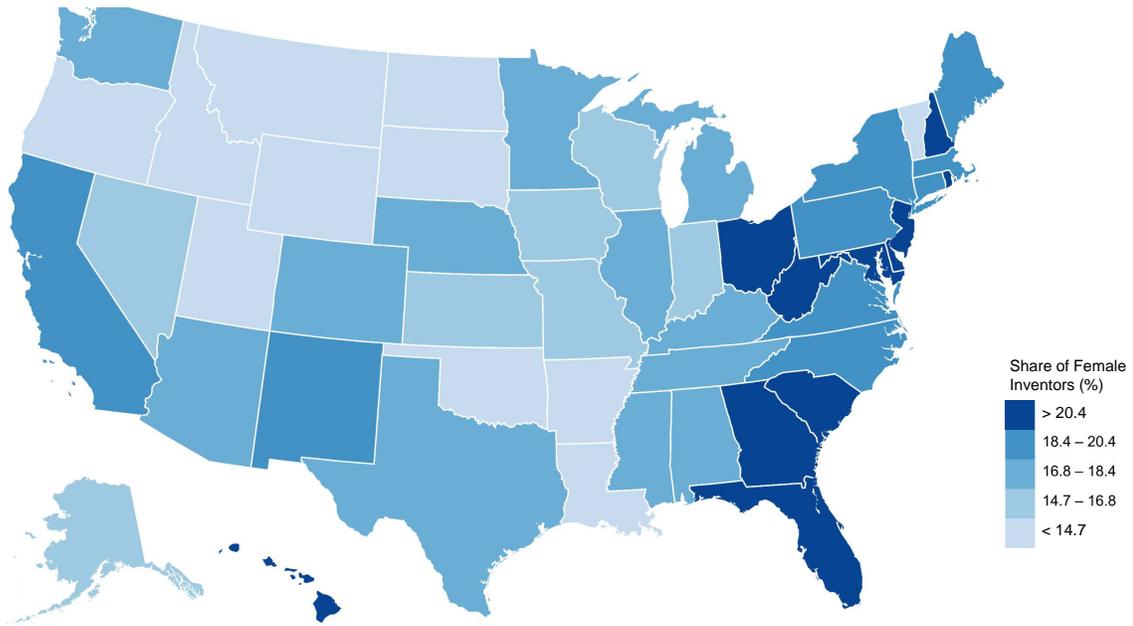
FIGURE VIII: Children's Patent Rates vs. Patent Rates of Workers in their Childhood CZ



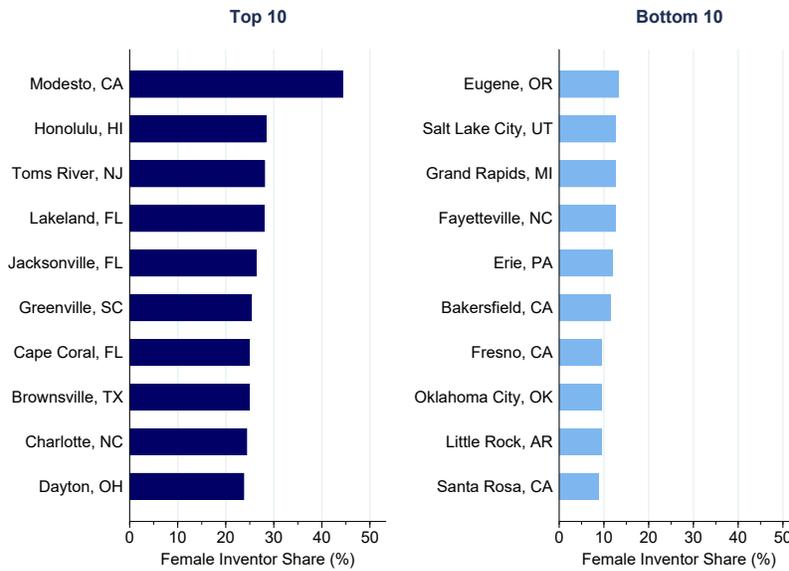
Notes: The figure plots the patent rates of children who grow up in a given CZ (constructed exactly as in Figure VII) vs. the patent rates of workers who live in that CZ. Patent rates of workers in each CZ are defined as the average number of patents per year issued to inventors residing in that CZ between 1980-1990 (based on the universe of USPTO data) divided by the CZ's population between the ages of 15-64 in the 1990 Census. We restrict the figure to the 100 CZs with the largest populations. The solid best-fit line is estimated using an unweighted OLS regression on these 100 observations (slope = 4.22, standard error = 0.40).

FIGURE IX: Geographical Variation in Gender Gaps in Patent Rates

A. Percent of Inventors who are Female by State where Child Grew Up



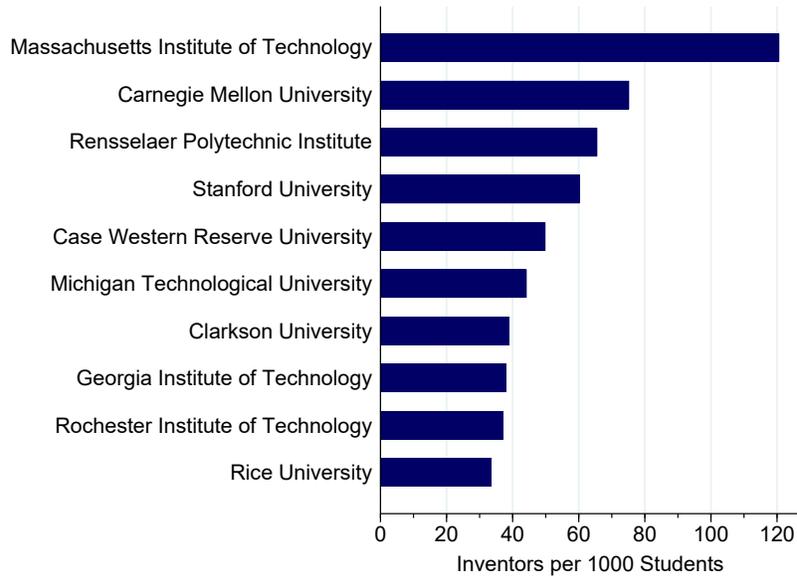
B. CZs with the Highest and Lowest Female Inventor Shares Among the 100 Largest CZs



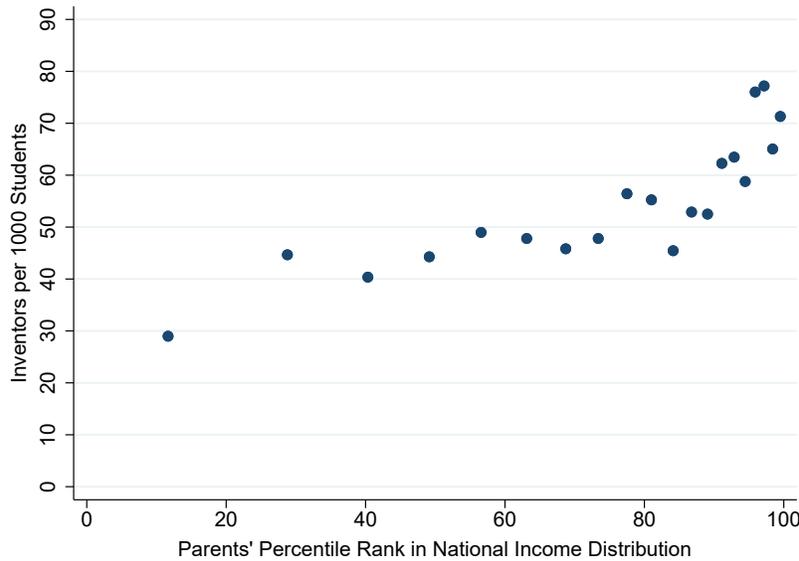
Notes: Panel A maps the percentage of female inventors by the state in which they grew up using our intergenerational sample (U.S. citizens in the 1980-84 birth cohorts). Each child is assigned a state based on ZIP code from which their parents filed their 1040 tax return in the year they were first claimed as dependents (which is typically 1996, as our data begin in 1996). The map is constructed by dividing the states into unweighted quintiles based on the female inventor share, with darker shades representing areas where women account for a larger share of inventors. Panel B lists the commuting zones (CZs) with the ten highest and lowest female inventor shares among the 100 CZs with the largest populations in the 2000 Census.

FIGURE X: Patent Rates by College

A. Colleges with the Highest Share of Inventors per Student

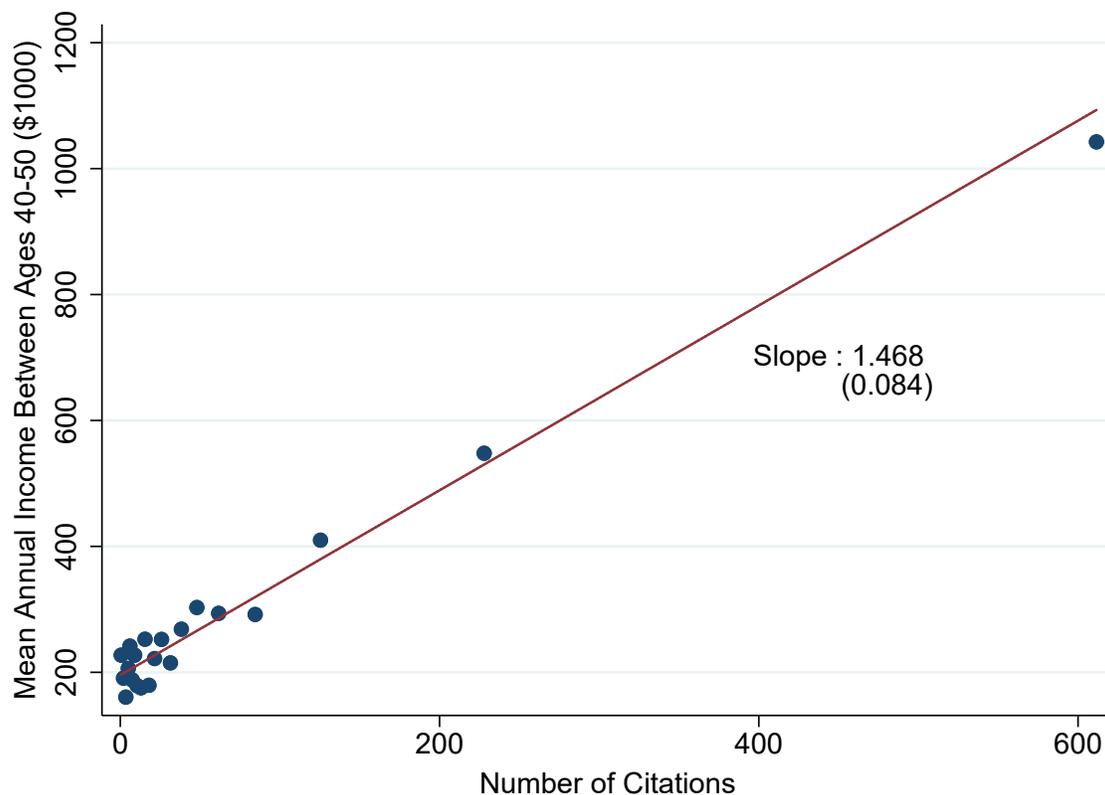


B. Patent Rates vs. Parent Income in the 10 Most Innovative Colleges



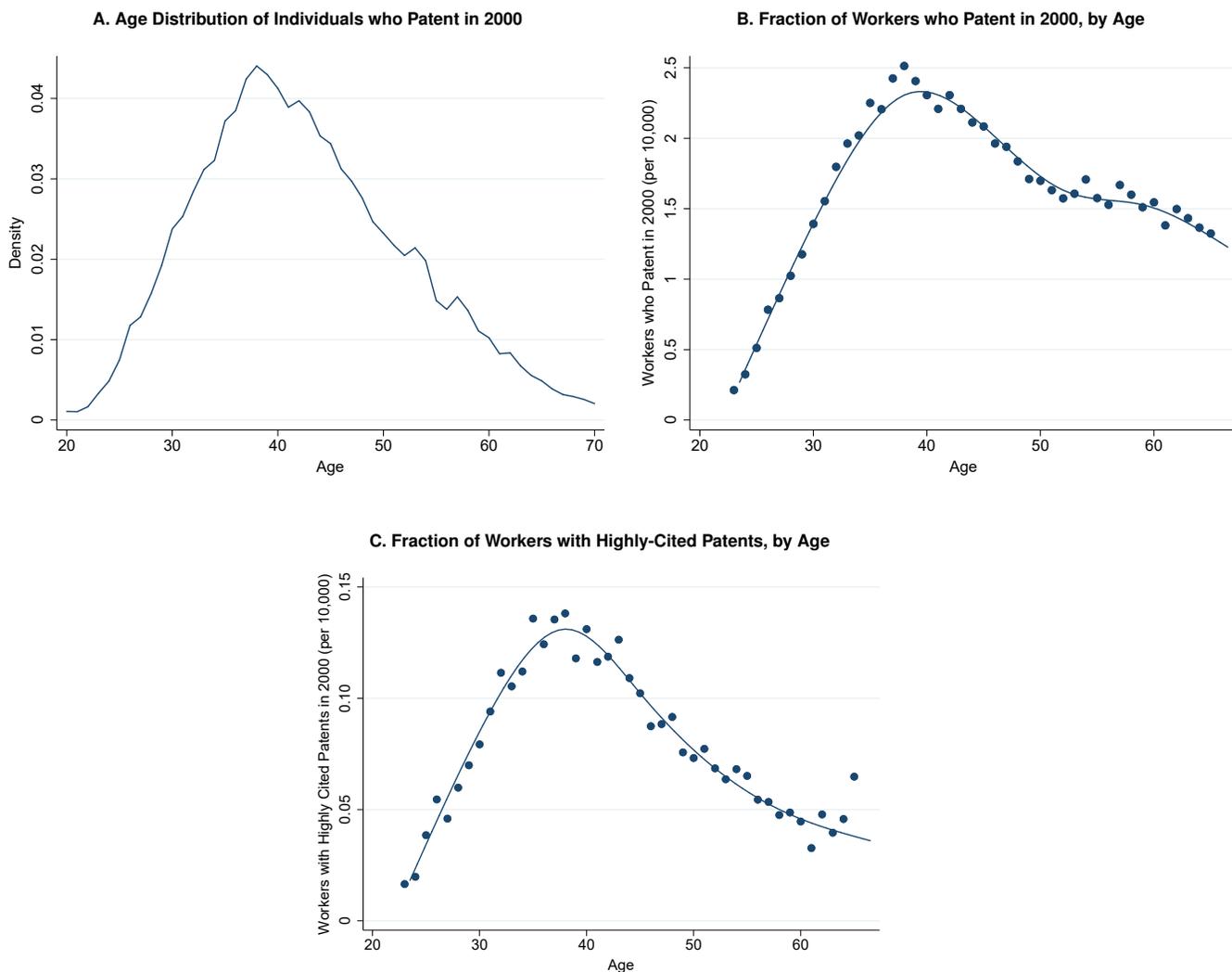
Notes: This figure presents data on the share of students who become inventors by 2014 (as defined in the notes to Figure I) by the college they attend. The sample consists of all individuals in the tax data in the 1980-84 birth cohorts who are linked to parents. Children are assigned to the college that they attend most frequently at age 19-22, following the methodology of Chetty et al. (2017). Panel A lists the ten colleges that have the highest fraction of students who become inventors, among colleges with at least 500 students per cohort. This figure is produced from the college-level estimates in Online Data Table 12 of Chetty et al. (2017). These college-level estimates are blurred to protect confidentiality using the procedure in Chetty et al. (2017, Appendix C). Panel B presents a binned scatterplot of patent rates vs. parental income for students who attended the 10 colleges listed in Panel A. It is constructed by binning parent income into 20 equal-sized bins (ventiles) and plotting the mean share of inventors (per 1000 students) vs. the mean parent rank in the national income distribution within each bin. There are fewer points on the left because there are fewer students from low-income families than high-income families at these colleges.

FIGURE XI: Inventors' Incomes vs. Patent Citations



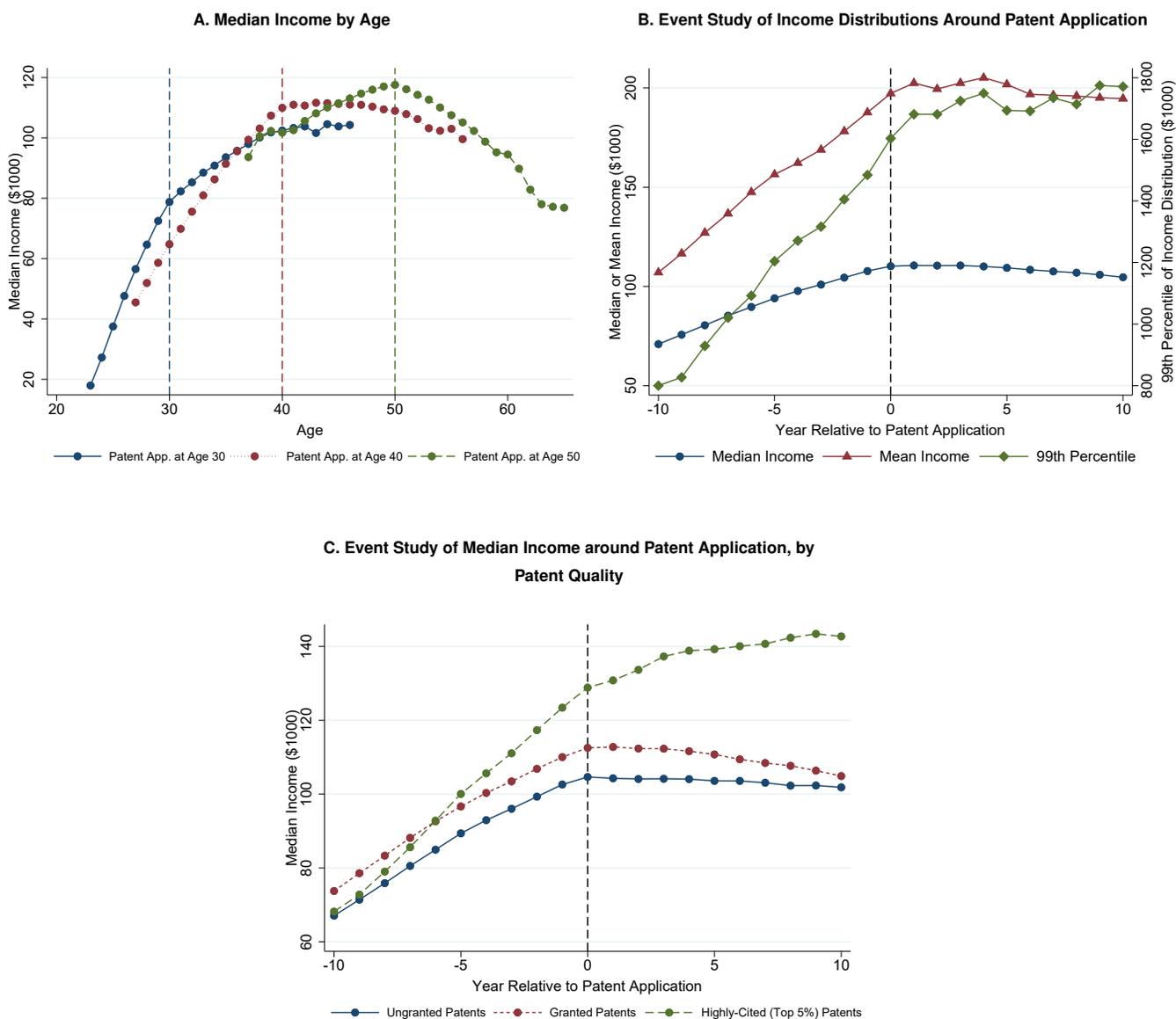
Notes: This figure presents a binned scatter plot of average annual income between ages 40 and 50 vs. the total number of citations an inventor obtains. The sample consists of individuals in our full inventors sample who were born in 1959-62, for whom we observe income at all ages between 40 and 50. We further limit the sample to the 13,875 individuals who applied for a patent in 1996 to maximize the time horizon over which we can measure future citations. This plot is constructed by dividing citations into 21 bins and plotting mean income vs. mean citations within each bin. The first 19 bins include inventors in the first 19 ventiles (5% bins) of the citations distribution, while the last two bins plot the same relation for the 95th to 98th percentiles and the 99th percentile of the citation distribution. The best fit line and slope shown on the figure are estimated using an OLS regression on the 21 points, weighted by the number of inventors in each bin. The standard error of the slope estimate is reported in parentheses.

FIGURE XII: Age Profile of Innovation



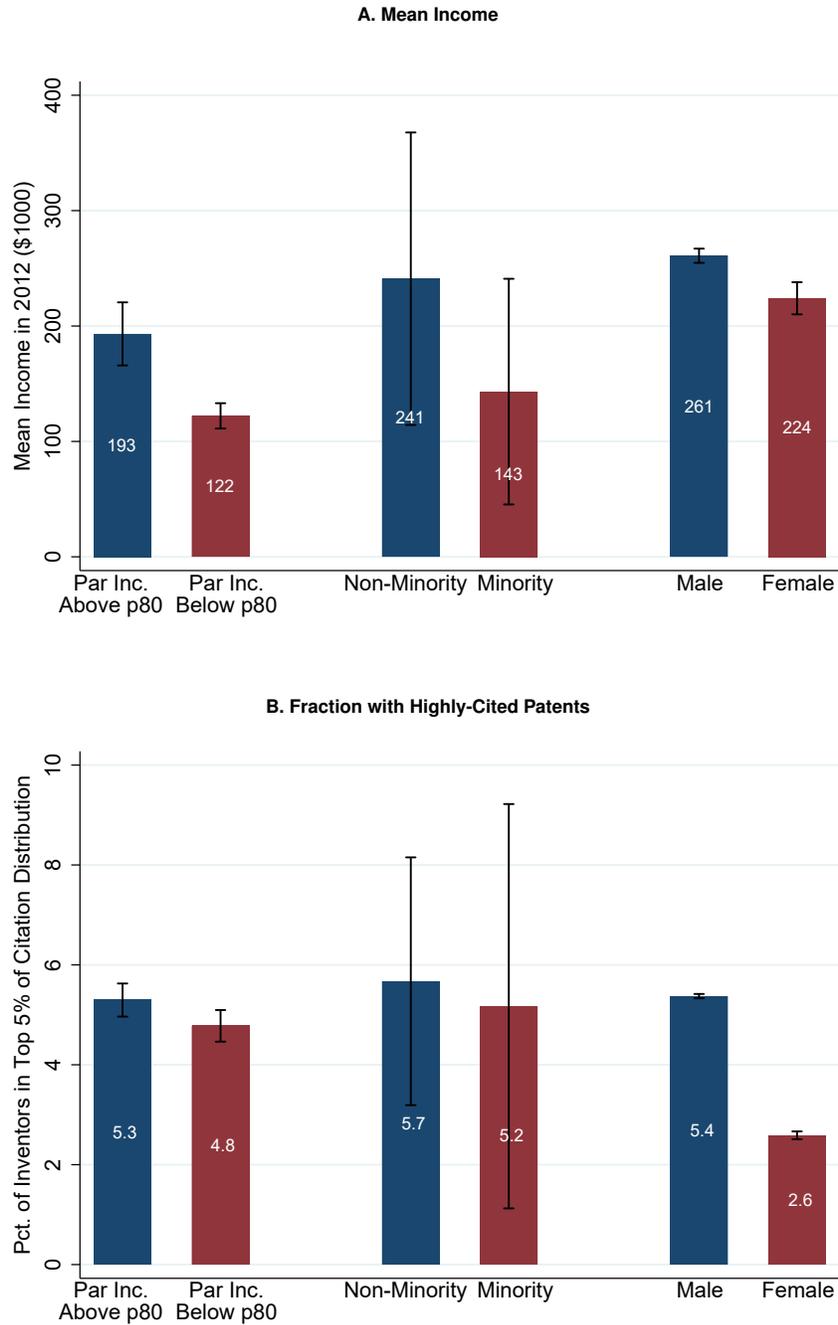
Notes: This figure examines the age distribution of inventors. The sample consists of individuals in our full inventors sample who applied for a patent in 2000 that was subsequently granted. Panel A presents a kernel density of the age distribution of these inventors. Panel B plots the fraction of workers (individuals with positive W-2 earnings) who patent in 2000 by age. Panel C plots the fraction of workers who filed a patent application in 2000 that went on to become highly-cited (in the top 5% of the distribution). The curves in Panels B and C are cubic splines.

FIGURE XIII: Income Profiles of Inventors



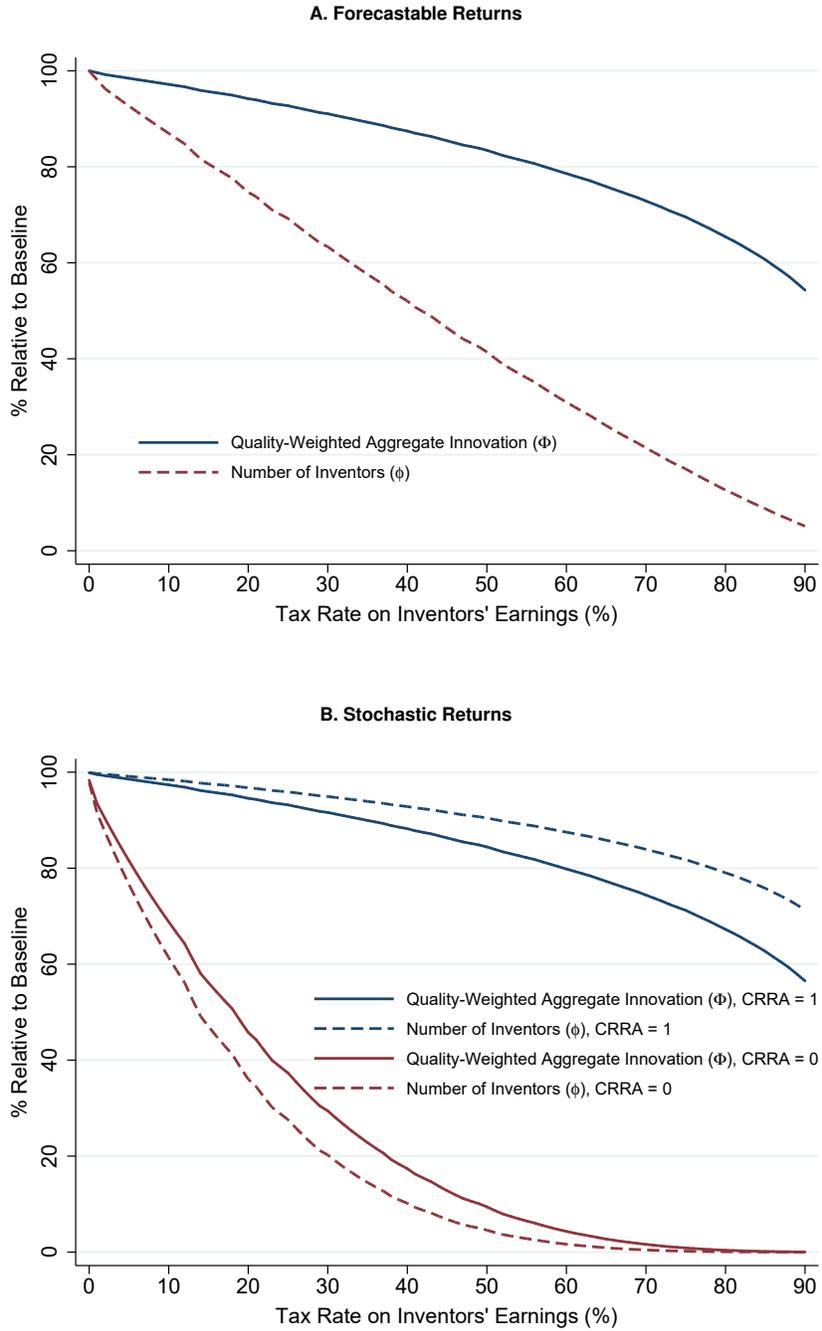
Notes: This figure plots the income profiles of inventors before and after they file patent applications using all individuals for whom the relevant data is available in our full inventors sample. Income is measured as total income (including wage earnings and capital income) at the individual level; see Section II.C for details. Panel A plots the median incomes by age of inventors who file a patent application at either age 30, 40, or 50 over the range of ages for which their incomes are observed (between 1996-2012). Panel B generalizes this analysis using an event study framework, defining year 0 as the year in which an individual files a patent application and other years relative to this reference year (e.g., +1 is the year after the application). In this figure, we limit the sample to individuals who file patent applications between ages 35 and 50. For individuals who file multiple patents in this age window, we choose one of the patents at random to define the reference year. We plot the mean and median (left y axis) and 99th percentile (right y axis) of the income distribution of inventors in each year relative to the event year. Panel C replicates the median income series in Panel B, separating inventors into three groups: those whose patent applications were not granted; those whose applications were granted; and those with patents granted that went on to have citations in the top 5% of the distribution relative to other patents granted in the same year.

FIGURE XIV: Income and Citations of Inventors by Characteristics at Birth



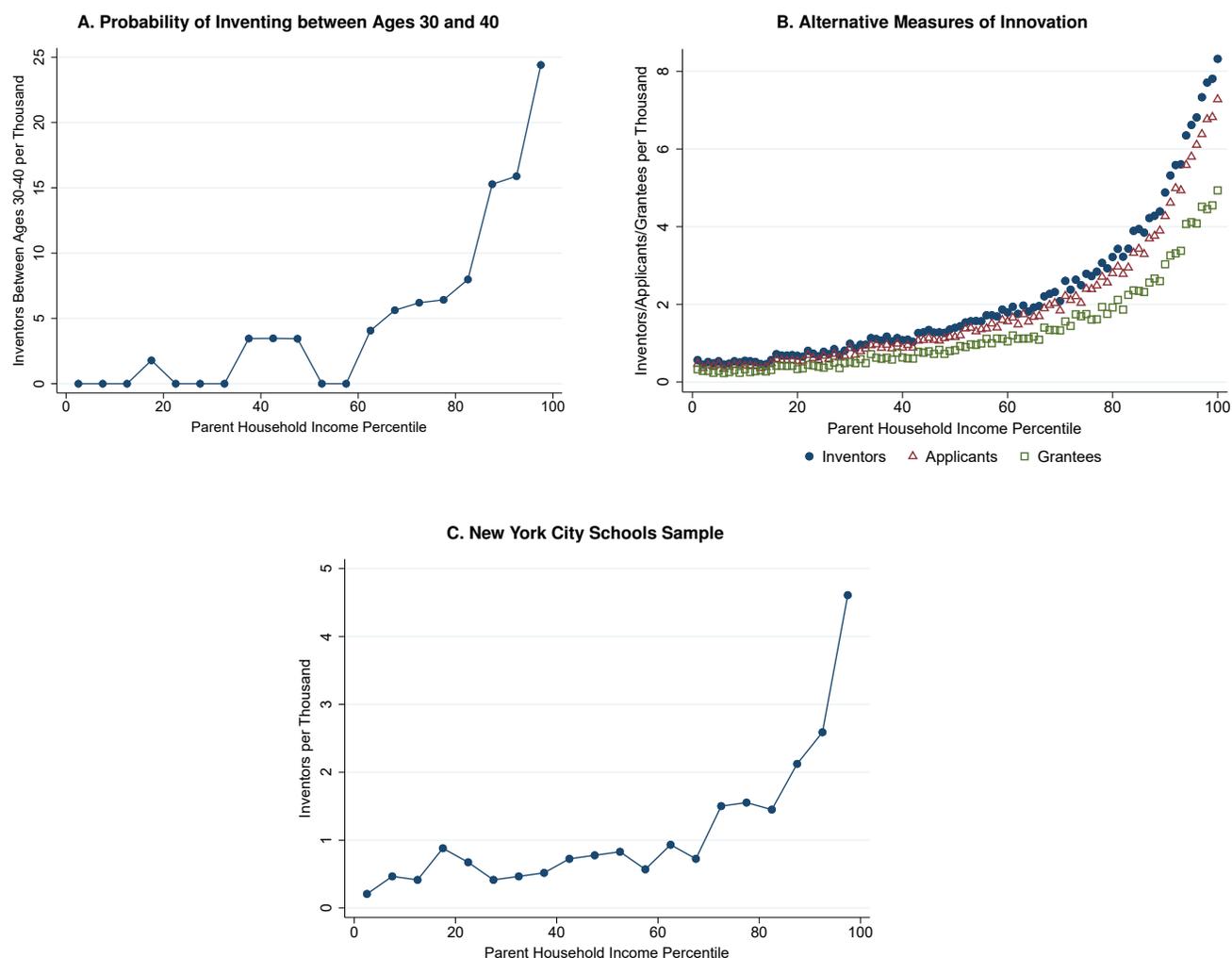
Notes: Panel A plots the mean incomes of inventors in 2012 by their parents' income, race/ethnicity, and gender. The first pair of bars uses our intergenerational sample (1980-84 birth cohorts), divided into two subgroups based on whether parents' household income is below or above the 80th percentile of the parent income distribution. The second pair of bars uses our New York City schools sample, divided into two subgroups based on race and ethnicity: minorities (Blacks and Hispanics) and non-Minorities. The third pair of bars uses our full inventors sample, divided by gender. The vertical lines depict 95% confidence intervals. Panel B replicates Panel A using the fraction of highly-cited inventors as the outcome. Highly-cited inventors are defined as inventors whose patents have citations per co-author in the top 5% of the distribution among those in their birth cohort.

FIGURE XV: Predicted Impacts of Tax Rates on Innovation



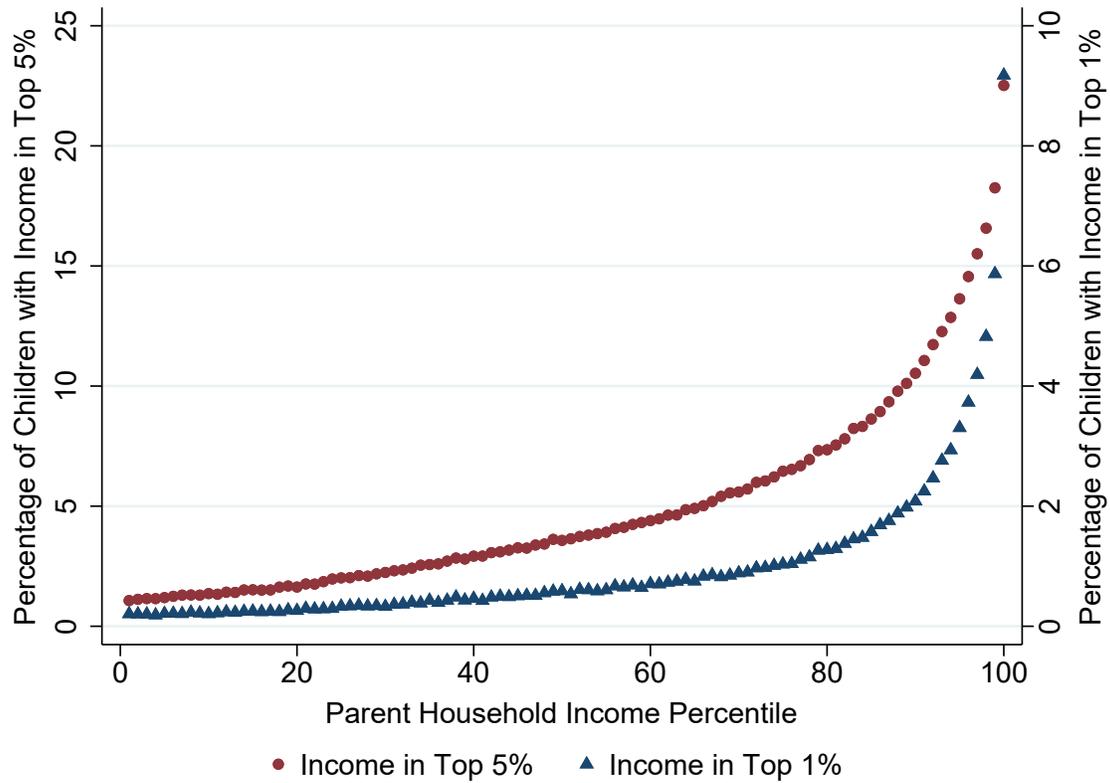
Notes: This figure plots the fraction of inventors (ϕ) and aggregate quality-weighted innovation (Φ) vs. the tax rate on inventors' earnings predicted by the model of career choice in Section VI. Panel A considers the case where the variation in private financial returns to innovation is driven purely by differences in ability across inventors and hence is perfectly forecastable at the time of career choice. Panel B considers the case where the variance in private returns come primarily from stochastic shocks, with a coefficient of relative risk aversion $\theta = 0$ (linear utility) or 1 (log utility). The shape parameters for the Pareto distributions of stochastic returns and innovation abilities, denoted β_π and β_α in the model, are chosen such that the inventors' earnings distribution generated by the model matches the Pareto shape parameter of 1.26 estimated using inventors' empirical earnings distribution in Online Appendix Figure VI, i.e. such that $\frac{\beta_\pi}{\beta_\pi - 1} \frac{\beta_\alpha}{\beta_\alpha - 1} = \frac{1.26}{1.26 - 1}$. In Panel B, stochastic returns account for 90% of total skewness in inventors' earnings, i.e. $s \equiv \left(\frac{\beta_\pi}{\beta_\pi - 1} - 1\right) / \left(\frac{\beta_\alpha}{\beta_\alpha - 1} - 1 + \frac{\beta_\pi}{\beta_\pi - 1} - 1\right) = 0.9$. In each series, the level of innovation is normalized to 100% when the tax rate is 0. The normalized values are invariant to the other parameters of the model ($w_{\bar{I}}$, λ , and θ in Panel A).

ONLINE APPENDIX FIGURE I: Patent Rates vs. Parent Income: Sensitivity Analysis



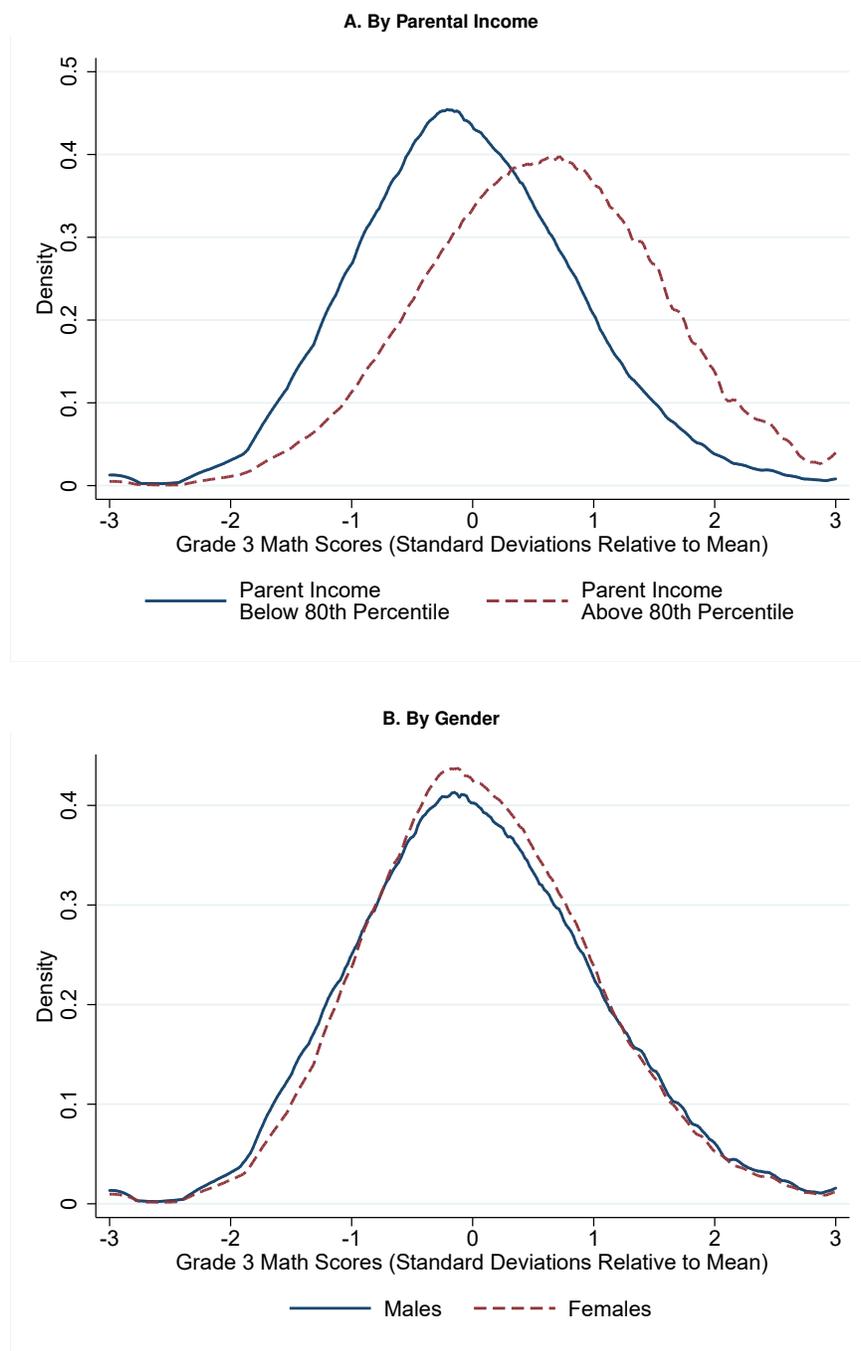
Notes: This figure replicates Figure Ia using alternative samples and definitions. Panel A uses data from the 1971-72 birth cohorts in the Statistics of Income sample, a 0.1% sample of tax returns (see Section II.B for details). This sample allows us to examine whether an individual filed a patent application or was granted a patent between the ages of 30 and 40, rather than just by their early thirties as in Figure Ia. Given the smaller sample size, we plot mean patent rates between the ages of 30-40 by parent ventile (20 bins) rather than percentiles in this figure. In Panel B, the series in circles replicates Figure Ia exactly, where inventors are defined as those who applied for a patent between 2001-2012 or were granted a patent between 1996-2014. The other two series in that figure show the fraction of individuals who applied for patents and the fraction who were granted patents separately. Panel C replicates the baseline series in Figure Ia (plotting the fraction of inventors) using the subset of children in the New York City public schools sample. In this figure, we rank parents within the NYC sample based on their household incomes and plot the fraction of children who become inventors by 2014 by parent income ventiles.

ONLINE APPENDIX FIGURE II:
Fraction of Children with Incomes in Upper Tail vs. Parent Income



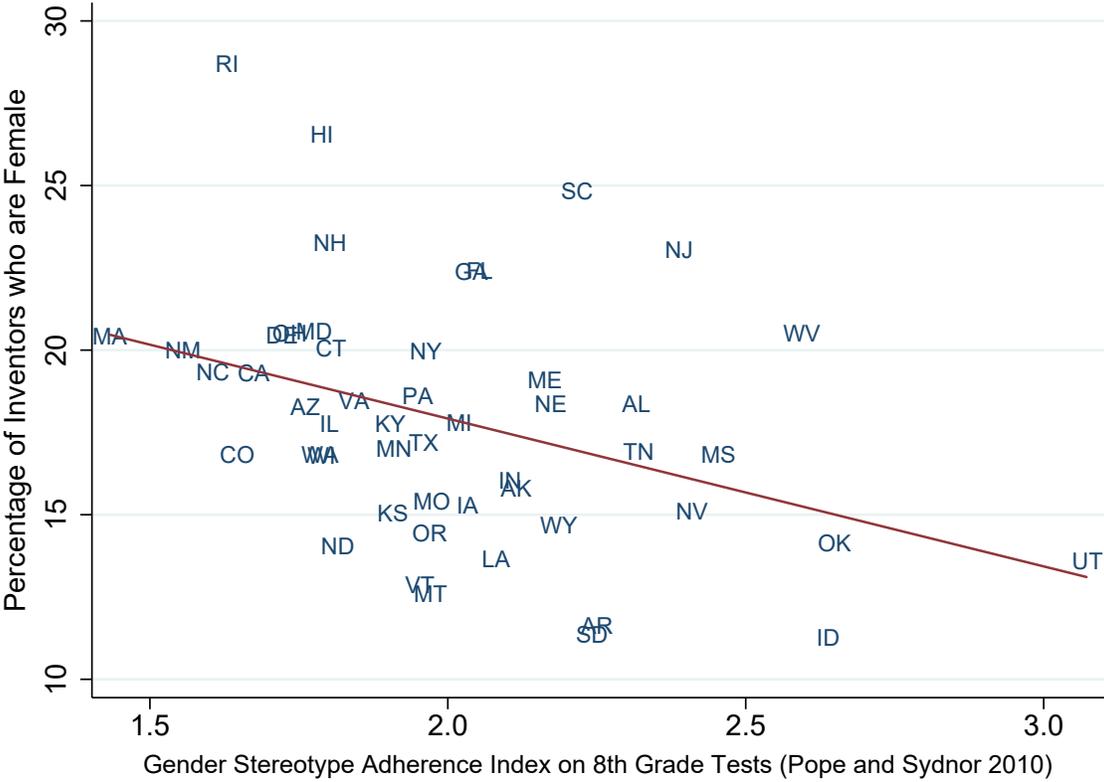
Notes: This figure replicates Figure I, replacing the outcome variable with an indicator for having mean individual income in 2011-12 in either the top 1% or top 5% of the income distribution among individuals in the same birth cohort. The sample is our intergenerational sample, which includes U.S. citizens in the 1980-84 birth cohorts.

ONLINE APPENDIX FIGURE III: Distribution of Math Test Scores in 3rd Grade



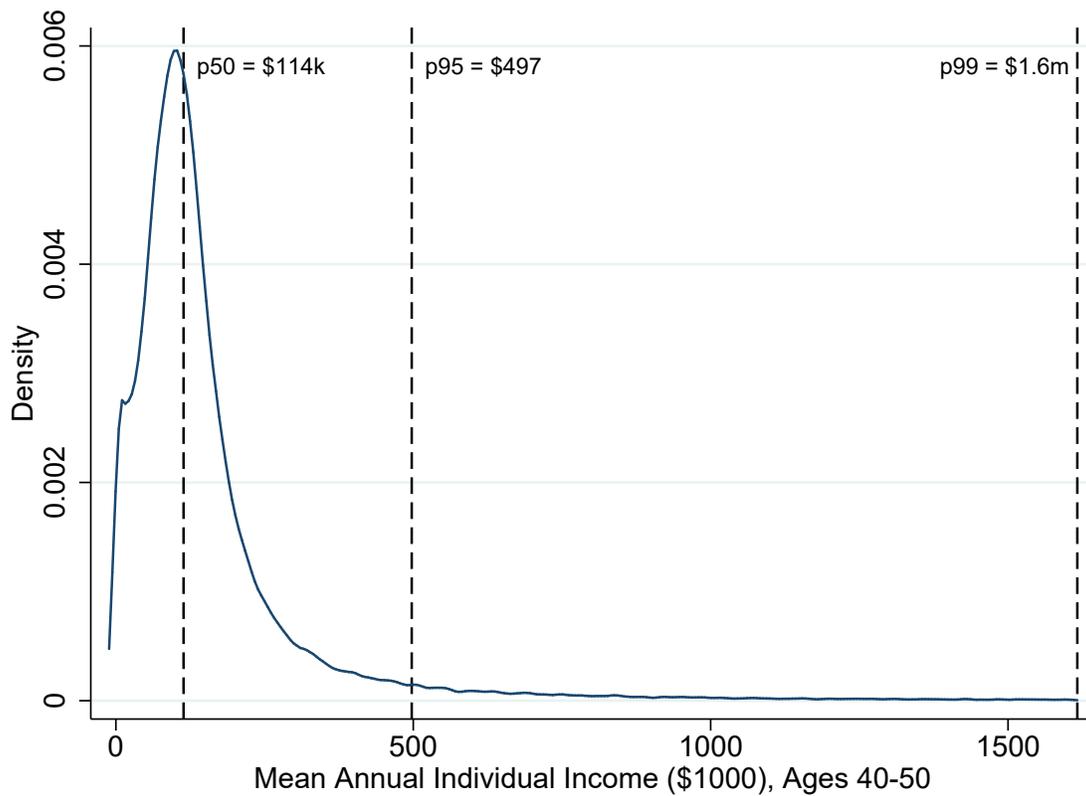
Notes: These figures present kernel densities of 3rd grade math test scores for children in the 1979-1985 birth cohorts who attended New York City public schools. Test scores, which are based on standardized tests administered at the district level, are normalized to have mean zero and standard deviation one by year and grade. In Panel A, we divide children into two groups based on whether their parents' incomes fall below the 80th percentile of the income distribution of parents' income in the New York City schools sample. Panel B compares boys and girls.

ONLINE APPENDIX FIGURE IV: Female Inventor Share and Gender Stereotype Adherence



Notes: This figure plots the share of inventors who are female vs. Pope and Sydnor’s (2010) gender stereotype adherence index by state. Female inventor shares are taken directly from Figure IXa; see notes to that figure for details. The stereotype adherence index is computed as $S = (N_{m,math\&science}/N_{f,math\&science} + N_{f,reading}/N_{m,reading})/2$, where $N_{g,s}$ denotes the number of students of gender $g \in \{m, f\}$ who score among the top 5% of students in their state in subject $s \in \{\text{math \& science, reading}\}$ in 8th grade. The index S measures the degree to which students adhere to the typical gender stereotype that boys do better at math/science and girls do better in reading; higher values represent greater adherence to this stereotype. The solid best-fit line is estimated using an unweighted OLS regression (slope = -4.49, standard error = 1.42).

ONLINE APPENDIX FIGURE V: Inventors' Income Distribution



Notes: This figure plots a kernel density of the distribution of inventors' income, measured as mean annual income over ages 40-50 in 2012 dollars. Income is measured at the individual level and includes both labor and capital income. The sample consists of all individuals in our full inventors sample born between the ages of 1959-1962, for whom we see income at all ages between 40 and 50. For scaling purposes, the top and bottom percentiles of the distribution are omitted. The dashed lines mark the median, 95th percentile, and 99th percentile of the distribution.