

The Lifecycle of Inventors*

Alex Bell, Harvard University
Raj Chetty, Stanford University and NBER
Xavier Jaravel, Harvard University
Neviana Petkova, Office of Tax Analysis, US Treasury
John Van Reenen, Centre for Economic Performance, LSE and NBER

This version: June 5th 2016

Abstract

We use administrative records on the population of individuals who applied for or were granted a patent between 1996 and 2014 to characterize the lives of more than 1.2 million inventors in the United States. We show that children of low-income parents are much less likely to become inventors than their higher-income counterparts (as are minorities and women). Decompositions using third grade and older test scores indicate that this income-innovation gap can largely be accounted for by differences in human capital acquisition while children are growing up. We establish the importance of “innovation exposure effects” during childhood by showing that growing up in an area with a high innovation rate in a particular technology class is associated with a much higher probability of becoming an inventor specifically in that technology class. Similarly, exposure to innovation from parents or their colleagues in specific fields is also associated with greater future innovation by children in those same technological fields. Inventors’ incomes are very skewed and uncertain at the start of their career. While our analysis does not directly identify the causal mechanisms that drive innovation, our descriptive findings shed light on which types of policy tools are likely to be most effective in sparking innovation. We extend rational sorting models of talent allocation to allow for imperfect information which generates predictions that are consistent with our data. Calibrations of this model suggest that “extensive margin” policies drawing more talented children from low-income families into the R&D sector have great potential to improve aggregate innovation rates.

*The opinions expressed in this paper are those of the authors alone and do not necessarily reflect the views of the Internal Revenue Service or the U.S. Department of the Treasury. We would like to particularly thank Philippe Aghion, with whom we started thinking about these issues in 2013, for inspiration and many insightful comments. We would like to also thank Daron Acemoglu, Ufuk Akcigit, Olivier Blanchard, Danny Kahnemann, Henrik Kleven, Richard Layard, Josh Lerner, Jim Poterba, Scott Stern, Otto Toivanen and participants in seminars in AEA, Brussels, Harvard, IFS/UCL, LSE, Manchester, Mannheim EEA, MIT, NBER Innovation and Labor groups and the OTA for helpful comments and discussions. Trevor Bakker, Augustin Bergeron, Jamie Fogel, Nikolaus Hidenbrand, and Benjamin Scuderi provided outstanding research assistance. This research was funded by the National Science Foundation, the Lab for Economic Applications and Policy at Harvard, the Economic and Social Research Council at CEP and the Kauffman Foundation.

I Introduction

Innovation is at the center of growth theory and advanced economies use a broad range of policies intended to spur innovation, ranging from subsidies for research and development (R&D) to investments in technical education. However, relatively little is known about the characteristics and life trajectories of inventors. Indeed, information on even the most basic demographic characteristics - such as the age distribution and income composition of patent holders - is scarce because existing databases do not record such information.

In this paper, we present the first comprehensive portrait of inventors in the United States by linking data on all patents granted between 1996 and 2014 to federal income tax returns. Our linked dataset contains information on over 1.2 million patent applicants or holders. We use this linked data to document a set of stylized facts about the lives of “inventors”¹ that inform current theoretical debates and identify new patterns to be explained by the next generation of models of innovation and growth. We structure our analysis around the chronology of an inventor’s life, starting with her family background and neighborhood at birth, then turning to her education and finally to her labor market career.

While our focus is primarily on descriptive facts rather than identification of causal mechanisms, the facts we document help discriminate between alternative theories in the literature and shed light on the types of policies that are likely to be most effective in sparking innovation. Most economic research on innovation policies focuses on what could be termed the “intensive margin” of innovation, namely getting more innovation out of the existing stock of inventors. For example, the US Research and Experimentation Tax credit reduces research costs relative to other forms of investment. Similarly, a leading argument for low top income tax rates is that they increase incentives for innovation and therefore boost growth (e.g. Mankiw (2013)). One potential problem with such policies is that they rely on those with the current potential to innovate to do more of it.² In light of these limitations, Romer (2000) recommends studying policies that focus on increasing the “extensive margin” of innovation - increasing the underlying supply of inventors. Our analysis of the lives of inventors yields a better understanding of what makes and constrains a potential innovator, which is an essential ingredient for developing such policies.

¹Patents have well-known pros and cons as indicators of invention. Not all innovations are patented and not all patents correspond to meaningful innovations. But we use the rich information on the patent documents to deal with these drawbacks, for example, using future citations received by a patent as a proxy for its quality.

²In fact, there is a risk that with an inelastic supply of innovators, subsidies to research will simply increase the equilibrium wage of R&D scientists, rather than stimulate a greater volume of innovation (e.g. Goolsbee 1998).

We begin by studying the birth and origin of potential inventors. The children of high-income parents are much more likely to be inventors: children born to parents in the top 1% of the income distribution are more than ten times as likely to become an inventor as children born to families with below-median income. Part of the relationship between parent income and children's patent rates could stem from children of the rich being born with higher ability than those of the poor and therefore being naturally more likely to make technological breakthroughs. Alternatively, part of the inventor-income relationship could be that even when children begin with identical traits, having low-income parents may hold children back from becoming innovators because of a lower quality education, neighborhoods, mentors, or jobs opportunities. If such barriers are important, this is very policy-relevant. It would mean a lack of opportunity for those who would have had a comparative advantage in innovation but for their financial position. It would mean a loss of innovation and output due to a misallocation of talent. How many "lost Einsteins" could there be due to inequality of opportunity (e.g. Celik (2014))?

We shed light on these alternative mechanisms by using data from all individuals who went through the New York City (NYC) public school system between 1989 and 2009, from which we have standardized data on test score results in grades 3 through 8. We show that only around 30% of the invention gap between rich and poor can be accounted for by third grade math test scores. We conduct a similar analysis to study the gap in innovation by gender and by race. Unlike the gap in innovation by parents' socio-economic status, only 3% of the gender gap in innovation can be explained by differences in math test scores. Comparing Blacks and Hispanics to Whites, we also find a large invention gap - only a small fraction of it is due to initial ability, similar to the gender gap. A much more substantial fraction is due to income differences.

Existing talent misallocation models such as Hsieh et al. (2013) are based on Roy models of occupational choice according to comparative advantage but with frictions that create additional costs to all those from disadvantaged groups. These "rational sorting" models imply that individuals from disadvantaged groups who do become inventors should have *higher* levels of human capital than their more advantaged counterparts. In fact, our data shows the opposite - inventors from disadvantaged groups do *not* appear more talented. The quality of their patents (as measured by citations) and their initial test scores are similar or worse than other inventors.

What alternative model could account for these findings? We emphasize two related phenomena at play as a child grows up prior to choosing a career. First, using the NYC data on test scores from grades 3 to 8 we show that a substantial difference in educational outcomes between rich

and poor families opens up as children progress through school. Using a wider sample we show that by the time we know the identity of the college attended, there is relatively little difference in whether a rich or poor child becomes an innovator. Second, we show that “exposure” to innovation in childhood has a strong association with the chances of growing up to be an inventor. Our measures of exposure include (i) whether the parent was an inventor; (ii) how innovative was the industry where a child’s parents worked and (iii) neighborhood characteristics such as innovation in the childhood Commuting Zone (CZ). Being exposed to innovation when a child is growing up is strongly associated with later becoming an inventor. On all of these measures we show that it is not only the amount of innovation, but exposure to *type* of innovation that matters by using detailed technology class information. It is not simply that children who grow up in the Bay Area are more likely to be inventors (even when they live elsewhere) - they are more likely to specialize in the technologies that are relatively successful in the Bay Area (like computer software relative to medical devices). This evidence suggests that mentoring effects or exposure to careers in science and innovation at young ages may play a key role in children’s later outcomes. Since children from low-income backgrounds are less likely to benefit from such exposure, this evidence reinforces the view that the innovation gap between the rich and the poor is driven by differences in environment and human capital accumulation, not intrinsic traits.

In the last part of the descriptive life-cycle analysis we look at the labor market and show that the returns to innovation appear highly skewed and uncertain, especially at the time of career choice. Returns often come later in life and are earned not just after the patent event, but during a broader period of several years leading up to patenting.

Motivated by these findings, we present a simple inventor lifecycle model that has barriers to human capital acquisition as rational sorting models of misallocation, but extends such models to allow for imperfect information over inventor careers. In particular, we argue that many individuals from disadvantaged groups may under-estimate the net benefits of an inventor career if they are not exposed to innovation during childhood. The model’s predictions broadly match the findings of our dataset in a way that existing models cannot. Simple calculations suggest that the returns of supply-side policies that would realistically reduce the innovation rate gap between privileged and disadvantaged groups would be extremely large, with the potential of increasing the population of inventors by over 30%. We also present a quantitative analysis of the effect of changing the top marginal tax rates on inventors’ incomes.³ We show that the effect of top tax rates on the key

³Note that we only consider increasing top tax rates on inventors, while holding the tax schedule fixed for other

individuals in the innovation process - the inventors themselves - is likely to be small due to the skewness and randomness of the payoffs. Our contribution is to calibrate this response using the skewness of the empirical earnings distribution of inventors. We show that even large cuts in top income tax rates on inventors will only induce a small change in the population of inventors.

We emphasize from a positive perspective that “extensive margin” innovation policies drawing talented individuals into the innovation sector may be very effective at increasing innovation, given that many talented low-income individuals currently do not make the choice of becoming inventors, in part due to the lack of exposure to this occupation as they grow up. The question of the optimal allocation of talent across sectors extends beyond the scope of this paper. Although we discuss how our results relate to common mechanisms in the “misallocation” literature, we do not draw normative conclusions about the observed distribution of talent across sectors (in particular, whether or not it is optimal to find so few talented low-income individuals among inventors). Rather, our results show from a positive perspective that exposure to innovation is an important driver of occupational choice or, in other words, that it has an important “allocation” effect. We view this finding as an important lesson for innovation policy, which is currently overwhelmingly focused on “intensive margin” incentives targeting individuals who are already part of the innovation sector. Naturally, exposure effects may be a key driver of occupational choice in other contexts as well (for example for doctors, lawyers, financiers, etc.). In this paper, we focus on the decision of becoming an inventor for two reasons. First, inventors play a key role for growth and determining which policies have the potential to allocate more individuals to this occupation is therefore potentially very important for social welfare. Second, studying inventors has various methodological advantages: we can precisely characterize exposure effects by exploiting variation across detailed technology classes, and the discrete nature of patent applications allows us to conduct events studies to measure inventors’ financial returns to innovation.

Our findings contribute to several vast literatures. The theoretical literature on the individual incentives to innovate are summarized by Scotchmer (2004). For example, there is extensive work on how different types of employment contracts will alter innovation incentives (e.g. Pakes and Nitzan (1983); Franco and Mitchell (2008)). Second, there is a growing literature on how misallocation can be a first order constraint on economic performance (e.g. Hsieh and Klenow (2009)).

agents in the economy. Our evidence does not shed light on broader effects of changes in tax rates on innovation dynamics. Indeed, beyond inventors, many other agents are involved in the innovation process, for instance firms and financiers, for whom the returns to innovation may not be analogous to a random draw (e.g. because they hold large and diversified portfolios of innovations). Moreover, our analysis doesn’t take general equilibrium effects into account.

Specifically, Hsieh et al. (2013) argue that 15-20% of US GDP per worker growth 1960-2008 can be explained by the improved allocation of talent by race and gender. We find evidence that innovation (and growth) are held back because highly talented children from low income families are not becoming innovators as quickly as their richer, but less talented, peers. Although the link between inter-generational inequality and misallocation has been frequently discussed, it has not to our knowledge ever been examined in a systematic statistical manner. We bridge the gap between the endogenous growth literature (Aghion and Howitt (1992)) and the reallocation literature (Hsieh et al. (2013)) by showing that allocation of talent affects the rate of innovation and therefore long-run growth. In contrast, the existing reallocation literature has focused on higher productivity levels as misallocation is reduced. Third, there is a literature on academic scientists (e.g. Azoulay et al. (2010a), Azoulay et al. (2010b)) where biographies can be more easily built up. There is related work looking at sub-sets of patentors, especially star scientists in bio- and nano-technology (e.g. Zucker et al. (1998)). In parallel work using the same patent- tax data merge, Jaravel et al. (2015) show that teamwork is critical in the typical inventor's career, with large spillover effects from peers across the skill distribution.

Finally, there is a literature on looking at the characteristics of patentors (see Jung and Ejermo (2014) for a survey). The classic study is Schmookler (1957) who examined 87 US patentors and the most comprehensive recent work is the PatVal-EU data (e.g. Giuri et al. (2007)) which covers 9,107 inventors filing at the European Patent Office. An issue with these studies is that sample responses are low and possibly non-random. In response to this issue researchers have recently started matching patent data to near population employer-employee administrative datasets. Toivanen and Vaananen (2012) match administrative wage data to 1,800 Finnish inventors at the US Patent Office (as do Depalo and Di Addario (2015) on matched Italian administrative data). Using this Finnish data combined with a distance to college instrument, Toivanen and Vaananen (2015) argue that access to schools offering post-graduate engineering training have a causal impact on becoming an inventor. Jung and Ejermo (2014) use Swedish patents to examine the issue of gender and age differences for just under 20,000 inventors and Dorner et al. (2014) match the IEB employer-employee data (see Card et al., 2013) to a cross section of German patents in 2002. The advantage of our data over these complementary papers is that (i) it is far larger than these other studies being at least one order of magnitude larger; (ii) we focus on the US as the country that is at the technology frontier in most industries and (iii) the US also has a relatively competitive labor market and so is less likely to depend on institutional idiosyncrasies. Finally, in terms of substantive questions, none

of these earlier papers has systematically investigated the relationship between parental income and children’s subsequent innovation. To our knowledge the only other paper to do this is the excellent (and complementary) paper by Aghion et al. (2015b) which looks at a similar set of issues extending the rich Finnish data (which also has IQ data on the male population).

The structure of the paper is as follows. The next section describes the data, Section III presents initial characteristics in early childhood; Section IV looks at schooling and exposure during later childhood and section V examines inventors in the labor market. Section VI discusses the model, its relationship to our stylized facts and its policy implications. Section VII concludes.

II Data

More details on the data are in Appendix A, but we sketch the main features in this section.

II.A Patent Data

We combine two sources of raw patent data. First, we use the several thousand weekly text and XML files of patent grant records hosted by [Google](#). The files on this page contain the full text of about 5 million patents granted from 1976 to today, extracted from the USPTO’s internal databases in weekly increments. We focus on the 1.7m patents that were granted between 1996 and 2014 to US residents. Second, we use data on 1.6m patent applications between 2001 and 2012 (Strumsky, 2014).⁴We use the names of all individuals denoted as inventors in the patent documents, not just those who are also assigned the intellectual property rights (i.e. the “self-assigned” holders of the patent rights). For example, if an individual is working for a firm, it is usually the company who will be the assignee rather than the employee who will still be named as the inventor. We define an individual as an inventor if he or she is named as such on the patent application or grant, have a US address and applied for the patent in the 1996-2012 period (to match the IRS data).⁵

II.B IRS Data

From Treasury administrative tax files, we collect information on inventors’ city/state, employer ID and adjusted gross income, as well as their current citizenship status and gender sourced from

⁴In 2001 the US moved into line with other patent offices and published patent applications 18 months after filing. Prior to this only successful applicants who were granted patents had their details published. For a fee, applicants can choose to have their filing kept secret and 15% of applicants choose to do so. The analysis in Graham and Hegde (2015) suggests that these (non-granted) applicants were of very low value. We show below the robustness of the results to considering only granted patents.

⁵So if a patent applied for in 2012 (or earlier) is granted by 2014, the individual is classified as an inventor.

Social Security records. Most data are available starting in 1996 (and currently ends in 2012). Wages and employer ID are available only starting in 1999.

II.C Matching

We match inventors to taxpayers using inventors' name, city, and state. Any inventor whose given address is outside of the United States is excluded from the matching process and dropped. We find equivalent information for taxpayers on 1040's, W2's, and other information return forms. The iterative stages of the match algorithm are described in more detail in Appendix A. We match approximately 88% of inventors of patents applied for in the last decade (the period in which information returns are most available) and slightly above 80% in the late 1990s.

We conduct various exercises to assess the quality of the match using additional data sources (e.g. data on inventor age from Jones (2010)).⁶ We also explore selection on observables and find no strong selection effects (see Appendix). Because the taxpayer data is a source of linked observations of variants of a person's name and cities of residence lived in over time, the matching process provides a simple way to link different patents filed by the same inventor even if the inventor's name differs across patents or he has moved cities.⁷

II.D Inter-Generational Analysis

For the analysis in which we study parental income, we can only look at a sub-sample of the IRS data for adults who are born after 1980, the earliest cohort for which we have sufficient records to match them to parents through 1040 forms that record dependents. For years when parents file tax returns, we calculate parents' income as the pre-tax household income; we gather parent income from W2 and other information returns in years when a parent does not file. Further details on the process of matching children to parents are outlined in Chetty et al. (2014b). Looking at the sample of individuals born 1980-84 who would be aged 28-32 in 2012 we still have a substantial sample of 45,083 inventors. This focus on "young" inventors may seem a disadvantage, but 13% of patents in our data are invented by those aged 32 or under. We also show our results are robust to using older cohorts from the Statistics of Income, which is a 0.1% sample of the IRS data available for years prior to 1996.

⁶Jones (2010) determines the ages of 55,000 inventors using name, zip code information, and a public Web site (<http://www.AnyBirthday.com>). Comparing the birth dates obtained from this website and from Treasury tax files for the inventors that are in both Jones' database and ours, we find close to an exact match.

⁷While the panel nature of the linked dataset allows us to see full income and patenting profiles for our sample, a drawback is that we cannot classify individuals as inventors who were active only prior to 1996.

II.E Test Score Analysis

When looking at the test scores of elementary school children, we further condition on a sub-sample in which we know whether children attended a New York City public elementary school between 1989 and 2009. We observe standardized test scores in grades 3 through 8 for these children, and limit our analysis to the approximately 250,000 of these students born between 1979 and 1985. Further details of the process by which these students were matched to taxpayers can be found in Chetty et al. (2014a).

II.F Initial Data Description

Table A1 contains descriptive statistics. In the whole sample there are a total of 1.2m inventors in our matched data (Panel A). The innovation data is highly skewed - the average inventor has 3.3 patents with 28.5 citations. But the standard deviation is enormous: 9.1 for patents and 154.6 for citations. The annual wage of an inventor is just over \$112,000 at the mean and \$81,000 at the median, with incomes of over \$178,000 (mean) and \$109,000 (median). These are higher than for US workers as a whole. Consistent with Hunt (2009), only 11.6% are women. The average age of an inventor is 45. Comparable data for the inter-generational analysis sub-sample and test score sub-sample are in Table A1 Panels B and C respectively.⁸

III Birth and Early Experience

III.A Parental income

Figure 1 shows (solid blue circles) the number of inventors per 10,000 individuals (left hand vertical axis) ordered by the parents' percentile position in the national income distribution (x-axis).⁹ We measure the latter as average household income 1996-2000, which Chetty et al. (2014b) have shown to be a good proxy for permanent income. A sharp, quite convex upward slope is apparent. Children born to the richest 1% of parents had invention rates of 8.3 in every 10,000, which is an order of magnitude higher than the proportion of inventors born in the bottom half of the income distribution (0.85). One hypothesis is that kids from richer parents are more likely to produce more low-value patents, but the distribution of more prolific inventors is not dependent on parent income. To test this hypothesis, we repeated the analysis using as an outcome whether an inventor

⁸The low levels of income and wages are because these are younger people who will usually be earning no income while at school.

⁹We use the 1980-82 birth cohort, but similar results are apparent using the 1980-84 or cohorts from individual years.

was in the top 5% of her age group’s lifetime citation count. The green triangles corresponding to the right-hand vertical axis show that the positive relationship between parental income and invention is just as strong for high-quality patents as it was for all patents.¹⁰ As noted above, we test for whether the results are specific to looking at young inventors by using the Statistics of Income 0.1% IRS sample. We take a cohort born in 1970-72 (ten years younger than those in Figure 1) and examine the fraction of inventors aged 30-40 (instead of aged 30-32). Figure A1 show that the strong gradient between parental income and inventor status is clearly visible in this older sample (although it is noisier due to smaller sample size).

This inventor-parent income relationship has never been comprehensively documented before to our knowledge. But one could regard the relationship in Figure 1 as deeply unsurprising. We would expect that parental income was positively associated with many other indicators of “elite” success such as becoming a successful lawyer, physician, hedge fund manager or economist, for example. 9.7% of children born to the richest 1% of parents stayed in the top 1%, compared to only 0.3% of children in the bottom 50% getting into the top 1% (see Figure A3). We focus on inventors because there is ample evidence that there are positive spillovers from innovation and therefore the social returns are greater than the private returns.¹¹ Hence, understanding barriers to the creation of more inventors is more important for public policy than the supply of hedge fund managers or lawyers. Furthermore, the focus on inventors also enables us to implement empirical strategies, such as the use of detailed technology class information in patent documents, to shed light on whether the relationship in Figure 1 is related to the childhood experience of potential innovators. Such strategies would be harder to implement with other professions.

There could be many reasons for the relationship in Figure 1. For example, since income is related to ability and this human capital is partly genetic, the relationship could be due to inherited ability. Alternatively, it may be that a poorer child begins life equally gifted as a richer one, but faces barriers to becoming an inventor. To investigate these issues we turn to the New York City (NYC) schools data where we know standardized statewide test scores of children in grades 3 to 8 in math and English. Figure A4 shows that the inventor-income gradient holds in this sub-sample (cohorts born in 1979-84). We use a parental income split at the 80th percentile and label the “rich” those above this and the “poor” those below. This is somewhat arbitrary, but other parental

¹⁰Similarly, Figure A2 shows that if we use alternative definitions of inventor status such as just patent grantees or just the post 2001 applicants data we again obtain a ratio of top 1% inventor rates to bottom 50% rates of about 10 to 1.

¹¹ For example, Bloom et al. (2013), Jones and Williams (1999) and Griliches (1992))

income splits produce similar qualitative results to everything we will show below.¹²

Figure 2 shows the kernel density of third grade math test scores for rich and poor kids.¹³ The rich kids' distribution is strongly shifted to the right as we would expect. Only 7% of poor kids are in the top decile of the math score distribution compared to 23% of rich kids. Panel A of Figure 3 shows the proportion of children who grow up to be inventors as a function of their math test scores in third grade. The pattern is striking. Children in the top 5% of the math distribution have a future innovation rate of over 5 in a 1,000 people whereas those in the next 5% have innovation rates of just over 2.5. Those in the bottom 90% have innovation rates of around 0.5.

These results can be shown in a regression setting (Table A2) where the dependent variable is whether the child grows up to be an inventor (scaled by 1,000). Column (1) shows that third grade math scores are highly predictive of becoming an inventor - a standard deviation increase in math test score is associated with a (highly significant) increase of 0.85 in 1,000 chance of becoming an inventor. Column (2) substitutes standardized grade 3 English test scores which enters with only a slightly lower coefficient of 0.68. Column (3) includes both math and English together and shows, interestingly, that conditional on math test scores (which remain highly significant) English test scores are insignificant. Column (4) has a less parametric version of column (3) which uses dummies for each vingtile of the English score and vice versa in column (5). Whereas math scores remain positive and significant in column (4) English scores are completely insignificant in column (6). This implies that early math ability is very informative for future inventor status whereas English performance is not. Note that this is not the case when re-estimating these equations but using a dummy for whether a child ended up in the top 1% of the income distribution as an outcome. The final three columns use this as the dependent variable and show that both English and math are significant in such a regression.¹⁴

Panel B of Figure 3 shows that there is a positive relationship between early math test scores and the chances of becoming an inventor for both rich and poor children and again, the relationship is noisy until we get to the top decile of the math test score distribution. It is striking, however, that for children in this top decile, rich kids have a much higher invention rate than poor kids. About 7 in every 1000 rich kids in the top 5% of math test scores become inventors whereas the

¹²All results available upon request.

¹³We are certainly not claiming that all third grade math scores are genetic. Rather we are using this as an indicator of early childhood environment and initial ability, which we want to distinguish from later childhood environment as children grow up.

¹⁴These results are also robust if we instrument one test score with another to address the issue of noise in the test scores. See Kahneman and Ghiselli (1962) for a discussion.

rate is less than half this for poor kids. This strongly suggests that the innovation-income gradient cannot solely be attributed to early test scores.

We can quantify the role of early test scores in accounting for the innovation-income relationship in different ways. Table 1 (Panel A) provides a calculation based on the full distribution decomposition methodology of DiNardo et al. (1996) (“DFL”). We give the children below the 80th percentile of parent income the same math test scores as those from richer families (specifically, we divide the distributions into 5% vingtiles bins to do this). It is important to do this across the whole distribution rather than just at the mean (as in a conventional Oaxaca-Blinder decomposition) because most of the increase in innovation probability is in the right tail of the ability distribution as already shown in Figure 3. We calculate that the test score difference in third grade accounts for just over 30% of the difference in innovation propensities in later life. So this is a sizable proportion, but obviously there are other factors that must account for the majority of the difference.¹⁵

III.B Gender

Our data shows that only a small proportion of inventors are female. In Figure 4 we show the fraction of female inventors by birth cohort. Clearly more women are becoming inventors over time: only 3% of inventors born in 1940 were female, whereas this had risen to 15% for the 1980s birth cohort. This is a very slow rate of convergence, however. Extrapolating this line forward suggests it will take 140 years before women reach parity with their male counterparts. There has been much recent discussion about the causes of this difference and its persistence.¹⁶

Figure 5 illustrates the relative similarity of the distribution of math test scores across both genders. Boys have a slightly higher score at the mean and are slightly thicker in both tails.¹⁷ We

¹⁵In Table A3 we show several alternative ways of quantifying the contribution of test scores to the inventor-income gradient which reach broadly similar conclusions. Looking across the first row for third grade test scores using the balanced panel leads to a contribution of 35.5% (a bit higher than our baseline contribution of 30.1%). Using a split at median parental income (instead of the 80th percentile) leads to a smaller, 27.8% contribution. The last three columns use a method of introducing 20 dummy variables for each vingtile of the test score distribution and observing by how much the coefficient falls on the “rich” income dummy variable. This produces somewhat larger estimates of 49.4% in the baseline, 41.5% for the balanced panel and 43% for splitting at median income.

¹⁶See Hunt (2009), Thursby and Thursby (2005), or Ding et al. (2006). There has been much controversy over this reported in the media. For example, the Nobel Laureate Tim Hunt argued in 2015 that women did not work well in the high-pressure culture of academic R&D labs (<http://www.dailymail.co.uk/news/article-3117648/Ban-women-male-labs-distracting-cry-criticised-says-Nobel-prize-winner-Sir-Tim-Hunt.html>). Larry Summers speculated that the lower proportion of female elite scientists could be because there was a greater variance in men’s intrinsic ability than in women’s (e.g. <http://blogs.scientificamerican.com/the-curious-wavefunction/why-prejudice-alone-doesnt-explain-the-gender-gap-in-science/>).

¹⁷This is consistent with Machin and Pekkarinen (2008), who show that boys’ test scores had significantly higher variance than girls’ scores. They examine the OECD’s standardized PISA math and reading tests taken by 15 year

conducted an analogous decomposition exercise for women as we did for parental income in Table 2. This shows that we can account for very little of the gender-innovation gap using math test scores. We account for nothing at all using third grade test scores and only about 3% using eighth grade scores. So, in stark contrast to income, the gender gap in innovation does not appear to be ability related.

III.C Race

There is a small literature documenting racial differences in invention rates by race (e.g. Cook and Kongcharoen (2010)). Figure 6 uses the NYC data (where we can observe race) to show that there are wide disparities in patenting rates by minority status.¹⁸ The first blue bar shows that white children have an inventor rate of 1.6 in 1,000, which is more than three times the rate for black kids (0.5) and eight times the rate for Hispanics (0.2). By contrast, Asian children are twice as likely to grow up to be inventors than whites (an inventor rate of 3.3). We can implement the DFL decompositions to see how much of these differences can be accounted for by third grade math test scores. The second red bar in Figure 6 does this for each racial group where we take white kids as the base and normalized to 1.6. We can see that the bars are not changed very much by this reweighting, with each gap shrinking by only 0.1. For example, the Black-White gap shrinks from 1.1 to 1.0, a change of under 10%. By contrast, re-weighting by income (the third green bar) makes a much bigger difference with the Black-White gap falling by almost half from 1.1 to 0.6. This suggests income is much more important than ability in accounting for the inventor difference between blacks and whites. Income makes less difference for the white-Hispanic gap, however, falling from 1.4 to 1.3. The White-Asian gap actually widens from 1.5 to 2.6 when we reweight by income as Asian parents in NYC public schools are on average poorer than white parents. Figure 7 shows the race results in another way. If we plot the inventor-test score gradient by racial group there are hardly any differences except in the top 15% of the ability range. It is amongst the most gifted at math that the differential invention rates by race become clearly visible. For third graders in the 10% of the math test score distribution, future inventor rates are over 8 for Asians, about 4 for Whites and about 1 for Blacks and Hispanics.

olds. In 37 of the 41 countries examined the boy-girl variance was significantly different for math.

¹⁸The share of students by race is: Asian 7.46%, Hispanic 32.81%, black 38.95%, white 20.78%.

III.D Implications for Models of Talent Allocation

One interpretation of these differences in invention rates by income, race and gender is that they reflect barriers to becoming an inventor over and above the intrinsic ability and preferences of individuals. In a recent contribution, Hsieh et al. (2013) develop a Roy model of heterogeneous occupations which have barriers to entry that differ by group (e.g. race and gender). These barriers or frictions are a combination of direct discrimination, in which individuals are paid less than their marginal product, and barriers to the acquisition of the type of skills that are useful to enter the occupation (e.g. a legal training to become a lawyer). This causes a misallocation in which a talented individual may not sort to the occupation that best fits her comparative advantage. The frictions cause misallocation and a loss of welfare. Calibrations in Hsieh et al. (2013) suggest that improvements in the allocation of talent (primarily from less barriers to women) were responsible for 15% to 20% of the growth in US output per worker between 1960 and 2008.

One might, however, be skeptical over whether these type of rational sorting models will generate first-order welfare losses. Consider such a model applied to our context of inventor careers. Suppose that children from poorer families have difficulties in accessing high-quality schools to improve their math ability, increasing the cost for them to become successful inventors. To the extent that such barriers exist, rational sorting models imply that they will dissuade only marginal inventors. The most talented individuals will still be prepared to make the effort to acquire skills necessary to become an inventor. It is not “Einsteins” who are lost to the R&D sector, but rather those of more mediocre ability. By contrast, we discuss below a model where disadvantaged groups are relatively less informed and underestimate the net benefits of an inventor career. In this case, even some very talented potential inventors will not pursue a career in innovation because they miscalculate the true expected cost-benefit ratio. There could be more first-order welfare losses in such a model as the breakthroughs of possible Einsteins may be lost.

To investigate whether the rational sorting model is the most plausible explanation of the patterns we observe, we implement a simple non-parametric test. This class of models has a clear prediction that conditional on becoming an inventor, individuals from the discriminated group should be, on average, of *higher* talent than the non-discriminated group. We investigate this in two ways in Figure 8. First, we use our third grade math test scores as a measure of ability. In Panel A we show the mean math grades for inventors split into (a) whether their parents were rich or poor, (b) whether they were white or from a minority group (black or Hispanic) and (c) whether

they were male or female. The results are striking - in no case are the mean test scores higher for the “discriminated group” as rational sorting models would suggest. If anything, it is the opposite with rich kids and whites having significantly higher scores conditional on being an inventor (the male-female gap is insignificant).

The second empirical way we implement our test of rational sorting is to use patent citations as a measure of inventor ability.¹⁹ We define a highly cited patent as one which is in the top 5% of future citations for its cohort (as in Figure 1). Again, we find no evidence for the basic rational sorting model, which predicts that the quality of innovation should be higher for the disadvantaged group. Conditional on patenting, mean citation rates are similar (or higher) for men, whites and those from high-income backgrounds.²⁰

The upshot of this discussion is that the rational sorting models may be underestimating the loss from misallocation in our context. To consider some alternative explanations of how the allocation of individuals across sectors can occur, we next turn to the conditions under which inventors grew up.

IV Inventors pre-labor market

We look at two aspects of the period of time between early school and joining the labor market for potential inventors: direct measures of later schooling (sub-section IV.A) and then “exposure” to innovation (sub-section IV.B).

IV.A Schooling

To investigate the importance of human capital acquisition we repeated the DFL decompositions in Table 1 but exploit information from later grades (NYC data goes through grade 8). Panel B shows the DFL results. As children get older, test scores account for more of the inventor-income gap. By 8th grade 53% of the gap is accounted for, compared to 30% in grade 3. On average an extra 4.4 percentage points of the gap is accounted for each year by test scores and the null hypothesis that there is no additional explanatory power of the later grades is rejected (p-value =

¹⁹A disadvantage of this first test is that it might be privately optimal for high-ability disadvantaged groups to go into other high-skilled occupations such as investment banking (we do not have an early measure of inventor-specific ability). The second test does not suffer from this drawback.

²⁰Note the smaller confidence bands on gender is because in Panel B we can use the entire matched IRS-patents database whereas in Panel A we just use the NYC data as we are restricted to using test scores. The large confidence bands for minorities is because there are very few highly cited patents in these cells and because minority status is only available for the NYC sample.

0.025).²¹ Low-income kids fall further behind richer ones as they progress through school. In Panel B of Table 2 we present decompositions for gender. Although test scores do become more important as children pass through school, the additional explanatory power of schooling is very low in explaining the gender gap for innovation. By 8th grade, test scores still only account for under 9% of the advantage of boys over girls. This suggests that explanations for gender differences in invention are not ability-related, at least as measured by test scores.

For the individuals who attended a US college between 1999 and 2012 (200,000 inventors) we can determine which college they attended. One striking fact is that a few select colleges produce an enormous fraction of US inventors. For example, the top 10 most innovative colleges (defined as the colleges whose graduates will be granted the most patents) account for only 3.7% of enrollments, but 15% of (citation-weighted) patents (Figure A5). In these top 10 colleges, 5.6% of students have a patent by age 30 (about 28 times the average for the population as a whole by this age). In the 10 most innovative colleges, the invention-parental income relationship is severely attenuated. Figure 9 shows that in this sample children whose parents were in the (nation-wide) top 1% of the income distribution produced 63 patents per 1,000 (compared to 0.1 per 1,000 in Figure 1), compared to 43 per 1,000 in the bottom half of the income distribution. This ratio of about 1.5 to 1 compares to 10 to 1 in Figure 1. Hence, by the time students graduate from a school like MIT or Stanford, the income of the student’s parents makes relatively little difference. The main role of higher income is increasing the chance of a child going to one of these colleges - Table A7 shows the sharp positive gradient between being born into a wealthy family and attending an innovative college.

IV.B Exposure to Innovation

We next explore the role of being *exposed* to innovation in childhood and growing up to be an inventor. We examine exposure along three dimensions: (a) whether an individual’s parent was an inventor (sub-section IV.B.1), (b) for children of non-inventors, what was the innovation in the industry of one’s father (sub-section IV.B.2) and (c) innovation in the area where the child grew up (sub-section IV.B.3). For all three dimensions we examine not just the overall level of innovation, but the *type* of innovation by exploiting the information on patent technology class. Patents can be classified into seven broad categories,²² thirty seven sub-categories and four hundred and forty five classes. Looking at whether exposure to specific technological fields (rather the innovation in

²¹The lower rows of Table A3 show this is equally true using other decomposition methods (up to 61% by grade 8 in the median split specification of column (6)).

²²The classes are Chemicals; Computers; Communications; Drugs and Medical; Electrical and Electronic; Mechanical and “Others”.

general) is correlated with becoming an inventor in that specific field is a sharper test of whether becoming an inventor is affected by the environment rather than just being an innate trait.

IV.B.1 Parental Innovation Status

We analyze innovation rates for 16 million children born between 1980 and 1984 for whom we know whether their parents themselves filed a patent since 1996. Amongst those whose parents were inventors, the patent rate was 11.1 per 1,000. By contrast, if a child’s parent was not an inventor then the patent rate was only 1.2 per 1,000. A fraction of these were children-parent teams on the same patent. However, even if we remove these observations, then it is still the case that the inventor rate is 8.5 for the children of inventors.

Of course, this relationship within the family could reflect a genetic predisposition to be an inventor. To address this we examine the patent class in which the parent invented: it is very unlikely that there is a gene that codes specifically for a type of technology class such as “modulators” (technology class 332) vs. “demodulators” (technology class 329) or synthetic resins (class 520) vs. natural resins (class 530). Conditional on inventing any technology, children of inventors are nine times more likely to invent in the same sub-class as their inventing fathers than they would be by random chance.²³

We also develop a closeness measure based on looking at whether a patenter in class A also patented in class B. This exploits the fact we have individual identifiers due to the IRS match and can examine cross-class patenting by individuals over time.²⁴ Table A4 gives an example. Starting with technology class 375 (“pulse or digital communications”) it has a distance of zero with itself by definition. For inventors who had a patent in pulse or digital communications, the next most popular class to patent in was demodulators so we give this a distance of 1. After this comes modulators (distance = 2) and so on. Using this information, Figure 10 shows the proportion of children of inventors patenting in the same technology class and then in other “close-by” classes. We calculate the proportion for each technology class and then average across all technology classes with weights proportional to the number of patents in each technology class. There is a clear spike in the probability of children of inventors inventing in the same technology class as their parent. On average, if an individual’s parent invented in a particular technology class, the child had over a 0.9 in a 1,000 chance of inventing in exactly the same class. The chance of inventing in the next

²³0.9% of children invent in the same class as their fathers compared to an expected 0.1% by random chance (see Table A9).

²⁴Other distance metrics in the literature include classes within the same sub-category and cross class citations (see Bloom et al. (2013)).

“closest” technology class (at distance = 1) was under 0.2, about a fifth as high. As the class becomes more distant, the child’s probability of investing in that class diminishes (graphically, the downward slope going from left to right of Figure 10). This evidence is consistent with the view that being brought up in a family where there is some very specific knowledge about a technology helps the younger member of that family to go and become inventors in that technology.

IV.B.2 “Mentors”: Innovation in father’s industry

The direct relationship between parent’s and children’s inventor class is striking, but of course this is focused on a narrow slice of the data: those with parents who have invented. Our second measure of exposure looks at characteristics in the (six-digit) industry where a child’s father worked. The idea is that the network of people in the firm and industry could influence what careers young people are interested in studying and pursuing in later life. Table 3 puts this results in a regression framework, where the data is constructed solely from all children whose parents were *not* inventors (in order to rule out the direct channels examined in the previous sub-section). Column (1) is run at the six-digit industry level. For each individual we calculate the proportion of workers in their father’s industry who were inventors (right hand side variable). The dependent variable is the proportion of children who became inventors (within a father’s industry). There is a strong positive and significant relationship between these two variables.²⁵ A one standard deviation increase in the fraction of inventors in the father’s industry (0.0023) is associated with an increase of 0.0006 inventors, or 25.3% at the mean of the dependent variable (0.0023).

In column (2) of Table 3 the unit of observation is an industry (345) by patent category (7) cell. We construct the data in the same way as before but calculate the proportion of inventors within an industry-category cell and then include category fixed effects in the regressions. Column (3) goes one step further and constructs cells defined by industry and sub-category (including 37 sub-category fixed effects in the regression) and column (4) is the most disaggregated using industry by technology class cells (with 450 technology class fixed effects). The same pattern emerges in columns (2)-(4) as in column (1). Children are much more likely to innovate within a narrowly defined technology area if their father has worked in an industry that has more inventors in this same area. Column (5) replicates column (4) but includes (i) the fraction of inventors in the same sub-category but in a different industry class and (ii) the fraction in the same sub-category but a different class and (iii) other categories. All coefficients in these variables are positive and two are

²⁵Figure A10 has a the scatterplot of this relationship.

significant (the own class coefficient drops slightly but remains significant).²⁶

IV.B.3 Innovation in childhood neighborhoods

Our third measure of exposure looks at the geography of innovation. The existing literature examines the effects of place on economic outcomes in general and the effects on innovation in particular. Marshall’s theory of industrial districts suggested that one advantage of geographical clusters was that there were “ideas in the air” and accounts of the success of Silicon Valley have also stressed the benefits of geographically localized innovation spillovers.²⁷ From our data we know where all individuals were grew up and in Figure 11 we present the invention rates by childhood Commuting Zone (CZ). Note that the figure does not present the patenting rates based on where inventors are *currently* residing (this is generally used in the literature because location at time of invention is available in the USPTO data), but rather the future patenting rates in commuting zones where inventors grew up. Figure 11 shows an interesting spatial pattern: innovation hotspots are in the Bay Area, Northeast and Great Lakes, which is expected as there are strong universities located there. But there are also many other innovation pockets around the country - in the North-West, Utah and Colorado, which may be less obvious hotspots. Very low innovation areas are found in the Deep South.²⁸

These relationships cannot be read as causal place-based effects as there are many other unobservable factors associated with place of birth and future outcomes. We probe this relationship more in Table 4 which is analogous to Table 3 and again focus on children whose parents were not inventors. Column (1) is the baseline regression where the dependent variable is the fraction of kids who lived in a commuting zone that grow up to be inventors and the key right-hand-side variable is the invention rate in the childhood commuting zone (as in Figure 11). There is a strong and significant relationship between the two (as also illustrated in the scatterplot in Figure 12 for the largest 100 commuting zones) showing that children who grow up in innovation-intensive neighborhoods are more likely to become innovators themselves. According to column (1) of Table 4, increasing the fraction of inventors in the childhood commuting zone by a standard deviation (0.0002) is associated with a 30% increase in invention rates.²⁹

²⁶Figure A11 shows a bar chart of these industry regression coefficients grouped by distances. The falloff from innovating in the exact same class to the next closest class is very striking.

²⁷There is a large literature on inventor mobility and agglomeration effects from localized knowledge spillovers, e.g. Kim and Marschke (2005).

²⁸The list of the top 10 and bottom 10 most innovative CZs are in Table A5.

²⁹ The mean of the dependent variable is 0.002, so using the coefficient estimate $(0.0002*2.9006)/0.002 = 0.30$.

Column (2) of Table 4 reports estimates at the individual level and includes 390 dummy variables for the commuting zone where the individual was living in 2012. The coefficient is essentially identical to column (1). This is effectively exploiting the origin-destination CZ matrix of where a child grew up and where they currently live (almost 390*390 cells). The coefficient indicates that, for example, for two adults currently living in Chicago, the fact that one grew up in high-innovation Cambridge makes it more likely she will be an inventor than another who grew up in lower-innovation Little Rock. In column (3) we use CZ by patent category level (analogously to Table 4 column (2)) and include seven category fixed effects. Column (4) disaggregates this further by the CZ where the child grew up (so a cell is the current child’s CZ by childhood parent CZ by category). Like column (2), this shows that there is a positive association with exposure even after controlling for the CZ where the child is currently living. Column (5) disaggregates the current child CZ by patent category and by father’s industry. We can control for industry dummies here to show that the CZ exposure is picking up more than simply the parent industry exposure that we analyzed in Table 3. Column (6) is like column (3) but at the CZ by sub-category level (and includes sub-category fixed effects) and columns (7) and (8) are at the CZ by technology class level (regressions include class fixed effects). This table illustrates that not only are kids who grow up near inventors more likely to be inventors themselves, but these children also have an increased propensity to innovate in the same types of fine-level technologies as they were exposed to during childhood.³⁰

IV.B.4 Exposure to female inventors

We can also use our empirical strategy of exposure to innovation in general to look more specifically at female inventors. We can calculate the proportion of female inventors in the state where children grew up and we present the map of this in Figure 13.³¹ The Northeast appears to have the most female inventors and the North and mid-West the least. The top and bottom 10 CZ for female inventors are in Table A6. These female invention rates are correlated with the Pope and Sydnor (2010) Gender Stereotype Adherence Index on 8th Grade Tests (see Figure A13). It seems plausible that areas where these cultural attitudes about women are stronger are less likely to generate female inventors.

³⁰Figure A12 shows a bar chart of these CZ regression coefficients grouped by distance. The fall off from innovating in the exact same class to the next closest class is very striking

³¹Unfortunately, there are not sufficient numbers of female inventors to do this reliably at the CZ level

IV.B.5 Summary of findings on Exposure

A clear pattern emerges from this section that across all three measures of exposure - parents, parents' colleagues and neighborhoods - growing up surrounded by more innovation is associated with an increased probability of becoming an inventor, even within a detailed technology class.

V Inventors in the labor market

V.A Wages and Income

The distribution of inventor income is given in Figure 14. It is heavily skewed with a median of \$114,000, a mean of \$192,000 and \$1.6m at the 99th percentile. It is far more skewed than occupations such as doctors, architects or lawyers. Only the financial sector has similar degrees of skewness (Lockwood et al., 2014). Examining the composition of income for inventors, we observe that it is dominated by salaries - like for most workers. Focusing on the top 1% of inventors by income, non-wage income is more important than wage income. However, even here patent royalty income is not so important: other forms of non-wage income dominate (Figure A14).

Next, we examine income dynamics around the time of invention in Figure 15. Panel A shows median income in the 10 years before and 10 years after the patent event. Interestingly, we do not see an increase in income after applying for a patent. In fact, there is a much stronger increase in income prior to patenting: afterwards income flattens off and falls a bit. Panel B shows the pattern for mean wages and Panel C for the 99th percentile to check if the median is missing out on higher quantiles. But the picture looks similar with the main returns occurring prior to the patent event (although the flattening out now happens about three years after the patent is applied for). Panel D examines the proportion of income that is non-wage. Here we do see a clear break in trend with the fraction rising afterwards much faster than it was before. Panel E looks at alternative measures of the quality of patents. The patterns look similar, although ungranted appear less valuable than granted and highly cited patents (i.e. those in the top 5% of the future citation distribution by cohort have faster growth of income in the pre-patent period and the fall off of income in the post-patent period is less severe).³²

The patterns in Figure 15 are for prime-aged workers (35-50 years old) to abstract away from the fast growth of earnings that occurs for younger workers. Figure 16 shows the same median income “event studies” for patents applied for at exactly the ages of 30, 40 and 50. The pattern of

³²Note that mean income is higher in general for inventors with more citations (see Figure A15). This is a new external corroboration of the value of citations as indicators of quality.

a fast growth prior to patenting followed by a leveling off/fall is apparent in these graphs as well, suggesting the dynamic patterns are occurring at all ages. It is not some conflation of lifecycle age-earning profiles with patent events.

Our interpretation of the data is that there are some individuals who are on a successful innovation streak, culminating in applying for a patent. But the patent event is not news to the firm or labor market. The eventual patentee is obtaining rewards prior to patenting rather than afterwards. This is consistent with the evidence from matched patents-income administrative data from Italy in Depalo and Di Addario (2014) who also found big increases in income prior to patenting, but not afterwards.

These patterns suggest a model of career choice with an uncertain chance of a major innovation, rather than a choice about how and when to patent *per se*. The model in Section 6 builds on this insight.

V.B Inventor Age

The most common age of patent applicants is approximately 39-40 (Figure A16 Panel A). However, this should be compared to the overall age distribution of US workers in the whole IRS population (Panel B). It is clear that most inventions tend to come from inventors who are older than the working population as a whole. Innovation still peaks around 40, but trails off more gradually than in Panel A (there is a 32% decline from 2.5 per 1,000 wage earners at age 40 to 1.71 at age 60). Does this over-estimate the abilities of the middle-aged? For example, could the patents of older workers be of lower quality and the most important work still be done when researchers are young? Panel C uses only highly cited patents and finds a similar pattern except with a sharper decline in innovation at older age (from 0.14 at age 40 to 0.06 at age 60: a decline of 57%).³³ Therefore, consistent with the “young and restless” hypothesis of Acemoglu et al. (2014) and the evidence in academia of “great” discoveries (e.g. Jones et al. (2014)) raw patents do under-estimate the importance of youthful innovation. However, there does seem to be a lot of innovation done at older ages too. Thus, the returns to innovation accrue to inventors relatively late during their career, which speaks to the importance of occupational choice early in life (extensive margin), as opposed to a choice “at the margin” to adjust one’s innovative effort during one’s career (intensive margin).

³³Figure A16 shows that overall patent citation rates are correlated with individual income which is another corroboration of the idea that citations are a useful measure of patent value.

V.C Firms

We know what firms inventors work in from the EIN number attached to their W2s. Corporate structures are complex, however - for example, a US parent company may have multiple affiliates, the ultimate headquarters may be in another country and there are partnership structures as well as firms. We simply work with the EIN numbers and discuss some of the data issues in Appendix A.³⁴ Figure A17 has the CDF of the size distribution for inventors. 10% of inventors do not receive a W2, presumably because they are self-employed. But the bulk of inventors work and mostly they work for larger firms. 70% of inventors work in firms with more than 100 employees. Hence, it is unlikely that the pattern in Figure A17 could be due to individuals simply being unable to afford the fees associated with applying for a patent. It is the companies that inventors work for who are applying and paying for these fees. And although in a small firm the inventor himself may have to contribute, this is highly unlikely to be the case in a larger firm. This suggests that the influence of parental income on inventor status is likely to come through it's effect on the child when growing up, rather than after starting work in the labor market. This is consistent with the evidence from the entrepreneurship literature. Hurst and Lusardi (2004), for example, find that the correlation between parental wealth and entrepreneurship is unlikely to operate via relieving liquidity constraints after leaving school.

V.D Summary

In summary, the key findings from our labor market analysis are (i) there are substantial returns before the patenting itself; (ii) patenting returns appear very skewed: there is a small chance of a very high payoff (cf. Hall and Woodward, 2010, on entrepreneurs) and (iii) many returns are late in an inventor's career (cf. Jones, 2009, 2010) . Hence pay-offs are highly uncertain when individuals are making initial career choices.

³⁴The distribution of employment by firm size class in the IRS dataset, where each firm is assumed to be represented by a unique EIN, is almost identical to the size distribution across the US economy from the Economic Census. Despite many reasons why the Census Bureau concept of an "enterprise" could be quite different from the tax-based EIN, the match with the firm size distribution is very good. There appears to be a slight under-representation of firms with under 100 employees in the IRS data compared to the Census, and a slight over-representation in the 100-1000 employee range, but almost identical proportions in other size bins compared to the Census. And even the discrepancies are small. Narrowing the IRS sample to be closer to the Census in terms of industry composition does not fundamentally change this picture.

VI A simple model of inventor careers

VI.A Basic Model

Motivated by the stylized facts from our analysis we develop an inventor lifecycle model that is detailed in Appendix B. The model has similarities to a two-sector version of Hsieh et al. (2013). In period 1 human capital (H) is determined, which will depend on initial talent (which is heterogeneous across individuals and groups) and schooling.³⁵ At the end of schooling, agents make occupational choices over whether to enter the R&D sector or non-R&D sector on the basis of their expected utility. The R&D sector is different from the non-R&D sector in three respects: (i) Income is stochastic: there is a base wage (as in the non-R&D sector) but there is also a chance (π) of making a successful innovation; (ii) those with high human capital have a comparative advantage in the R&D sector - formally, we model this as allowing the probability of innovating to be higher when human capital is higher ($\pi(H), \pi'(H) > 0$); (iii) individuals have idiosyncratic preferences over the two sectors (hence, there will typically be some mass of agents who go into the R&D sector even if their expected monetary returns are lower). There is a two-part tax regime with a standard marginal tax rate up to a threshold and then a high “top” tax rate.

We consider that there are several disadvantaged groups, g , in the population. The canonical example we focus on are children born to low income families (so the groups are rich vs. poor), but the framework is equally applicable to considering men vs. women or whites vs. blacks. We allow for these groups to face additional costs in the acquisition of human capital (which reduces their likelihood of entering the R&D sector).³⁶

We depart from the standard rational sorting models by adding imperfect information. We assume that individuals with probability λ are correctly informed about the net returns to an inventor career, but others (with probability $1 - \lambda$) under-estimate the true returns from the R&D sector. The idea is that many children do not often come into contact with inventors via their parents, family networks or neighborhoods. Hence these less “exposed” groups underestimate the net benefits of choosing an inventor career (e.g. Hoxby and Turner, 2014). The precise way we formalize this notion is that these less-informed agents underestimate the degree of complementarity

³⁵In the baseline model we keep the human capital acquisition process exogenous (e.g. high income families “buy” more educational quality for their children). But we extend the model to allow for endogenous educational choice. This reinforces the misallocation results as disadvantaged groups choose to invest less in human capital as they are less likely to want an inventor career.

³⁶Hsieh et al. (2013) model this as a “tax friction” on spending goods in obtaining human capital. Another friction is direct discrimination in the labor market by paying disadvantaged groups a lower wage than their marginal product (in the R&D sector compared to the non-R&D sector). This is less likely for low income groups, but could be the case for women or minorities. In the context of their model these are observationally equivalent.

between human capital and innovation. We show a number of intuitive results using this model.

First, agents with higher human capital are more likely to enter the R&D sector (this follows directly from their comparative advantage). Second, individuals who are more exposed to inventors (our proxy for a higher value of λ) are more likely to enter the R&D sector. Third, disadvantaged groups are less likely to enter the R&D sector. This is because (i) they face barriers to human capital acquisition and (ii) they may begin with a lower ability level (for example, we saw the 3rd grade math scores of low income groups were lower), and (iii) they have worse information about the net benefits of an inventor career.

A fourth result is that conditional on being in the R&D sector, agents from disadvantaged groups are likely to have lower levels of initial ability. This is the opposite prediction from the standard rational sorting models which predict higher levels of average ability for disadvantaged groups in the R&D sector, because only the most talented will overcome the barriers to entry. By contrast, in this model agents from disadvantaged groups are imperfectly informed about the complementarity between human capital and innovation. Therefore, the high ability agents in these groups do not always go into the R&D sector, whereas for privileged groups the high ability agents do.

The empirical evidence lines up well with these predictions. We find that people with higher human capital (test scores and elite college attendance) are more likely to be inventors, and that disadvantaged groups are less likely to become inventors. Section 4 presented much evidence that early exposure mattered a lot for future innovation. Finally, Figure 8 showed that conditional on being an inventor, disadvantaged groups did not appear more talented than other groups. This is consistent with our imperfect information model but not with the basic rational sorting model

The normative implications of our model is that there is the potential for considerable welfare loss. First, there is a “level effect” - too few individuals from disadvantaged backgrounds enter the R&D sector. This effect is similar to Hsieh et al. (2013). Second, and in contrast with standard rational sorting models, there is also a “composition” effect: the “wrong” individuals from disadvantaged groups may enter the R&D sector due to information frictions (i.e. not the individuals with the highest comparative advantage for the R&D sector). And these effects are compounded because the disadvantaged groups will likely face other barriers. There is a loss both from fewer (externality generating) inventors and from an inferior allocation of talent compared to the first best.

VI.B *Alternative Interpretations of Exposure Measures*

We have interpreted our empirical measures in terms of exposure influencing λ . But one alternative is that the exposure measures actually improve inventor-specific human capital as the young person grows up. This would be an environmental effect, but the welfare effects are somewhat different from the model of the previous section. At the time of occupational choice, those exposed in childhood would have higher human capital rather than more information (λ). Hence, there would not be obviously greater talent losses in this model than in the rational sorting model.

As a way to investigate this we examine the income effects of exposure. If exposure generates inventor-specific human capital then we would expect productivity (and therefore wages) to be higher for those choosing an inventor career if they are more exposed at an earlier age to science, even conditional on early test scores. However, we could find no evidence of substantially higher wages for these groups more exposed to innovation. In our model the effects of exposure on inventor income are ambiguous. Although agents are better sorted to their area of comparative advantage as information improves, they may earn a lower wage in the R&D sector as there are compensating differentials to the non-pecuniary advantages of the R&D sector (\tilde{w}).³⁷

Another interpretation of the exposure “effect” is that it could cause a change in preferences rather than in information. The wage effects of this are also ambiguous, so it is difficult to empirically distinguish this from our information story. In terms of policy, if there are externalities from innovation then increasing exposure to inventors might still be highly valuable even if it is only about changing preferences. Note that the pure preference shift story would not predict that exposure should be more important for high-ability children in their likelihood of growing up to be an inventor. It is hard to check this directly in our data because there is insufficient variation in exposure in the NYC data to include interactions between exposure and early test scores in the “inventor equation”³⁸, but this would be a good avenue for future work.

There are, of course, other ways of interpreting our results but overall, the simple lifecycle model we sketch seems reasonably consistent with the data.

³⁷For example, Stern (2004) shows that life science post-graduates take about a 20% loss in income by taking a job in academic science compared to industry.

³⁸A positive coefficient on the interaction would be consistent with our model, but not with the pure preference model.

VI.C Policies to improve entry into innovation

If policy interventions could improve the position of potentially high-ability kids from poor backgrounds (as well as women and minorities) this would bring a whole new margin of individuals into the inventor pool. Card et al. (2013) report evidence from “gifted and talented” randomized control trials that suggest that although these programs do not work well for the typical student, those from poorer backgrounds do appear to particularly benefit. Changing to such policies has a near-zero financial cost.³⁹ This suggests that such interventions could have very large benefits in terms of growth as well as equity.

Our estimates can be used to assess the potential gains from such supply-side educational policies. These policies can have a *level* effect on innovation by increasing the rate of entry into innovation of children from low-income families, up to a level closer to that of children from high-income families. In addition, the policies may have a *composition* effect by affecting which children decide to enter the R&D sector within each income group - our results suggest that talented children from low-income families are less likely to enter the R&D sector, which could potentially be affected by policy). Appendix B4 discusses these calibrations in more detail, but we sketch the findings briefly here.

Regarding the level effect, we have shown that children born to families in the top 10% of income are ten times more likely to become inventors than children born to families of below-median income (Figure 1). We have documented that innate ability differences (which by definition cannot be affected by policy) are unlikely to explain more than a third of this difference (Table 1). Moreover, we have found that the differential innovation rates across technology classes is also a ten-to-one ratio (Figure 10), which suggest that exposure effects play a key role.

We consider a benchmark scenario assuming that supply-side policies providing such exposure effects could close a fifth of the total innovation gap between children with parents below the 90th percentile of income (1.6 inventors per 1,000) and children with parents in the top 10% (6.7 inventors per 1,000). Under these assumptions and using our data on the propensity of children to become inventors across the income distribution, the increase in the number of inventors induced by the policy is a staggering 30% of the current inventor population. The details of the calculation are reported in Appendix B5, where we also report the calibrated effect of the policy under other assumptions about the share of the innovation gap across the income distribution that can be closed by policy. The sensitivity analysis shows that the effect is large under a very wide set of parameter

³⁹Personal communication with David Card.

values. We have also considered a scenario based on the distribution of innovation-income gaps across US states.⁴⁰ If the mean innovation-income gap (similar to Michigan) could be lowered to the level of the fifth percentile of the distribution (similar to states like New Hampshire), then the increase in the number of inventors would be equal to 38% of the current inventor population. Similar calibrations suggest the composition effect also matters, but by less than the levels effect (under 6%).

By contrast to these “extensive margin” policies, policies on the intensive margin may be much less effective. For example, although existing estimates of R&D tax credits do suggest increases in innovative activity, they reach only a small fraction of the population. We examine the effects of income tax on innovation next.

VI.D How much do tax policies on inventor’s income affect innovation?

A much-discussed alternative set of policies to stimulate innovation are levels of top income tax rates (Akcigit et al. (2015)). We shed light on this issue using a quantitative theoretical exercise based on our data and the lifecycle model introduced in sub-section VI.A (more details are in Appendix B.3). In short, we find that changing marginal tax rates on high incomes would *not* substantially affect the occupational choice to become an inventor. The small effect of top income marginal tax rates on inventors’ behavior is driven by three considerations. First, marginal utility decreases with income, due to risk aversion. Second, inventors’ earnings are highly skewed - as documented in Section V. Third, the rank of an inventor in the earnings distribution has a large random element, especially within the top tail. We find that income is difficult to predict for inventors, especially in the upper tail, therefore we model it as a random draw from the empirical income distribution for all inventors.⁴¹

Put another way, the intuition for the results is that entering the innovation sector is like buying a lottery ticket. With concave utility and expected utility maximization, whether an agent has a

⁴⁰All calculations are based on state of birth, not on state of residence.

⁴¹This is a strong assumption, but various features of the data support its validity to a first approximation. First, we have shown in Section V that the returns to patenting are very skewed and that they often occur late in an inventor’s career, suggesting that the returns to innovation are very uncertain at the time of occupational choice. Second, the distribution of income for inventors is very skewed even conditional on having a very highly cited patent: the 99th percentile of income for people in the top 5% in the citation distribution is \$10 million, while the median is about \$200,000. In other words, even in the subgroup of inventors who produced high-quality innovations, the mean return is coming almost entirely from the upper tail. Third, we have checked that an inventor’s earnings in mid-career, between ages 40 and 45, are very difficult to predict based on this inventor’s earnings and patenting record (number of patents and citations) in his early career, between ages 27 and 32. Specifically, regressing mean earnings between ages 40 and 45 on a flexible polynomial of mean earnings, patent count and citation-weighted patent count between ages 27 and 32 yields a R^2 below 0.03.

small chance of winning \$20 million (e.g. in an economy with no extra taxation of top incomes) or an equally small chance of winning \$10 million (e.g. in an economy with a higher marginal tax rate of top incomes) makes little difference to buying the lottery ticket. In other words, occupational choice over whether to become an inventor should not respond much to top tax rates. Our contribution is to calibrate this response using the skewness of the empirical earnings distribution of inventors. It is important to note that our approach does not say anything about the broader effect of top tax rates on the rate of innovation in the economy.⁴²

For the calibration, we consider a variety of tax regimes. We use a stylized version of the US Federal tax schedule where the top marginal tax rate is 40% above \$439,000 and the marginal tax rate below this threshold is 28.5% (“standard rate”). We then consider the impact on innovation of increasing the top rate by a percentage point to 41%, but keeping the standard rate the same at 28.5%. Since the benefit of the policy is to raise revenue for public goods, we have to benchmark this in some way. So we consider a “benchmark” policy of raising the standard rate by a percentage point (to 29.5%) but keeping the top rate the same. We then calculate the fall in the fraction of the population of workers becoming inventors per dollar of tax revenue in both cases.⁴³ This is equivalent to a (marginal) deadweight cost per tax dollar. We denote the loss of inventors per tax dollar due to the higher top rate policy as $\gamma(\tau^1)$ and the corresponding deadweight for the benchmark policy of raising the standard rate as $\gamma(\tau^B)$. We then repeat these calibrations under various assumptions about the utility function.⁴⁴

As is standard in public finance, the deadweight cost crucially depends on a behavioral elasticity, which in our context captures the extent to which number of people choosing an inventor career responds to the change in the certainty equivalent wage induced by changes in the tax system. However, this elasticity cancels out when we express the relative innovation loss of any policy change relative to the benchmark policy change described above. In other words, we focus on a summary statistic $\gamma = \frac{\gamma(\tau^1)}{\gamma(\tau^B)}$, which we can calibrate based only on the empirical income distribution of inventors, without knowledge of the behavioral elasticity, and which captures the relative efficiency loss from increasing the top marginal tax rate compared to our policy benchmark of the change in

⁴²Indeed, beyond inventors, many other agents are involved in the innovation process, for instance firms and financiers, for whom the returns to innovation may not be analogous to a random draw (e.g. because they hold large and diversified portfolios of innovations). Moreover, our analysis does not take general equilibrium effects into account.

⁴³We compute the fall in the fraction of the population of workers becoming inventors based on the change in the certainty equivalent implied by changes in the tax system. Appendix B describes this and all other steps of the calibration in detail.

⁴⁴Specifically, we consider CRRA utility functions with coefficients of relative risk aversion equal to 0 (i.e linear utility), 0.5, 1 (i.e. log utility), 1.5 and 2, respectively.

the standard tax rate (Appendix B.3.1 provides a full derivation of these results).

We present the value of γ in Figure 17 for different levels of the coefficient of relative risk aversion in the utility function. The first bar is 100% for a CRRA parameter equal to 0 (linear utility): when people are risk neutral, there is no difference in changes in top and standard tax rates. As we move to the right in Figure 18, we consider increasing levels of risk aversion. The height of the bars falls indicating the welfare loss from raising top taxes is diminishing as risk aversion increases. It is about a quarter of the innovation loss of the benchmark policy when $\text{CRRA} = 0.5$ and only 2% of the benchmark policy when $\text{CRRA} = 1.5$.⁴⁵

The intuition behind the result is clear. The innovation loss from high top tax rates is to discourage entry into the riskier R&D sector - if an inventor is lucky and wins a very valuable innovation, she will be penalized more by high tax rates. But since the returns to innovation are so skewed, most inventors obtain only low returns. There can be considerable differences in income in the tails, but the utility difference will be much less due to the concavity of the utility function.

The generality of this result should not be overstated. It is a theoretical result based on calibrated parameters in a standard expected utility set-up, which crucially depends on the assumption that the income process for inventors is as good as random. If inventors are not fully rational expected utility maximizers, they may well behave differently. Furthermore, even in the context of the model it may be that potential inventors are better at predicting their expected income than we assume. If there was no uncertainty about the returns to innovation, then the results would no longer hold.⁴⁶ Finally, our evidence does not shed light on broader effects of changes in tax rates on innovation dynamics.⁴⁷

While these caveats must be kept in mind, our point is still an important one: we show that the

⁴⁵ Appendix B shows some other illustrations of this effect when looking at larger changes in top taxes and additional thresholds. We show that the skewness of the earnings distribution drives our quantitative results: in the presence of concave utility and stochastic payoffs, taxes have a lower efficiency cost for very skewed payoff distributions compared with less skewed ones.

⁴⁶ There is a recent empirical literature on the impact of top tax rates on inventor mobility. Akcigit et al. (2015) find that differences in top marginal tax rates across countries effect mobility decisions of inventors. However, whilst statistically significant the magnitude of the effect is quantitatively small in line with our calibrations. Moretti and Wilson (2015) find larger effects when looking at inventor mobility across states in response to changes in top income tax rates. This within US analysis is interesting but it does not speak to the issue of whether people are more likely to choose an inventor career, as the empirics work from mobility once people have already chosen to become inventors.

⁴⁷ In particular, it could be the case that the returns to innovation are much less skewed for a number of agents who play a crucial role in the innovation process and hold a large and diversified portfolio of innovations, for instance investors (venture capital firms, etc.) and R&D lab managers or CEOs. In other words, it could be that an increase in top tax rates would lead these agents to exert less effort at the margin, because there is a tight link between their effort level and their financial rewards (in contrast with what the data suggests for inventors themselves). For example, Aghion et al. (2015a) suggest that the returns to innovation are broadly shared, much beyond inventors, and that innovation has contributed to the increase in top income inequality in recent decades.

effect of top tax rates on the key individuals in the innovation process - the inventors themselves - is likely to be small due to the skewness and randomness of the payoffs.

VII Conclusion

In this paper we describe the lifecycle of inventors using a match between 1.2m patentors during 1996-2014 with administrative tax data from the IRS. We are able to follow potential inventors from their conditions at birth (parental income, gender and race), while growing up (neighborhood, school test scores, college attended) and finally in the labor market (their income profiles). We show that parental characteristics matter a lot: the rich, white and male are much more likely to grow up to be inventors than the poor, female and black. Early math test scores account for a third of the inventor gap for income, a tenth for the black-white difference and almost nothing for gender. Focusing on the differences by parental income, two elements prove to be key. First, the innovation-relevant human capital gap between rich and poor opens up as they go through school, and by the time they reach college parental income makes relatively little difference. Second, children who are most exposed to innovation through their parents, mentors or neighborhoods are much more likely to become inventors.

These findings suggest there are environmental barriers to disadvantaged groups in acquiring the human capital that is complementary with innovation. The predictions of standard “rational sorting” models that inventors from disadvantaged groups should be of higher quality (in terms of patent citations and early ability) is not born out by the data. We discuss an extension to the sorting model of inventor careers in which some individuals underestimate the benefits of a career as an inventor through less exposure and show that its predictions are broadly consistent with the data. Policies that reduce the income-related test score divergence (particularly for children in the top decile of early math achievement) are beneficial from an equity stance, but might also be beneficial from an efficiency perspective as they have the potential to increase the degree of innovation and growth in the US economy. Through a series of calibrations, we have shown that such “extensive margin” innovation policies drawing talented low-income individuals into innovation may be very effective at stimulating innovation, probably more than top income tax policies.

There are many more aspects of the data we have discussed here that could be used to inform the next generation of innovation models. How important is entrepreneurship in enabling innovators to successfully monetize high-quality inventions? Can we bring more structural models to bear on the inventor career model in order to consider counter-factual policies? How important are firm

policies in better attracting and incentivizing inventors? And could the same forces which influence whether a person becomes an inventor hold true for other elite occupations? These are some of the fascinating questions our research agenda opens up for the next generation of innovation theories and policies.

References

- Acemoglu, D and J. Robinson.** 2012. *Why Nations Fail*.
- 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. and P. Howitt.** 1992. “A Model of Growth Through Creative Destruction.” *Econometrica*, 60(2), 323-51.
- Aghion, P., N. Bloom, R. Blundell, R. Griffith, and Peter Howitt.** 2005. “Competition and Innovation: An Inverted U Relationship?” *Quarterly Journal of Economics*, 120(2), 701-728.
- Aghion, P., U. Akcigit, A. Bergeaud, R. Blundell, and D. Hemous.** 2015. “Innovation and Top Income Inequality.” *LSE mimeo*.
- Aghion, P., U. Akcigit, A. Hyttinen, and O. Toivanen.** 2015. “Living the American Dream in Finland?” *LSE mimeo*.
- Akcigit, U. and W. Kerr.** 2013. “Growth through Heterogeneous Innovations.” *Bank of Finland Discussion Paper No. 1328*.
- Akcigit, U., S. Baslandze, and S. Stantcheva.** 2015. “Taxation and the International Mobility of inventors.” *NBER Working Paper No. 21024*.
- Azoulay, P., Graff Zivin J.S., and Manso G.** 2010. “Incentives and Creativity: Evidence from Academic Life.” *RAND Journal of Economics*.
- Azoulay, P., J. S. Graff Zivin, and J. Wang.** 2010. “Superstar Extinction.” *Quarterly Journal of Economics*, 125 (2), 549-589.
- Bloom, N., M. Schankerman, and J. Van Reenen.** 2013. “Technology Spillovers and Product Market rivalry.” *Econometrica* (2013) 81 (4) 1347–1393.
- Bloom, N., P. Romer, Terry S., and Van Reenen J.** 2014. “A Trapped Factors Model of Innovation.” *NBER Discussion Paper* <http://cep.lse.ac.uk/pubs/download/dp1261.pdf>.
- Blundell, R, R. Griffith, and J. Van Reenen.** 1999. “Market Shares, Market Value and Innovation in a Panel of British Manufacturing Firms.” *Review of Economic Studies*, 66, 529-554.

- Breschi, S. and F. Lissoni.** 2005. “Cross-firm inventors and social networks: localized knowledge spillovers revisited.” *Annales d’économie et de statistique* 79/80, 189-209.
- Breschi, S. and F. Lissoni.** 2009. “Mobility of skilled workers and co-invention networks: an anatomy of localized knowledge flows.” *Journal of Economic Geography* 9(4), 439–468.
- Card, D. and L. Giulano.** 2014. ““Does Gifted Education Work? For Which Students ?”.” *NBER Working Paper 20453*.
- Card, D., J. Heining, and P. Kline.** 2013. “Workplace heterogeneity and the rise of West German wage inequality.” *The Quarterly Journal of Economics*, 128(3): 967-1015.
- Celik, M.** 2014. “Does the Cream Always Rise to the Top ? The Missallocation of Talent and Innovation.” *University of Pennsylvania, mimeo*.
- Chetty, R. and N. Hendren.** 2015. “The Impacts of Neighborhoods on Intergenerational Mobility: Childhood Exposure Effects and County-Level Estimates.” (19843).
- Chetty, R., J. N. Friedman, and J. E. Rockoff.** 2014. “Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates.” *American Economic Review*, 104 (9): 2593–2632.
- Chetty, R., N. Hendren, P. Kline, and E. Saez.** 2014. “Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States.” (19843).
- Cook, L. and C. Kongcharoen.** 2010. “The Idea Gap in Pink and Black.” *NBER Working Paper 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), 1001-44.
- Ding, W., F. Murray, and T. Stuart.** 2006. “Gender differences in patenting in the academic life sciences.” *Science*, 313, 665–556.
- Dorner, M., D. Harhoff, and K. Hoisl.** 2014. “Patterns and Determinants of Inventor Mobility – Evidence from the Employment Biographies of Inventors in Germany.” *mimeo*.

- Fallick, B., C. A. Fleischman, and J. B. Rebitzer.** 2006. “Job-hopping in silicon valley: some evidence concerning the microfoundations of a hightechnology cluster.” *The Review of Economics and Statistics* 88, 472–481.
- Fernandez-Zubieta, A., A. Geuna, and C. Lawson.** 2013. “Researchers’ Mobility and its Impact on Scientific Productivity.” *University of Turin Working paper No. 13/2013*.
- Franco, A. M. and M. F. Mitchell.** 2008. “Covenants not to compete, labor mobility, and industry dynamics.” *Journal of Economics and Management Strategy* 17, 581–606.
- Fryer, R.** 2014. “Injecting Charter School Best Practices into Traditional Public Schools: Evidence From Field Experiments.” *Quarterly Journal of Economics*, 129(3), 1355-1407.
- Giuri, Paola, Myriam Mariani, Stefano Brusoni, Gustavo Crespi, Dominique Francoz, Alfonso Gambardella, Walter Garcia-Fontes, and Aldo 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.
- Graham, Stuart and Deepak Hegde.** 2015. “Disclosing patents’ secrets.” *Science*, 347 (6219): 236–237.
- Griliches, Zvi.** 1992. “The Search for R&D Spillovers.” *Scandinavian Journal of Economics*, 94, S29-47.
- Grossman, G. and E. Helpman.** 1991. “Quality Ladders and Product Cycles.” *Quarterly Journal of Economics*, 106: 557-586.
- Hall, B., A. Jaffe, and M. Trajtenberg.** 2001. “The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools.” *NBER WP 8498*.
- Hall, B., A. Jaffe, and M. Trajtenberg.** 2005. “Market Value and Patent Citations: A First Look.” *Journal of Economics*, 36, 16-38.
- Hall, B. and J. Lerner.** 2009. “The Financing of R&D and Innovation.” *NBER Working Papers* 15325.

- Hoisl, K.** 2007. "Tracing Mobile Inventors - The Causality between Inventor Mobility and Inventor Productivity." *Research Policy*, 36: 619-636.
- Hoxby, C. and S. Turner.** 2015. "What High-Achieving Low-Income Students Know About College." *American Economic Review*, 105(5) 514-17.
- Hsieh, C. and P. Klenow.** 2009. "Misallocation and Manufacturing TFP in China and India." *Quarterly Journal of Economics*, 124(4), 1403-1448.
- Hsieh, C., Hurst E. Jones C. and P. Klenow.** 2013. "The Allocation of Talent and U.S Economic Growth." *Mimeo, Stanford*.
- Hunt, J.** 2009. "Why Don't Women Patent?" *NBER Working Paper No. 17888*.
- Hurst, E. and A. Lusardi.** 2004. "Liquidity Constraints, Household Wealth and Entrepreneurship." *Journal of Political Economy* 112(2), 319-347.
- Jaffe, A. and M. Trajtenberg.** 2002. *Citations and Innovations*.
- Jaravel, Xavier, Neviana Petkova, and Alex Bell.** 2015. "Team-Specific Capital and Innovation." SSRN Scholarly Paper ID 2669060, Social Science Research Network, Rochester, NY.
- Jones, Benjamin F.** 2010. "Age and Great Invention." *Review of Economics and Statistics*.
- Jones, B.F., E. J. Reedy, and B.A. Weinberg.** 2014. "Age and Scientific Genius." *NBER Working Paper No. w19866*.
- Jones, C. and J. Williams.** 1999. "Too Much of a Good Thing?" *Journal of Economic Growth* 5(1) 65-85.
- Jung, T. and O. Ejermo.** 2014. "Demographic Patterns and trends in patenting: Gender, age and education of inventors." *Technological Forecasting and Social Change* 86, 110-124.
- Kahneman, D. and E. Ghiselli.** 1962. "Validity and Nonlinear Heteroscedastic Models." *Personnel Psychology* (15) 1-11.
- Kim, J. and G. Marschke.** 2005. "Labor mobility of scientists, technological diffusion, and the firm's patenting decision." *RAND Journal of Economics* 36, 298-317.

- Krakel, M. and D. Sliwka.** 2009. “Should you allow your employee to become your competitor? On noncompete agreements in employment contracts.” *International Economic Review* 50, 117–141.
- Lai, R., A. D’Amour, D. M. Doolin, G.-C. Li, Y. Sun, V. Torvik, and A. Yu.** 2014. “Disambiguation and co-authorship networks of the U.S. patent inventor database (1975-2010).” *Research Policy, Elsevier*, 43 (6): 941–955.
- Machin, S. and Pekkarinen.** 2008. “Global Sex Differences in Test Score Variability.” *Science*, 322 (5906): 1331–1332.
- Mankiw, N. Gregory.** 2013. “Defending the One Percent.” *Journal of Economic Perspectives*, 27 (3): 21–34.
- Marx, M., D. Strumsky, and L. Fleming.** 2009. “Mobility, Skills, and the Michigan Non-Compete Experiment.” *Management Science*, 55 (6) 875-889.
- McNally, S.** 2013. “Student Awareness of Costs and Benefits of Educational Decisions: Effects of an Information Campaign.” *Centre for Economics of Education Discussion Paper 139*.
- Moen, J.** 2005. “Is mobility of technical personnel a source of R&D spillovers?” *Journal of Labor Economics* 23, 81–114.
- Nelson, R. and E. Phelps.** 1966. “Investment in humans, technological diffusion, and economic growth.” *American Economic Review*: 51 (2) 69-75.
- Pakes, A.** 1986. “Patents as Options: Some Estimates of the Value of Holding European Patent Stocks.” *Econometrica*, 54(4), 755-84.
- Pakes, A. and S. Nitzan.** 1983. “Optimum Contracts for Research Personnel, Research Employment, and the Establishment of Rival Enterprises.” *Journal of Labor Economics*, 1(4), 345-65.
- Pope, D. and J. Sydnor.** 2010. “Geographic Variation in the Gender Differences in Test Scores.” *Journal of Economic Perspectives*, 24(2), 95-108.
- Romer, P.** 2000. “Should the Government Subsidize Supply or Demand in the Market for Scientists and Engineers?” *NBER Chapters, in : Innovation Policy and the Economy, Volume 1*, 221-252.

Table 1: Fraction of Gap in Patenting by Parental Income accounted for by differences in Math Test Scores

Panel A: Decomposition Using 3rd Grade Math Score

	Patent Rate (per 1000 Individuals)	Gap Relative to Above p80 Group
Above 80 th Percentile of parent income	1.93	
Below 80 th Percentile	0.52	1.41
Below 80 th Percentile. (Reweighting Scores)	0.95	0.97
% of gap accounted for by 3 rd grade scores		30.9% (s.e. = 8.5%)

Panel B: Decomposition using later grades

% of gap accounted for by 4 th grade scores	36%
% of gap accounted for by 5 th grade scores	39%
% of gap accounted for by 6 th grade scores	45%
% of gap accounted for by 7 th grade scores	51%
% of gap accounted for by 8 th grade scores	53%
Average percentage point change per grade	4.6

Notes: This is a DiNardo, Fortin and Lemieux (1996) decomposition of the difference in innovation rates among children of the "rich" (top quintile of income) and all others (see text) using the NYC test score data combined with our IRS-USPTO match. We divide the math distribution into vintiles and give the low-income Children the test score distribution of the rich and re-calculate the implied innovation rates (re-weighting scores).

Table 2: Fraction of Gap in Patenting by Gender accounted for by differences in Math Test Scores

Panel A: Decomposition Using 3rd Grade Math Score

	Patent Rate (per 1000 Individuals)	Gap Relative to Above p80 Group
Men	1.36	
Women	0.47	0.89
Women (Reweighting Scores)	0.50	0.86
% of gap accounted for by 3 rd grade scores		3.7%

Panel B: Decomposition using later grades

% of gap accounted for by 4 th grade scores	2%
% of gap accounted for by 5 th grade scores	2.9%
% of gap accounted for by 6 th grade scores	4.5%
% of gap accounted for by 7 th grade scores	6.4%
% of gap accounted for by 8 th grade scores	8.7%
Average percentage point change per grade	1.7%

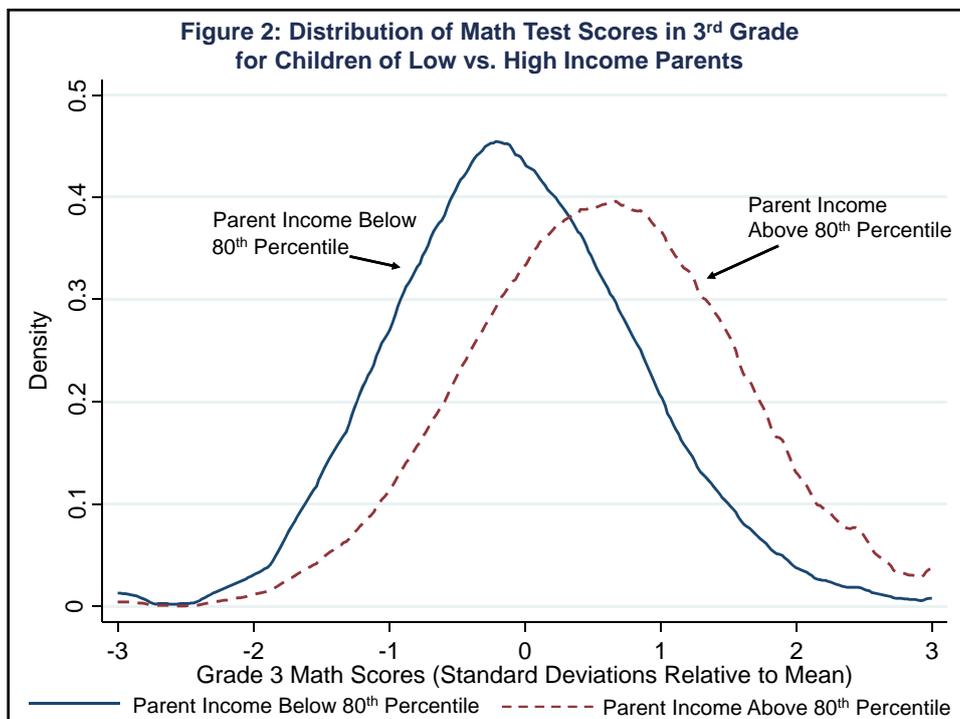
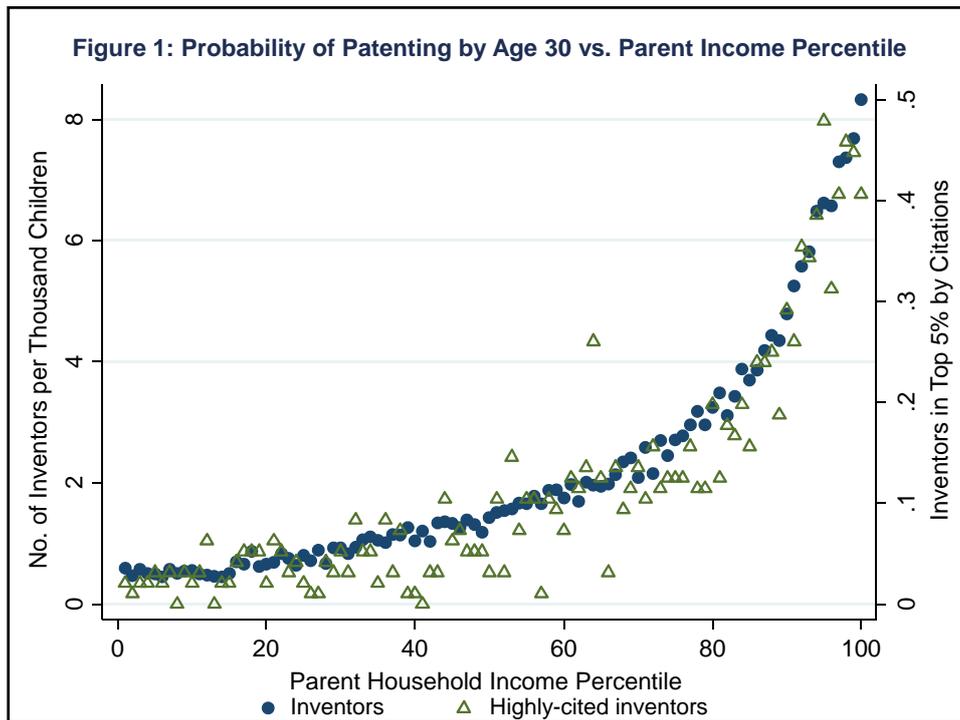
Notes: This is a DiNardo, Fortin and Lemieux (1996) decomposition of the difference in innovation rates among boys and girls (see text) using the NYC test score data combined with our IRS-USPTO match. We divide the math distribution into vintiles and give the girls the test score distribution of the boys and re-calculate the implied innovation rates (re-weighting scores).

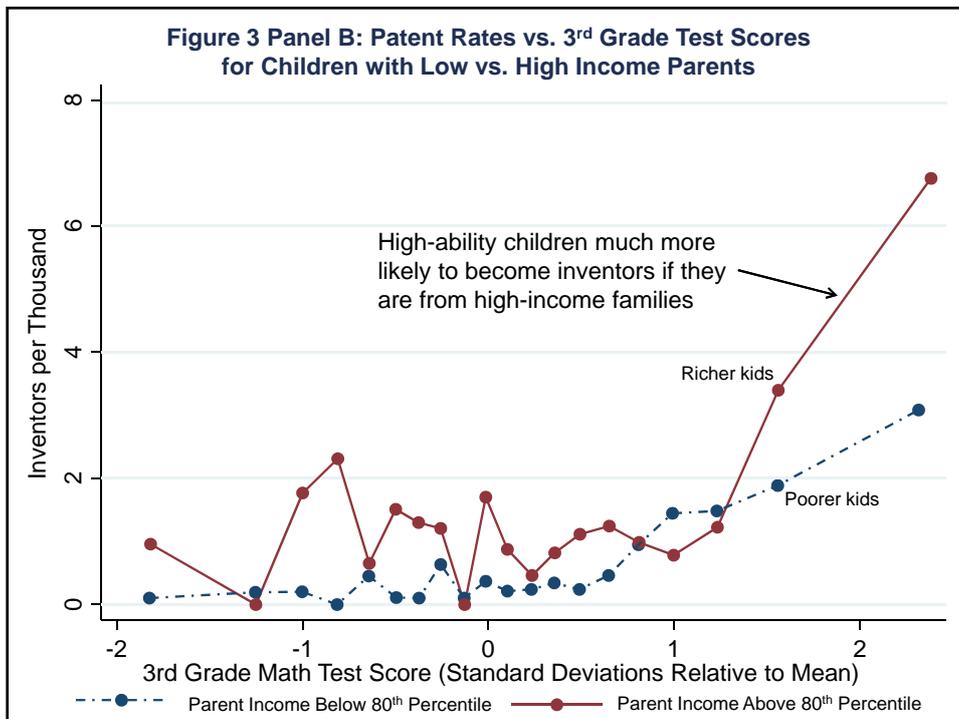
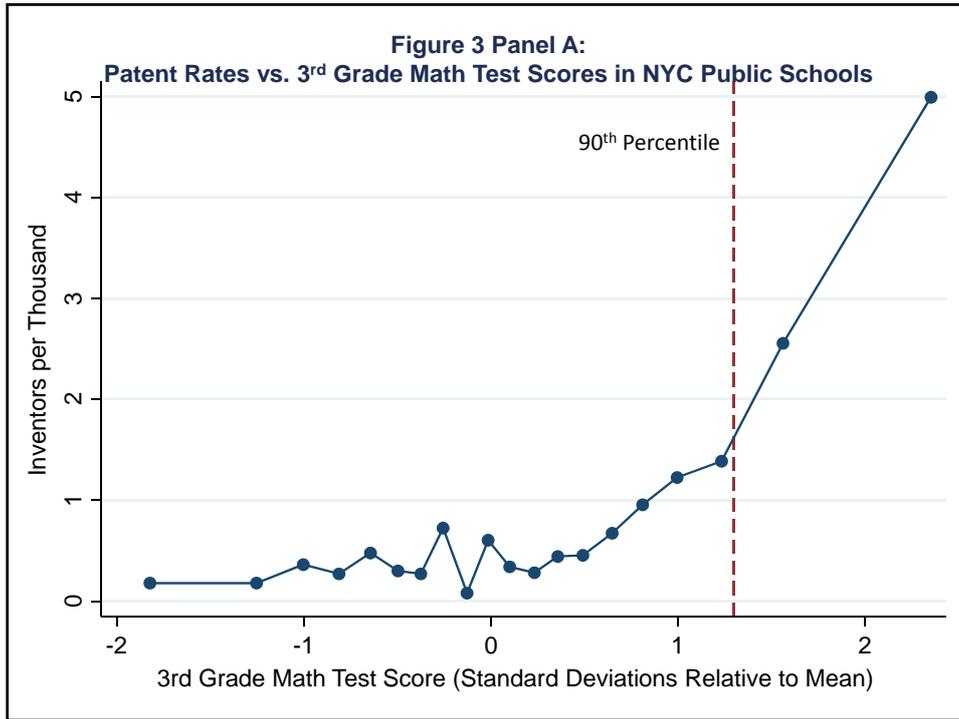
Dependent variable:	(1) Fraction Inventors	(2) Fraction Inventing in Category	(3) Fraction Inventing in Sub-Category	(4) Fraction Inventing in Class	(5) Fraction Inventing in Class
Fraction Inventors in Father's Industry	0.250*** (0.0276)				
Fraction in Category in Father's Industry		0.162*** (0.0166)			
Fraction in Sub-Category in Father's Industry			0.154*** (0.0168)		
Fraction in Class in Father's Industry				0.0780*** (0.0136)	0.0601*** (0.0129)
Fraction in same Sub-Category but other Class					0.00438*** (0.00083)
Fraction in same Category but other Sub-Cat.					0.00006 (0.00040)
Fraction in other Category					0.00021*** (0.00005)
Cells	Father's industry	Father's industry*category	Father's industry*sub-category	Father's industry*technology class	Father's industry*technology class
Observations	345	2,415	12,765	153,525	153,525

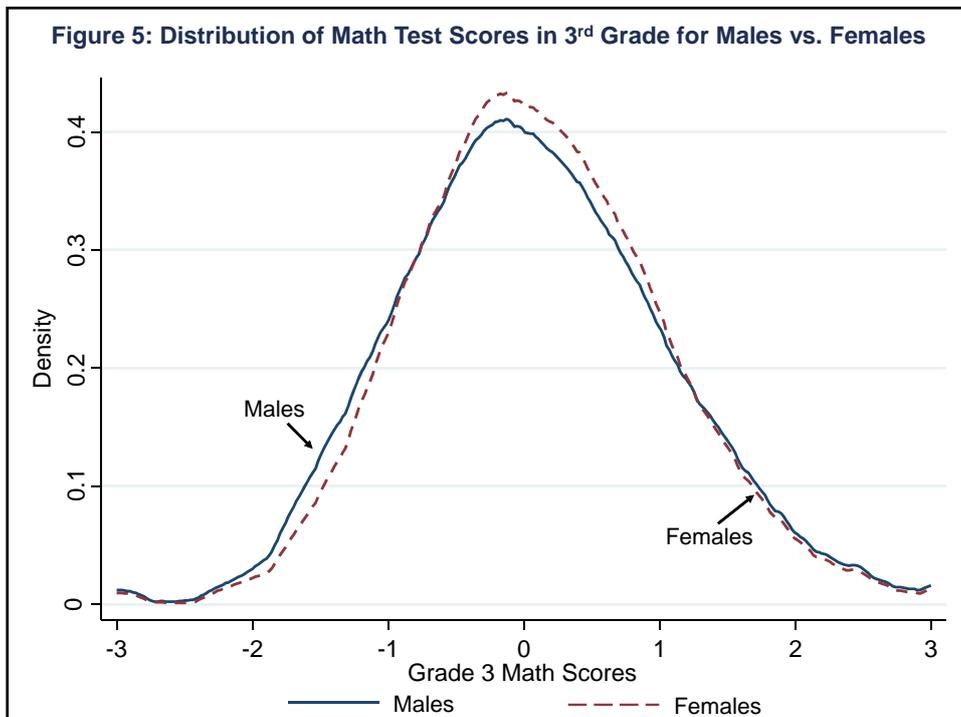
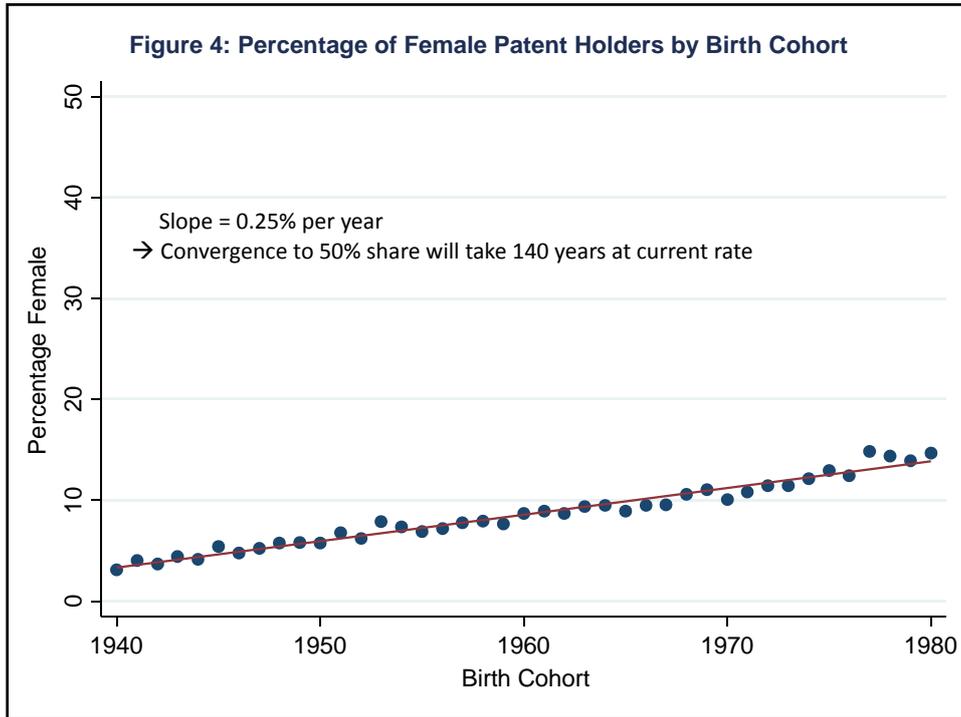
Notes: Standard errors are clustered by 345 industries. Column (2) includes 7 category fixed effects; column (3) includes 37 sub-category fixed effects; columns (4) and (5) include 450 technology class fixed effects. The sample is children whose parents are not inventors.

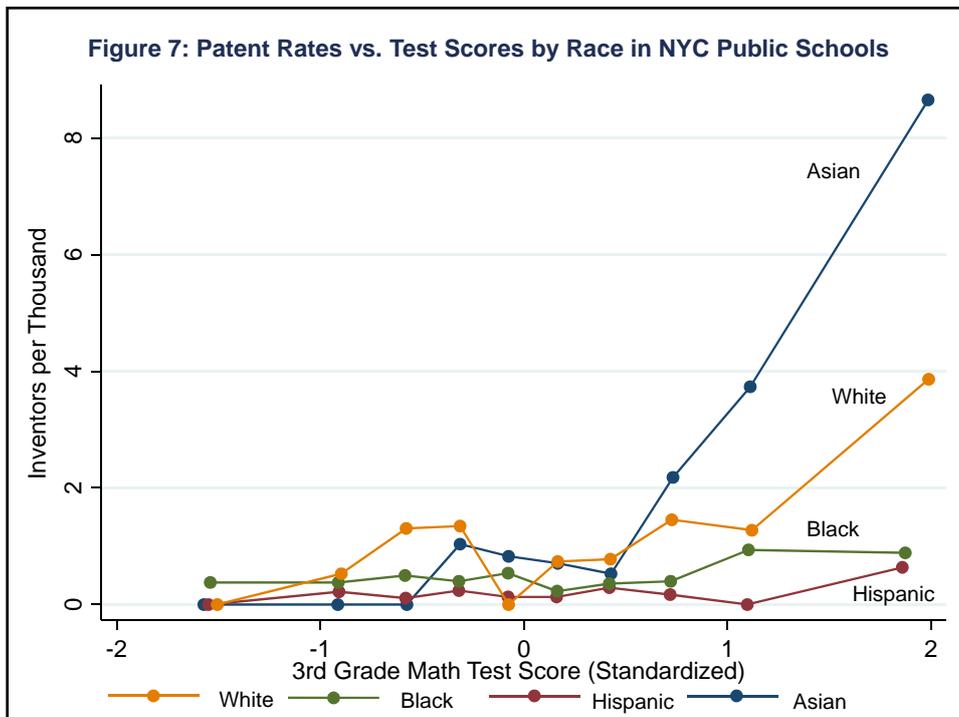
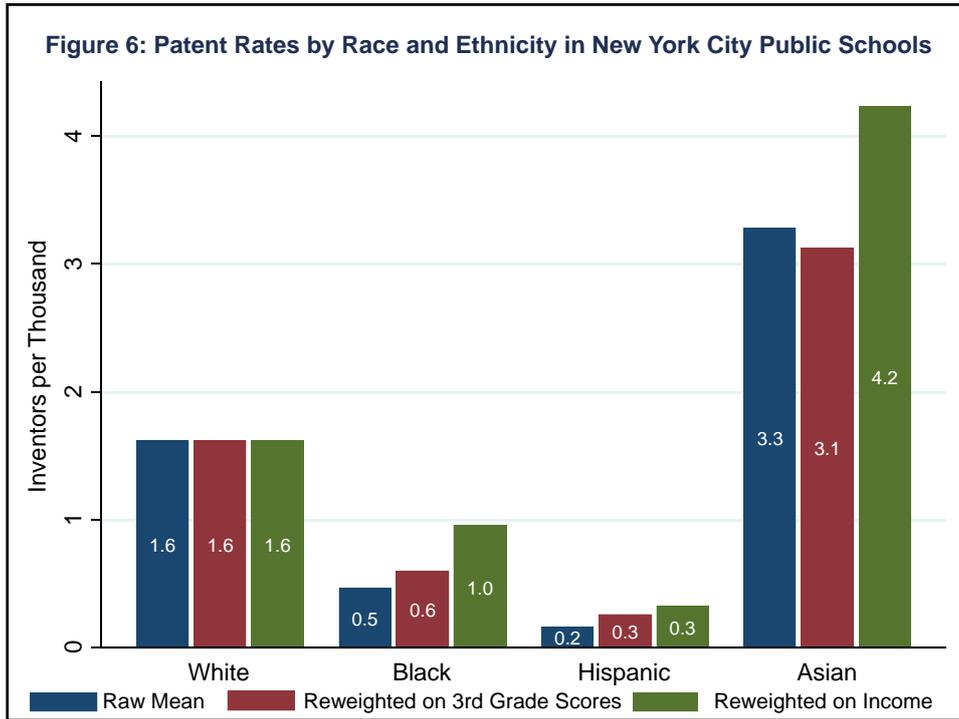
Dependent Variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Fraction Children inventing	Fraction inventing in CZ	Dummy if individual invented	Fraction Inventing in in cell	Fraction Inventing in cell	Fraction Inventing in in cell	Fraction Inventing in cell	Fraction Inventing in cell	Fraction Inventing in cell
Parent's CZ patent rate	2.906*** (0.435)	3.098*** (0.572)						
Parent's CZ patent rate in Category			1.724*** (0.409)	1.378*** (0.396)	1.955*** (0.448)			
Parent's CZ patent rate in Sub-Category						1.509*** (0.386)		
Parent's CZ patent rate in Technology Class							1.136*** (0.194)	1.050*** (0.173)
Parent's CZ patent rate in same Sub-Category, other class								-0.0005 (0.0064)
Parent's CZ patent rate in same category but other Sub-Category								-0.0018 (0.0027)
Parent's CZ patent rate in other category								0.0053*** (0.0007)
Fixed effects	None	Child CZ	Category	Child CZ & Category	Parent NAIC & Category	Sub-Category	Technology Class	Technology Class
Cells	Parent CZ	Individual	Parent CZ*category	Child CZ*Parent CZ*Category	Parent CZ*Parent NAIC*Cat.	Parent CZ*Sub-category	Parent CZ*technology class	Parent CZ*technology class
Observations	390	5,452,642	2,730	1,554,973	1,642,193	14,430	173,550	173,550

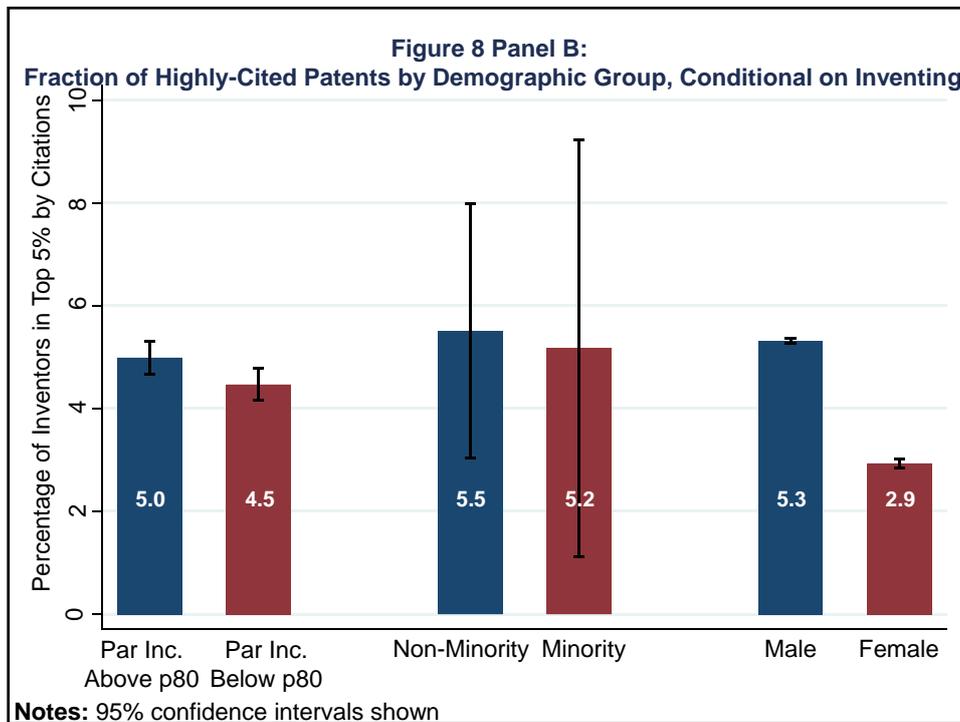
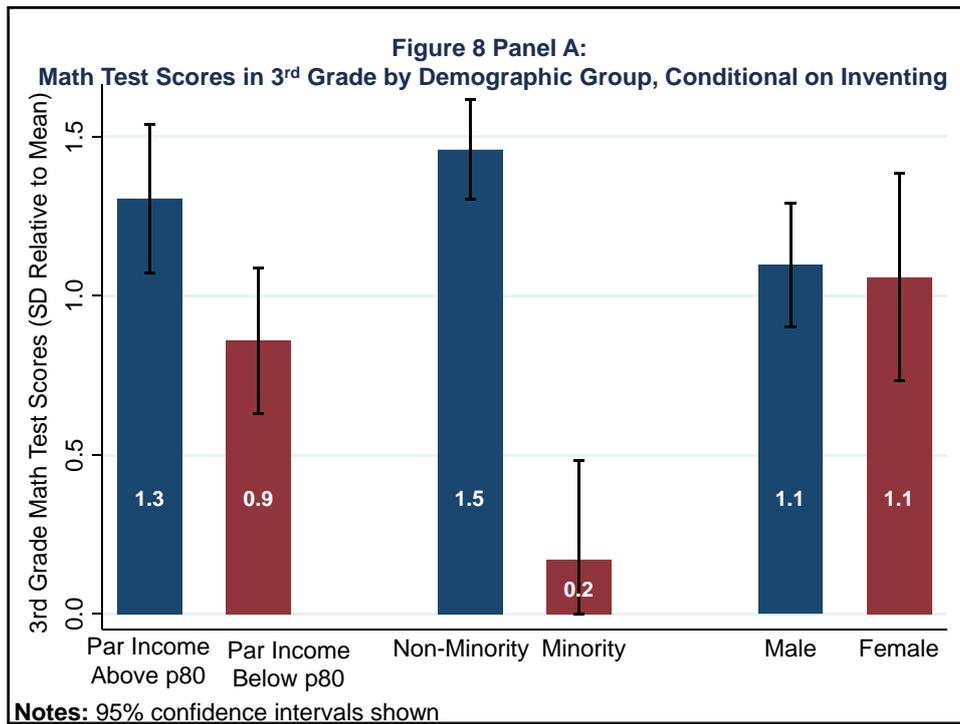
Notes: Standard errors clustered by 390 Commuting Zones (CZ) where children grew up ("parent's CZ"). Columns (2) and (4) include current CZ as of 2012 ("child's CZ") fixed effects. Sample is children whose parents were not inventors.

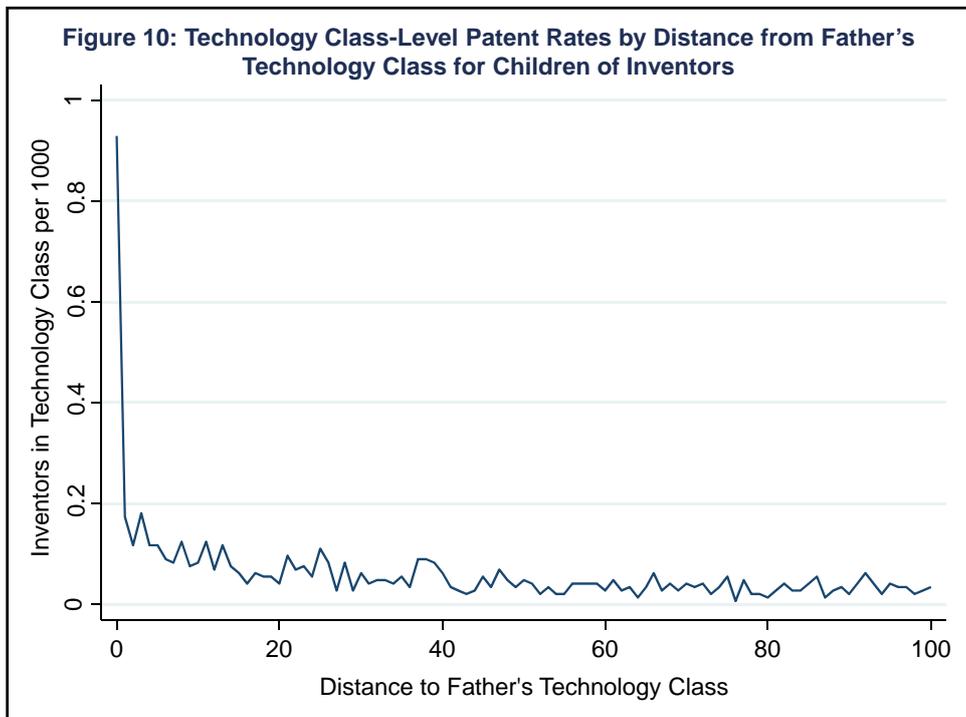
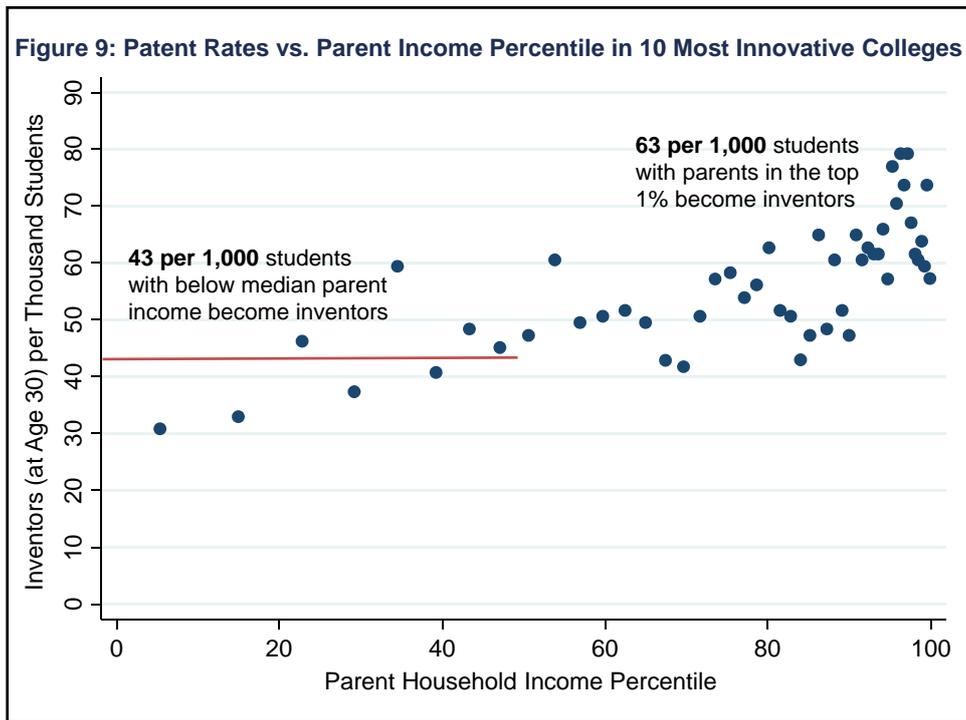












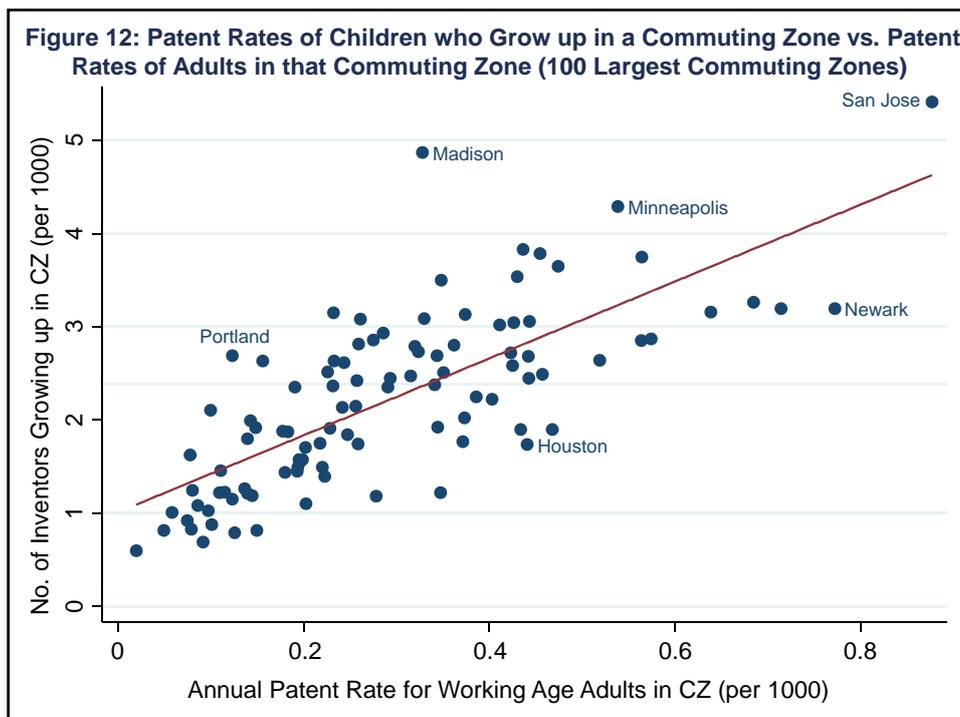
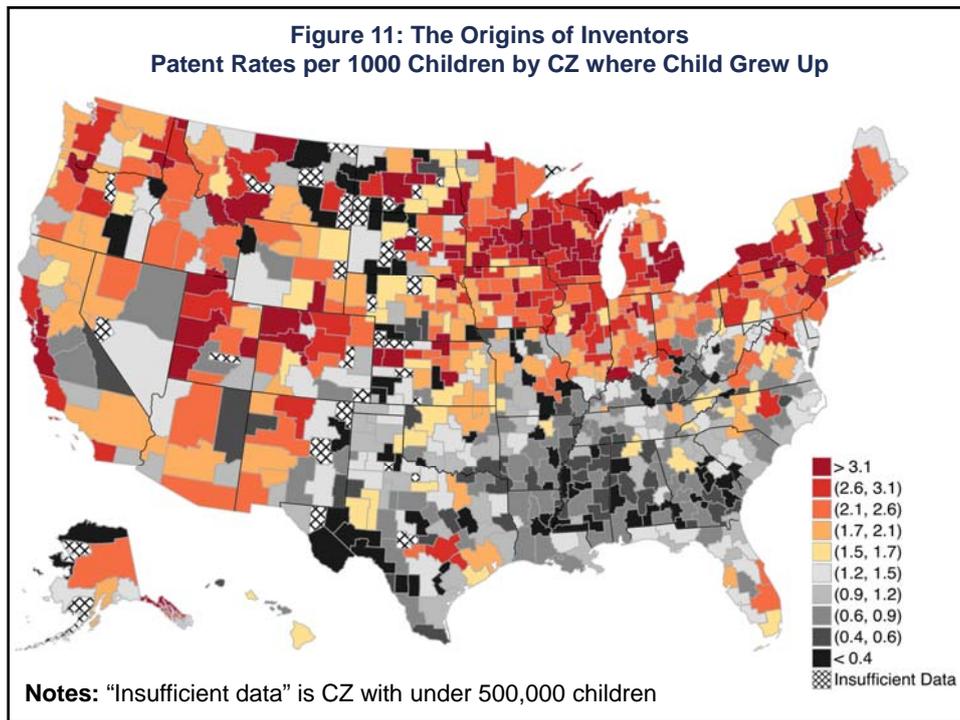


Figure 13: Percent of Inventors who are Female by State where Child Grew Up

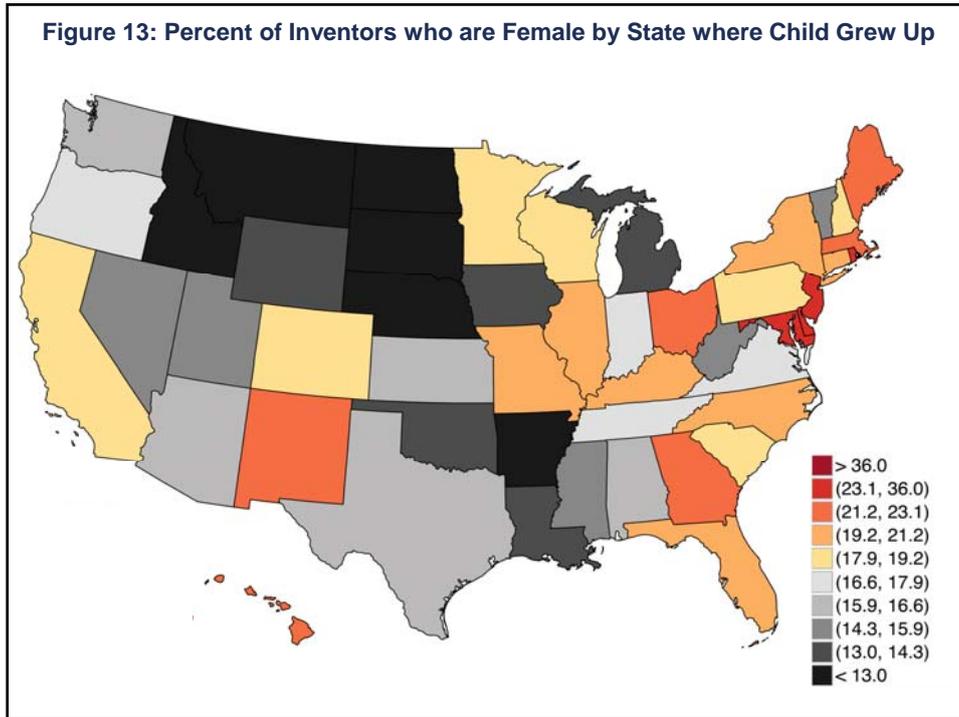
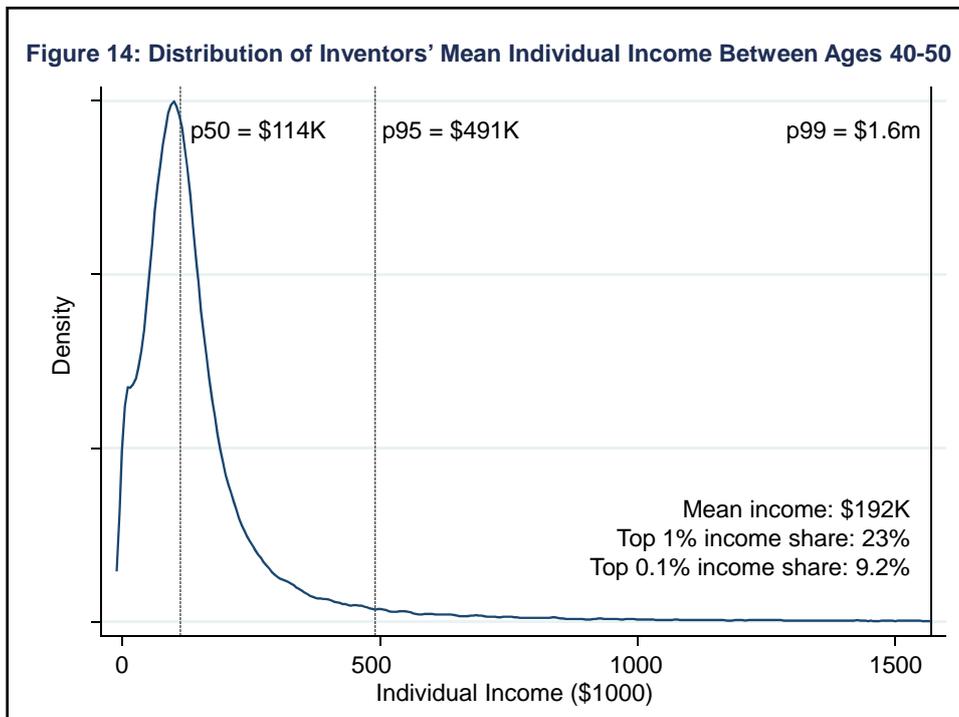
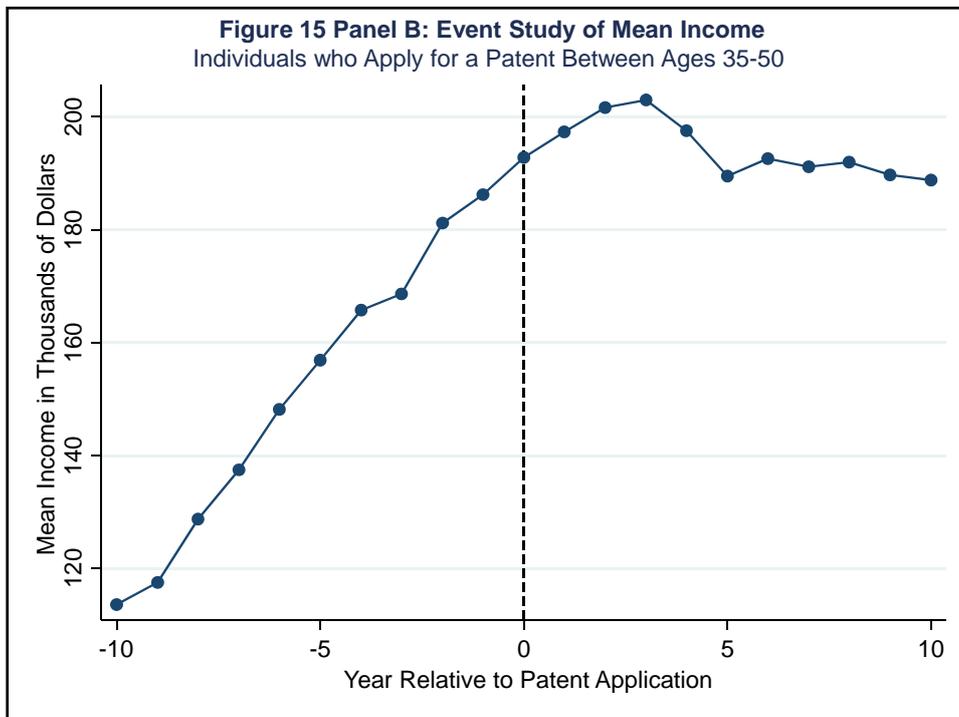
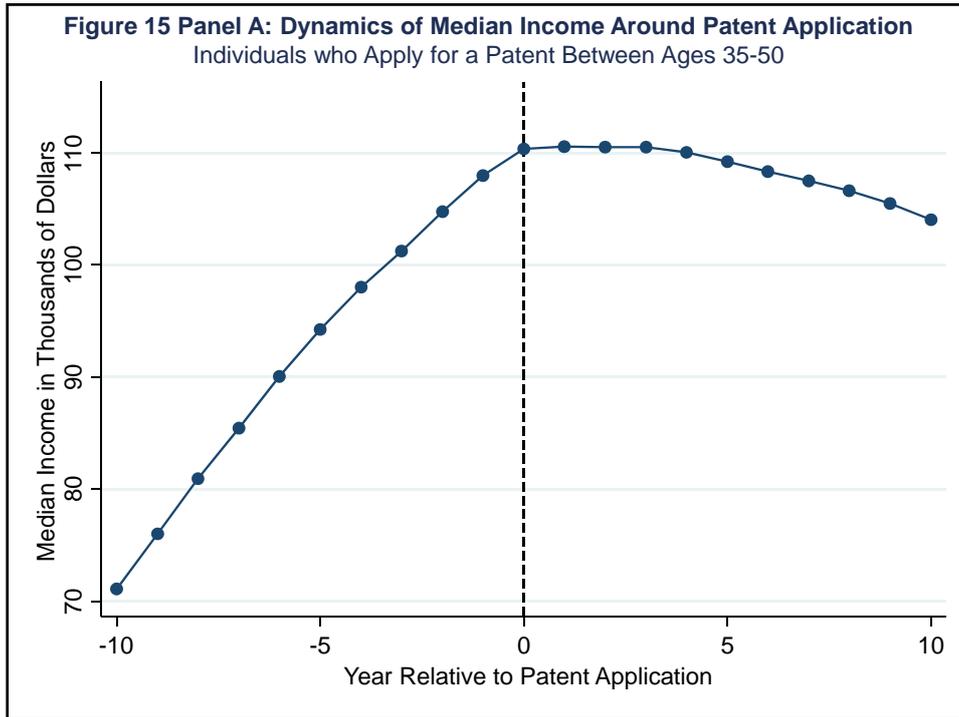
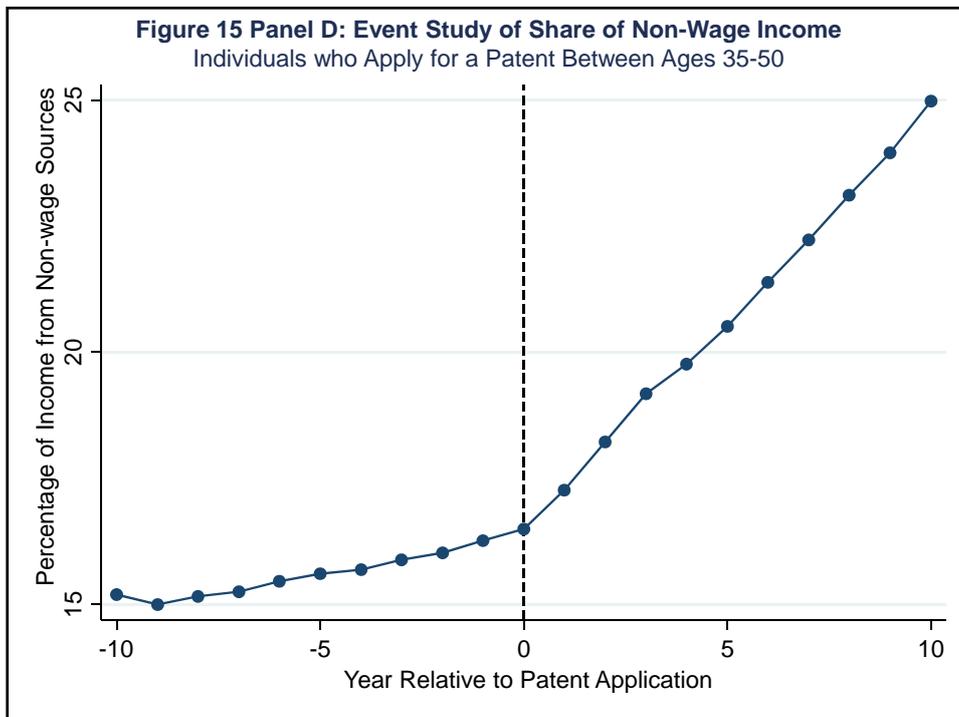
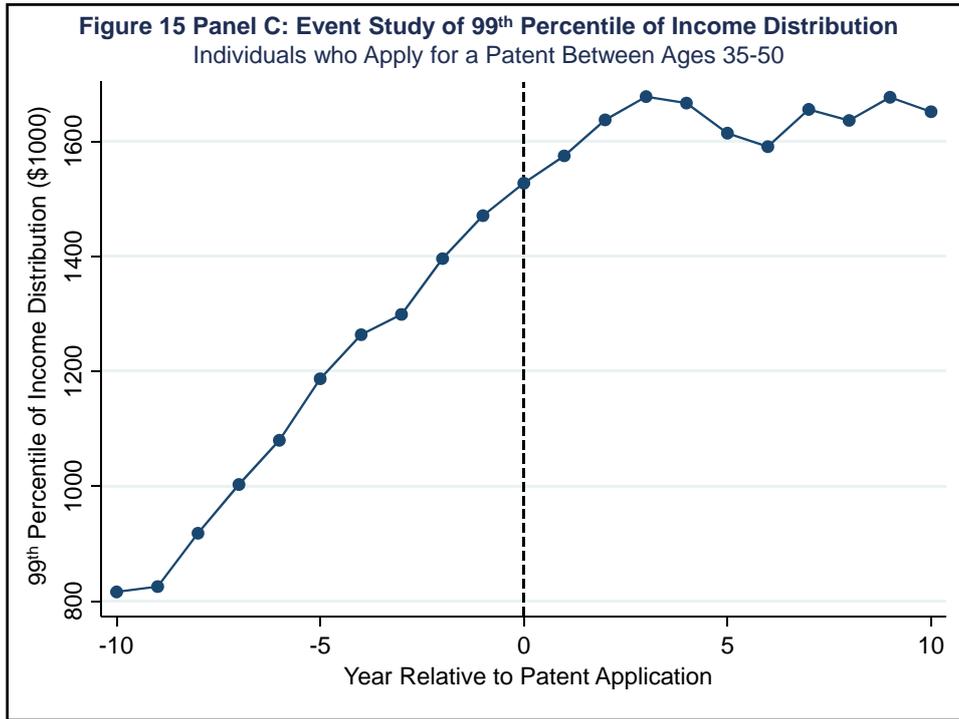
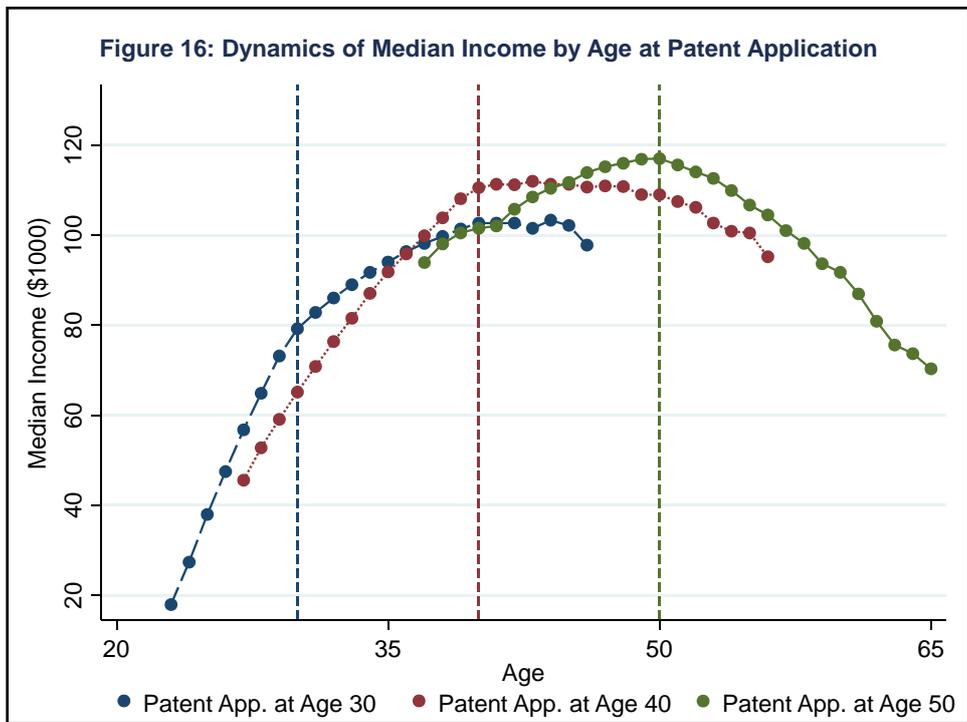
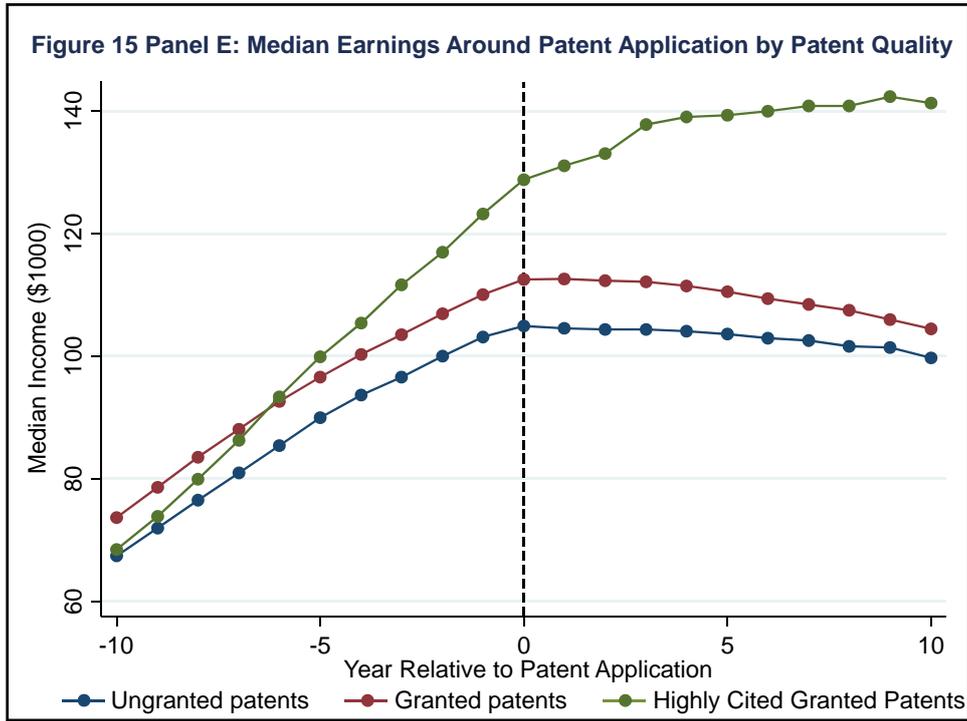


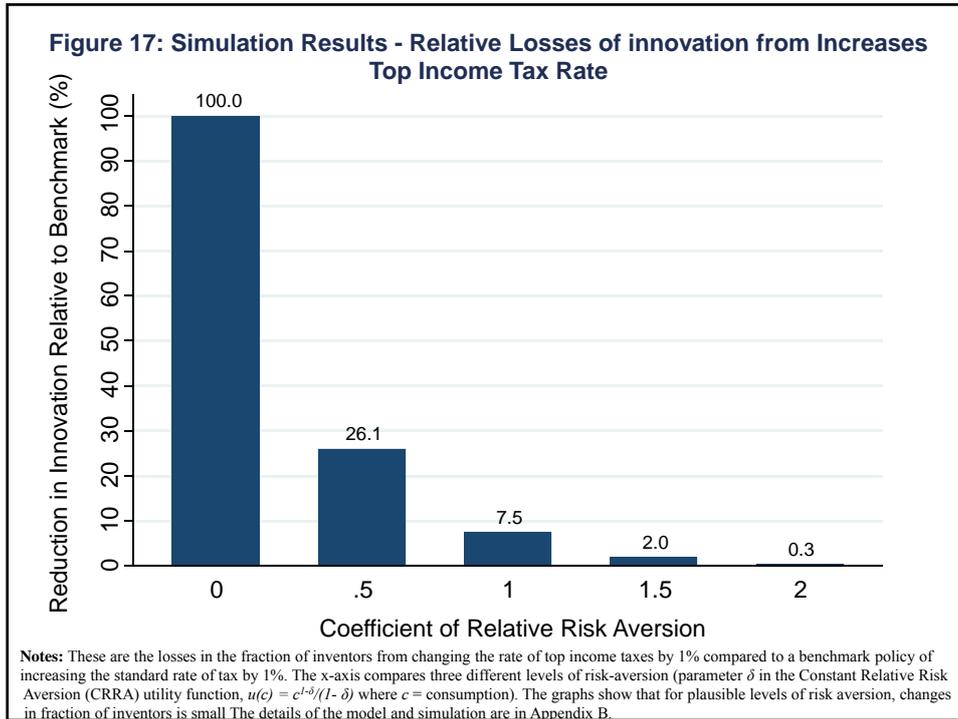
Figure 14: Distribution of Inventors' Mean Individual Income Between Ages 40-50











APPENDICES (NOT INTENDED FOR PUBLICATION UNLESS REQUESTED)

APPENDIX A: DATA

A.1 Data Preparation

- **Suffix Standardization.** Suffixes may appear at the end of taxpayers’ first, middle, or last name fields. Any time any of these fields ends with a space followed by “JR”, “SR”, or a numeral I-IV, the suffix is stripped out and stored separately from the name⁴⁹.
- **First name to imputed first/middle name.** The USPTO separates inventor names into “first” and “last,” but the IRS often separates names into first, middle, and last. In practice, many inventors do 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 outlined below.

A.2 Pseudocode for Match on Name and Location

The exact matching stages are as follows. We conduct seven progressive rounds of matching. 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 noted, with an impressive 49% being exact matches on the first stage.

- The matching algorithm takes as input a relation of inventor data and five relations of IRS data:
 - Input relations:
 - * Inventors(inv_id, first, last, imputed_first, imputed_middle, suffix) - directly from USPTO
 - * NamesW2(irs_id, first, middle, last, suffix) - all names used by individual on W2 information returns; name field is recorded as first, middle, and last
 - * Names1040(irs_id, first, middle, last) - all self-reported names from 1040 forms⁵⁰

⁴⁹Numerals I and V are only permissive suffixes at the end of a last name field, as these may be middle initials in a middle name field.

⁵⁰We only take names off of 1040s for those who file singly because it proved difficult to parse names of those list them jointly

- * NameInW2(irs_id, fullname) - all names from W2, but a separate variable not recorded as first, middle, last that was more frequently present
 - * CitiesW2(irs_id, city, state) - all cities reported on W2
 - * Zips1040(irs_id, name) - all zip codes reported on 1040
- Output relation:
- * Unique-Matches (inv_id, irs_id)
- **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 match some city and state reported by that taxpayer exactly.
 - 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 firstname is the same as the taxpayer’s combined first and middle name.
 2. Inventor’s imputed firstname matches taxpayer’s and middle names match on initials.
 3. The inventor has no middlename data, but inventor’s firstname is the same as the taxpayer’s middle name.
- As always, 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 1040 names instead of names from W2’s.
 - 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 W2 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:** Relaxed match using W2 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.

A.3 Final matched sample

As noted in the text we matched 88% of inventors in the 200s and 80% in the 1990s. The match seemed balanced by characteristics of the patents such as citation rates, number of claims, technology class, zip code, etc. The match rate is slightly lower among self-assigned patents.

APPENDIX B: LIFECYCLE MODEL OF INNOVATION

B.1 Basic Set up

B.1.1 Model

We sketch a simple inventor lifecycle model. Consider a two period model. In period 1, individual i begins with an endowment of human capital we call “ability” (a_i). Their human capital then evolves over the course of their schooling increasing by s_i which is determined by family, school and neighborhood quality. Initial ability is complementary with future learning and parental income is positively associated with both initial ability and subsequent inputs. We initially assume these are exogenous, but later allow s_i to be endogenous. After acquiring human capital, $H_i = H(a_i, s_i)$, individuals enter the labor market in period 2 and choose an occupation. For now we consider just two possible occupations: the R&D sector and the Non-R&D sector.

Wages in the non-R&D sector are $\tilde{w} + \rho H_i$.⁵¹ The expected utility (V^N) for working in the non-R&D sector will depend on idiosyncratic tastes for working in this sector which we denote in utility terms as v_i^N . There is also a tax schedule $T(\cdot)$ which we will allow to be non-linear with a higher marginal rate above an upper threshold. So $V_i^N = u[T(\tilde{w} + \rho H_i)] + v_i^N$ where $u[\cdot]$ is a utility function over consumption.

In the R&D sector workers also receive a deterministic base wage ρH_i but also have a chance of receiving an additional stochastic reward from innovation. We assume that those with more human capital have a higher chance of successfully innovating, so the additional (potential) innovation reward is $\pi(H_i)$, with $\pi'(H_i) > 0$. For simplicity we parameterize this as $\pi(H_i) = \pi H_i$. We allow a taste term for the R&D sector v_i^R which implies that the expected value of choosing to work in the R&D sector is:

$$V_i^R = E_\pi(u[T(\rho H_i + \pi H_i)]) + v_i^R$$

where $E_\pi(\cdot)$ is the expectations operator taken over the stochastic innovation variable π .

We now introduce the idea that some individuals who are less exposed to inventors underestimate the real returns to invention. The *perceived* value of a career in the R&D sector is:

$$V_i^{PR} = \lambda_i V_i^R + (1 - \lambda_i) E_\pi(u[T(\rho H_i + \eta)]) + v_i^R$$

where $0 \leq \lambda \leq 1$ and $\eta < \pi H_i$. With probability λ an individual has the correct (full information) on the true value of being in the R&D sector, V_i^R . But with probability $1 - \lambda$ the individual believes that the chances of innovation are lower than they actually are.⁵² The idea is that greater exposure to innovation will increase the chance that an individual believes $\lambda = 1$. If this exposure is greater for some groups than others (such as rich vs. poor), this will be a cause of there being fewer poor (but equally able) inventors.

Define the difference between the perceived value of the R&D sector vs. the non-R&D sector as:

$$\varphi_i = V_i^{PR} - V_i^N = \lambda_i E_\pi(u[T(\rho H_i + \pi H_i)]) + (1 - \lambda_i) E_\pi(u[T(\rho H_i + \eta)]) - u[T(\tilde{w} + \rho H_i)] + \tilde{v}_i$$

⁵¹This \tilde{w} could in principle be negative but the evidence in Stern (2004) suggests that it is positive when looking at multiple job offers for post-doctoral biologists. Scientists “pay” about 15-20% of their entry salaries to be scientists.

⁵²Formally, the agent underestimates the extent of the complementarity between innovation and human capital. If the formulation was ηH_i with $\eta < \pi$ this would be like a “tax” on groups who receive less than their productivity in an occupation. There would still be rational sorting for less informed agents in this world (although less of it) because the very talented would still be prepared to go into the R&D sector as their wage would be sufficiently high to compensate them for the tax.

with $\tilde{v}_i = v_i^R - v_i^N$.

Whether an individual chooses the R&D sector will depend on the sign of φ_i . If we define $I(\varphi_i)$ an indicator function equal to one if $\varphi_i > 0$ and zero otherwise, then to calculate the number of inventors we simply have to integrate $I(\varphi_i)$ across individuals. To calculate the fraction of a group (low income kids, minorities or women) who become inventors we integrate over the individuals in the relevant group.

The comparative statics of the model are straightforward.

1. *Inventor probability increases with the correct information/exposure to the R&D sector*

$$\frac{\partial \varphi_i}{\partial \lambda_i} = E_\pi(u[T(\rho H_i + \pi H_i)]) - E_\pi(u[T(\rho H_i + \eta)]) > 0$$

2. *The probability of being an inventor increases with human capital (since $\frac{\partial \varphi_i}{\partial H_i} > 0$ and $\lambda \geq 0$)*

This is because of the complementarity between human capital and the probability of innovation (i.e. $\pi > 0$). A corollary is that inventor probability also increase with initial ability so long as $\frac{\partial H_i}{\partial a_i} > 0$

3. *Inventor probability increases with the relative preference for the R&D sector*

$$\frac{\partial \varphi_i}{\partial \tilde{v}_i} > 0$$

4. *On the margin, high human capital agents will be more likely to enter the R&D sector as their information improves than low human capital agents*

$$\frac{\partial^2 \varphi_i}{\partial H_i \partial \lambda_i} > 0$$

Corollary A. Groups with higher average values of information, initial ability, human capital and/or a preference for the R&D sector will have a higher fraction of inventors.

Corollary B. A group which is badly informed is less positively selected (in terms of human capital) into the R&D sector

Notice that are two types of welfare gains in the model from increasing λ . First, if there are simply too few inventors from the Social Planner's perspective (e.g. because of knowledge externalities) then a higher λ will help address this. Second, a higher λ should improve the *composition* of inventors as the new inventors will come disproportionately from the high human capital agents. This is the essence of Corollary B which gives a different result from the basic rational sorting model.

To take an example, say that there are two groups, rich and poor. The rich have $\lambda = 1$ and the poor have $\lambda = 0$. Human capital and preferences for working across the sectors are heterogeneous across individuals but identically distributed in the two groups. Since some of the poor will have a strong preference for the R&D sector some will become inventors, but they will be on average of

lower human capital compared to inventors from rich families. As we increase λ for the poor, the individuals who choose to enter the R&D sector will be the most talented of the poor. At some point the average human capital of inventors will be the same for the two groups. So an increase in λ not only increases the number of inventors it also increases the average quality, which will have a stronger effect on successful innovation than simply adding another inventor of the same average incumbent ability.

B.1.2 Empirical Implications

We find support for several of the predictions of the model. First, the evidence of the strong correlation between early math test scores and inventor status is consistent with result 2, as is the college-innovation relationship. Second, we find that people more exposed to innovation (even by class of technology) are more likely to become inventors which is consistent with result 1.

A criticism of this result is that the exposure measure may reflect other mechanisms. For example, rather than increasing λ , exposure might work through increasing \tilde{v}_i , the preference for an R&D career. In terms of welfare, if the main concern is insufficient numbers of people (of all ability) choosing to be inventors then it does not matter too much whether a policy of exposure works through information or preferences. However, it is unclear that just shifting preferences will have any implications for talent misallocation.

B.2 Extensions

There are many extensions that can be made to the basic model which we now consider.

B.2.1. Endogenous acquisition of Human Capital

In stage 1 we can allow agents to invest in their own human capital. Following Hsieh et al. (2013) we can model this as a “goods tax” (τ_g) on human capital investment that is higher for some groups than others.⁵³ This is a reduced form way of capturing things like poorer schools in poor neighborhoods. Those who have a low value of λ will rationally choose to invest less in their human capital all else equal. So this will compound the degree of misallocation. Disadvantaged young people perceive a lower return from their talents, invest less and so make their miss-perceptions self-fulfilling. There is evidence that experiments to improve the information of disadvantaged kids can make them more likely to attend college (e.g. Hoxby and Turner (2015); McNally (2013)). In the context of our model this mean they have a greater chance of becoming an inventor.

⁵³They also allow for direct discrimination in the labor market whereby some disadvantaged groups obtain a lower wage for their marginal product than others. This is unlikely to be an issue in our context for children from low income families, but it might be for women and minorities. In any case, this turns out to be observationally equivalent to (τ_g) .

B.2.2. Multiple Occupational Sectors

We have simplified our model into having two sectors, but there is no difficulty in allowing multiple sectors. One complication arises between general and occupational specific human capital in such a model, however. In our set-up general human capital gives agents a comparative advantage in the innovation sector. In a multi-occupational Roy model it makes more sense to distinguish between different talents in different sectors. Agents can be born with an initial draw of such talent and will allocate themselves (possibly with endogenous skill acquisition) across these sectors.

B.3 Top Rates of Income Tax: A Simple Calibration

In this sub-section we use the model developed above to consider a quantification of the impact of changing top tax rates on the incentives to become an inventor. We simulate this using our empirical data. Increasing the top tax rate has a benefit that it brings in more revenue to spend on public goods, so when considering such a tax policy change we have to benchmark this in some way to make it revenue neutral. We benchmark against increasing the standard rate of tax so that we consider how raising a dollar through increasing the top tax rate compares with raising a dollar through increasing the standard rate. Our calculations are equivalently to thinking of how cost effective it would be to incentive innovation by reducing the top rate of tax.

B.3.1 Calibration Framework

Step 1. We start with a nonparametric estimate of the pre-tax empirical earnings distribution in the innovation sector. Our proxy for the “permanent income” of an inventor is the average “adjusted gross income” (reported on 1040 Forms) of an inventor between age 40 and 45, minus their spouse’s wage income (reported on the spouse’s W2 Forms). This measure include the inventor’s wage and non-wage income, such as royalties. Our estimate of the empirical earnings distribution is thus simply the percentiles of the observed income distribution. For the rest of the calibration, we work with these 100 cells.

Step 2. Next, we compute expected utility in the innovation sector under various assumptions about the utility function and about the tax regime. Specifically, we consider various CRRA utility functions, $u(c) = \frac{c^{1-\delta}}{1-\delta}$ where c = consumption and the relative risk aversion parameter is δ . We examine δ of 0 (i.e linear utility), 0.5, 1 (i.e. log utility), 1.5 and 2 respectively. We also consider various tax regimes. The status-quo tax regime $\bar{\tau}$ approximates the US tax system, with a tax rate of 28.5% below \$439,000 and of 40% above. We then consider a tax regime which keeps the same top tax rate (40%) but increases the standard tax rate by one percentage point to 29.5% - we will refer to this scenario as the “benchmark policy change” (τ^B) relative to the status quo. We

then consider tax regime τ^1 with the same standard tax rate (28.5%) but increasing the top tax by one percentage point to 41%. We will make comparisons between the various tax regimes in terms of “fall in innovation per dollar of revenue raised”, so that any increase in top taxes is revenue neutral compared to the alternative policy of simply increasing the standard rate. In other tax policy experiments we consider tax regimes similar to the status-quo tax regime but introducing an additional top tax threshold with a tax rate of 60% beyond this threshold.

We use the estimated empirical earnings distribution (the 100 cells from Step 1) and apply the tax regime to obtain post-tax earnings at each point of the income distribution. We then apply the utility function (assuming that consumption is equal to earnings, i.e. there are no savings) to obtain utility in each of the 100 cells, and finally we average over all cells to obtain expected utility. Likewise, we obtain expected tax revenue under each tax regime.

Step 3. For each tax regime and degree of risk aversion, we compute the “certainty equivalent” (post-tax) wage. The certainty equivalent wage governs the inventor’s decision of whether to enter the (risky) R&D sector, as opposed to joining the (safe) non-innovation sector. We assume that there is uncertainty over earnings only in the innovation sector and that each inventor has a “safe” outside option, in the form of a fixed (post-tax) wage in the non-innovation sector. The certainty equivalent is the fixed wage level such that the agent is indifferent between getting this wage for certain (in the non-innovation sector) and drawing from the empirical earnings distribution in the innovation sector. The assumption that tax policy changes do not affect wages in the non-R&D sector is made for analytical simplicity⁵⁴ and show in sub-section B.3.3 that this can be relaxed and does not drive our quantitative results (in sub-section B.3.3. we adopt a more parametric approach and specify an earnings distribution for inventors both inside and outside of the R&D sector).

In Step 3, the certainty equivalent is obtained very simply by starting from the expected utility computed in Step 2 and inverting the utility function to recover the certainty equivalent (post-tax) wage.

Step 4. Given the results from step 3, we compute the change in certainty wage equivalent $\frac{dW}{d\tau}$ for each tax regime τ , relative to the status quo tax regime.

Step 5. Finally, we compute the change in the fraction of people becoming inventors in response to a change in taxation. Formally, this is equivalent to the (marginal) deadweight cost of taxation

⁵⁴In other words, the certainty equivalent responds to tax policy changes only to the extent that they affect expected utility in the innovation sector. This is in line with our general focus on the effect of taxes on inventors and it greatly simplifies the calibration because we do not need to estimate the (counterfactual) earnings distribution the inventors would have obtained had they worked outside of the innovation sector. One way to motivate this assumption is that the changes in tax policy we consider are in a range that is well above the certainty equivalent wage for inventors.

per dollar of tax revenue raised by the government (denoted $\gamma(\tau)$) when switching from the status quo policy $\bar{\tau}$ to the new tax regime τ . We normalize the total labor force to 1 and denote by ϕ the fraction of the labor force choosing to work in the R&D sector. We also compute the expected tax revenue from inventors under each tax regime.

We want to estimate the following quantity:

$$\begin{aligned}\gamma(\tau) &= \frac{\frac{d\phi}{d\tau}}{\phi E[R^\tau - R^{\bar{\tau}}]} \\ &= \epsilon \frac{\frac{dW}{d\tau}}{W^{\bar{\tau}}} \frac{1}{E[R^\tau - R^{\bar{\tau}}]}\end{aligned}$$

where $\frac{d\phi}{d\tau}$ is the change in the fraction of the labor force in the R&D sector as a result of the policy change, and R^τ and $R^{\bar{\tau}}$ are the amount of revenue raised under the new and status quo tax regimes, respectively. The expectation is taken with respect to the empirical earnings distribution. $\frac{dW}{d\tau}$ is the percentage change in the certainty equivalent wage (W) between the new tax regime and the status quo, and ϵ is the elasticity of occupational choice with respect to the change in the certainty equivalent. By definition, $\epsilon = \left(\frac{d\phi}{d\tau}\right) / \left(\frac{dW}{W^{\bar{\tau}}}\right)$, i.e. if the certainty equivalent wage decreases by 1% in the new tax regime, the fraction of the labor force going into innovation decreases by $\epsilon\%$.

As is standard in public finance, the deadweight cost (innovation impact in our application) of tax crucially depends on a behavioral elasticity, here denoted ϵ . To make the point that $\gamma(\tau)$ is small for increases in top tax rates, we do not need to estimate the value of ϵ . Instead, we express everything *relative to the benchmark policy change*, denoted τ^B (increasing the tax rate below \$439,000 from 28.5% to 29.5%). Under the assumption of a constant elasticity, we obtain:

$$\frac{\gamma(\tau)}{\gamma(\tau^B)} = \frac{\frac{dW}{d\tau} E[R^{\tau^B} - R^{\bar{\tau}}]}{\frac{dW}{d\tau^B} E[R^\tau - R^{\bar{\tau}}]}$$

We know $\frac{dW}{d\tau}$ from Step 3 and $\frac{E[R^{\tau^B} - R^{\bar{\tau}}]}{E[R^\tau - R^{\bar{\tau}}]}$ from Step 2. Hence, we can summarize the impact of a top tax change on the amount of inventors, by examining how the tax change effects the utility (certainty equivalent) of going into the innovation sector.

To summarize, the calibration is based on the following steps:

1. Get the 100 percentiles of the distribution of earnings for inventors between age 40 and 45 from 1040 Forms.
2. Calculate expected utility and expected tax revenue in the innovation sector under the various tax regimes and utility functions.

3. Calculate the certainty equivalent (post-tax) wage under the various tax regimes and utility functions.
4. Compute the change in the certainty equivalent (post-tax) wage relative to the status quo, for each tax regime and utility function.
5. Using the previous estimates, compute the fall in innovation per dollar of tax revenue raised for each tax policy change, relative to the benchmark policy change. Repeat for the various utility functions.

Sub-section B.3.2 below reports these results and discusses the role of risk aversion. Section B.3.3 moves to a more parametric framework to illustrate the role of skewness of the earnings distribution in generating these results.

B.3.2 Magnitude of the Innovation effect and The Role of Risk Aversion

Figure 17 reports the value of the ratio $\frac{\gamma(\tau^1)}{\gamma(\tau^B)}$, defined as above, where τ^1 is the tax policy regime increasing the top tax rate, but keeping the standard rate fixed (i.e. tax rate of 28.5% below \$439,000 and a 41% tax rate above this threshold). Under risk neutrality (linear utility), by definition the two policies have the same deadweight cost because the behavioral response is the same (the inventor values an additional dollar the same at any point of the earnings distribution). The figure shows that the relative efficiency cost steeply declines with the coefficient of relative risk aversion. For standard values of the coefficient of relative risk aversion above 1, the efficiency cost of the increase in top tax rates is one order of magnitude smaller than the efficiency cost of the benchmark policy change. For a coefficient of relative risk aversion of 2, there is essentially no effect on innovation.

Figure A18 illustrates the innovation loss relative to the benchmark for a number of other policy changes, increasing the top tax rate above \$439,000 to 41%, 45%, 50%, 60%, etc. up to 95%. The innovation loss is of course increasing in the tax rate, but the figure shows that the magnitude of the efficiency loss always remains very small (below 12% of the loss in the benchmark policy change) for reasonable values of the coefficient of relative risk aversion, above 1.

Figure A19 takes this point one step further by showing the effect of introducing a new top tax bracket with a 60% top tax rate above a variety of thresholds (\$439,000, \$800,000, \$1.6m, \$4m, \$10m, \$20m, and \$30m). The figure shows that “millionaire taxes” have extremely small innovation losses, equal to at most a couple of percents of the innovation loss of the benchmark policy change. The intuition behind this result is that changing the probability of extremely high payoffs does not

affect the certainty equivalent wage by much, due to the concavity of the utility function.

All of this results are quantitatively equivalent when winsorizing the inventors' empirical earnings distribution to \$100,000. In other words, the results are not driven by the fact that the benchmark policy would have a comparatively large effect because it increases the tax rates on much lower income levels (which are very rare levels of income in the population of inventors).

B.3.3 The Role of Skewness of income returns to being an inventor

To investigate the role of skewness in more detail we adopt a more parametric approach, specifying an earnings distribution for inventors outside of the R&D sectors, in order to show that for very skewed pay-off function, taxes matter less than for less skewed ones in the presence of concave utility. Consider lowering marginal tax rates on high income earners. Assume that the CDF of the returns to innovation is $F(\pi)$ and the CDF of the taste for the two sectors is $G(v)$. Since the baseline wage is common in both sectors except for a shift factor that is common across individuals (above denoted \tilde{w}), we abstract away from this. The benefit of working in the R&D sector is:

$$\varphi = F_0(\pi, v) = F(\pi(1 - \tau)) - G(v)$$

Consider the effect of changing tax on the probability of going into the R&D sector:

$$\begin{aligned} \frac{\partial \phi}{\partial(1 - \tau)} &= f_0 E[u'(\pi(1 - t))\pi] \\ &= f_0 \frac{E[u'\pi]}{E\pi} E\pi = f_0 \bar{u}' I \end{aligned}$$

where f_0 is the PDF of F_0 and \bar{u}' is the profit weighted mean of marginal utility (u'). When $F(\pi(1 - \tau))$ gets more skewed \bar{u}' will get small, so the marginal effects of tax on entering the R&D sector will also become small holding innovation revenue (I) fixed.

The key aspect of the result is that for very skewed pay-off functions, taxes will matter less than for less skewed ones in the presence of concave utility. To examine this further we calibrate the model to some empirical distributions. To do this we put some more structure on the model. We assume that innovation returns are Pareto distributed, tastes are log normally distributed and utility takes the CRRA form (as above). We draw a million individuals from these distributions for different levels of the Pareto tail parameter, α between 1.1 and 2. This determines the skewness of the distribution ($\frac{\alpha}{\alpha-1}$), with lower levels denoting a more skewed ("thick tailed") distribution. We then consider the effects of a one percentage point change in the top marginal tax rate from

its current level US level (as discussed above). The mean effect of lowering marginal taxation for different levels of skewness of the innovation returns is shown in Figure A20 (where we have used a risk aversion parameter to be $\delta = 1.5$).

Lowering taxes always has a positive effect on innovation which is why all the values on the y-axis are above zero. R&D falls by about 1% when skewness is 3 ($\alpha = 1.5$), for example. However, as skewness increases, the marginal effect of taxation falls. When skewness is 5 ($\alpha = 1.25$) the marginal effect of innovation is about -0.20% and effectively zero for very high levels of skewness. Our analysis of inventor careers showed heavy levels of skewness is consistent with other papers on the distribution of patent values using other methods of valuing patents such as future citations, patent renewal fees, licensing revenues or surveys of inventors.

B.4 Gifted and Talented Programs for the Disadvantaged

A supply side policy would be to increase human capital so that there are more highly skilled people who could become inventors. We have seen that children born to poor parents appear to be at a particular disadvantage in growing up to become inventors in later life and have argued that this is, in large part due to their slower acquisition of human capital during school years rather than “initial ability”. Test scores at grade 3 only accounted for 30% of the lower probability of children from poorer families growing up to be inventors, whereas education by 22 years of age could account for virtually all of the income-invention gap.

Card and Giulano (2014) finds evidence that a Gifted and Talented program particularly benefited high ability children from lower income families. Tracking such children and putting them in separate classrooms within public schools raised math and reading by 0.3 of a standard deviation. Looking at the evidence from Figure 3, this implies that such a program would roughly double the future invention rate among the treated group. Since the cost of this program was effectively zero, such policies would seem very desirable on grounds of both equity and growth.⁵⁵ By contrast, reducing top marginal tax rates is likely to have some cost in terms of lost revenue to the government.

The pay-off in terms of innovation for such educational programs are long-term. The impact of cutting marginal tax rates is somewhat speedier, although note that our model assumes that this would effect the flow of new graduates into the R&D sector, rather than immediately affecting the stock of inventors (it is different than an R&D tax credit in this respect).

⁵⁵Another set of supply side policies would be the spreading of high performance school practices as detailed by Fryer (2014) and many others. These seem to have particular benefits for disadvantaged children.

B.5 A Simple Calibration of the Effect of Educational Supply-Side Policies

As discussed in Section VI.C, our estimates can be used to assess the potential gains from supply-side (“extensive margin”) policies. Three sets of policy parameters are needed to compute the effect of such policies on innovation. First, we must determine a “reference income group” for which the rate of innovation is policy invariant. Both intuitively and on the basis of the evidence presented in Card and Giulano (2014), supply-side policies will mostly affect disadvantaged students. Accordingly, we consider three policy scenarios with various reference income groups: the policy will reduce a certain fraction of the innovation gap between these (high-income) groups and the rest of the population. Supply-side policies affect only students in families below the 90th percentile of the income distribution in the first scenario we consider, below the 80th percentile in the second scenario, and below the median in the third scenario. Second, we must determine the fraction of the innovation gap between the reference group and the rest of the population that is due to ability. We consider a scenario in which ability accounts for 30% of the gap (in line with our empirical estimates of Table 1) and another scenario in which it accounts for half of the gap (which we view as an upper bound). Finally, we must determine the fraction of the innovation gap - after adjusting for ability - that can be closed by supply side policies. We consider a series of scenarios in which the innovation gap between the reference group and the rest of the population can be closed by between 5% and 50%, respectively.

Appendix Table B1: A Calibration of the Potential Gains from Supply-Side Policies, % Increase in Inventor Population

Reference Group:		Richest 10%		Richest 20%		Richest 50%	
Share of Innovation Gap From Ability:		30%	50%	30%	50%	30%	50%
	5%	7.5%	5.4%	5.2%	3.7%	2.1%	1.5%
	10%	15.1%	10.8%	10.5%	7.5%	4.2%	3.0%
Fraction of the Ability-Adjusted Innovation Gap Closed by Policy:	20%	30.2%	21.6%	21.1%	15.1%	8.5%	6.1%
	30%	45.4%	32.4%	31.6%	22.6%	12.7%	9.1%
	40%	60.5%	43.2%	42.2%	30.1%	16.9%	12.1%
	50%	75.6%	54.0%	52.7%	37.7%	21.2%	15.1%

Appendix Table B1 reports the percentage increase in the population of inventors under these various scenarios - the gains are in general very large. For each percentile of the income distribution,

we compute the implied increase in the percentage of inventors and then average over the relevant range of the income distribution. Alternatively, for each scenario the calculation can be carried out in one step as follows:

$$\% \Delta \text{Inventors} = \frac{s \times (1 - a) \times (R^I - R^{\bar{I}}) \times (1 - P^I)}{R^I \times P^I + R^{\bar{I}} \times (1 - P^I)} \times 100$$

where s is the share of the (ability-adjusted) innovation gap closed by the policy, a is the share of the initial innovation gap accounted for by ability (which by definition cannot be closed by policy), R^I is the average rate of inventors in reference income group I , $R^{\bar{I}}$ is the average rate of inventors in the rest of the population, and P^I is the share of the reference income group in the total population.

The benchmark scenario discussed in Section VI.C uses the top 10% as the reference group, assumes that ability accounts for 30% of the innovation gap, and that policy can close 20% of the ability-adjusted innovation gap. The increase in the number of inventors induced by the policy is equal to over 30% of the inventor population.

Another useful benchmark to consider is based on the distribution of the innovation-income gaps across states (where we consider state of birth). On average, the rate of inventors among children in the top 20% of the family income distribution is 177% higher than for children in the bottom 80%, but there is a lot of variation across states. The 5th percentile of the distribution of innovation gaps across states is 132% (e.g. New Hampshire) while the 95th percentile is 222% (e.g. Georgia). The 5th percentile of the distribution can be considered to be a “feasible benchmark” that other states could potentially converge to.⁵⁶ In other words, the innovation-income gap could be reduced by $\frac{177-132}{177} = 25.4\%$. Using a calculation similar to above⁵⁷, this corresponds to a 38.4% increase in the overall inventor population.

Regarding the composition effect, the data underlying Figure 2 shows that there is a 44% gap between the 3rd grade test scores (expressed in standard deviations relative to the mean) of inventors from high income families and those of inventors from families below the 80th percentile in the income distribution. In Section VI.A, we have discussed that this difference can result from the fact that high-ability children from low-income families are less likely to enter the R&D sector

⁵⁶We have checked that, when expressed in percentages, the state income-innovation gap is not correlated with either the state population or with the number of patents per resident. These results are available from the authors upon request.

⁵⁷The formula is

$$\frac{0.254 \times (R^I - R^{\bar{I}}) \times (1 - P^I)}{R^I \times P^I + R^{\bar{I}} \times (1 - P^I)} \times 100$$

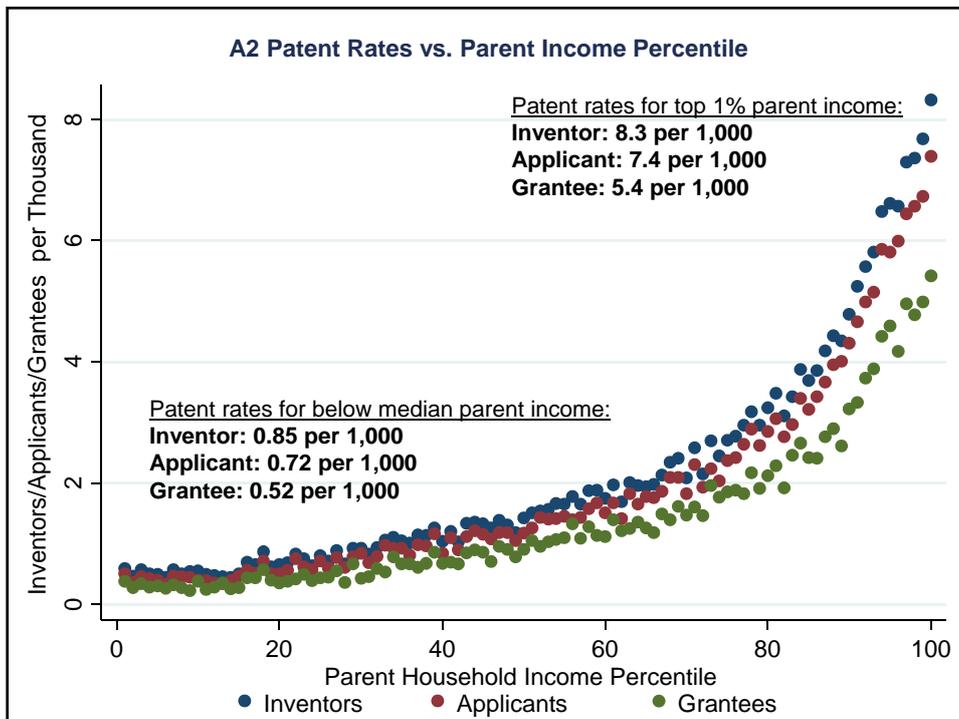
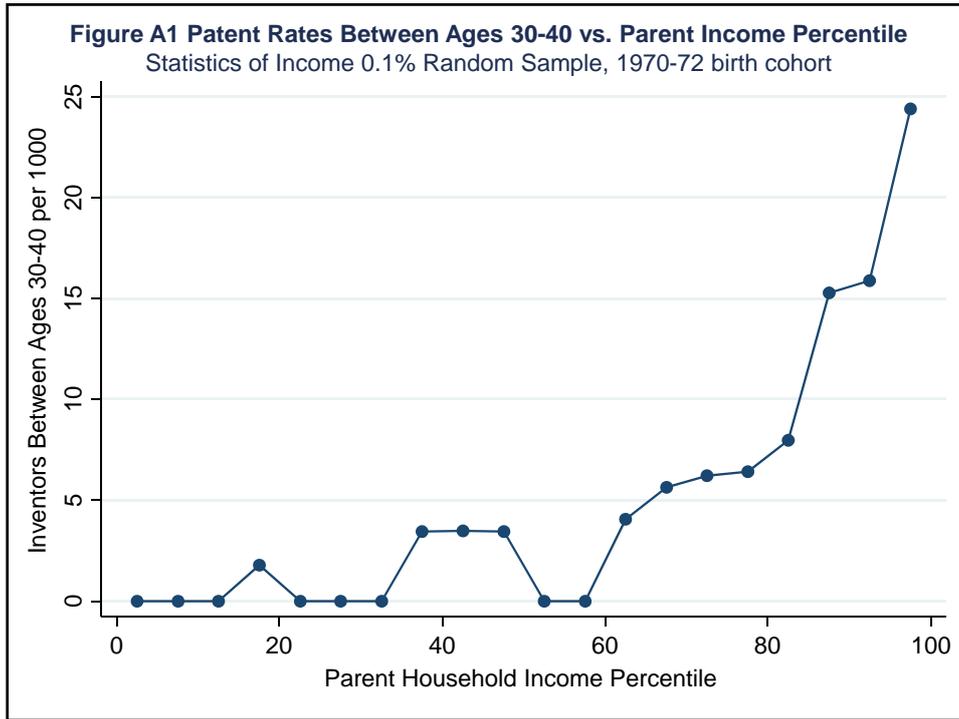
where I is the top 20% and \bar{I} the bottom 80%.

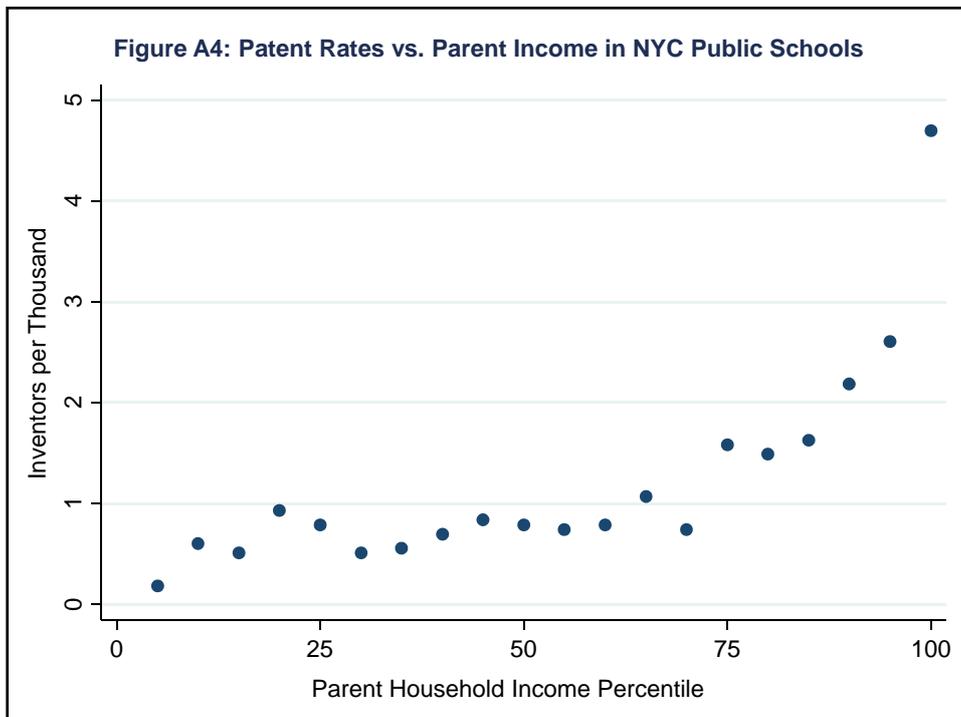
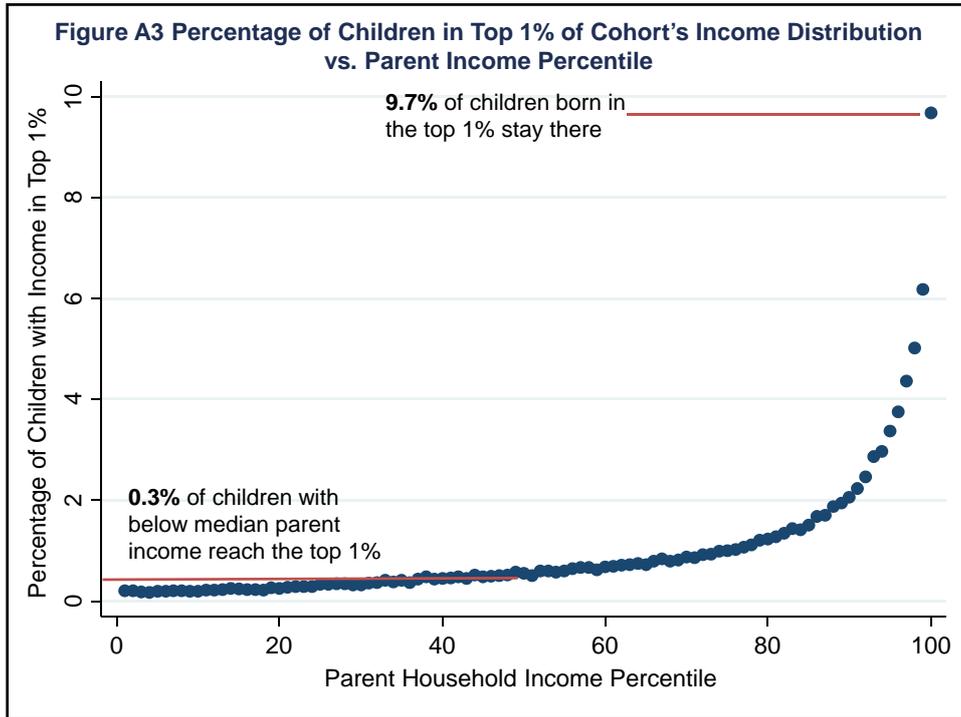
than their high-income counterparts because they have had less exposure to innovation. One way of calibrating the magnitude of the composition effect is to consider hypothetical policies that would keep constant the total number of inventors from low-income families going into the R&D sector but that would bring into this sector a higher share of high-ability low-income children and a lower share of low-ability low-income, compared to the current equilibrium. The composition can have an effect on the overall rate of innovation (and growth) since higher ability individuals will produce better innovations. The details of the calculation are reported below. We estimate the relevant parameters and find that the composition effect is smaller than the level effect. The reason for this is that the level effect is very large: intuitively, we found that innovation rates are very different across groups (a tenfold difference), while test scores conditional on innovation (innate ability of individuals going into innovation) are much more similar.

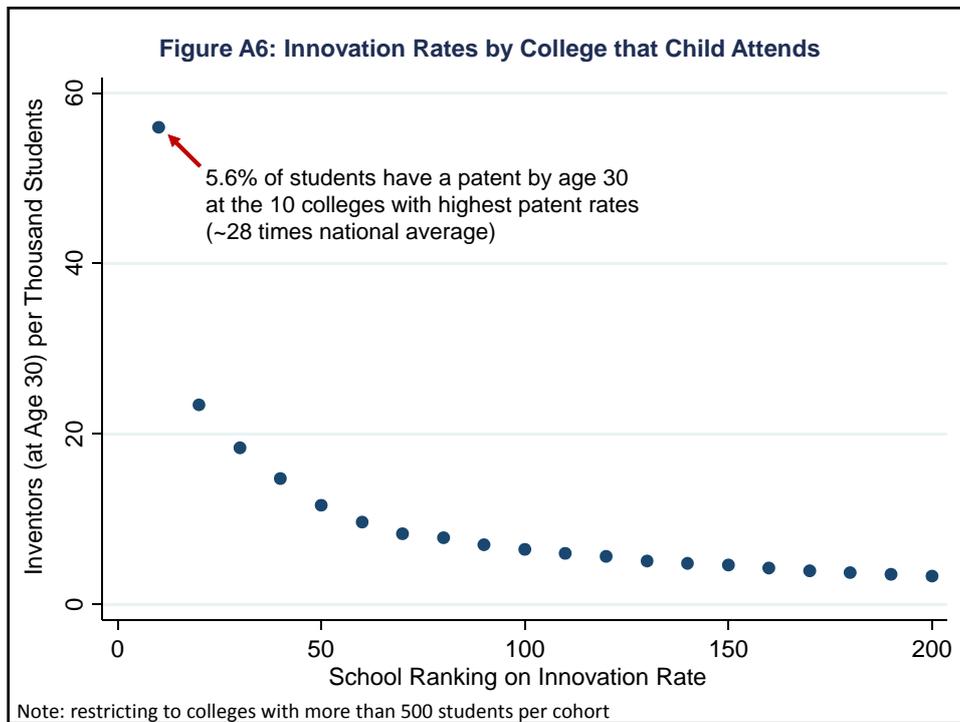
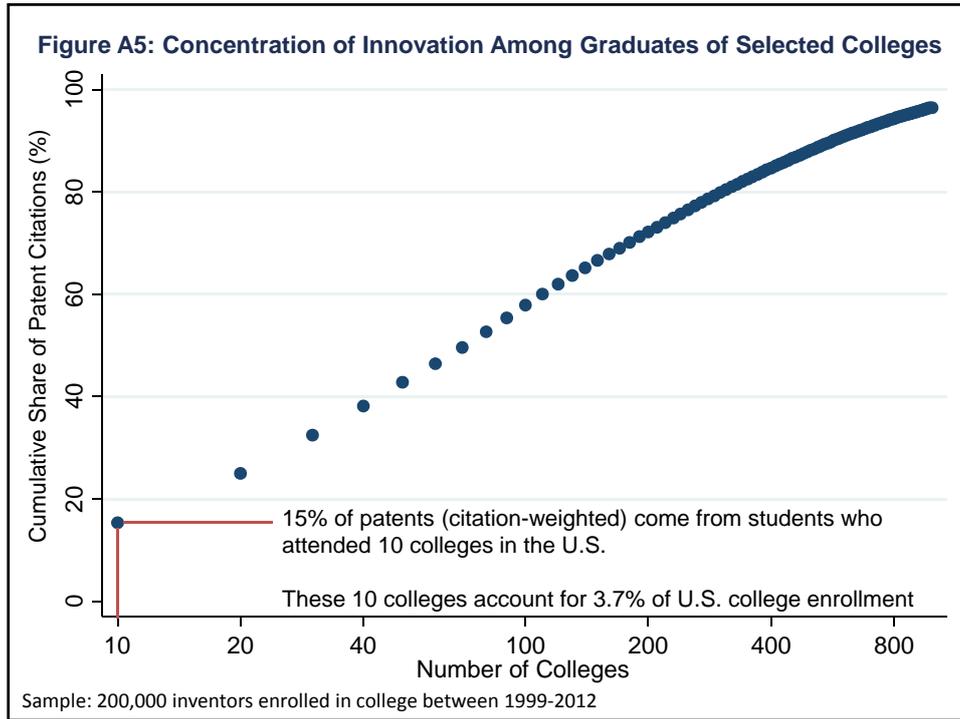
The magnitude of the composition effect can be calibrated as follows:

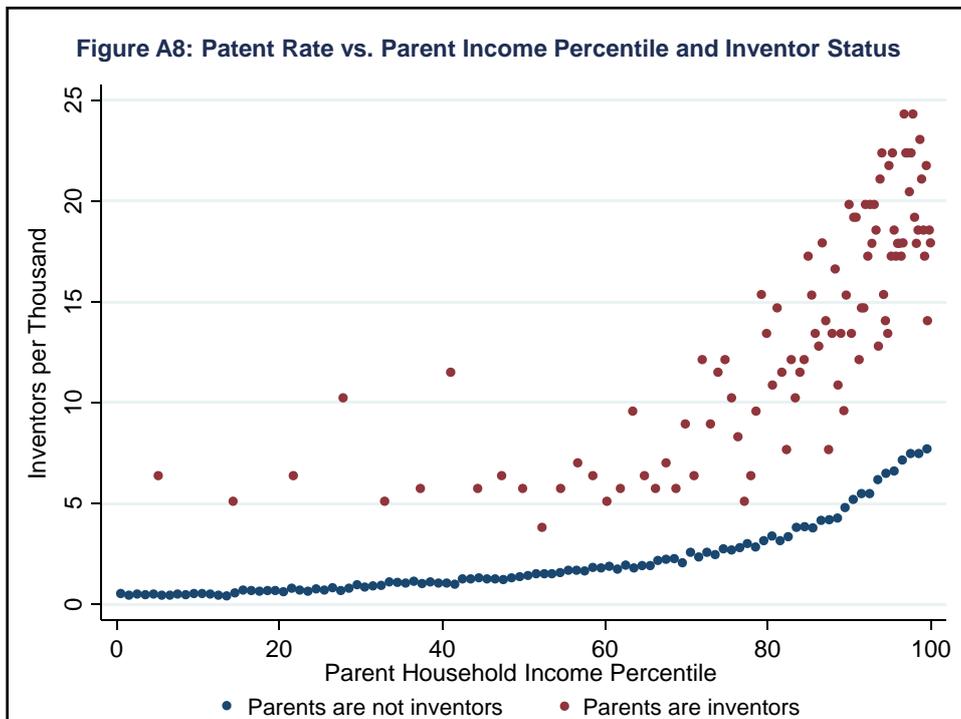
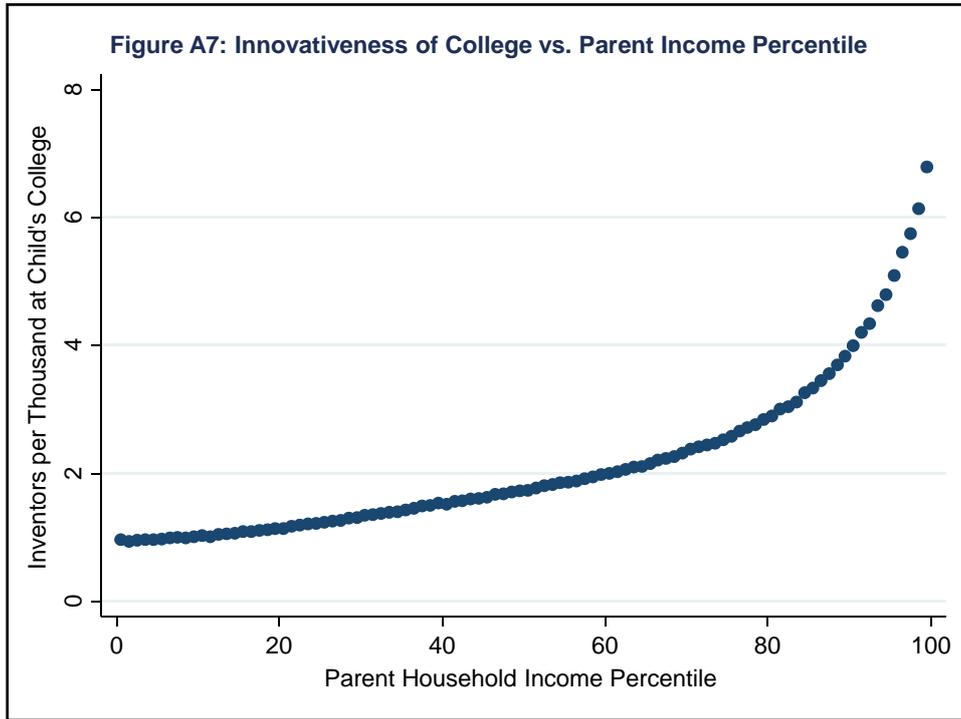
$$\% \Delta Innovation = \frac{\tilde{s} \times (Q^I - Q^{\bar{I}})}{Q^{\bar{I}}} \cdot \frac{Q^{\bar{I}} \times (1 - P^I)}{Q^I \times P^I + Q^{\bar{I}} \times (1 - P^I)} \times 100$$

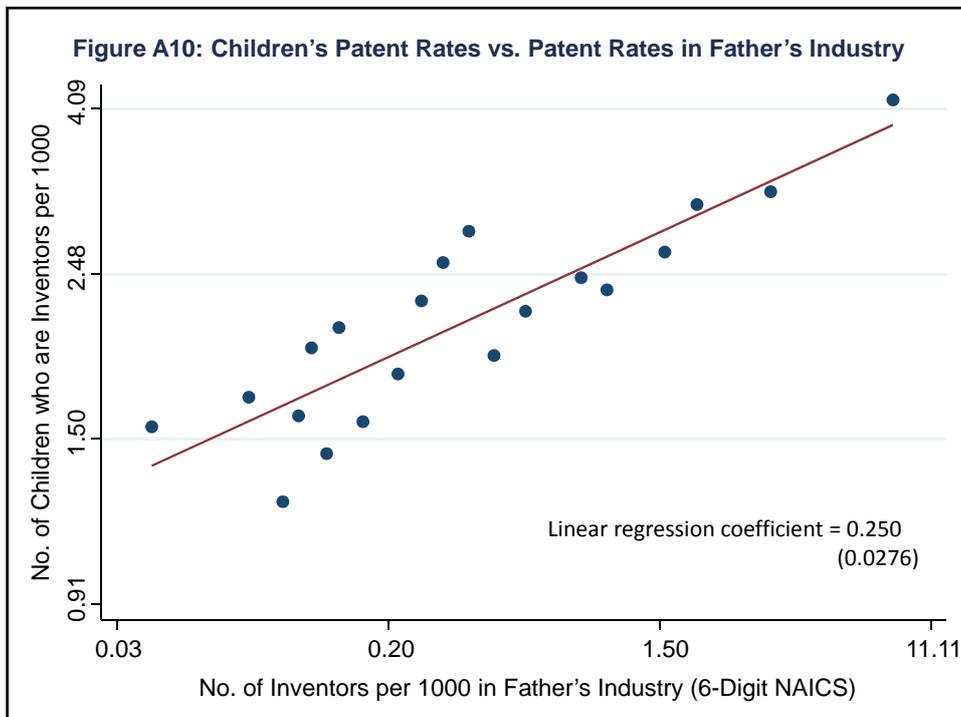
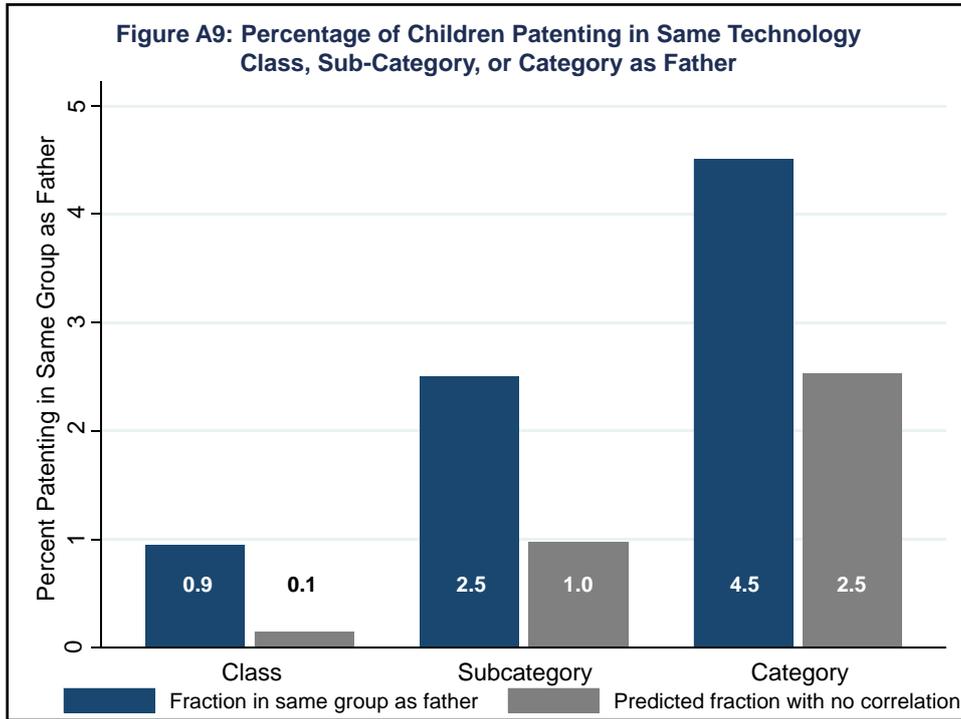
where Q^I is the expected amount of (quality-adjusted) innovations over the course of the career of an inventor in the reference income group I ($Q^{\bar{I}}$ is the same in the rest of the population) and \tilde{s} is the share of the “quality-adjusted innovation gap” ($Q^I - Q^{\bar{I}}$) between the two income groups that can be closed by policy. Thus, the first term in the formula is the share of the quality-adjusted innovation gap that can be closed by policy and the second term is the share of total quality-adjusted innovations that is accounted for by the low-income group. Consider an example where the reference group I is children from families in the top 10% of the income distribution. In the data, I represents 34% of quality-adjusted innovation (as measured by citation-weighted patents). We also observe that these inventors have third-grade test scores that are 44% higher than inventors in the rest of the distribution. Under the assumptions (i) that a percent increase in third-grade test score corresponds to a one percent increase in expected quality-adjusted innovations during an inventor’s career, and (ii) that supply-side policies can close 20% of the gap in third-grade test scores, then such policies would increase innovation by just under 6% ($0.20 \times 0.44 \times (1 - 0.34) = 5.81\%$).

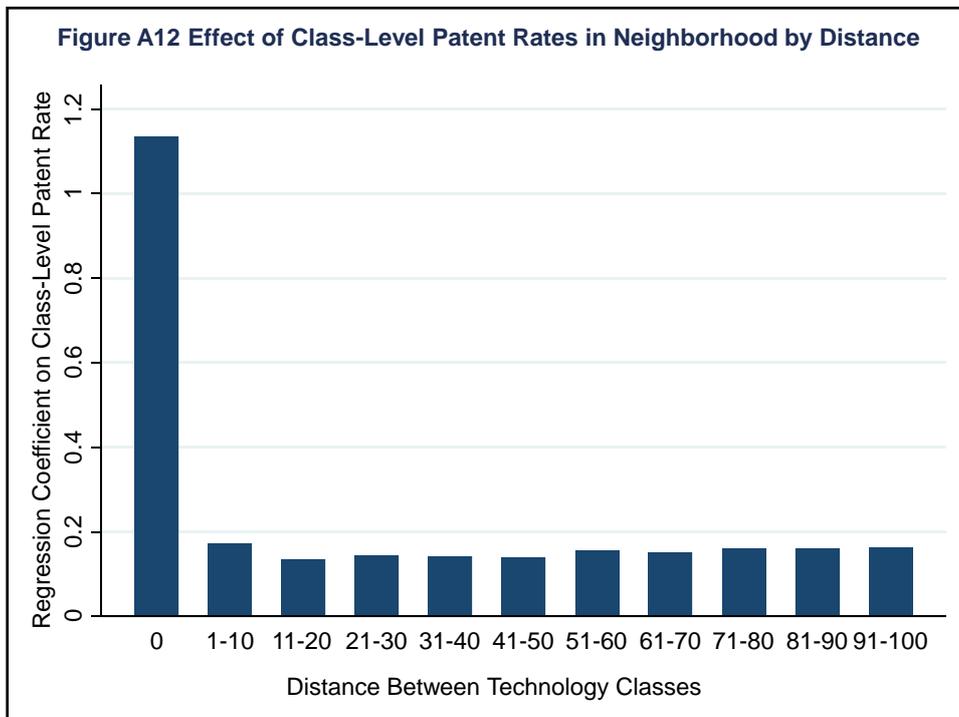
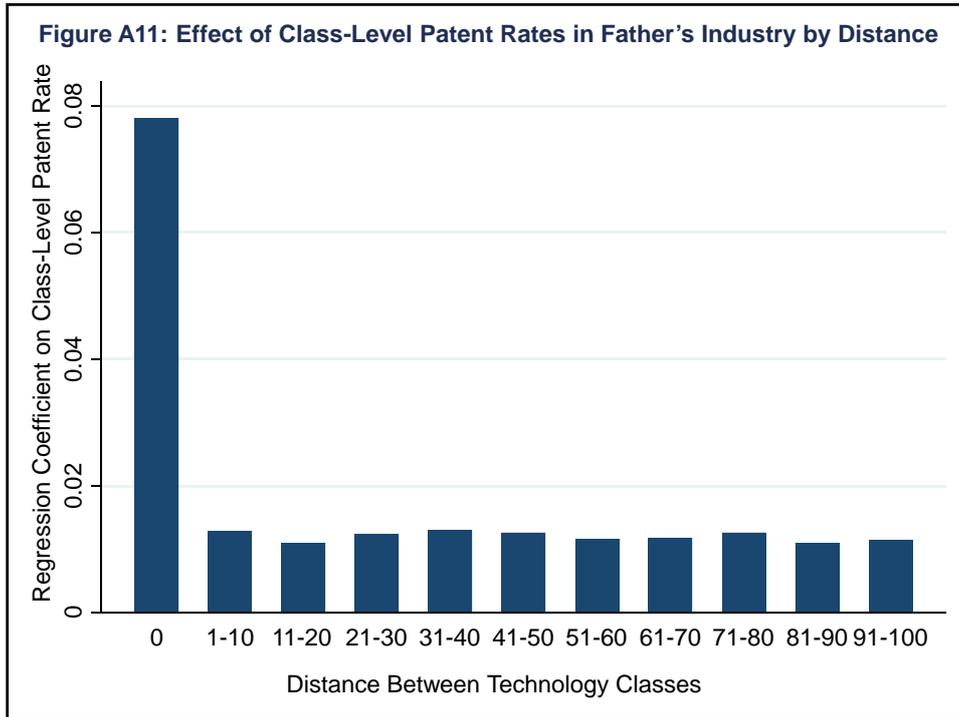


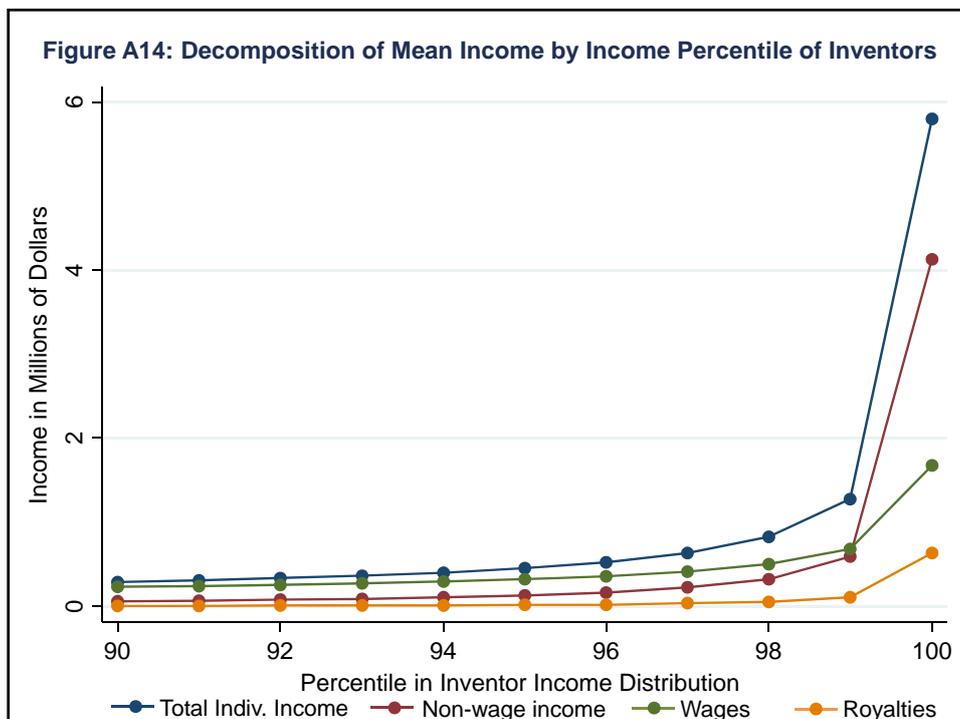
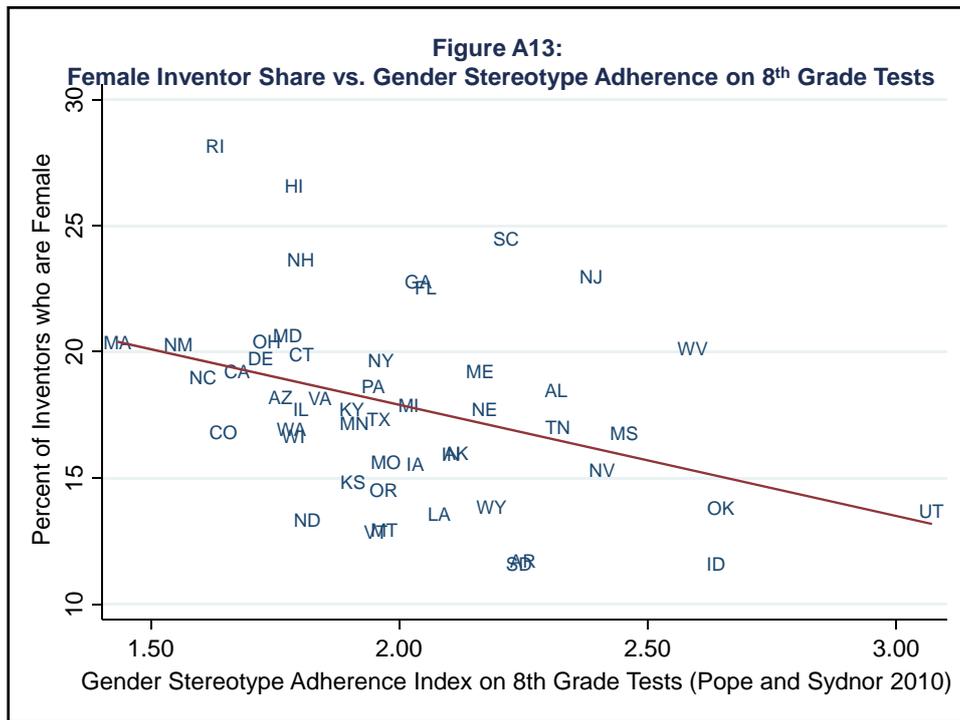


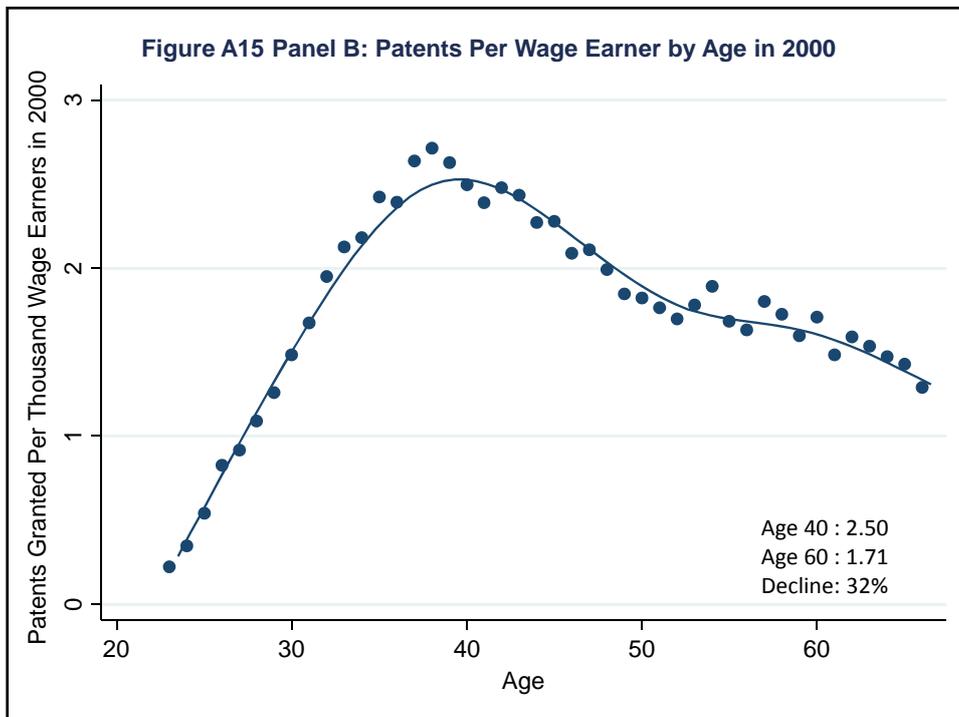
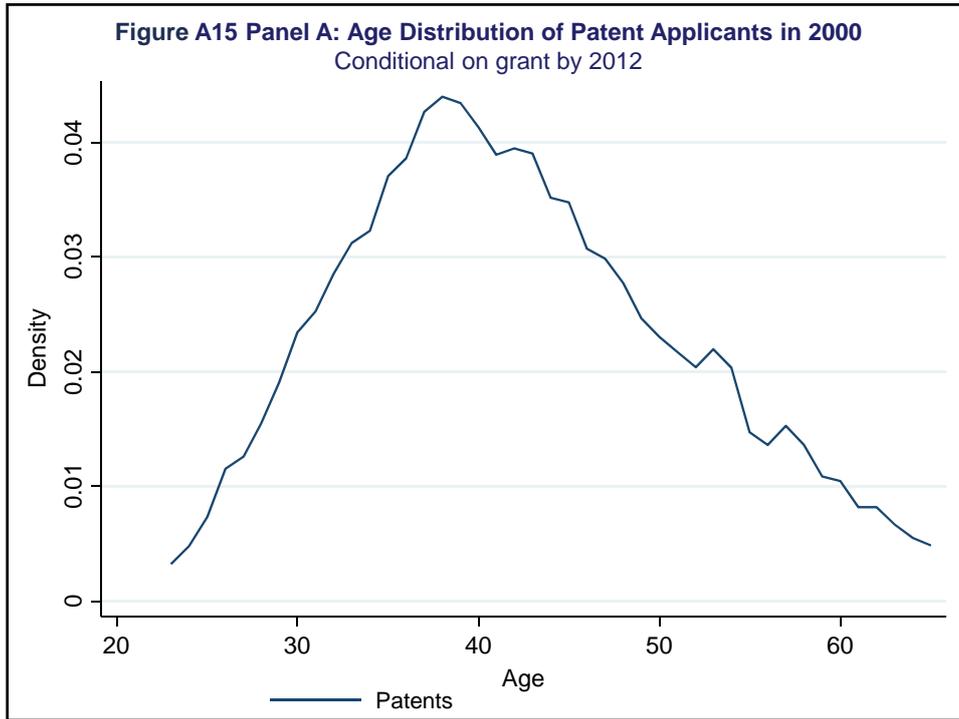


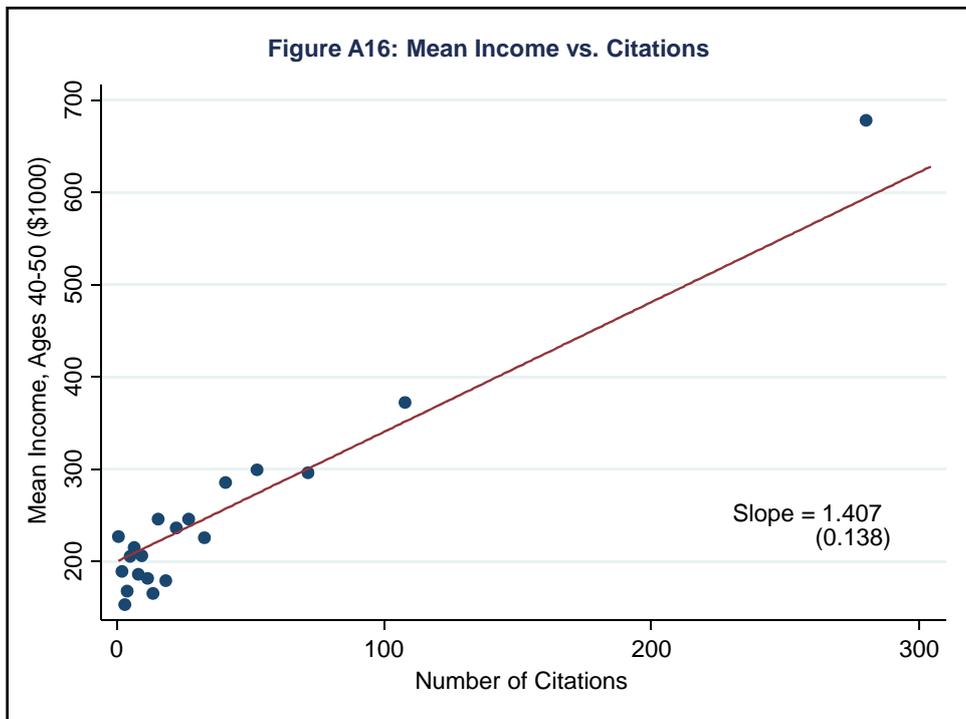
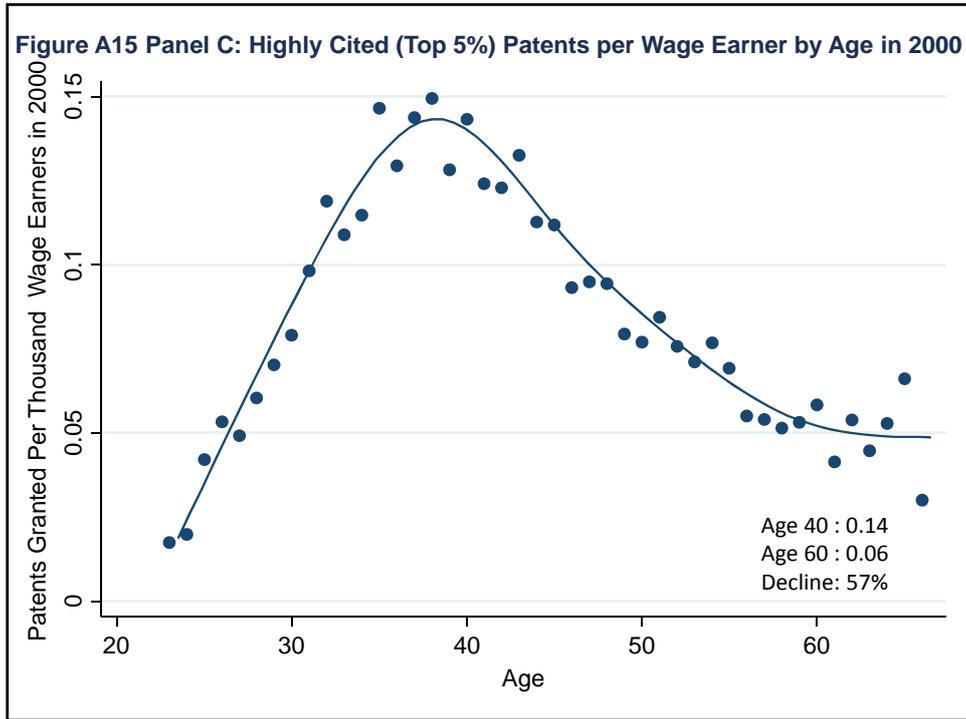












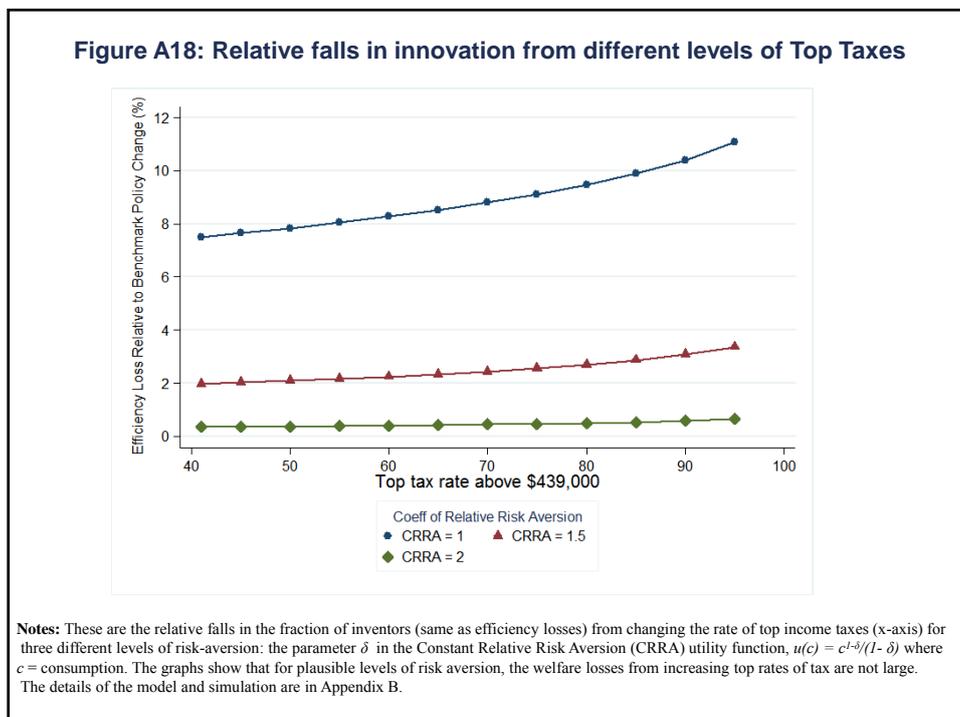
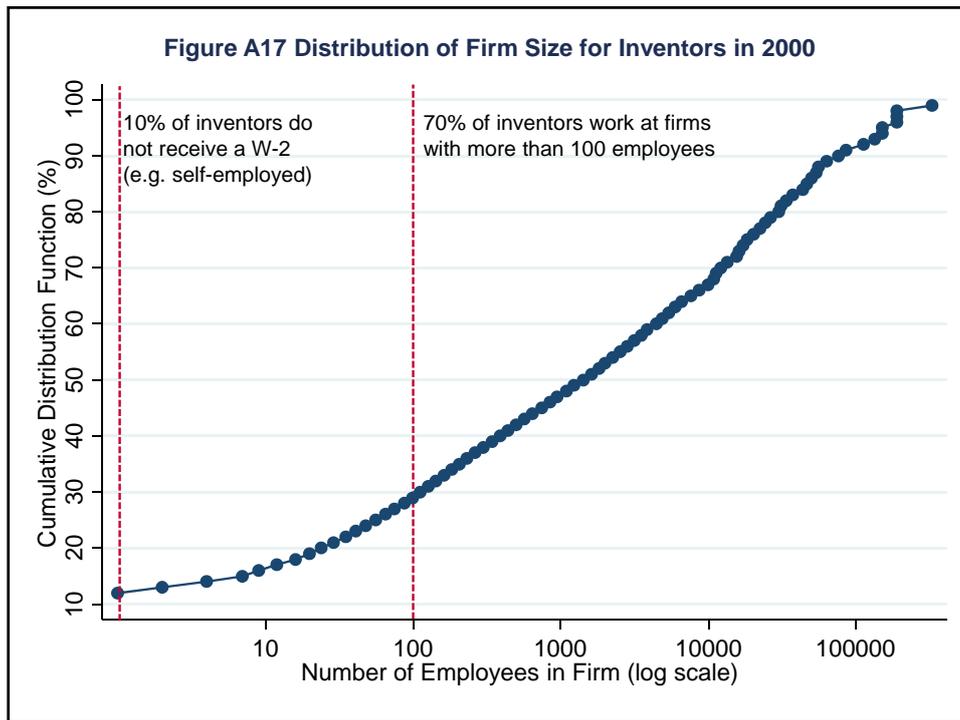
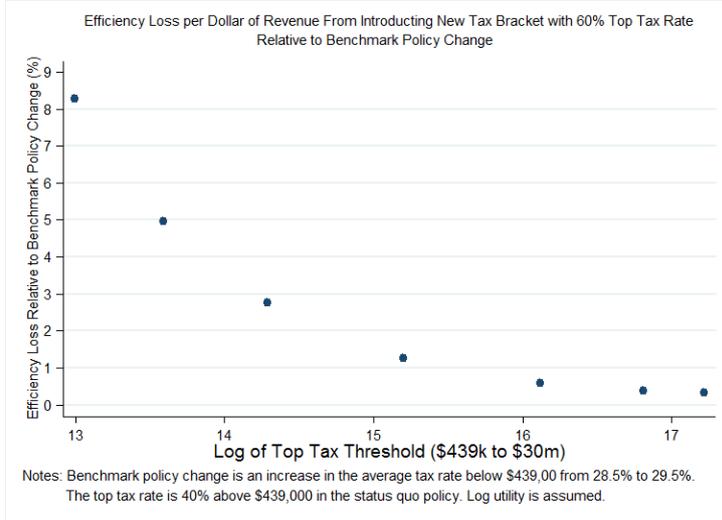
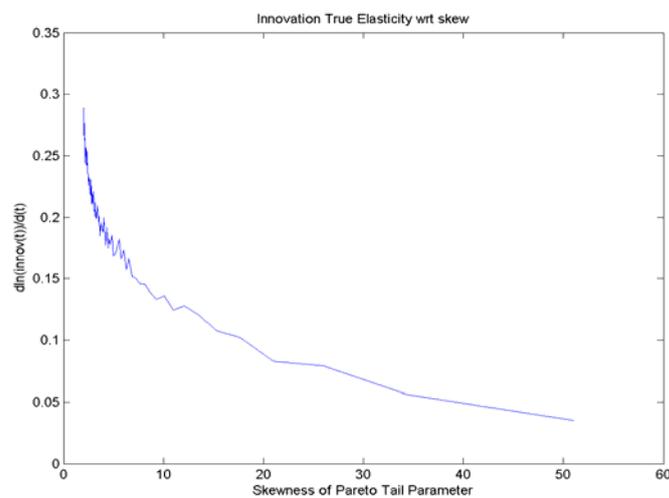


Figure A19:
Relative falls in innovation from Introducing an Additional Top Tax Threshold



Notes: These are the efficiency losses from introducing a new “millionaires’ tax” of 60%. We consider different thresholds from the current level of the existing threshold at \$439,000 at the far left of the x-axis to \$30million on the far right. The details of the model and simulation are in Appendix B.

Figure A20: Changes in innovation elasticity with respect to top taxes as the Skewness of wage returns to innovation rises



Notes: This graph illustrates how the elasticity of the fraction of people entering the R&D sector changes with marginal top tax rates. The graph shows that as we increase the skewness of the returns to innovation the marginal impact declines. CRRRA $\delta = 1.5$; skewness = $\alpha / (\alpha - 1)$ where α is The Pareto tail parameter. Details in Appendix B.