The Long-Term Impact of Business Cycles on Innovation:

Evidence from the Massachusetts Institute of Technology

Pian Shu*

January 6, 2012

Abstract

I explore a novel channel through which short-term economic fluctuations affect the long-run innovative output of the U.S. economy: college graduates' initial career choices. I develop a two-period Roy-style model to show that shocks to initial career choices could affect long-term patent production by changing graduates' long-term occupational affiliation or changing their acquisition of inventive human capital. Using a newly constructed data set on the patenting history of all individuals obtaining a bachelor's degree from the Massachusetts Institute of Technology (MIT) between 1980 and 2005, I find that cohorts graduating during economic booms produce significantly fewer patents over the subsequent two decades. A one percentage point decrease in the unemployment rate in the year of scheduled graduation on average decreases the future annual patent output of a cohort by around 5%, or approximately 2.5 patents per year for an average-size cohort. Economic conditions at the time of graduation do not affect the number of graduates who patent or their characteristics. The decrease in patent output of cohorts graduating during booms is a result of lower inventive output from inventors with relatively low GPAs, and marginal patents receive fewer citations than the average and median patents. I find no evidence that initial economic conditions affect inventors' long-term occupational affiliation, suggesting that the effect on patent production is primarily due to differences in inventors' long-term level of inventive human capital.

*Department of Economics, MIT. Email: pshu@mit.edu. I am very grateful to my advisers David Autor, Daron Acemoglu, Scott Stern, and Lee Fleming for their generous guidance and detailed feedback on this project. I also thank Pierre Azoulay, Panle Barwick, JB Doyle, Joseph Hadzima, Jaimie Lee, Roman Lubynsky, Hong Luo, Attoinette Schoar, David Simon, Heidi Williams, and numerous seminar participants at MIT and NBER Productivity lunch for helpful comments and suggestions. I'm grateful to Suzanne Berger, Maggy Bruzelius, Claude Canizares, Daniel Hastings, Deborah Liverman, Joseph Recchio, Ri Romano, Lydia Snover, and especially Gregory Harris for help with data collection. All errors are my own.

1 Introduction

Allocating talent to innovative activities is key to promoting a country's long-term economic growth (Baumol, 1990; Murphy et al., 1991). However, empirically we know little about what factors affect talented individuals' innovative output. Do short-term shocks to individuals' career choices have a long-term impact on innovation? Who are the people most affected? In this paper, I provide empirical evidence to answer these questions by exploring one particular source of exogenous variation: economic conditions at the time of college graduation. Using individual patent output as a measure of innovative activities, I estimate the causal impact of initial labor market conditions on the long-term patent production of a sample of highly skilled individuals: the alumni of the Massachusetts Institute of Technology (MIT).

To show how initial labor market conditions could affect an individual's long-term patent production, I develop a two-sector two-period model. My model combines features from standard static Roy models (Roy, 1951; Willis and Rosen, 1979; Gould, 2002; Heckman and Honore, 1990; Mulligan and Rubinstein, 2008) as well as models with occupation-specific and task-specific human capital (Neal, 1995; Gibbons and Waldman, 2004). The theory indicates that, by changing initial career choices, initial economic conditions could affect future patent production in two ways. First, if individuals acquire occupation-specific human capital on the job, initial market conditions could affect individuals' long-term occupational affiliations.[1] Second, by altering graduates' career paths, initial economic conditions could affect their future level of human capital, even when there is no effect on long-term occupational affiliations.

I examine the empirical implications of my model using a newly constructed longitudinal data set on the patenting history of everyone who received a Bachelor's degree

[1]Oyer (2006, 2008) provides empirical evidence on this observation. There are also other causes for sticky jobs, such as search frictions and employer's uncertainty about the workers' skill (Gibbons et al., 2005; Oreopoulos et al., 2012). I discuss their implications in Section 2.

from MIT between 1980 and 2005. I match the alumni to the U.S. inventor database from Lai et al. (2011b) based on names and locations.[2] My data include 27,145 graduates with over 475,000 person-year observations. Around 16% of the graduates have produced at least one patent in the years I study. Overall, the inventors have produced nearly 25,000 patents and received over 323,000 patent citations by the end of 2010. I link the patent data to individual-level administrative records on demographics and academic performance at MIT to control for a rich set of characteristics in my empirical analysis.

Since MIT is one of the major technology-based universities, MIT alumni are particularly suited for the purpose of this study. My sample has nearly 24,000 engineering and science graduates, who constitute 0.24% of the total number of engineering and science bachelor's degree recipients between 1980 and 2005 (NCES, 2011).[3] Previous studies have shown that firms founded by MIT alumni generate hundreds of billions of dollars in revenue and hundreds of thousands of jobs in the U.S. (Roberts and Eesley, 2009). It is thus not surprising that many alumni are productive inventors. The MIT alumni in my sample have produced around 1.2% of the utility patents with U.S. origin granted between 1981 and 2010 (USPTO, 2011). An average patent produced by the MIT graduates in my sample receives approximately 1 citation per year since the year of patent application, which is twice as much as an average U.S. utility patent produced during the same period.

I find that adverse labor market conditions at the time of college graduation lead to an *increase* in the future patent production of MIT alumni. A one-percentage-point

[2] The raw patent data from the United States Patent and Trademark Office (USPTO) do not provide unique identifiers for inventors, making it difficult to track the output of an inventor over time. Lai et al. (2011b) provide a solution by employing a Bayesian supervised learning approach to match inventors to U.S. patents.

[3] Over 17,000 graduates in my sample are engineering majors, equivalent to 0.6% of the total engineering graduates between 1980 and 2005. There are over 4000 degree-granting institutions in the National Center of Education Statistics (NCES) population.

increase in the national unemployment rate in the year of scheduled graduation increases the average graduate's annual patent output by around 5%, or approximately 2.5 patents per year for an average size cohort of 1000 graduates. A 1.25 standard deviation decrease in the equity market return during the students' sophomore and junior years has a similar effect. The effect of initial economic conditions on patent production increases over time and is largest between 10 and 20 years after graduation, which are also graduates' peak inventive years. Meanwhile, economic fluctuations have no measurable effect on the contemporaneous innovative output of graduates.

There are two possible explanations for these findings, which are not mutually exclusive. First, more graduates may become inventors as a result of graduating in a worse economy (changes at the extensive margin). Second, inventors who graduate in a worse economy may be more productive (changes at the intensive margin). Comparing the patent production of the 1980-1995 cohorts during their first 15 years after graduation, I find no evidence for changes at the extensive margin. Inventors from recession cohorts are not ex ante more likely to patent, where I use their cumulative grade point average (GPA) at MIT as a measure of their inventive ability at the time of graduation.[4] Thus, graduates who become inventors would most likely patent regardless of initial labor market conditions, but graduating in a worse economy increases the number of patents they produce.

The increase in patent production due to graduating in adverse labor market conditions comes primarily from science majors working in non-software-engineering sectors, such as chemical, drugs and medical industries. Initial economic conditions have no significant effect on the distribution of the inventors' long-term sector. Furthermore, graduating in a worse economy has a significantly negative effect on the time that an inventor takes to produce their first patent. Taken together, these results suggest that

[4] I normalize the GPA by major and cohort. The normalized GPA significantly predicts future patent production.

the accumulation of human capital is likely to be the main channel through which initial labor market conditions affect long-term patent production. The most plausible explanation is that inventors from recession cohorts either start working in patent producing sectors sooner or are more likely to go to graduate school, though my data do not allow me to determine the relative importance of the these two channels.

I show that initial conditions affect the patent production of inventors with relatively low GPAs. Graduates with the highest inventive ability upon graduation do not seem to be affected. Consistent with the finding that the relatively less inventive individuals produce the marginal patents, those patents also receive slightly fewer citations than the average and median patents in my sample.[5] These results suggest that there exists positive sorting into patent production, where the most talented inventors produce the same patents regardless of their graduating economic conditions.

My results have several important implications. First, I provide some of the first empirical evidence on how talented individuals invent. Despite the large number of studies on the patent production of firms,[6] very few papers examine the determinants of patent production at the individual level. Compared with previous studies, my data have distinct advantages, as I observe a large group of potential innovators with homogenous training, characteristics, and abilities. Previous work such as Amesse et al. (1991), Kerr

[5]It is important to note that the marginal patents still receive more citations than the average and median of all U.S. utility patents.

[6]This literature provides ample evidence on the economic value of patented inventions. At the firm-level, the number of patents produced by a firm strongly and positively correlates with its research and development (R&D) expenditure, and this relationship holds across different industries (Griliches, 1990). Other inputs such as venture capital funding also significantly increase patents (Kortum and Lerner, 2000). Different measures of patent production, such as citation-weighted patent count, number of patents per R&D dollar, citations per patent, and a weighted index of multiple indicators of patent quality, are all found to boost firms' market value and productivity (Comanor and Scherer, 1969; Trajtenberg, 1990; Bloom and Van Reenen, 2002; Hagedoorn and Cloodt, 2003; Lanjouw and Schankerman, 2004; Hall et al., 2005). Studies also directly calibrate the economic value of patents using patent renewal data and surveys, and show that the value increases in the number of citations received (Schankerman and Pakes, 1986; Lanjouw et al., 1998; Harhoff et al., 1999). Although not all inventions are patentable, firms patent most of the inventions that can be patented, even in industries where patent protection is relatively unimportant (Mansfield, 1986; Cohen et al., 2000).

(2008), and Jones (2009) primarily uses samples of only inventors to study their behavior and characteristics.[7] Without a comparison group of non-inventors, these studies do not shed light on vital issues such as what leads talented individuals to invent. My data also include a rich set of individual characteristics, which helps to determine the factors that affect patent production.[8] Furthermore, my results are relevant for understanding how top engineering and science students in the U.S. innovate, which has key policy implications.[9] For instance, Romer (2001) argues that a top priority of the innovation policy in the U.S. should be to increase the supply of engineers and scientists. My results provide a first step towards quantifying the actual return, in terms of producing patented inventions, of a potential policy that provides incentives for engineering and science students to pursue careers in innovative sectors.

This study also presents some of the first evidence on the links between business cycles, talent allocation, and long-term innovation at the micro-level. A literature has analyzed the *contemporaneous* relationships between business cycles and relevant outcomes, including labor productivity (Bernanke and Parkinson, 1991; Goldin, 2000), technological progress (Field, 2003; Nicholas, 2003, 2008), and venture capital investment (Nanda and Rhodes-Kropf, 2011). My study complements this literature by showing that through changing talent allocation, business cycles could also have a *dynamic* effect on future inventive output. Since adverse labor market conditions change the relative demand for labor of different sectors, my results suggest that sectors producing more patented

[7] One exception is the study by Ding et al. (2006), which finds that female life scientists are less likely to patent than male life scientists.

[8] It also helps verify the accuracy of my matching procedure. For example, although I do not use majors at MIT in my matching, the engineering and science majors in my sample are significantly more likely to patent than the other majors.

[9] Compared with other top engineering and ivy league universities, MIT admits similar students based on standardized test scores(Grove, 2011). The 25th and 75th percentiles of the SAT math score of admitted students at MIT are 740 and 800, respectively. Other top engineering programs such as California Institute of Technology and the engineering school of Cornell University have similar score range for their admitted students. The 25th percentile of the SAT math score at ivy league universities and other top universities such as Stanford University and University of Chicago is around 680, and the 75th percentile is around 770.

inventions are less cyclical, and that increasing labor demand from highly pro-cyclical sectors could potentially have a negative impact on long-term innovation. For example, graduating during a recession leads to higher graduate degree attainment and higher enrollment in PhD programs in science and engineering (Bedard and Herman, 2008; Kahn, 2010). In contrast, finance is a prominent example of a highly pro-cyclical sector (Oyer, 2008). Kedrosky and Stangler (2011) make the alarming observation that the increase in financial employment over the last several decades parallels the decline in new firm founding activities in the United States. However, the *causal* effect of going to finance (or graduate school) right after college on an individual's long-term innovative output remains unexamined. Although I do not directly estimate such effects in this paper, my results suggest that shocks to initial career choices could have long-term effect on producing innovations, pointing to the importance of potential follow-on research that addresses these open questions.

An influential line of work shows that, because innovation generates positive spillovers, innovators receive inadequate compensation relative to their contribution to society. As a result, the equilibrium level of innovation is less than optimal (Nelson, 1959; Arrow, 1962). While the patent system is designed to help inventors capture at least some of the benefits from their innovations, previous empirical studies suggest that there are large, positive social externalities to the creation of new ideas that are not fully internalized by the patent system (Mansfield et al., 1977; Jaffe, 1986; Trajtenberg, 1989; Jaffe et al., 1993; Caballero and Jaffe, 1993; Nadiri, 1993; Cockburn and Henderson, 1994; Hall, 1996; Jones and Williams, 1998; Hall et al., 2001; Bloom et al., 2010).[10] Since wages do not perfectly measure inventors' marginal product of labor, the growing literature that

[10]It is possible that patented inventions also generate negative externalities through patent races or patent blocking, but this large body of empirical evidence suggests that the positive spillovers outweigh the negative. Even with patent protection, research and development still creates sizable social returns that are at least twice as large as the private returns. Furthermore, patent citations are often used as a direct measure of knowledge spillovers (Caballero and Jaffe, 1993; Jaffe et al., 1993; Hall et al., 2005).

examines the effect of graduating economic conditions on private returns does not have clear implications for social welfare. For example, graduating in adverse labor market conditions has a negative long-term impact on the earnings of college graduates (Kahn, 2010; Oreopoulos et al., 2012), the career development of aspiring investment bankers (Oyer, 2008), and the productivity of economists(Oyer, 2006). In contrast, this paper is one of the first to focus on an outcome that generates potentially large social externalities.[11] My results suggest that a thorough welfare analysis of the impact of adverse labor market conditions should account for the potential social gains of the increased innovative output.[12]

My study also contributes to this literature by analyzing the relevance of sorting, which is important for both interpreting empirical results as well as conducting potential welfare analysis. For instance, Oyer (2008) finds that the MBA graduates who enter investment banking as a result of graduating in a booming stock market are more likely to stay in the industry. He argues that this is evidence for the existence of occupation-specific human capital in investment banking. However, without studying selection, he cannot exclude the possibility that graduates with higher innate ability could self-select into banking during booms.[13] The main difficulty in empirically identifying selection is the lack of a good measure of ex ante ability. In my data, I am able to use cumulative grade point average (GPA) at MIT as a uniform measure of innate ability to invent,

[11] Another study that looks at a socially important outcome is Schoar and Zuo (2011), who show that CEOs who start their careers during recessions have more conservative management styles.

[12] To perform such a welfare analysis, one would also need to observe several additional outcomes such as wages inclusive of non-pecuniary benefits and other measures of innovation like academic publications and new firm founding activities.

[13] Boehm and Watzinger (2011) is one of the first studies that examine sorting. They find that economics PhD candidates who graduate in a recession positively select into academia, in the sense that the average graduate staying in academia in a recession is better. But they use ex post publication records to measure ex ante ability. In contrast, Oyer (2006) finds that economics PhD graduates who enter the labor market in a recession tend to get jobs at lower-ranked schools, and consequently produce less research. This suggests that initial labor market conditions could have opposite effect on long-term outcomes through sorting. Genda et al. (2010) and Oreopoulos et al. (2012) find low ability workers to be more sensitive to initial career shocks than high ability workers. This is not evidence for positive sorting into high paying sectors since the distributions of low and high ability workers are unknown.

which is is fully determined by the time of graduation.[14] I show that my results are not driven by selection, and that it is not the case that more skilled graduates become inventors as a result of graduating in bad economic conditions. Furthermore, I find that initial economic conditions only affect the patent production of inventors with relatively low GPAs, suggesting that this group may be particularly sensitive to changes in the relative incentives of going into different sectors upon graduation.

The paper proceeds as follows. Section 2 derives the theoretical predictions using a simple two-period Roy-style model. Section 3 discusses the data and the patent matching procedure. I present the estimates of the main effect of initial labor market conditions on future patent production in Section 4, and decompose the effect in Section 5. Section 6 concludes.

2 Conceptual Framework

By affecting initial career choices, initial labor market conditions could affect long-term patent production in two ways: changing the level of human capital accumulated over time, and changing the future occupational affiliation. In this section, I formalize this idea using a two-sector two-period Roy-style model. Compared to a standard static Roy model, my model has two distinct features. The first is that I allow individuals to switch to a different sector after they enter the labor market. I define the path of human capital accumulation, and discuss the scenarios in which individuals may have the incentive to switch sectors even at the cost of losing their accumulated occupation-specific human capital. The second feature is that I specify the externality of patent production. As a result, the career paths individuals choose to maximize their own utility could differ from the social optimum. I use the model to show how a temporary shock to initial career choices could affect patent production in the future. I also discuss

[14] I normalize GPA by major and cohort. The normalized GPA strongly predicts future patent production.

the relevance of self-selection.

2.1 Assumptions

I consider the career choice problem of a single graduating cohort with P individuals. The economy contains two sectors of production, inventive ("I") and non-inventive ("N"). Empirical examples of patent-producing sectors include graduate school in science or engineering and industries such as bio-technology and electrical engineering. Examples of non-patent-producing sectors include finance and management consulting. I assume that individuals live for two periods. During each period, a person is able to choose her sector of employment after observing the state of the economy. Let $Q_i^t = I$ or N denote the sector chosen by individual i in period $t = 1, 2$. There are four possible career paths: $(Q_i^1, Q_i^2) \in \{(I, I), (I, N), (N, I), (N, N)\}$. An example of (N,I) is working in finance or consulting for two years before going back to graduate school in science or engineering.

An individual chooses a career path that maximizes her total utility:

$$U_i(Q_i^1, Q_i^2) = W_i^1(Q_i^1) + \beta W_i^2(Q_i^2|Q_i^1) + \delta \sum_j Pat_{j \neq i}$$

where $W_i^1(Q_i^1)$ is her wage in sector Q_i^1 in period 1, $W_i^2(Q_i^2|Q_i^1)$ is her wage in sector Q_i^2 in period 2 conditional on working in sector Q_i^1, and $\beta \leq 1$ is her discount rate. I introduce the externality of patented inventions by assuming that each person's utility function depends on other individuals' patent production. The weight δ thus captures, in reduced form, the magnitude of the externality from creating new patented technology. If $\delta > 0$, externality is positive, and if $\delta < 0$, it is negative.

Earnings and human capital

Each graduate i is endowed with sector specific human capital, denoted $h_{i,I}$ and $h_{i,N}$, which determine her initial wages. For simplicity, wages are linear in human capital and depend on the state of the economy. In period 1,

$$W_i^1(I) = h_{i,I};$$
$$W_i^1(N) = h_{i,N} + s$$

where s is the change in the wage in the non-inventive sector that depends on the state of the economy in period 1. I assume that s is constant across individuals and only affects wages in the non-inventive sector. The latter is without loss of generality, as only the change in the relative wage between the two sectors matters for equilibrium patent output. I discuss the equilibria in two cases: $s = 0$ and $s > 0$.

In addition to their initial endowments, workers develop occupation-specific human capital on the job, denoted k_I and k_N. Thus wages in period 2 are

$$W_i^2(I|I) = h_{i,I} + k_I;$$
$$W_i^2(N|I) = h_{i,N};$$
$$W_i^2(I|N) = h_{i,I};$$
$$W_i^2(N|N) = h_{i,N} + k_N. \tag{2.1}$$

Notice that I assume the shock s is temporary, only affecting wages in period 1. To focus on the effect of initial economic conditions, I assume that the economic conditions in period 2 do not affect wages in period 2.

Patent production

Since graduates rarely produce patents right after graduation, I assume that patent production only occurs in period 2. Let Pat_i be the patent production of individual i in period 2. Define

$$Pat_i(Q_i^2|Q_i^1) = \begin{cases} W_i^2(I|Q_i^1) & \text{if } Q_i^2 = I; \\ 0 & \text{if } Q_i^2 = N. \end{cases}$$

That is, all patents are produced by the inventive sector. Notice that an individual is more inventive if she works in the inventive sector throughout her career (i.e. choosing (I, I)) than if she only works in the inventive sector in period 2 (i.e. choosing (N, I)). In order to obtain a closed-form solution, I assume that a graduate's patent production equals her inventive human capital in period 2. The results would be qualitatively the same if patent production were some other weakly increasing function of inventive human capital.

2.2 Career path and patent production

Benchmark case: no shock

I start with the benchmark case where $s = 0$. An individual i chooses the career path (Q_i^1, Q_i^2) that maximizes $U_i(Q_i^1, Q_i^2)$. Since she cannot affect others' patent production, maximizing $U_i(Q_i^1, Q_i^2)$ is equivalent to maximizing $\widetilde{U}_i(Q_i^1, Q_i^2) = W_i^1(Q_i^1) + \beta W_i^2(Q_i^2|Q_i^1)$. We have

$$\widetilde{U}_i(I, I) = h_{i,I} + \beta\left(h_{i,I} + k_I\right); \tag{2.2}$$

$$\widetilde{U}_i(I, N) = h_{i,I} + \beta h_{i,N}; \tag{2.3}$$

$$\widetilde{U}_i(N, I) = h_{i,N} + \beta h_{i,I}; \tag{2.4}$$

$$\widetilde{U}_i(N, N) = h_{i,N} + \beta\left(h_{i,N} + k_N\right). \tag{2.5}$$

Proposition 1. *(Competitive Equilibrium) An individual chooses (I, I) if $z_i \geq \frac{-\beta k}{1+\beta}$ and (N,N) otherwise, where $z_i = h_{i,I} - h_{i,N}$ and $k = k_I - k_N$.*

Proof. By comparing Equation (2.2) to Equation (2.5), it follows that

(1) An individual chooses (I,I) if and only if $z_i \geq max\left\{-k_I, -\beta k_I, \frac{-\beta k}{1+\beta}\right\}$.

(2) An individual chooses (I,N) if and only if $z_i < -k_I$ and $z_i \geq max\{0, \beta k_N\}$.

(3) An individual chooses (N,I) if and only if $z_i \geq k_N$ and $z_i < min\{0, -\beta k_I\}$.

(4) An individual chooses (N,N) if and only if $z_i < min\left\{k_N, \frac{-\beta k}{1+\beta}, \beta k_N\right\}$.

Since k_I and k_N are non-negative, both (I,N) and (N,I) are implausible. Moreover, $\frac{-\beta k}{1+\beta} \leq \beta k_N \leq k_N$ and $\frac{-\beta k}{1+\beta} \geq -\beta k_I \geq -k_I$. Thus, an individual chooses (I, I) if $z_i \geq \frac{-\beta k}{1+\beta}$ and (N,N) otherwise. □

Proposition 1 shows that individuals work in the inventive sector in period 1 if and only if the premium of working in I over N is sufficiently high. When $s = 0$, they have no incentive to switch to a different sector in period 2. It follows that patent production is

$$Pat_i = \begin{cases} (h_{i,I} + k_I) & \text{if } z_i \geq \frac{-\beta k}{1+\beta}; \\ 0 & \text{if } z_i < \frac{-\beta k}{1+\beta}. \end{cases}$$

Social optimum in the benchmark case

Consider the problem of a social planner who chooses the career paths of all individuals to maximize social welfare, $\sum_i U_i(Q_i^1, Q_i^2)$. Because patent production directly enters agents' utility, this is equivalent to choosing (Q_i^1, Q_i^2) to maximize

$$SP_i(Q_i^1, Q_i^2) = W_I^1(Q_i^1) + \beta W_i^2(Q_i^2|Q_i^1) + \delta(P-1)Pat_i(Q_i^2|Q_i^1).$$

The following proposition shows that when patented inventions generate a positive externality, the equilibrium size of the inventive sector in period 2 is less than socially optimal.

Proposition 2. *When $\delta > 0$ (i.e. positive externality), the social optimum has higher* *total patent production than the competitive equilibrium.*

Proof. Those who choose (I,I) do not change their path in the social optimum. Since $\delta(P-1)Pat_i(I|I) = \delta(P-1)W_i^2(I|I) > 0$, it is easy to see that individuals with $-\delta(P-1)W_i^2(I|I) \leq \widetilde{U}_i(I,I) - \widetilde{U}_i(N,N) < 0$ choose (N, N) while the social planner would choose (I, I) or (N,I) for them. Thus, more individuals would work in the inventive sector in the second period and the total patent production is higher.[15] $\square$

Temporary shock to the wage in the non-inventive sector

Now, consider a temporary shock to the economy in period 1 that changes the wage in the non-inventive sector, $W_i^1(N)$, from $h_{i,N}$ to $h_{i,N} + s$. Suppose $s > 0$, so the non-inventive sector becomes temporarily more attractive in period 1.

Proposition 3. *(Competitive Equilibrium) There are two cases when $s > 0$.*
Case I: $s < k_N + \beta k_I$. An individual chooses (I,I) if $z_i \geq \frac{s-\beta k}{1+\beta}$ and (N,N) otherwise, where $z_i = h_{i,I} - h_{i,N}$ and $k = k_I - k_N$.
Case II: $s \geq k_N + \beta k_I$. An individual chooses (I,I) if $z_i \geq \frac{s-\beta k}{1+\beta}$, (N,I) if $k_N \leq z_i < \frac{s-\beta k}{1+\beta}$, and (N,N) otherwise.

Proof. Compared to the benchmark case, $\widetilde{U}_i(N,I) = h_{i,N} + s + \beta h_{i,I}$ and $\widetilde{U}_i(N,N) = h_{i,N} + s + \beta\left(h_{i,N} + k_N\right)$. The derivation then follows the same steps as the proof to Proposition 1. $\square$

Graduates who choose (N,N) before have no incentive to change their career path. As the return to working in the non-inventive sector in period 1 increases, some who choose (I,I) in the benchmark case may switch to (N,I) or (N,N). Notice that initial jobs are sticky because of positive human capital accumulation. If $k_I = k_N = 0$, then the

[15] I do not explicitly solve for the social optimum here since it depends on both z_i and $h_{i,I}$. It also requires more assumptions on $\delta(P-1)$, k_I, and k_N. For instance, given certain $\delta(P-1)$, k_I, and k_N, it is possible for social planner to choose (N,I). I skip deriving the specific conditions here since they are irrelevant to the empirical predictions.

temporary shock in period 1 has no effect on individuals' occupation in period 2. Because starting in the non-inventive sector helps a graduate to gain specific human capital in N, some of the individuals that switch to N in period 1 stay in N in period 2. However, if the incentive to working in N is sufficiently high, it would attract workers with high premium of working in I over N to temporarily work in N in period 1, but switch back to the inventive sector in period 2.

Consider the case where $s \geq k_N + \beta k_I$, the change in an individual's patent production relative to the benchmark case is

$$\triangle Pat_i = \begin{cases} -\left(h_{i,I} + k_I\right) & \text{if } \frac{-\beta k}{1+\beta} \leq z_i < k_N; \\ -k_I & \text{if } k_N \leq z < \frac{s-\beta k}{1+\beta}; \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, a shock that temporarily makes working in the non-inventive sector more profitable in period 1 affects patent production in period 2. In the presence of such a shock, patent production is lower as fewer graduates work in the inventive sector and some of those who still do have less inventive human capital. There are two channels for the effect. When $\frac{-\beta k}{1+\beta} \leq z_i < k_N$, an individual switches from (I,I) to (N,N), and her patent production decreases from $(h_{i,I} + k_I)$ to 0. I refer to this as the "occupational choice channel", where an individual's patent production changes as a result of working in a different sector. When $k_N \leq z < \frac{s-\beta k}{1+\beta}$, an individual switches from (I,I) to (N,I), and her patent production decreases from $(h_{i,I} + k_I)$ to $h_{i,I}$. I call this the "human capital channel" since her occupational affiliation in period 2 does not change.[16] The key difference between the occupational choice and the human capital channels is that in the former, the effect results from graduates switching sectors in the long-run, whereas in

[16]The human capital channel only occurs when s is sufficiently large and $k_I > 0$. If $k_I = 0$, then there is obviously no change in patent production.

the latter it does not.[17] Similarly, if $s < 0$, individuals who work in (N,N) may switch to (I,I) or (I,N). Total patent production would be higher than in the benchmark case. Without knowing the sign of s, the effect of adverse labor market conditions at the time of graduation on future patent production could be either positive or negative and remains an empirical question.

Sorting

As the model shows, initial labor market conditions are unlikely to affect every graduate's patent production. I thus classify the workers into three groups: "non-inventors", "marginal inventors", and "infra-marginal inventors". Regardless of graduating economic conditions, the non-inventors never patent and the infra-marginal inventors always patent. However, the marginal inventors produce patents only if they graduate in certain economic conditions. Importantly, the infra-marginal inventors' patent production could change at the intensive margin. Hence, the total change in patent production, due to a change in initial economic conditions, has two components: patents from the marginal inventors (changes at the extensive margin) and patents from the infra-marginal inventors (changes at the intensive margin).[18]

For any potential policy analysis, it is important to understand where the marginal

[17] There are several ways to extend the predictions of the model. First, one can easily add labor market frictions, such as search costs, but they would only complement the effect of the human capital accumulation on occupational choice. Substituting labor market frictions for human capital accumulation is equivalent to assuming that k_I and k_N only matter for earnings but not patent production. In this case, patent production only changes through the occupational choice channel. Second, one can include non-pecuniary returns of working in different occupations. See (Bayer et al., 2011; D'Haultfoeuille and Maurel, 2011) for examples of estimating the non-pecuniary returns in an extended Roy model. It is possible that individuals fully internalize the externality of their patent production by deriving enough non-pecuniary benefits from innovating. In this case, the equilibrium patent production is the same as in the social optimum. Finally, switching jobs could be more likely than what the stylized model predicts. For instance, if individuals get positive non-pecuniary returns from trying different jobs or if they have uncertainty about the returns from working in a particular sector.

[18] Since I assume that the non-inventive sector has no patent production, changes in patent production at the intensive margin could only happen through the human capital channel. In practice, there are different occupations with different levels of patent production, so both channels could potentially change patent production at the extensive and intensive margins.

patents come from. Are marginal inventors more or less skilled than infra-marginal inventors? Do marginal patents have higher or lower quality than average patents? Suppose a policy aims to increase patent production by temporarily rewarding individuals for entering the inventive sector upon college graduation. This is equivalent to introducing a negative s in the simple model. If the most skilled workers are already working in the inventive sector, the return to such a policy, in terms of the increase in the production of innovation, is decreasing in the size of the inventive sector.

To see this mathematically, I consider the special case where $\beta = 1$ and $k_I = k_N$. I also assume that

$$\begin{pmatrix} h_{i,I} \\ h_{i,N} \end{pmatrix} \sim N\begin{pmatrix} \mu & \sigma_I^2 & \sigma_{IN}^2 \\ & , & \\ \mu & \sigma_{IN}^2 & \sigma_N^2 \end{pmatrix}, \text{ and } z_i = h_{i,I} - h_{i,N} \sim N(0, \sigma_z^2).$$

Proposition 4. *$E(h_{i,I}|z_i = s)$ is decreasing in s if and only if $\rho_{Iz} = Corr(h_{i,I}, z_i) < 0$* (i.e., $\sigma_I^2 > \sigma_{IN}^2$).

Proof. Given the distributional assumptions, $E(h_{i,I}|z_i = s) = \mu_I + \rho_{Iz}\sigma_I s$. Thus $\frac{\partial E(h_{i,I}|z_i=s)}{\partial s} = \rho_{Iz}\sigma_I < 0$ if and only if $\rho_{Iz} < 0$. □

Without any shock, $E(h_{i,I}|Q_i^2 = I) = E(h_{i,I}|z_i > 0) = \mu_I + \rho_{Iz}\sigma_I\lambda(0)$, where $\lambda(0) = \phi(0)/\Phi(0)$ is the inverse mills ratio. Thus, $\rho_{Iz} > 0$ implies that there is positive sorting into the inventive sector, defined as $E(h_{i,I}|Q_i^2 = I) > E(h_{i,I})$. In other words, the average inventive skill (h_I) of the inventors is higher than the average for the entire cohort.[19] When s becomes marginally negative, individuals with $z_i = s \approx 0$ switch from (N,N) to (I,I). The gain in total patent production is $E(h_{i,I}|z_i = s) + k_I$. Proposition 4 shows that, as s becomes more negative, the gain in patent production is decreasing if and only if there is positive sorting into the inventive sector.

[19]Negative sorting implies that the most inventive individuals are inclined to work in the *non-inventive* sector. This is counter-intuitive, but theoretically possible, for instance if the return to the inventive skill is higher in the non-inventive sector than the inventive sector.

In practice, it is possible that a shock changes the composition of workers in each sector through sorting even without affecting the size of either sector. For instance, in response to a shock, people with high inventive ability may select out of the inventive sector and get replaced by less skilled inventors.[20] In the empirical section, I identify the nature of sorting by examining how the distribution of inventors' ex ante ability changes with graduating economic conditions. These results are discussed in detail in Section 5.

3 Data

3.1 Sample Construction

3.1.1 Data from MIT

MIT Office of the Registrar and the Alumni Association have generously provided individual-level data on every student that received a bachelor's degree from MIT between 1980 and 2005.[21] I observe basic demographic information such as gender and ethnicity, as well as information about their degree such as year of graduation, major(s) and cumulative grade point average. I group the graduates into three fields based on their major: Engineering; Science; and Others ("Non-SE").[22] I also observe the current employer of the alumni as self-reported on Infinite Connection in June 2011.[23] For those with available information on their employer, I assign their currently employed sector as Technology & Industrial, Academia, or Non-Science and Non-Engineering ("Non-SE"). Appendix A.1 explains how I assign the sectors based on the employers.

[20] This could happen in a standard Roy model if the return to skill in the inventive sector decreases. See Gould (2002); Mulligan and Rubinstein (2008) for detailed discussions on similar models.

[21] Although I have information on individuals that received a graduate degree from MIT during the same time period, I exclude them from the analysis and focus only on the Bachelor's population. The graduate population is much more heterogeneous than the undergraduate population.

[22] For the few incidences of double majors, I use whichever major declared first.

[23] Infinite Connection is an online alumni directory hosted by the MIT Alumni Association.

3.1.2 Patent Matching

To match graduates to their patents, I employ a separate database containing the full names of all alumni in the base sample as well as their addresses at the city level. For names, I observe both the Registrar's records and the ones currently used by the Alumni Association to contact the alumni. For locations, I observe the last two work addresses and home addresses reported on Infinite Connection. Each graduate has at least one home address in the data. For those that have never updated their alumni profiles on Infinite Connection, the home address information is from their Registrar's records at MIT.

I match this data to the U.S. patent inventor database from Lai et al. (2011b). The U.S. Patent and Trademark Office (USPTO) does not provide unique identifiers for inventors, making it difficult to track all the patents produced by the same inventor. Lai et al. (2011b) apply a Bayesian supervised learning approach and match inventors across all the U.S. utility patents granted between 1975 and 2010. Compared to the raw patent data from USPTO, the database from Lai et al. (2011b) allows me to match the alumni to inventors rather than patents.[24] Since the amount of individual information provided by an inventor could differ across patents, matching the alumni to inventors increases the likelihood that all patents from the same alumni inventor are included.[25]

I explain the matching procedure in Appendix A.2. Although matching errors are inevitable, they are unlikely to cause serious concerns for my empirical analysis. First, the summary statistics reported in Table 2 show that the science and engineering students

[24]Lai et al. (2011b) present two sets of results using different blocking rules. One minimizes the probability of lumping multiple individuals as one inventor, while the other minimizes the probability of splitting one individual into multiple inventors. See Lai et al. (2011a) for a detailed explanation of their procedure. Since each alumnus can be matched to multiple inventors, errors from splitting are less of a concern than errors from lumping. Thus I use the former set of results in my matching. My findings change little if I use the other.

[25]For instance, an inventor with a middle name may list the full middle name on some patents but only the initial on the others.

have much higher inventive output than the non-SE students. Since the matching errors should be randomly distributed across different majors, one can use the patent output of the non-SE students as an upper bound for the amount of false positives. Under the extreme assumption that the non-SE students should not produce any patents, there is still a significant amount of patent production from the science and engineering students. Second, the errors are not correlated with economic conditions at time of graduation and thus do not cause omitted variable bias (Figure A.1). Finally, as patent output is the dependent variable, measuring it with classical measurement errors could only increase the variance of the residual without generating any bias in the estimator.

I use the matching results to construct the patenting history of the alumni in my sample. For each graduate in each year after graduation, I calculate the number of granted patents which were applied for in that year and the subsequent number of citations received for those patents. The MIT Office of Provost Institutional Research then links the patent data to the base sample with individual characteristics.

3.2 Descriptive Statistics

Characteristics

Table 1 shows the mean characteristics for four groups: everyone, engineering majors, science majors, and inventors.[26] 68% of all the alumni in my sample are male; 58% are white; and 41% went to high school in the northeast of the United States. 63% of the graduates majored in engineering and nearly 25% majored in science; only 13% of the graduates majored in non-science and non-engineering ("non-SE") fields.[27] The engineering majors have less females, while the science majors have more females. The science majors also have more white and Asian American students.

Around 16% of all the graduates in my sample are inventors, that is, they have

[26] An "inventor" is someone with positive patent production since graduation.

[27] Around half of the graduates with non-SE majors majored in economics or management.

positive patent production since graduation. The inventors are more likely to be male, engineering majors and Caucasian. Less than 4% of the inventors majored in non-SE fields. The inventors have above average GPA, where the GPA is normalized by major and year of graduation. Specifically, the average normalized GPA of the inventors is around 0.22 standard deviations higher than the sample average.

About 78% of the sample report their employers on Infinite Connection, based on which I assign a sector.[28] Approximately 45% of the alumni with non-missing employer information currently work in industries that are generally related to engineering and science, such as technology and industrial. 16% work in academia, and 26% work in non-SE industries such as finance and consulting. Around 13% of the alumni work for firms where I cannot immediately assign a sector based on the firm name.[29] The inventors have a higher proportion that currently works in technology and industrial industries and a lower proportion that works in non-SE industries.

Patent Production

Figure 1 plots the average number of patents produced and the number of citations received in each year since graduation against two x-axes: year since graduation and year of patent application. All of the four series have an inverse "U" shape. Patent production increases over time in the first 15 years after graduation, which is consistent with the assumption in the conceptual framework that individuals accumulate inventive human capital from experience. There is also a upward trend in patent production before 2000, which is consistent with the aggregate trend in national patent statistics (Hall et al., 2001; USPTO, 2011). Since some of the patent applications in more recent years are still under review and I only observe the granted patents, the data are truncated from

[28]The employers are as of the last time they updated their alumni profile on Infinite Connection before June 2011. Although it is possible that the alumni have switched jobs and not updated on Infinite Connection, I assume that they stay in the same sector, and use the reported employers to determine the current sector.

[29]These would include, for example, small firms that do not indicate what they do in the company name.

the right. Thus, there are downward trends in patent production after 2000. While this is a relevant concern for the interpretation of descriptive statistics, I will control for the truncation in my regression analysis.

Table 2 shows summary statistics for annual patent production. The unit of observation is person by year. In an average year, an average cohort with a size of 1000 graduates produces around 52 patents. These patents together receive on average 681 citations by the end of 2010. Engineering graduates on average produce more patents than science and non-SE graduates, and non-SE graduates are the least likely to participate in inventive activities. An average patent from an MIT alumni inventor receives around 1 citation per year, which is twice as much as an average patent produced between 1981 and 2010. Since inventors only produce patents in some years, their annual patent production can be zero.

Table 3 shows the distribution of the patents' technology fields by inventors major. Following Hall et al. (2001), I classify patents into four technology fields based on their primary class: 1) Computer and Communications; 2) Electrical & Electronic and Mechanical; 3) Chemical and Drugs & Medical; and 4) Others.[30] Nearly half of all patents are from graduates majoring in electrical engineering and computer science (EECS), who patent mostly in computer and communications. Not surprisingly, alumni inventors tend to patent in their field of study. For example, graduates who majored in mechanical engineering and material science patent more in hardware engineering, while those that studied chemical engineering, chemistry or biology in college patent more in bio-tech related field. Few patents are in the "Others" field.

[30]Hall et al. (2001) has six technology categories. I combine Electrical & Electronic and Mechanical into one field. I also combine Chemical and Drugs & Medical into one field.

4 Initial Labor Market Conditions and Patent Output

4.1 An Illustrative Example: MIT Class of 1983 versus Class of 1984

Before presenting the regression estimates, I first discuss an illustrative example, which compares the patent output of the 1983 versus the 1984 graduate cohorts. The two classes have similar characteristics, though the class of 1983 has slightly fewer engineering and science majors (Table A.2). They overlapped for 3 years at MIT and experienced largely the same economic environment during college. By far, the most substantial difference between the two classes was the state of the economy at the time of their graduation. The class of 1983 graduated at the end of a recession. The annual unemployment rate was 9.6%, and the 2-year equity market return from the Center for Research in Security Prices (CRSP) before their senior year was 7.8%. By contrast, class of 1984 graduated during a recovery period when the annual unemployment rate was 7.5%, and the 2-year CRSP market return was 50%. Figure 2 plots each cohort's average patent output by year of graduation and year of application. In almost every year, the patent output of the class of 1983 surpasses the output of the 1984 cohort. The differences are especially large between 10 and 20 years after graduation. In total, the graduates in the class of 1983 have produced 2022 patents while the class of 1984 have produced only 1602 patents in their first 25 years after graduation. Table A.2 shows that both classes have similar proportion of inventors. The class of 1983 alumni are slightly more likely to work in the technology and industrial sector and less likely to work in the non-SE sector, although the differences are not statistically significant in the t-test.

4.2 Baseline Regression

Specification

For graduate i in year t, I observe the number of patents she produced that year, denoted by Pat_{it}, where this is computed as the number of patents she applied for in year t that were ultimately granted at some time prior to the end of my sample. I estimate the following equation:

$$(Pat)_{it} = G\left(\theta R_j + \delta(Controls)_{ijt} + \epsilon_{it}\right) \tag{4.1}$$

where j denotes the year of graduation, and R_j is either the national unemployment rate in year j or the CRSP stock market return from September $j-3$ to September $j-1$ or both. Since my outcome, the number of patents produce by each individual in each year, is non-negative and discrete, I estimate Equation (4.1) using quasi-maximum likelihood Poisson model and $G(.)$ denotes the likelihood function. .[31] All standard errors are corrected for heteroskedascity and clustered by cohort and application year.

Following the literature, I use the national unemployment rate in the year of college graduation as my main measure of initial labor market conditions.[32] Although MIT alumni are generally unlikely to be unemployed upon graduation, the aggregate economy still affects the availability and payoff of certain jobs. As a result, those graduating in a recession may pursue different career paths than those graduating in a booming economy. For instance, the MIT Class of 2009, who graduated during the financial crisis, still had high job placement rate comparable to the previous classes. But they have a higher proportion going to graduate school and a lower proportion entering the financial sector

[31] Compared to alternative count models such as negative binomial, the Poisson model has the advantage of being robust to model mis-specification (Cameron and Trivedi, 2001; Wooldrige, 2002). The quasi-ML Poisson model also accounts for any over dispersion in the data.

[32] Examples of other studies using the same measure include Kahn (2010); Genda et al. (2010); Oreopoulos et al. (2012).

(Hastings et al., 2010). Unfortunately, I do not observe the initial career choices of the 1980-2005 cohorts, and thus cannot estimate the effect of initial economic conditions on selecting into different initial placements.

I also control for a rich set of characteristics including gender, age, ethnicity/citizenship, high school region, dummies for fields of study (Engineering, Science, and Non-SE), and GPA standardized by major and cohort. The log of the federal research and development expenditure as a ratio of U.S. GDP in the year of college graduation controls for the demand for engineers and scientists (Goolsbee, 1998; Ryoo and Rosen, 2004; Majumdar and Shimotsu, 2006; NSF, 2010). To control for the potential nonlinear effects of patent application year, I include dummies for the application years. I also control for experience dummies, which are indicator variables that equal 1 for each year since graduation.

Results

Table 4 shows the coefficient estimates for the two measures of initial labor market conditions, using quasi-maximum likelihood Poisson models with different levels of controls. The effect of the unemployment rate is robust to alternative controls and the inclusion of the stock market return. Column (3) shows that a one percentage point increase in the national unemployment rate at the time of graduation increases the expected annual patent output of that cohort by almost 5.4%. Since an average cohort with a size of 1000 graduates produces 52 patents, a 5.4% change is equivalent to 2.8 patents per year. The effect of the stock market return is only significant when all controls are included. Column (6) shows that a one standard deviation decrease in the stock market return, equivalent to around 18%, increases the expected future annual patent output of the graduating cohort by 4.2%. However, it does not have any additional effect on patenting once the unemployment rate is included (since these are alternative measures of economic conditions). Table A.3 reports the coefficient estimates for individual characteristics from Column (7). As the descriptive statistics suggest, engineering and science majors are significantly more likely to produce patents than the non-SE majors. Engineering and science majors with higher GPA produce significantly more patents. Female graduates are less likely to produce patents.[33]

Figure 3 plots the persistence of the impact of graduating conditions on future patent production. I interact R_j with the experience dummies and plot the coefficients of the interaction terms against year since graduation for 25 years after graduation.[34] The effect of unemployment rate is insignificant in early years but becomes significant and persistent between 10 and 20 years after graduation. The effect of stock market return peaks around 13 years after graduation but is largely insignificant.

Robustness Checks

Table A.4 in the Appendix shows a set of robustness checks. Panel A shows the results in OLS and 2SLS using birth year dummies as the instruments for graduating economic conditions. Panel B restricts the sample to balance panels from the 1980-1995 cohorts on patent production in the first 15 years after graduation or between 2000 and 2010. Panel C excludes the top inventors in two ways: first by using an indicator variable that equals 1 if the annual number of patents produced is greater than zero as the dependent variable; and second by excluding the 100 graduates with the most patent production. The results that initial labor market conditions significantly affect future patent production are robust across all the alternative specifications. Table A.5 in the Appendix shows that initial labor market conditions do not change the students' choice of major at the time of college graduation.

[33]Ding et al. (2006) find that there exists gender difference in the tendency to patent among the life scientists. Female scientists are less likely to patent than male scientists.

[34]Only the early few cohorts are observed 25 years after graduation.

4.3 Initial Conditions versus Current Conditions

To see whether contemporaneous economic conditions affect patent output in addition to economic conditions at the time of graduation, I use the following specification:

$$(Pat)_{it} = G\left(\theta R_j + \beta R_t + \delta (Controls)_{it} + \epsilon_{it}\right) \tag{4.2}$$

where I include R_t , the labor market condition at time t, as well its 1 or 2-year lag in various specifications. Since the lagged current conditions are just the initial conditions for recent graduates, I exclude all the observations where t (the observation year) is less than $j+2$ when I include one lag and $j+3$ when I include two lags . Since I can no longer control for application year fixed effects, I include the application year and cohort year trends.Estimation of this model on the full sample is no longer possible since patents applied for more recently are less likely to show up in the data due to the lag between patent application and patent grant. To ensure that there is no spurious correlation caused by data truncation, I run the regression only on the sample with observation-years before 2000. Table 5 shows that the coefficients on initial conditions do not change from the previous table. Contemporaneous economic conditions have no significant effect on patent output.

5 Understanding the Effect of Initial Labor Market Conditions

The unemployment rate at the time of college graduation has a positive and significant impact on a graduate's patent production over the next 20 years. Two important questions remain. Which individuals' patent production is most affected by initial labor market conditions? Do marginal patents have higher or lower quality than average patents? As discussed in Section 2, the answers to these questions have important implications for potential welfare and policy analysis. In order to frame the discussion, I define the following terminology.

a) “Marginal inventors” are graduates whose decision to become inventors (i.e., produce at least one patent) is affected by graduating economic conditions.

b) “Infra-marginal inventors” are graduates who become inventors regardless of initial labor market conditions.

c) “Marginal patents” are patents whose production is contingent upon labor market conditions at graduation. Marginal patents could be produced by either marginal inventors or infra-marginal inventors.

The characteristics of the marginal inventors and the marginal patents are of primary interest to me as they provide vital information about the impact of initial labor market conditions. In this section, I provide empirical analysis in two steps. First, I identify which of the marginal patents come from marginal inventors and which come from infra-marginal inventors. Second, I study the characteristics of the marginal patents and how they differ from the average patents. Since the analysis is cross-sectional (i.e., at the inventor-level or patent-level), it is impossible to control for application year fixed effects as in the panel data. In order to accommodate the fact that each cohort has experienced a different number of post-graduation years, I use a balanced panel that includes the 1980-1995 cohorts, observed for the first 15 years after graduation.[35] Hence, all the cohorts have the same amount of time to produce patents. Column (B1) in Table A.4 also confirms that the results from the previous section hold for a balanced panel. A one percentage point increase in the unemployment rate at the time of graduation increases the average annual patent production in the first 15 years after graduation by around 4.2% for the 1980-1995 cohorts. The average number of patents produced per person per year is 0.054. Thus, a 4.2% change in the annual patent production for a cohort is equivalent to around 34 patents in 15 years.

[35]excluding the year of graduation.

5.1 Decomposition

Entry into Invention

I first estimate the effect of the initial unemployment rate on the probability of becoming an inventor in the first 5, 10, or 15 years after graduation. To test whether there are more graduates producing patents from recession cohorts, I estimate the following Linear Probability Model:[36]

$$Pr(D_i = 1) = \theta R_j + \delta(Controls)_{ij} + \epsilon_i. \tag{5.1}$$

where R_j is the national unemployment rate at the time of graduation, $D_i = 1$ for all graduates who patent in the first 5, 10, or 15 years after graduation, and observations are at the individual level. I control for a linear and quadratic cohort graduation year trend since different cohorts experience different aggregate patenting trends in their first 15 years. I also control for observed individual characteristics.

Table 6 reports the estimated effects. Although the coefficient estimates are small and positive for 5 and 10 years after graduation, they are not significant once I control for individual characteristics (Columns (2) and (4)). The coefficient estimates for 15 years are also very small and insignificant (Columns (5) and (6)). Thus, the initial unemployment rate has no significant effect on the probability of becoming an inventor.

Distribution of Inventors' Ability and Characteristics

As shown in the conceptual framework, if initial economic conditions change the nature of sorting into inventive careers, it is possible that the average inventors from a recession cohort have higher innate ability even when the number of inventors stays

[36] Probit and Logit models give almost identical estimates.

the same.[37] To examine whether the distribution of inventors' ability changes with initial labor market conditions, I estimate the following equation *only on the sample of inventors:*

$$(ability)_i \quad = \quad \theta R_j + \delta(Controls)_{ij} + \epsilon_i. \tag{5.2}$$

where *ability* is measured by GPA, and I control for the linear and quadratic cohort graduation year trend. Since the regression is estimated only on the sample of inventors, θ is the effect of the initial unemployment rate on the ability of the average inventors.[38] If the inventors from a recession cohort have higher ability than the inventors from a boom cohort, θ from Equation (5.2) should be positive. As an alternative to GPA, I also consider separately a dummy for both engineering and science majors on the left-hand side. Panel A from Table 7 suggests that the average inventors' GPA does not vary with graduating economic conditions, and that the average inventors' tendency to major in engineering or science is unaffected as well. To uncover the effect on the distribution of inventors' GPA, I also estimate Equation (5.2) with a quantile regression. Figure 4 plots the coefficient estimates with 95% confidence intervals for different quantiles. The estimates are generally insignificant and close to zero. Taken together, Panel A from Table 7 and Figure 4 suggest that the initial unemployment rate does not affect which graduates become inventors, at least in terms of their GPA and majors.

Since I find no evidence of changes at the extensive margin or in the nature of sorting, at least the majority of the change in patent production is at the intensive margin.

[37] In a standard Roy model, this would happen if the return to skill in the more inventive sector increases.

[38] This set-up is similar to the reduced-form specification that tests the difference between the marginal and average outcomes in Gruber et al. (1999) and Chandra and Staiger (2007). In theory, I can use the log of risk-adjusted proportion of inventors, instrumented by R_j, on the right-hand side to test whether the average ability of inventors changes with any change in the size of the inventor population induced by initial labor market conditions. However, doing so requires a first stage where R_j significantly affects the proportion of inventors, which does not exist in the data.

Thus, there are no marginal inventors. But there are two types of infra-marginal inventors: those that produce the marginal patents, and those unaffected by initial economic conditions. As discussed in Section 2, there are two possibilities:

1. Initial economic conditions do not affect a graduate's long-term occupational affiliation, but graduating in a worse economy increases an individual's accumulation of inventive human capital over time.[39] In this case, I expect initial economic conditions to have no effect on an inventor's sector.

2. Initial economic conditions change a graduate's long-term occupational affiliation. Graduating in a worse economy leads more graduates to work in patent-producing sectors.

To test whether the change happens through the occupational choice channel, I estimate the following equation using a Linear Probability Model *at the inventor level:*

$$Pr(Field_i = k) = \theta R_j + \delta(Controls)_{ij} + \epsilon_i \tag{5.3}$$

where $Field_i$ is the technology field in which inventor i patents the most, and $(Controls)_{ij}$ include linear and quadratic cohort graduation year trends as well as individual characteristics.[40] As an alternative measure of inventors' long-term occupation, I use the sector of employment reported on Infinite Connection as of June 2011. Panels B and C from Table 7 report the OLS estimates.[41] None of the estimates are statistically significant, suggesting that initial labor market conditions do not affect inventors' long-term occupational affiliation.[42] These results suggest that the change in patent production is

[39] For instance, by increasing graduate school enrollment.

[40] 73% of the inventors patent in only one field.

[41] Logit and Multinomial Logit regressions produce very similar results.

[42] Since the classification of sectors is fairly coarse, it is possible that initial labor market conditions change the inventors' sub-sector or firm. However, the differences in the mean level of patent production should be much larger across the general sectors than within a sector.

most likely caused by a change in inventors' post-graduation human capital accumulation rather than their long-term occupational choice.

As a robustness check, I also consider the effect of initial economic conditions on the time it takes an inventor to produce her first patent after graduation. If human capital accumulation is important, then one would expect inventors from recession cohorts to patent sooner. In order to evaluate whether or not this is true, I estimate the following equation by OLS:

$$T_i = \theta R_j + \delta(Controls)_{ij} + \epsilon_i. \tag{5.4}$$

where T_i is the number of years between the first patent and the time of graduation. In addition to the cohort trend and demographics, I also include dummies for the inventor's technology field to control for differences in the mean time to patent across fields. Panel D from Table 7 reports the OLS estimates.[43] Columns (D1) to (D3) show that a one percentage point increase in the unemployment rate at the time of graduation significantly decreases the time to the first patent by around 0.1 years. Given that the average of T_i is 7.95, the magnitude of the effect is very small.

Taken together, these results suggest that initial labor market conditions do not affect inventors' long-term occupation. Thus, the most likely hypothesis is that initial labor market conditions affect inventors' post-graduation human capital accumulation by affecting their *initial* career choices. For instance, a graduate may go directly into graduate school in science or engineering if she graduates in a recession, while in a boom she may initially work in a non-patent-producing sector such as finance or management consulting. Even though she could end up being an engineer in 10 years in both cases, in the former case she is likely to develop more skills that are relevant for inventing. It is important to note that human capital accumulation could occur if an inventor starts her career in a patent-producing sector such as high-technology, or goes directly

[43]Quasi-maximum likelihood Poisson regressions have almost identical estimates.

to graduate school. Unfortunately, without observing graduates' initial career choices, I cannot estimate the return to going to graduate school (or starting in an inventive sector) in terms of increased patent production.

5.2 Sorting

Inventors' Ability

The evidence suggests that initial economic conditions do not affect the probability of becoming an inventor in the first 15 years after graduation. Thus, to identify the nature of sorting, I compare the marginal *patents* to the average *patents*. I estimate the following equation by OLS at the patent level:

$$(ability)_p \quad = \quad \theta R_j + \delta(j, j^2) + \epsilon_p. \tag{5.5}$$

where *ability* is defined as in Equation (5.2). The key difference from before is that Equation (5.5) is *estimated at the patent level.* One can think of the dependent variable as the inventors' average ability weighted by how many patents they produce. Thus θ estimates the effect of initial economic conditions on the patent-weighted average GPA. A negative θ implies that the inventors who produce the marginal patents have lower GPAs than the average inventors; a positive θ suggests the opposite. Panel A from Table 8 shows that the national unemployment rate at the time of graduation has no effect on the average patent-weighted GPA of the inventors. Columns (A2) and (A3) suggest that the inventors who produce the marginal inventions are more likely to be science majors and less likely to be engineering majors.

Although there is no change in the average patent-weighted GPA, it is possible that there is a change in its distribution. Figure 5 presents two plots on the distribution of patent-weighted GPA. The left panel of Figure 5 plots the coefficient estimates with 95%

confidence intervals by quantile from a quantile regression of Equation (5.5). The positive coefficients at around the 20% quantile and the negative coefficients around the 60%-80% suggest that a disproportionate share of patents created by inventors graduating in a bad economy are from those with relatively low GPAs. This is also consistent with the right panel of Figure 5, which plots the kernel density of patent-weighted GPA separately for the cohorts above and below the median initial unemployment rate (7%). The patents produced by cohorts graduating in a bad economy are more likely to be from those inventors with GPAs around or below the median (0.48). Hence, initial labor market conditions affect the patent production of the inventors with relatively low GPAs.[44] But inventors with the highest GPAs are unaffected. There are two explanations for these findings. The first is that initial labor market conditions only affect the initial career choices of the inventors with lower GPAs. For example, it could be that the students with the best GPAs go to graduate school regardless of the economic conditions. The second possibility is that initial labor market conditions affect everyone's initial career choices, but initial career choices do not affect human capital accumulation for the most able inventors.

Technology Field

Before analyzing the quality of the marginal patents, it is important to know which technology fields they are from since the tendency to cite differs across fields. [45] I examine the change in the distribution of technology field at the patent level and estimate the following equation:

$$Pr(Field_p = k) \quad = \quad \theta_1 R_j + \theta_2 R_j * Science_i + \delta(Controls)_p + \epsilon_t \qquad (5.6)$$

[44]Notice that inventors on average have higher GPAs than non-inventors, so a relatively low GPA for an inventor is still around the mean of the population (around 0).

[45]For instance, computer and communications patents on average receive significantly more citations than mechanical patents.

which is similar to Equation (5.3) but estimated *at the patent level*. Since the marginal patents are likely from science majors, I also interact R_j with an indicator variable for being a science major, allowing the effect of graduating conditions to differ for the science majors. I also control for a cohort graduation year trend, an application year trend, and individual characteristics. Panel B from Table 8 reports the OLS estimates. An increase in the national unemployment rate at the time of graduation does not have a significant effect on the technology field of the patents from the average inventors. But it does have a significantly different effect on the technology field of the patents from inventors with science majors. In particular, the patents from science majors graduating in worse economic conditions are more likely to be in the chemical, drugs and medical field.[46] This is consistent with the finding that the marginal patents are likely from science majors, who are also more likely to patent in the chemical, drugs and medical field than engineering majors (Table 3).

Citations

I measure the quality of a patent using the number of patent citations it received by the end of 2010. I estimate the following equation at the patent level:

$$(Citations)_p = \alpha + \theta R_j + \delta (Controls)_p + \epsilon_p \tag{5.7}$$

where I control for inventor characteristics, linear and quadratic cohort graduation year trends, dummies for technology field, and dummies for year of patent application. Following the same logic as in Equation (5.5), θ in Equation (5.7) measures the change in the average quality of the patents as a result of a change in the national unemployment

[46] There is a particular concern that software patents have negative externalities due to the patent war in the industry. The results here suggest that the marginal patents are not software patents.

rate in the year of graduation.[47] A negative θ implies that the marginal patents have lower quality than the average patents.

The mean citations received in the overall sample is 17.72 and the median is 7. Since the distribution is skewed, I estimate Equation (5.7) using both OLS and median regressions. Panel A from Table 9 reports the OLS estimates, which are negative but insignificant. Panel B reports the estimates from median regressions, in which the effect is negative and marginally significant. These results suggest that initial economic conditions have no significant effect on the mean or median quality of patents.

Similar to Figure 5, Figure 6 shows two plots on the distribution of citations. The left panel shows the coefficients from the quantile regression using the specification in Column (B3). The coefficients are significantly negative between the 55% and 85% quantile, suggesting that the quality of the marginal patents is likely below the median. This is consistent with the right panel of Figure 6, which plots the kernel density of the risk-adjusted citations. The risk-adjusted citations are the residuals from regressing citations on the list of controls in Equation (5.7). The residuals adjust for the effect of other covariates, such as application year, on the distribution of citations. Based on the figure, it is clear that the quality of the marginal patents is below the median and the mean. Together with Figure 5, the results suggest that the marginal patents are of below median quality and are produced by inventors with median ability.

6 Conclusion

In this paper, I explore a novel channel through which short-term economic fluctuations affect the long-run innovative output of the U.S. economy: college graduates' initial career choices. Using a newly constructed data set on the patenting history of

[47]Note that it is possible though less likely for initial labor market conditions to influence the quality of the patents without changing the number of patents produced by an inventor. This does not affect the interpretation of the empirical results. One can just re-define the marginal patents to be the ones whose existence as well as quality are affected by inventors' graduating economic conditions.

MIT alumni, I find that cohorts graduating during economic downturns produce significantly more patents over the subsequent two decades. This effect stems from initial career choices; economic fluctuations have no measurable effect on the contemporaneous innovative output of graduates during their peak inventive years. Graduating in bad economic conditions leads inventors to select career paths that help them accumulate more inventive human capital. Consequently, they take less time to start producing patents, and produce more patents over the 20 years after graduation. I show that the inventors who produce more patents as a result of graduating in adverse labor market conditions are likely to be science graduates who work in non-software-engineering sectors such as bio-technology. My results also suggest that there exists positive sorting into inventing: graduates who are ex ante more inventive are also more likely to self-select into producing patents regardless of initial labor market conditions.

There are several promising directions for future research. Compared to the average engineering and science student population, MIT graduates are expected to have higher ability.[48] On one hand, MIT graduates are potentially less sensitive to labor market shocks if they have more skills (Oreopoulos et al., 2012). On the other hand, they may also be more productive at innovating, so any small change in their initial career choices could lead to relatively large changes in innovative output. Thus, it is not clear whether my results would generalize to the average college population. Studying the effect of initial labor market conditions on the patent production of other populations of engineering and science students would be a valuable extension. Second, since I do not directly observe initial career choice or graduate school enrollment in my data, I cannot estimate the causal impact of working in a certain sector or going to graduate school on long-term patent production. Future work identifying the return (in terms of innovative output) to different initial career choices would have important policy implications. Finally, my

[48]For instance, Grove (2011) shows that students accepted by MIT have higher SAT scores than those accepted by public universities.

results are not a welfare analysis of the impact of initial labor market conditions. A comprehensive welfare analysis that accounts for wages as well as the externalities of patented invention would be very informative.

References

Amesse, Fernand, Claude Desranleau, Hamid Etemad, Yves Fortier, and Louise Seguin-Dulude, "The Individual Inventor and the Role of Entrepreneurship : A Survey of the Canadian Evidence," *Research Policy*, February 1991, *20* (1), 13–27.

Arrow, K., "Economic Welfare and The Allocation of Resources for Invention," in "The Rate and Direction of Inventive Activity," Princeton University Press and NBER, 1962, pp. 609–25.

Baumol, W. J, "Entrepreneurship: Productive, Unproductive, and Destructive," *The Journal of Political Economy*, 1990, *98* (5-1).

Bayer, P., S. Khan, and C. Timmins, "Nonparametric Identification and Estimation in a Roy Model with Common Non-pecuniary Returns," *Journal of Business and Economic Statistics*, 2011, *29* (2).

Bedard, Kelly and Douglas A. Herman, "Who Goes to Graduate/Professional School? The Importance of Economic Fluctuations, Undergraduate Field, and Ability," *Economics of Education Review*, April 2008, *27* (2), 197–210.

Bernanke, Ben S. and Martin L. Parkinson, "Procyclical Labor Productivity and Competing Theories of the Business Cycle: Some Evidence from Interwar U.S. Manufacturing Industries," *Journal of Political Economy*, June 1991, *99* (3), 439–459.

Bloom, N., M. Schankerman, and J. Van Reenen, "Identifying Technology Spillovers and Product Market Rivalry," *Working Paper*, August 2010.

Bloom, Nicholas and John Van Reenen, "Patents, Real Options and Firm Performance," *The Economic Journal*, March 2002, *112* (478), C97–C116.

Boehm, Michael J and Martin Watzinger, "The Allocation of Talent: Evidence from the Market of Economists," March 2011.

Caballero, Ricardo J. and Adam B. Jaffe, "How High Are the Giants' Shoulders: An Empirical Assessment of Knowledge Spillovers and Creative Destruction in a Model of Economic Growth," *NBER Macroeconomics Annual*, January 1993, *8*, 15–74.

Cameron, A. Colin and Pravin K Trivedi, "Essentials of Count Data Regression," in "A Companion to Theoretical Econometrics A Companion to Theoretical Econometrics," Blackwell Publishing Ltd, 2001, pp. 331–348.

Chandra, A. and D. Staiger, "Productivity Spillovers in Health Care: Evidence from the Treatment of Heart Attacks," *Journal of Political Economy*, 2007, *115*, 103–140.

Cockburn, Iain and Rebecca Henderson, "Racing To Invest? The Dynamics of Competition in Ethical Drug Discovery," *Journal of Economics & Management Strategy*, September 1994, *3* (3), 481–519.

Cohen, Wesley M., Richard R. Nelson, and John P. Walsh, "Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not)," *National Bureau of Economic Research Working Paper Series*, 2000, *No. 7552.*

Comanor, William S and Frederic M Scherer, "Patent Statistics as a Measure of Technical Change," *Journal of Political Economy*, 1969, *77* (3), 392–98.

D'Haultfoeuille, X. and A. Maurel, "Inference on an Extended Roy Model, with an Application to Schooling Decisions in France," *Economic Research Initiatives at Duke (ERID) Working Paper No. 101*, 2011.

Ding, W.W., F. Murray, and T.E. Stuart, "Gender Differences in Patenting in the Academic Life Sciences," *Science*, 2006, *313* (5787), 665.

Field, Alexander J., "The Most Technologically Progressive Decade of the Century," *The American Economic Review*, 2003, *93* (4), 1399–1413.

Genda, Yuji, Ayako Kondo, and Souichi Ohta, "Long-Term Effects of a Recession at Labor Market Entry in Japan and the United States," *Journal of Human Resources*, December 2010, *45* (1), 157 –196.

Gibbons, Robert and Michael Waldman, "Task-Specific Human Capital," *American Economic Review*, May 2004, *94* (2), 203–207.

_ , Lawrence F. Katz, Thomas Lemieux, and Daniel Parent, "Comparative Advantage, Learning, and Sectoral Wage Determination," *Journal of Labor Economics*, October 2005, *23*, 681–724.

Goldin, Claudia, "Labor Markets in the Twentieth Century," in "The Cambridge History of the United States," stanley engerman and robert gallman ed., Vol. III, Cambridge: Cambridge University Press, 2000, p. 549 – 623.

Goolsbee, A., "Does Government R&D Policy Mainly Benefit Scientists and Engineers?," *The American Economic Review*, 1998.

Gould, E. D, "Rising Wage Inequality, Comparative Advantage, and the Growing Importance of General Skills in the United States," *Journal of Labor Economics*, 2002, p. 105–147.

Griliches, Z., "Patent Statistics as Economic Indicators: A Survey," *Journal of Economic Literature*, 1990, *28*, 1661–1707.

Grove, Allen, "SAT Scores for the Top Engineering Schools - Compare SAT Scores for the Top Engineering Schools in the United States," http://collegeapps.about.com/ 2011.

Gruber, J., P. Levine, and D. Staiger, "Abortion Legalization and Child Living Circumstances: Who is The "Marginal Child"?*," *Quarterly Journal of Economics*, 1999, *114* (1), 263–291.

Hagedoorn, John and Myriam Cloodt, "Measuring Innovative Performance: Is There an Advantage in Using Multiple Indicators?," *Research Policy*, September 2003, *32* (8), 1365–1379.

Hall, B. H, "The Private and Social Returns to Research and Development," *Technology, R&D, and the Economy*, 1996, *140*, 162.

_ , A. B Jaffe, and M. Trajtenberg, "The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools," *NBER Working Papers 8498, National Bureau of Economic Research, Inc*, 2001.

Hall, Bronwyn H., Adam Jaffe, and Manuel Trajtenberg, "Market Value and Patent Citations," *The RAND Journal of Economics*, April 2005, *36* (1), 16–38.

Harhoff, Dietmar, Francis Narin, F. M. Scherer, and Katrin Vopel, "Citation Frequency and the Value of Patented Inventions," *Review of Economics and Statistics*, 1999, *81* (3), 511–515.

Hastings, Daniel, Steven Lerman, and Melanie Parker, "The Demand for MIT Graduates," *MIT Faculty Newsletter*, February 2010, *XXII* (3).

Heckman, J. J and B. E Honore, "The Empirical Content of the Roy Model," *Econometrica*, 1990, *58* (5), 1121–1149.

Jaffe, Adam B., "Technological Opportunity and Spillovers of R & D: Evidence from Firms' Patents, Profits, and Market Value," *The American Economic Review*, December 1986, *76* (5), 984–1001.

_ , Manuel Trajtenberg, and Rebecca Henderson, "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," *The Quarterly Journal of Economics*, 1993, *108* (3), 577 –598.

Jones, B. F, "The Burden of Knowledge and the "Death of the Renaissance Man": Is Innovation Getting Harder?," *Review of Economic Studies*, 2009, *76* (1), 283–317.

Jones, Charles I. and John C. Williams, "Measuring the Social Return to R&D," *Quarterly Journal of Economics*, 1998, *113* (4), 1119–1135.

Kahn, L. B, "The Long-Term Labor Market Consequences of Graduating from College in a Bad Economy," *Labour Economics*, 2010, *17* (2), 303–316.

Kedrosky, Paul and Daniel Stangler, "Financialization and Its Entrepreneurial Consequences," *Kauffman Foundation Research Series: Firm Formation and Economic Growth*, March 2011.

Kerr, William R., "Ethnic Scientific Communities and International Technology Diffusion," *The Review of Economics and Statistics*, 2008, *90* (3), 518–537.

Kortum, S. and J. Lerner, "Assessing the Contribution of Venture Capital to Innovation," *RAND Journal of Economics*, 2000, *31* (4), 674–692.

Lai, R., A. D'Amour, A. Yu, Ye Sun, Vetle Torvik, and L. Fleming, "Disambiguation and Co-authorship Networks of the US Patent Inventor Database," July 2011.

Lai, Ronald, Alexander D'Amour, Amy Yu, Ye Sun, and Lee Fleming, "Disambiguation and Co-authorship Networks of the U.S. Patent Inventor Database (1975 - 2010)," 2011.

Lanjouw, Jean O and Mark Schankerman, "Patent Quality and Research Productivity: Measuring Innovation with Multiple Indicators," *The Economic Journal*, April 2004, *114* (495), 441–465.

_ , Ariel Pakes, and Jonathan Putnam, "How to Count Patents and Value Intellectual Property: The Uses of Patent Renewal and Application Data," *The Journal of Industrial Economics*, December 1998, *46* (4), 405–432.

Majumdar, S. and K. Shimotsu, "Enrollment Responses to Labour Market Conditions: A Study of the Canadian Market for Scientists and Engineers," *Queen's Economics Department Working Paper No. 1105*, 2006.

Mansfield, Edwin, "Patents and Innovation: An Empirical Study," *Management Science*, February 1986, *32* (2), 173–181.

_ , John Rapoport, Anthony Romeo, Samuel Wagner, and George Beardsley, "Social and Private Rates of Return from Industrial Innovations," *The Quarterly Journal of Economics*, May 1977, *91* (2), 221 –240.

Mulligan, C. B and Y. Rubinstein, "Selection, Investment, and Women's Relative Wages Over Time," *Quarterly Journal of Economics*, 2008, *123* (3), 1061–1110.

Murphy, K. M, A. Shleifer, and R. W Vishny, "The Allocation of Talent: Implications for Growth," *The Quarterly Journal of Economics*, 1991, *106* (2), 503.

Nadiri, M. Ishaq, "Innovations and Technological Spillovers," *National Bureau of Economic Research Working Paper Series*, August 1993, *No. 4423.*

Nanda, Ramana and Matthew Rhodes-Kropf, "Investment Cycles and Startup Innovation," *Harvard Business School Entrepreneurial Management Working Paper No. 12-032*, October 2011.

NCES, "Integrated Postsecondary Education Data System (IPEDS) Completions Survey, 1980-2005," https://webcaspar.nsf.gov/ 2011.

Neal, Derek, "Industry-Specific Human Capital: Evidence from Displaced Workers," *Journal of Labor Economics*, 1995, *13* (4), 653–77.

Nelson, R. R, "The Simple Economics of Basic Scientific Research," *Journal of Political Economy*, 1959, *67.*

Nicholas, Thomas, "Why Schumpeter was Right: Innovation, Market Power, and Creative Destruction in 1920s America," *The Journal of Economic History*, 2003, *63* (04), 1023–1058.

_ , "Does Innovation Cause Stock Market Runups? Evidence from the Great Crash," *American Economic Review*, 2008, *98* (4), 1370–96.

NSF, "Federal R&D Funding by Budget Function: Fiscal Years 2009-2011," http://www.nsf.gov/statistics/nsf10323/ October 2010.

Oreopoulos, Philip, Till von Wachter, and Andrew Heisz, "Short- and Long-Term Career Effects of Graduating in a Recession," *American Economic Journal: Applied Economics*, 2012, *Forthcoming.*

Oyer, P., "Initial Labor Market Conditions and Long-Term Outcomes for Economists," *The Journal of Economic Perspectives*, 2006, *20* (3), 143–160.

_ , "The Making of an Investment Banker: Stock Market Shocks, Career Choice, and Lifetime Income," *Journal of Finance*, 2008, *63* (6), 2601–2628.

Roberts, Edward B. and Charles E. Eesley, "Entrepreneurial Impact: The Role of MIT," *Report to the Kauffman Foundation*, February 2009.

Romer, Paul M., "Should the Government Subsidize Supply or Demand in the Market for Scientists and Engineers?," in "Innovation Policy and the Economy," MIT Press, January 2001, pp. 221–252.

Roy, A. D., "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers, New Series*, 1951, *3* (2), 135–146.

Ryoo, J. and S. Rosen, "The Engineering Labor Market," *Journal of Political Economy*, 2004, *112* (S1).

Schankerman, Mark and Ariel Pakes, "Estimates of the Value of Patent Rights in European Countries During the Post-1950 Period," *The Economic Journal*, December 1986, *96* (384), 1052–1076.

Schoar, Antoinette and Luo Zuo, "Shaped by Booms and Busts: How the Economy Impacts CEO Careers and Management Style," *National Bureau of Economic Research Working Paper Series*, 2011, *No. 17590.*

Trajtenberg, M., "The Welfare Analysis of Product Innovations, with an Application to Computed Tomography Scanners," *The Journal of Political Economy*, 1989, *97* (2), 444–479.

Trajtenberg, Manuel, "A Penny for Your Quotes: Patent Citations and the Value of Innovations," *The RAND Journal of Economics*, April 1990, *21* (1), 172–187.

USPTO, "Calendar Year Patent Statistics," http://www.uspto.gov/ 2011.

Willis, R. J and S. Rosen, "Education and Self-Selection," *Journal of Political Economy*, 1979, *87* (5 pt 2).

Wooldrige, JH, *Econometric Analysis of Cross Section and Panel Data*, 2 ed., Cambridge, MA: The MIT Press, 2002.

Tables and Figures

Table 1: Mean Characteristics for Person-level Data

	All	**Engineering**	**Science**	**Inventors**
Proportion Female	0.319	0.267	0.398	0.152
Age at Graduation	22.58	22.621	22.376	22.56
GPA (Normalized)	-0.006	-0.005	-0.003	0.214
Engineering	0.626	1	0	0.784
Science	0.245	0	1	0.177
Non-SE Majors	0.128	0	0	0.039
Ethnicity/Citizenship				
White	0.584	0.577	0.627	0.676
Asian American	0.197	0.189	0.207	0.167
International	0.079	0.085	0.064	0.077
Other Minorities	0.140	0.149	0.102	0.080
High School Region				
Northeast	0.406	0.393	0.427	0.421
Midwest	0.133	0.135	0.136	0.151
South	0.183	0.187	0.175	0.165
West	0.143	0.146	0.140	0.137
International	0.136	0.139	0.123	0.126
N	27,145	17,002	6,662	4,356
Currently Employed Sector				
Tech. and Industrial	0.450	0.542	0.294	0.631
Academia	0.157	0.104	0.328	0.125
Non-SE	0.260	0.217	0.270	0.112
Unassigned	0.134	0.137	0.107	0.132
N	21,178	13,576	4,896	3,836

Notes: This table reports the mean of individual characteristics by person. An "inventor" is anyone that has produced at least one patent since graduation. "Currently Employed Sector" is assigned from alumni's current employer as reported on Infinite Connection in June 2011 (see Appendix A.1); missing values are excluded.

Table 2: Patent and Citation Statistics for Person*Year-Level Data

	Mean	Std.Dev.	Min.	Max.
	Panel A: All Fields (N=475,636)			
Num. of Patents	0.052	0.448	0	48
Num. of Citations	0.681	11.956	0	2557
	Panel B: Engineering (N=303,506)			
Num. of Patents	0.064	0.501	0	48
Num. of Citations	0.859	12.848	0	2095
	Panel C: Science (N=114,593)			
Num. of Patents	0.040	0.387	0	27
Num. of Citations	0.477	12.138	0	2557
	Panel D: Non-SE (N=57,537)			
Num. of Patents	0.012	0.191	0	14
Num. of Citations	0.144	4.119	0	500
	Panel E: Inventors (N=89,435)			
Num. of Patents	0.276	1.004	0	48
Num. of Citations	3.620	27.378	0	2557

Notes: This table reports the summary statistics of the patents and citations at the person*year level. Note that inventors could produce zero patents in some years since they are defined as anyone with positive patent production since graduation.

Table 3: Patents' Technology Fields By Inventors' Major

	Computer & Communications	Electrical & Electronic Mechanical	Chemical Drugs & Medical	Others	Total
Engineering					
EECS	26.55%	12.61%	2.87%	1.26%	43.29%
Mechanical	3.95%	6.31%	4.84%	2.08%	17.17%
Chemical	0.93%	1.69%	3.27%	0.71%	6.60%
Material	0.67%	3.35%	1.18%	0.36%	5.57%
Aeronautics	1.76%	1.74%	1.03%	0.53%	5.07%
Other	0.41%	0.26%	0.21%	0.18%	1.05%
All Engineering	34.28%	25.95%	13.40%	5.12%	78.75%
Science					
Physics	2.10%	3.89%	1.35%	0.20%	7.54%
Chemistry	0.47%	0.76%	2.88%	0.17%	4.28%
Mathematics	2.12%	0.58%	0.23%	0.10%	3.03%
Biology	0.60%	0.38%	1.65%	0.11%	2.74%
Other	0.43%	0.33%	0.01%	0.03%	0.79%
All Science	5.72%	5.94%	6.12%	0.60%	18.39%

Notes: This table reports the fraction of patents in a specific technology field produced with inventors with a specific major. Technology fields are compiled from the six "Technology Categories" defined in Hall et al. (2001). Other engineering majors include Civil & Environmental; Ocean; and Nuclear. Other science majors include Brain and Cognitive Sciences; Earth, Atmospheric, and Planetary Sciences.

Table 4: Panel Estimates of the Impact of Graduating Conditions on Patent Production (Dep.Var. = Num. of Patents, Mean = 0.052)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Unemployment Rate	0.100***	0.056***	0.053***				0.044***
	(0.023)	(0.011)	(0.011)				(0.013)
Marginal Effect	*0.105*	*0.058*	*0.054*				*0.045*
2 Year Market Return				0.236	-0.244***	-0.263***	-0.105
				(0.201)	(0.074)	(0.070)	(0.082)
Marginal Effect				*0.266*	*-0.216*	*-0.231*	*-0.100*
Log (Fed R&D/GDP)	No	No	Yes	No	No	Yes	Yes
Characteristics	No	Yes	Yes	No	Yes	Yes	Yes
Experience Dummies	No	Yes	Yes	No	Yes	Yes	Yes
Current Year Dummies	No	Yes	Yes	No	Yes	Yes	Yes
N	475636	475636	475636	475636	475636	475636	475636

Notes: Person-year-level observation. All estimates are from quasi-maximum likelihood Poisson models. Sample includes all person-years from the year after graduation to 2010 for the 1980-2005 cohorts. Robust standard errors clustered at the cohort-year level are shown in parentheses. *: p< 0:10; **: p< 0:05; ***: p< 0:01. Dependent variable is the number of granted patents a graduate applies for in the current year. *"Unemployment rate":* the annual unemployment rate in the year of graduation. *"2 year market return":* the CSRP market return during the sophomore and junior years. *"Log (Fed R&D/GDP)"*: the log of federal r&d expenditure as a ratio of U.S. GDP in the year of graduation. *"Characteristics":* include age, cumulative GPA standardized by major and cohort, indicator variables for gender, race, engineering or science student, high school region. *"Experience dummies"*: 0/1 indicator variables for the difference between the current year and year of graduation. *"Current year dummies"*: 0/1 indicator variables for the current year.

Table 5: Panel Estimates of the Impact of Current Economic Conditions on Patent Production (Dep.Var. = Num. of Patents)

	(1)	(2)	(3)	(4)	(5)	(6)
	R = Unemployment Rate			R = Stock Return		
R_j (*Initial*)	0.060***	0.057***	0.058***	-0.356***	-0.363***	-0.372***
	(0.013)	(0.013)	(0.013)	(0.094)	(0.095)	(0.095)
R_t (*Current*)	-0.041	-0.074*	-0.011	-0.082	-0.113	-0.112
	(0.028)	(0.040)	(0.061)	(0.120)	(0.132)	(0.178)
R_{t-1}		0.039	-0.068		0.040	-0.009
		(0.038)	(0.078)		(0.164)	(0.187)
R_{t-2}			0.074			0.060
			(0.047)			(0.163)
Log (Fed R&D/GDP)	Yes	Yes	Yes	Yes	Yes	Yes
Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Experience dummies	Yes	Yes	Yes	Yes	Yes	Yes
Current Year Trend	Yes	Yes	Yes	Yes	Yes	Yes
Cohort Trend	Yes	Yes	Yes	Yes	Yes	Yes
N	198,804	178,990	160,228	198,804	178,990	160,228

Notes: Person-year-level observation. All estimates are from quasi-maximum likelihood Poisson models. Column (1) and (4) includes all person-years from two years after graduation to 2000 for the 1980-1998 cohorts. Column (2) and (4) includes all person-years from three years after graduation to 2000 for the 1980-1997 cohorts. Column (3) and (6) includes all person-years from four years after graduation to 2000 for the 1980-1996 cohorts. Robust standard errors clustered at the cohort-year level are shown in parentheses. *: p< 0:10; **: p< 0:05; ***: p< 0:01. Dependent variable is the number of granted patents a graduate applies for in the current year. R_j is the annual unemployment rate in the year of graduation or the CSRP market return during the sophomore and junior years. R_t is the annual unemployment rate in the current year or the CSRP market return in the two years before. *"Log (Fed R&D/GDP)"*: the log of federal r&d expenditure as a ratio of U.S. GDP in the year of graduation. *"Characteristics":* include age, cumulative GPA standardized by major and cohort, indicator variables for gender, race, engineering or science student, high school region. *"Experience dummies"*: 0/1 indicator variables for the difference between the current year and year of graduation. *"Current year trend" and "cohort trend"*: current year variable and cohort variable.

Table 6: Cross-sectional Estimates of the Impact of Graduating Conditions on Entry into Invention

	5 Years		10 Years		15 Years	
	(1)	(2)	(3)	(4)	(5)	(6)
Unemployment Rate	0.0017	0.0015	0.0024**	0.0022	0.0006	0.0004
	(0.0011)	(0.0010)	(0.0011)	(0.0014)	(0.0013)	(0.0015)
Characteristics	No	Yes	No	Yes	No	Yes
Cohort Trend	Yes	Yes	Yes	Yes	Yes	Yes
N	16610	16610	16610	16610	16610	16610

Notes: Person-level observation. All estimates are from ordinary-least-squares (OLS) models. Dependent variable is 0/1 indicator variable for becoming an inventor in 5 years (Column (1) and (2)), 10 years (Column (3) and (4)), or 15 years (Column (5) and (6)) after graduation. Sample includes all graduates from the 1980-1995 cohorts. Robust standard errors clustered at the cohort level are shown in parentheses. *: p< 0:10; **: p< 0:05; ***: p< 0:01. *"Unemployment rate"*: the national unemployment rate in the year of graduation. *"Characteristics":* include age, cumulative GPA standardized by major and cohort, indicator variables for gender, race, engineering or science student, high school region. *"Experience dummies"*: 0/1 indicator variables for the difference between the current year and year of graduation. *"Cohort trend"*: cohort variable and its square.

Table 7: Cross-sectional Estimates of the Impact of Graduating Conditions on Inventor Characteristics, Long-term Sector, and Time to First Patent

	Panel A: Characteristics (N=2,828)		
	(A1) GPA	(A2) Engineering	(A3) Science
Unemployment Rate	-0.0036	0.0064	-0.0032
	(0.0143)	(0.0100)	(0.0071)
Cohort Trend	Yes	Yes	Yes
	Panel B: Technology Field (N=2,828)		
	(B1) Computer & Communications	(B2) Electrical & Electronic; Mechanical	(B3) Chemical; Drugs & Medical
Unemployment Rate	0.003	0.001	-0.005
	(0.011)	(0.008)	(0.004)
Characteristics	Yes	Yes	Yes
Cohort Trend	Yes	Yes	Yes
	Panel C: Currently Employed Sector (N=2,538)		
	(C1) Tech. & Industrial	(C2) Academia	(C3) Non-SE
Unemployment Rate	0.004	-0.007	-0.001
	(0.008)	(0.005)	(0.004)
Characteristics	Yes	Yes	Yes
Cohort Trend	Yes	Yes	Yes
	Panel D: Time to First Paten (N=2,828)		
	(D1)	(D2)	(D3)
Unemployment Rate	-0.109*	-0.111**	-0.105**
	(0.052)	(0.051)	(0.048)
Cohort Trend	Yes	Yes	Yes
Technology Field Dummies	No	Yes	Yes
Characteristics	No	No	Yes

Notes: Person-level observation. All estimates are from ordinary-least-squares (OLS) models. Sample includes graduates from the 1980-1995 cohorts who have produced at least one patent in the first 15 years after graduation. Dependent variable in Panel A is: GPA (A1), 0/1 indicator variable for being an engineering major (A2) or science major (A3). Dependent variable in Panel B is 0/1 indicator variable for being in one of the three technology fields listed in the column names. Dependent variable in Panel C is 0/1 indicator variable for being in one of the three currently employed sectors listed in the column names. Dependent variable in Panel D is the number of years between the year of graduation and the year of application for the first granted patent. Robust standard errors clustered at the cohort level are shown in parentheses. *: p< 0:10; **: p< 0:05; ***: p< 0:01. *"Technology field":* the technology field in which the inventors patent the most. *"Currently employed sector"*: assigned from the current employer reported on Infinite Connection as of June 2011; missing values are excluded. *"Time to First Patent":* the number of years between year of graduation and year of patent application for the first granted patent. *"Unemployment rate"*: the national unemployment rate in the year of graduation. *"Characteristics":* include age, cumulative GPA standardized by major and cohort, indicator variables for gender, race, engineering or science student, high school region.*"Cohort trend"*: cohort variable and its square.

Table 8: Cross-sectional Estimates of the Impact of Graduating Conditions on Patent Characteristics

	Panel A: Inventor Characteristics		
	(A1) GPA	(A2) Engineering	(A3) Science
Unemployment Rate	0.0061	-0.0107	0.0120
	(0.0198)	(0.0103)	(0.0078)
Cohort Trend	Yes	Yes	Yes
	Panel B: Technology Field		
	(B1) Computer & Communications	(B2) Electrical & Electronic; Mechanical	(B3) Chemical; Drugs & Medical
Unemployment Rate	0.006	0.005	-0.012
	(0.018)	(0.012)	(0.012)
Unemployment*Science	-0.031	0.002	0.034***
	(0.023)	(0.027)	(0.011)
Inventor Characteristics	Yes	Yes	Yes
Cohort Trend	Yes	Yes	Yes
Application Year Trend	Yes	Yes	Yes

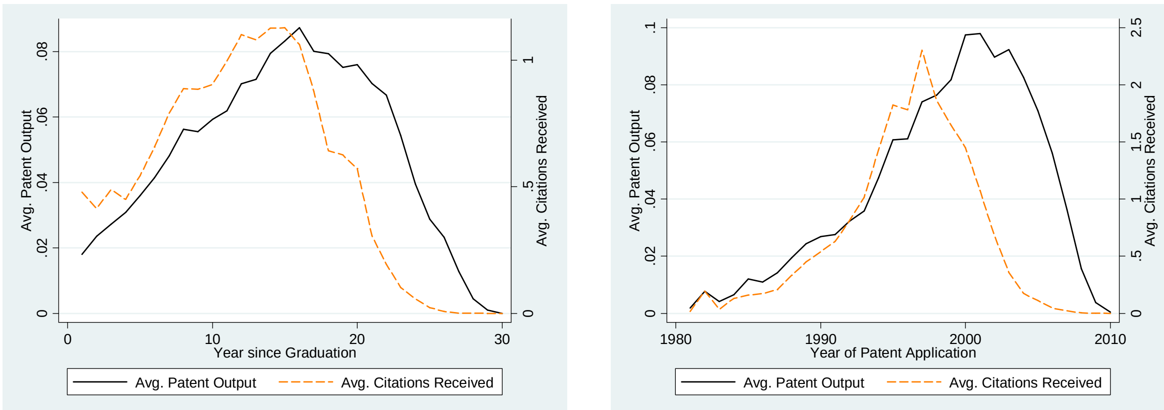
Notes: Patent-level observation (N=13,336). All estimates are from ordinary-least-squares (OLS) models. Sample includes all patents produced by the 1980-1995 cohorts in the first 15 years after graduation. Dependent variable in Panel A is: GPA (A1), 0/1 indicator variable for being an engineering major (A2) or science major (A3). Dependent variable in Panel B is 0/1 indicator variable for being in one of the three technology fields listed in the column names. Robust standard errors clustered at the cohort level are shown in parentheses. *: p< 0:10; **: p< 0:05; ***: p< 0:01. *"Unemployment rate"*: the national unemployment rate in the year of graduation. *"Unemployment*Science":* the interaction term of "unemployment rate" and the 0/1 indicator variable for being a science major. *"Inventor Characteristics":* include age, cumulative GPA standardized by major and cohort, indicator variables for gender, race, engineering or science student, high school region. *"Cohort trend"*: cohort variable and its square. *"Application year trend":* application year variable and its square.

Table 9: Cross-sectional Estimates of the Impact of Graduating Conditions on Patent Citations

	Panel A: OLS		
	(C1)	(C2)	(C3)
Unemployment Rate	-0.202	-0.672	-0.655
	(0.694)	(0.442)	(0.415)
Application Year Dummies	No	Yes	Yes
Technology Field Dummies	No	Yes	Yes
Inventor Characteristics	No	No	Yes
Cohort Trend	Yes	Yes	Yes
	Panel B: Median Regression		
	(D1)	(D2)	(D3)
Unemployment Rate	-0.273	0.001	-0.139
	(0.182)	(0.084)	(0.086)
Application Year Dummies	No	Yes	Yes
Technology Field Dummies	No	Yes	Yes
Inventor Characteristics	No	No	Yes
Cohort Trend	Yes	Yes	Yes

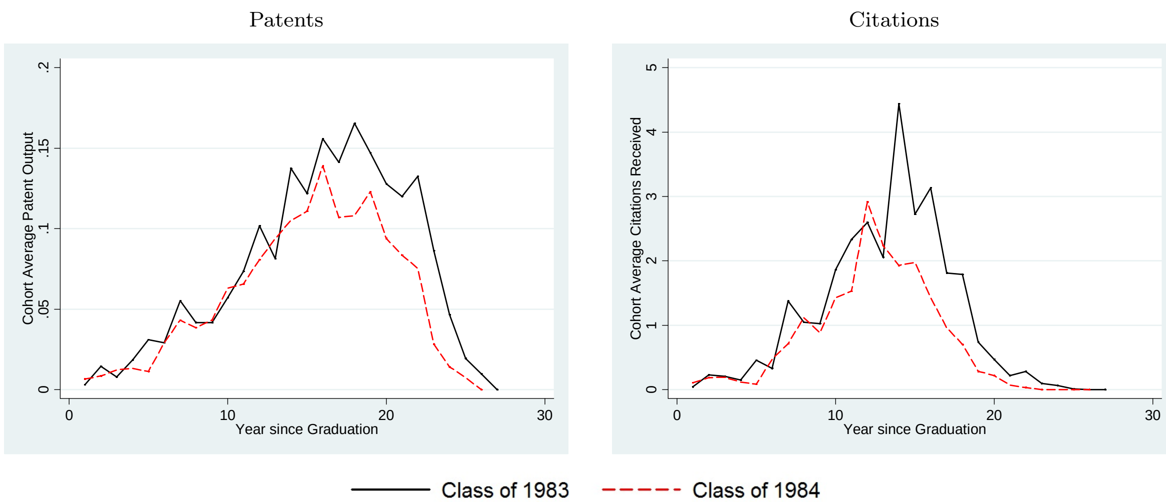
Notes: Patent-level observation (N=13,336). Estimates in Panel A are from ordinary-least-squares (OLS) models. Estimates in Panel B are from quantile regressions estimated at the median. Sample includes all patents produced by the 1980-1995 cohorts in the first 15 years after graduation. Dependent variable is the number of citations received by the end of 2010. Robust standard errors clustered at the cohort level are shown in parentheses. *: p< 0:10; **: p< 0:05; ***: p< 0:01. *"Unemployment rate"*: the national unemployment rate in the year of graduation. *"Inventor Characteristics":* include age, cumulative GPA standardized by major and cohort, indicator variables for gender, race, engineering or science student, high school region. *"Cohort trend"*: cohort variable and its square. *"Application year dummies":* the set of 0/1 indicator variables for each application year. *"Technology field dummies":* the set of 0/1 indicator variables for each technology field.

Figure 1: By Year: Average Patents and Citations Per Person



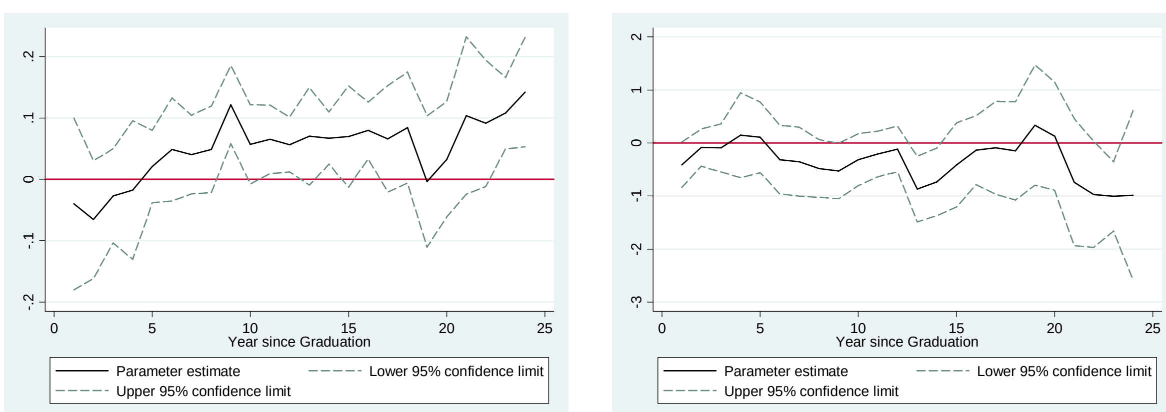
Notes: This figure plots the average patent output in the each year, by year since graduation (on the left) and year of patent application (on the right).

Figure 2: Class of 1983 VS 1984: Patent Output by Year



Notes: Outcome plotted is, by year since graduation, the average patent output in the each year(on the left) and the average citations received for patents produced in that year (on the right). The black line is Class of 1983, and the red dashed line is Class of 1984.

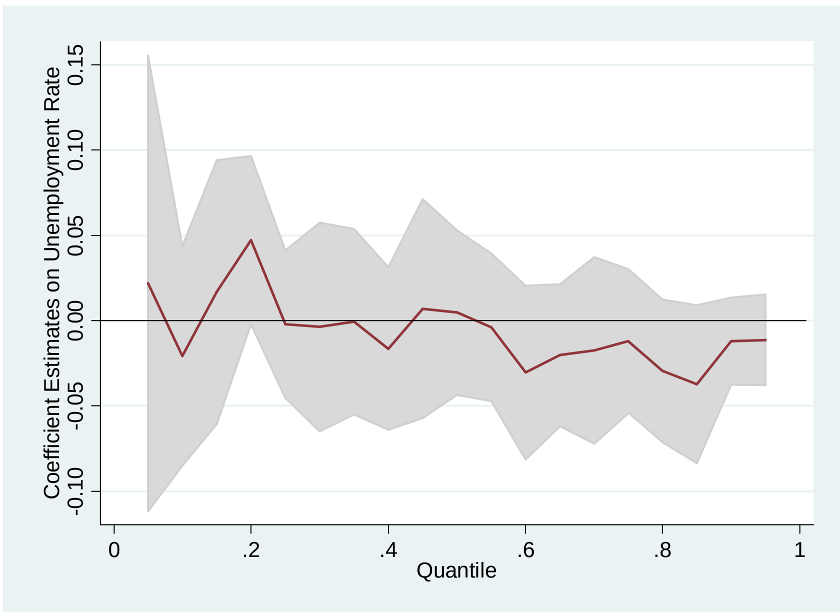
Figure 3: Persistent Effects of Graduating Conditions on Patent Output



Unemployment Rate in the Year of Graduation — Stock Return during Sophomore and Junior Years

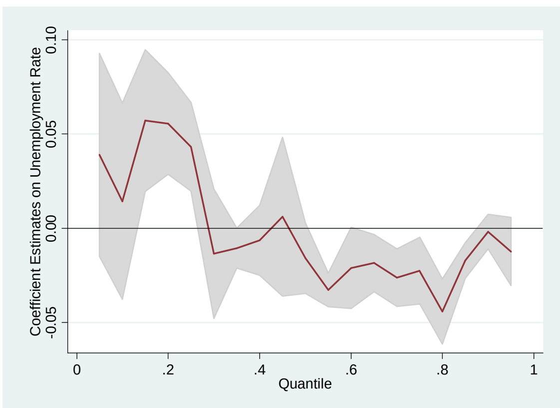
Notes: Person*year-level observation. These figures plot the coefficient estimates and confidence interval of the interaction term between the shock and the dummy for each year since graduation from a quasi-ML Poisson model. Dep. Variable = Number of patents produced in a year. On the left: shock measured by unemployment rate in the year of graduation. On the right: shock measured by the stock return during the sophomore and junior years.

FIGURE 4: BALANCED PANEL: INITIAL CONDITIONS AND INVENTORS' GPA†

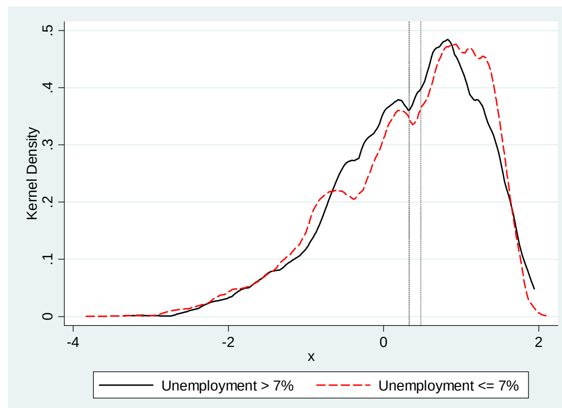


Notes: Person-level observation. This figure plots the coefficient estimates and 95% CI from the Quantile regression. Dep. Var. = GPA. Independent variable plotted: national unemployment rate in the year of graduation. Standard errors are bootstrapped with 2000 repetitions. Sample includes all the individuals from the 1980-1995 cohorts that have produced at least one patent within 15 years after graduation.

Figure 5: Balanced Panel: Initial Conditions and Patent-Weighted Inventors' GPA



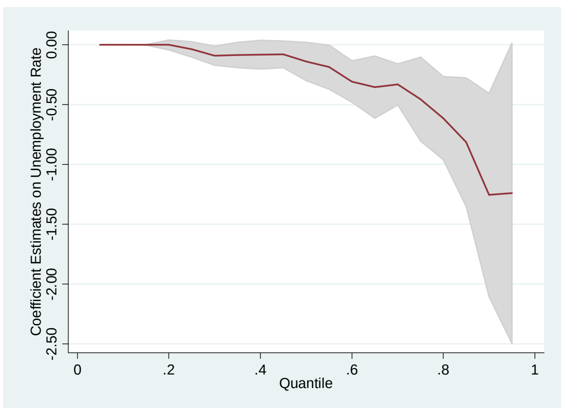
(A) Quantile coefficient estimates



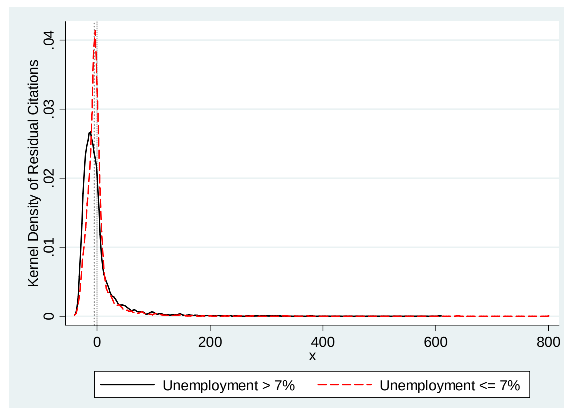
(B) Kernel density

Notes: Patent-level observation. Sample includes all the patents produced by the 1980-1995 cohorts within 15 years after graduation. On the left: the coefficient estimates from the Quantile regression. Dependent variable is the (patent-weighted) GPA of the inventor. Independent variable plotted: national unemployment rate in the year of graduation. Standard errors are bootstrapped with 2000 repetitions. On the right: the kernel density of inventor's GPA. Black line: the sample of patents produced by the inventors who graduated with the national unemployment higher than 7%. Red dashed line: the sample of patents produced by the inventors who graduated with the national unemployment rate lower than 7%. The two gray vertical lines are the mean and median of the whole sample.

Figure 6: Balanced Panel: Initial Conditions and the Distribution of Citations†



(A) Quantile coefficient estimates



(B) Kernel density of Residual Citations

Notes: Patent-level observation. Sample includes all the patents produced by the 1980-1995 cohorts within 15 years after graduation. On the left: the coefficient estimates from the Quantile regression. Dependent variable is the number of citations. Independent variable plotted: national unemployment rate in the year of graduation. Standard errors are bootstrapped with 2000 repetitions. On the right: the kernel density of residual citations. Residual citations are the residuals from regressing citations on cohort trend, inventor characteristics, application year dummies and technology field dummies. Black line: the sample of patents produced by the inventors who graduated with the national unemployment higher than 7%. Red dashed line: the sample of patents produced by the inventors who graduated with the national unemployment rate lower than 7%. The two gray vertical lines are the mean and median of the whole sample.

A Appendix

A.1 Assigning Sector based on Employer

Based on the employer reported on Infinite Connection, I assign the graduates to three sectors: technology and industrial (anything that generally involves patent production); academia; and non-science and non-engineering ("non-SE", including finance, consulting, law, real estate, and government). I determine the sector in two ways. The first is by firm. For instance, Google is in the first group whereas Goldman Sachs is in the third group. However, this is only plausible for large firms. Since the graduates work for a very wide range of firms (more than 10,000 unique names), it would be too time consuming to go through all the firms and determine their sectors. Thus, the second way to assign sector is based on keywords. For instance, any firm with "semiconductors" or "pharma" in its names is assigned to the first group; any employer with "university" or "college" is assigned to the second group; any firm with "holding" or "consult" is assigned to the third group. Although doing so inevitably allows more measurement errors than assigning sector by firm, it is more efficient and covers most of the sample. Only 13% of the reported employers are unassigned. They are generally small firms such as start-ups.

Samples of keywords used to identify each sector are:

1. Technology & Industrial: tech machine syst dynamics scien research communication devic wire manufact telecom syst soft defense instrument engineer space material equipment aircraft energ motor electr industri robot network chemical conduct comput auto mobile product info elevat data design media petro oil engrg solution innovat power metal analysis utilit diagnosti metric engine digita activ internet intranet atomic aviation cemex cement oceanograph analyt telegraph nuclear pharma therapeut molecu biomed cure cancer;

2. Academia: universi college "medical school" "business school";

3. Non-SE: consult capital trading asset invest securitie bank venture finance financial wealth holding fund insurance broker architect hotel society foundation entertain picture embassy community school academy ministry teach healthcare airline propert program practice clinic attorney realestate marketing adverti realty.

A.2 Patent Matching

My matching procedure has two steps. In the first step, I match the alumni to the inventors that have the same first and last names, and drop those with different non-missing middle names or initials. It is possible that this first step could drop a small set of patents produced by the alumni, if a) the names are misspelled on the patent grants, or b) the alumni use new names on the patents but have not reported the name change to the Alumni Association.

In the second step, I assign each alumnus-inventor pair an integer score out of 10 based on a) how well the middle names match, b) how well the locations match, and

c) how rare the first and/or last names are. Table A.1 provies a summary of the score assignment. In the middle score category, the full score, 3, is when the full middle names, including when there are no middle names, are matched between the alumnus and the inventor. A score of 2 is when the initials match or when the middle name is missing for the inventors but not the alumni. Since it is not required for inventors to report their full names, not having a middle name listed on the patent does not imply there is no middle name. A score of 1 is when the middle name is missing for the alumnus but not the inventor. It is less likely but still possible that some alumni do not report their middle names to the Registrar's Office and the Alumni Association. It is also possible that the alumnus has added a middle name since graduation and listed it on the patents. In the location category, the full score, 4, is when one of the work or home city-level addresses reported by the alumnus perfectly matches the city of the inventor. 3 is when the states match, and 1 is when the countries match. In the name rarity category, there are two sub-criteria: how rare the names are among the MIT population (full score 2), and how rare the names are among the inventors population (full score 1). On the first criterion, the full score 2 is when the first or last name is very rare (less than 10 people with the same first or last name); 1 is when the first or last name is fairly rare (less than 100 people with the same first or last name); the rest are 0. On the second criterion, 1 is when there are less than 15 unique inventors with the same first and last name, and 0 otherwise.

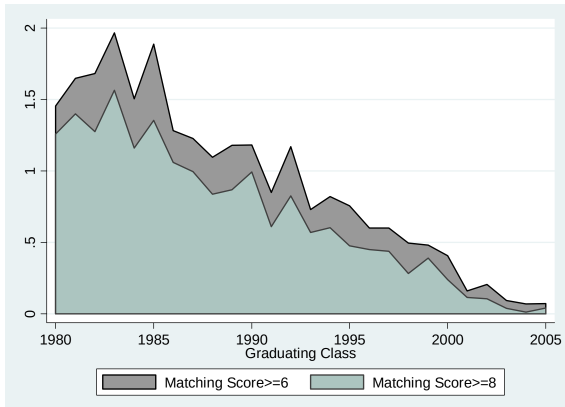
A higher score implies a greater likelihood that the matching is correct. For example, an alumnus-inventor pair scores 10 when the graduate and the inventor have a rare first or last name and are exactly matched on middle name and city of residence. A score of 3 means the two have common names, live in different states, and the middle name is missing in the alumni records. Since an inventor may provide different information across patents, the matching score could also differ across patents. In this case I use the highest score for each alumnus-inventor pair. In the very few cases where two alumni are matched to the same inventor, I look both up on Google or LinkedIn and determine the correct match based on their years of graduation and where they have worked.

The matching score is not a perfect criteria due to obvious data limitations. If an alumni inventor has moved many times, then the location match would not be perfect. On the other hand, even within a city, there are people with the same names. A score lower than 6 means the alumnus-inventor pair fails to fully satisfy at least two out of the three criteria listed above. In this case, it is hard to distinguish whether the low score is from not observing the correct address or the match is a false positive. A score above 8 means that the alumnus-inventor pair is a perfect match by multiple criteria, but not all the correctly matched pairs would score this high. Thus, scores that are sufficiently low should be dropped, but there is a trade-off between Type I and Type II errors. I use 6 as the main threshold to gain more statistical power in my analysis, but I also use 8 as a robustness check. After dropping all the pairs that score below 6, the final sample includes over 4,500 alumni inventors with more than 25,000 total patents granted by the end of 2010. These patents have received over 300,000 citations in total by the end of 2010. Restricting to those scoring above 8 still leaves over 3,400 alumni and 19,300 patents with nearly 250,000 citations. I exclude all the patents that were applied for before and during the year of graduation. Figure (A.1) shows that the variations in the cohort-level patent production do not depend on the score[49].

[49]Since the later cohorts have less time to invent, there is a natural downward slope in both patent and citation output.

A.3 Additional Tables and Figures

Figure A.1: Average Patent Output by Cohort



Patent

Notes: This figure plots, by cohort, the number of patents produced per person since the year after graduation.

Table A.1: Assigning Matching Score

	Strong	**Medium**	**Weak**
Middle Name	[3] Exact match	[2] Initial; Inventor missing	[1] Alumnus missing
Location	[4] City	[3] State	[1] Country
Name Rarity	[3] Very rare	[2] Rare	[1] or [0] Not rare

Table A.2: 1983 VS 1984 Characteristics†

Class	**1983**	**1984**
Female	0.2091	0.215
Age at Graduation	22.677	22.576
Inventor	0.251	0.251
Ethnicity/Citizenship		
White	0.797	0.773
Asian Am.	0.044**	0.068
Other Minorities (US)	0.092	0.073
International	0.068	0.086
Highschool Region		
Northeast	0.504*	0.465
Midwest	0.137	0.127
South	0.151	0.146
West	0.101	0.112
International	0.108***	0.151
Field of Study		
Engineering	0.683	0.701
Science	0.208	0.223
Non-SE	0.109**	0.077
N (Person)	1033	1065
Current Sector		
Tech. and Industrial	0.520	0.500
Academia	0.131*	0.161
Non-SE	0.186	0.202
Unassigned	0.163	0.138
N (Person)	861	908

Notes: Statistical significance reported for the T-test of equal means. * $p<0.10$; ** $p<0.05$; *** $p<0.01$.

Table A.3: Baseline Coefficient Estimates: Characteristics

Female	-1.129***	Asian American	0.199***
	(0.044)		(0.041)
Age at Graduation	-0.022***	International	-0.083*
	(0.008)		(0.047)
Engineering	1.431***	Other Minorities	-0.380***
	(0.072)		(0.058)
Science	1.042***	HS Midwest	0.063
	(0.075)		(0.042)
GPA*Engineering	0.378***	HS South	-0.100***
	(0.068)		(0.035)
GPA*Science	0.339***	HS West	-0.055
	(0.069)		(0.035)
GPA (Non-SE)	-0.043	HS International	0.068*
			(0.041)

Notes: Coefficients reported from Column (7) of Table 4. Base groups are: non-SE; white; northeast highschool. * p<0.10; ** p<0.05; *** p<0.01.

Table A.4: Robustness Checks: The Impact of Graduating Conditions on Patent Production

	Panel A: OLS & 2SLS			
	OLS		2SLS	
	(A1)	(A2)	(A3)	(A4)
Unemployment Rate	0.004***		0.003***	
	(0.001)		(0.001)	
2 Year Market Return		-0.015***		-0.021***
		(0.004)		(0.005)
N	475,636	475,636	475,636	475,636
	Panel B: Balanced Panel			
	Sample I		Sample II	
	(B1)	(B2)	(B3)	(B4)
Unemployment Rate	0.042***		0.069***	
	(0.014)		(0.018)	
2 Year Market Return		-0.436***		-0.410***
		(0.085)		(0.124)
N	249,150	249,150	182,710	182,710
	Panel C: Excluding Top Inventors			
	Dependent: (Pat>0)		Top Inventors Excluded	
	(C1)	(C2)	(C3)	(C4)
Unemployment Rate	0.039***		0.060***	
	(0.008)		(0.009)	
2 Year Market Return		-0.129***		-0.309***
		(0.053)		(0.082)
N	474,588	474,588	474,763	474,763

Notes: Person-year-level observations. Dependent variable is number of patents produced in a year except for (C1) and (C2). Robust standard errors are corrected clustered at the cohort-year level are shown in parentheses. * p<0.10; ** p<0.05; *** p<0.01. Panel A: Estimates in (A1) and (A2) are from ordinary-least-squares (OLS) models. Estimates in (A3) and (A4) are from two-stage-least-squares (2SLS) models. Sample includes the 1980-2005 cohorts observed for years between the year after graduation and 2010. Panel B: Estimates are from quasi-maximum likelihood (QML) Poisson models. Sample I includes 1980-1995 cohorts observed for the first 15 years after graduation; Sample II includes 1980-1995 cohorts observed between 2000 and 2010. Panel C: Estimates in (C1) and (C2) are from Logistic regressions. Estimates in (C3) and (C4) are from QML Poisson models. Dependent variable in (C1) and (C2) is 0/1 indicator variable for positive patent production. Sample in (C3) and (C4) excludes the most productive inventors with more than 50 lifetime patents. *"Unemployment rate":* the annual unemployment rate in the year of graduation. *"2 year market return":* the CSRP market return during the sophomore and junior years. All the regressions include the following controls: the log of federal r&d expenditure as a ratio of U.S. GDP in the year of graduation; age, cumulative GPA standardized by major and cohort, indicator variables for gender, race, engineering or science student, high school region; 0/1 indicator variables for the difference between the current year and year of graduation; and 0/1 indicator variables for the current year.

Table A.5: Robustness Checks: The Impact of Graduating Conditions on Selection into Majors

Dependent	*Engineering=1*		*Science=1*	
	(1)	(2)	(3)	(4)
Unemployment	0.004	0.003	-0.004	-0.002
	(0.004)	(0.004)	(0.003)	(0.003)
Characteristics	No	Yes	No	Yes
Cohort Trend	Yes	Yes	Yes	Yes
N	27,145	27,145	27,145	27,145

Notes: Person-level observations. Coefficients reported are marginal effects from Logistic models. Dependent variable is 0/1 indicator variable for being an engineering major ((1) and (2)) or science major ((3) and (4)). Robust standard errors are corrected clustered at the cohort level are shown in parentheses. * p<0.10; ** p<0.05; *** p<0.01. *"Unemployment rate":* the annual unemployment rate in the year of graduation. *"Characteristics":* include age, cumulative GPA standardized by major and cohort, indicator variables for gender, race, engineering or science student, high school region. *"Cohort trend"*: cohort variable.