Age and Great Invention*

Benjamin F. Jones†

April 2005

Abstract

Great achievements in knowledge are produced by older innovators today than they were a century ago. Using data on Nobel Prize winners and great inventors, I find that the age at which noted innovations are produced has increased by approximately 6 years over the 20th Century. This trend is consistent with a shift in the life-cycle productivity of great minds. It is also consistent with an aging workforce. The paper employs a semi-parametric maximum likelihood model to (1) test between these competing explanations and (2) locate any specific shifts in life-cycle productivity. The productivity explanation receives considerable support. I find that innovators are much less productive at younger ages, beginning to produce major ideas 8 years later at the end of the 20th Century than they did at the beginning. Furthermore, the later start to the career is not compensated for by increasing productivity beyond early middle age. I show that these distinct shifts for knowledge-based careers are consistent with a knowledge-based theory, where the accumulation of knowledge across generations leads innovators to seek more education over time. More generally, the results show that individual innovators are productive over a narrowing span of their life cycle, a trend that reduces – other things equal – the aggregate output of innovators. This drop in productivity is particularly acute if innovators' raw ability is greatest when young.

JEL Codes: O31, J24, I2, B3, C14
Keywords: innovation, age of achievement, Nobel Prize, life-cycle productivity, educational attainment, semi-parametric maximum likelihood estimation

---

*I would like to thank Daron Acemoglu, Abhijit Banerjee, Leemore Dafny, James Dana, Ben Olken, Mike Mazzeo, Joel Mokyr, Sendhil Mullainathan, Scott Stern, and Peter Temin for helpful comments and suggestions. Alexander Karlan and Gowoon Lee provided excellent research assistance. All errors are my own.

†Kellogg School of Management and NBER. Contact: 2001 Sheridan Road, Room 628, Evanston, IL 60208. Email: bjones@kellogg.northwestern.edu.

Age is, of course, a fever chill
that every physicist must fear.
He's better dead than living still
when once he's past his thirtieth year.

– Paul Dirac, 1933 Nobel Laureate in Physics

1 Introduction

It is widely perceived that great innovations are the provenance of the young. The sentiments of Dirac expressed above have been shared by Einstein, von Neumann and many other eminent scientists and mathematicians (Zuckerman & Merton, 1973; Simonton, 1988). Empirical investigations of this view, usually undertaken within the fields of psychology and sociology, tend to support the idea that innovative activity is greater at younger ages, although great achievement before the age of 30 is not typical. Rather, a researcher's output tends to rise steeply in the 20's and 30's, peak in the late 30's or early 40's, and then trail off slowly through later years (Lehman, 1953; Simonton, 1991).

While many great insights do occur at younger ages, it is also clear that innovators spend a large number of their early years undertaking education.[1] Indeed, human capital investments dominate the early part of the innovator's life-cycle. Learning a subset of the skills, theories, and facts developed by prior generations seems a necessary ingredient to innovative activity. Newton acknowledged as much in his famous letter to Hooke, "if I have seen farther it is by standing on the shoulders of giants". Dirac and Einstein, who produced major contributions at the age of 26, first went through significant educational periods and then built directly on existing work. Dirac built on Heisenberg's uncertainty principle and Hamiltonian mechanics. Einstein's early insights built

[1] Research in the psychology literature suggests that substantial training periods – ten years at minimum – are a prerequisite to expertise in many fields, from science to sports, music, medicine, and chess (see Ericsson and Lehmann 1996 for a review).

on the work of Planck and Maxwell. Certainly, innovation would be a very difficult enterprise if every generation had to reinvent the wheel.

These two observations suggest an intriguing tradeoff. If innovators are especially productive when young, but education is an important preliminary input to innovation, then the opportunity cost - the time spent - in education may be significant. Moreover, how innovators approach this tradeoff may evolve as the economy evolves. As one example, the accumulation of knowledge across generations may impose an increasing educational burden on successive cohorts, so that the expanding time costs of education will delay the onset of active innovative careers. This possibility poses a problem for innovation as it reduces, ceteris paribus, the lifetime output of individual innovators, especially if their potential is greatest when young.

In this paper, I show that the great achievements in knowledge of the 20th Century occurred at later and later ages. The youth of Einstein and Dirac is far more unusual today than it was at the time of their great insights. The mean age at great achievement for both Nobel Prize winners and great technological inventors rose by about 6 years over the course of the 20th Century. Moreover, this trend in knowledge-based achievement is distinctive: the age of great achievement in athletics has shown no trend.

Having presented these age trends in Section 2, I then develop hypotheses to explain them. In one set of hypotheses, the productivity of innovators over the life-cycle may have shifted on specific dimensions. For example, increasing educational attainment at younger ages may delay the onset of active innovative careers. Alternatively, innovator productivity may have increased at more advanced ages due to improved health or an increased role for experience. Determining the location of any shift in innovator productivity will substantially constrain the set of theories that can explain the general upward trend in age.

In another hypothesis, the upward age trend in the data could simply reflect underlying demographic shifts. Great innovations at a point in time are selected from a set of innovations produced by a range of innovators. Since the population has become substantially older with time, we are more likely to draw older innovators today than we were at the beginning of the 20th century. Put another way, if people lived shorter lives in the past, then innovators in the past will also appear younger.

Section 3 develops a formal econometric model of the age at great achievement to (1) test between these competing explanations and (2) locate any specific shifts in life-cycle productivity. The model is estimated using semi-parametric maximum likelihood methods, with the results presented in Section 4. I find substantial shifts in life-cycle productivity beyond any demographic effect. Specifically, there has been a large upward trend in the age at which innovators begin their active careers; meanwhile, there is no compensating shift in the productivity of innovators beyond middle age. The estimates suggest that, on average, the great minds of the 20th Century began innovating at age 23 at the start of the 20th Century, but only at age 31 at the end - an upward trend of 8 years.

Section 5 considers both implications of these patterns for aggregate innovative output and possible reasons for the upward trend in the early life cycle. I show that the delayed start to the career is captured partly by an increasing age at Ph.D.. Using a simple model for the educational decisions of innovators, I then show precise circumstances in which innovators choose to expand educational attainment. Interestingly, while general increases in the value of education may increase an innovator's output, they can discourage increased educational attainment. An expansion of foundational knowledge, on the other hand, can produce the empirical features we see. Section 6 concludes.

2 Age and Great Achievement

This section presents benchmark facts about the age of individuals at the time of their great achievements. Two types of achievements are considered. The first group considers knowledge-based achievements: research that leads to Nobel Prizes in Physics, Chemistry, Medicine, and Economics, as well as technological achievements presented in almanacs of the history of technology. The second group considers physical achievements: world-record breaking events in track and field, as well as Most-Valuable Player awards in baseball. The athletes are included for comparison, allowing us to highlight what is distinctive about knowledge-based achievement.

The data set uses established sources to identify great achievements in knowledge. Nobel Prizes are determined by committees of experts and are given in principle for a distinct advance. The technological almanacs compile key advances in technology, by year, in several different categories such as electronics, energy, food & agriculture, materials, and tools & devices. The year (and therefore age) of great achievement is the year in which the key research was performed. For the technological almanacs, this is simply the year in which the achievement is listed. For Nobel Prizes, which are retrospective, the year of achievement was determined by consulting various biographical resources. The Data Appendix describes the data collection and sources in further detail.

As a first look at the data, Figure 1 presents ages at great innovation, considering all 20th Century observations together. Three features are of immediate note. First, there is a large variance in age. The largest mass of great innovations in knowledge came in the 30's (42%), but a substantial amount also came in the 40's (30%), and some 14% came beyond the age of 50. Second, there are no observations of great achievers before the age of 19. Dirac and Einstein prove quite unusual, as only 7% of the sample produced a great achievement at or before the age of 26. Third, the age distribution for the Nobel Prize winners and the great inventors, which come from independent sources, are extremely similar over the entire distributions. Only 7% of individuals in the data appear in both the Nobel Prize and great inventors data sets.

The most surprising aspect of these data, however, becomes apparent when we consider shifts in this age distribution over time. To start, I run the following regression:

$$a_i = \alpha + \beta t_i + \gamma X_f + \varepsilon_i \tag{1}$$

where a_i is the age of individual i at the time of the great achievement, t_i is the year of the great achievement, and X_f are fixed effects for the field of the achievement and the country of the individual's birth. Results of this regression are presented in Table 1. We see that the mean age at great achievement is trending upwards by 5 or 6 years per century. These trends are highly significant and are robust to field and country of birth controls. Indeed, the controls cause the time trend to strengthen, rising to about 8 years over the course of the 20th century. The strengthening effect of the controls on the trend suggests a compositional shift in great innovation towards fields and countries that favor the young.

These trends can be seen in greater distributional detail in Figure 2, which presents the raw data again but divides the 20th Century into three chronological periods: from 1900-1935, 1935-1960, and 1960 to the present. This figure combines all unique individuals in the Nobel Prize and great inventors data sets. Here we observe a general shift of the age distribution away from younger ages. There is a distinct drop in the presence of those in their 20's and an increased presence of those in later middle age.

One obvious hypothesis for this outward age shift is a decline in the productivity of younger innovators in favor of older innovators. Given that the early part of an innovator's career is dominated by education, one natural reason for a decline in early innovative potential may be an increase in the time spent in training. More generally, there may be relative increases in the productivity of *older* innovators directly. As one example, if we believe that raw ability declines over the life-cycle while experience increases, then the shift in the distribution may suggest the rising importance of experience over ability. Alternatively, improved health technology may result in increased ability and effort at later ages.

But we must be careful in how we interpret the distributional shifts we see. A second hypothesis for the outward age shift is a simple demographic effect. If the underlying population of innovators is getting older, then older innovators will be more likely to produce substantial innovations, even if the relationship between age and innovative potential is fixed. The greater the ratio of 50-year-old innovators to 25-year-old innovators, the more likely the Nobel Prize winning invention or greatest technological insight to come from one of the 50-year-olds. Such demographic effects may be important: certainly, life-expectancy and the average age of the population have risen substantially over the 20th century.

One reduced-form way to test between these ideas is a difference-in-difference style analysis. If we view the scientific/technological innovators as a "treatment" group experiencing effects peculiar to knowledge-based careers, then we might profitably attempt a comparison with "control" groups that are claimed on a priori grounds to be immune to such knowledge effects. An obvious choice for a control group is great achievements in athletics. Age-achievement profiles in athletics will also be influenced to some extent by demographic effects, but athletes presumably do not face increasing training demands over time – the rules of their games are both straightforward and fixed. In fact, as shown in Table 2, the age of great achievements in athletics is not showing an upward trend. Figure 3 compares the underlying distribution of MVP winners in baseball before and after 1960, dividing the data in half. We see that the entire distribution appears essentially stationary.

The substantial upward trend in the age of knowledge-based achievement is entirely absent in physical achievement. Moreover, if general demographic effects explain the age trends among great innovators, then they should also be felt to some degree among the athletes. But comparing knowledge-achievements and athletic achievements along these lines cannot be wholly satisfactory. The age distribution of athletes clearly favors the young to a degree that knowledge-based achievement does not.[2] Shifts in the population density beyond middle age will therefore not influence the age distribution of athletic achievement, while they still might influence the age distribution of great achievements in knowledge. More generally, this comparison does not help us pinpoint any distinct shifts in life-cycle productivity for knowledge-based careers.

The following section develops a formal econometric model to identify specific shifts in innovation potential, controlling for demographic effects. With this econometric model, we can ask two questions explicitly. First, is the upward trend in the age of great achievement simply a demographic effect, or is it driven by shifts in innovator's life-cycle productivity? Second, if life-cycle productivity is shifting over time, is this due to effects at the beginning of the life-cycle, the end of the life-cycle, or both? The methods to answer these questions are developed in Section 3, and the results are presented in Section 4.

3 Econometric Model

This section presents a simple stochastic model to define the probability that witnessed innovations are produced by innovators at particular ages. The empirical analysis in Section 4 will then use this model to determine, mechanically, where the upward trend in the age of great achievement is coming from.

Formally, consider a population N. Given that we have witnessed an innovation, the probability that the innovation was produced by an individual i is defined by:

$$\Pr(i) = \frac{x_i}{\sum_{\{i \in N\}} x_i}$$

[2] No one has ever won the MVP in baseball or broken a track & field world record after the age of 40. But 41% of Nobel Prize winning research and the 20th Century's great technological inventions came beyond that age.

where x_i represents the innovation potential of person i. Innovation potential measures the relative innovative strength of an individual.[3]

It will be useful to consider the model in terms of cohorts of equally-aged individuals. First, define the set of cohorts as A, where $a \subset A$ represents the cohort with age a. Furthermore, let the set of individuals in this cohort be $N_a \subset N$, and let the number of individuals in such a set be defined as $|N_a|$. Then the probability that a witnessed innovation is produced by an individual in the cohort with age a is,

$$\Pr(a) = \frac{\sum_{\{i \in N_a\}} x_i}{\sum_{\{i \in N\}} x_i} = \frac{|N_a|\, \bar{x}_a}{\sum_{\{a \in A\}} |N_a|\, \bar{x}_a}$$

where $\bar{x}_a$ is the average innovation potential of individuals in the cohort with age a. Dividing top and bottom by the size of the entire population, $|N|$, and defining the age distribution of the population as $p_a = |N_a| \,/\, |N|$, we can rewrite this expression into a particularly useful form,

$$\Pr(a) = \frac{p_a \bar{x}_a}{\sum_{\{a \in A\}} p_a \bar{x}_a}$$

Our notation to this point has ignored the possibility that both the age distribution of the population, p_a, and the average innovation potential of individuals of age a, $\bar{x}_a$, may be changing over time. Investigating this possibility is the ultimate purpose of our estimation exercise and we acknowledge it explicitly by writing,

$$\Pr(a|t) = \frac{p_a(t)\bar{x}_a(t)}{\sum_{\{a \in A\}} p_a(t)\bar{x}_a(t)} \tag{2}$$

[3] Innovation potential can be interpreted as the instantaneous probability that person i produces an innovation, but it need not be given such a stochastic basis. One may think of innovators as being drawn, with replacement, from a box of names. A particular person's innovation potential then represents the frequency with which his or her name appears in the box, where we imagine that innovators with higher ability or effort level appear more often.

Any variation over time in the expected age at innovation is then determined explicitly:

$$E[a|t] = \sum_{\{a \in A\}} a \Pr(a|t)$$

Several useful points can now be made. First, it is now clear that trends in the age of great innovation must be driven either by shifts in the population age distribution, $p_a(t)$, or by shifts in the average innovation potential of various age groups, $\bar{x}_a(t)$. Second, any presumption that the innovators' upward age trends are driven mechanically by increasing life expectancy may be misleading. An increasing density of very old individuals is unlikely to be an important force, simply because the innovation potential, $\bar{x}_a$, of those in their later years is likely to be low – if only because people retire. Furthermore, demographic events like the post-war baby boom in the U.S. can create periods where the working population is getting younger even though life expectancy is rising. Third, the stochastic process represented in equation (2) can produce innovators with a large variance in age, as demonstrated in Figure 1. Finally, it is worth noting that this stochastic model makes few assumptions. While we will make further assumptions in how we define $p_a(t)$ and $\bar{x}_a(t)$, the model to this point is quite general.

Equation (2) is the central vehicle for the maximum likelihood estimation presented in Section 4. In particular, given data for the population distribution, $p_a(t)$, and a series of year-age observations for great achievements, we can use (2) to test hypotheses about the shape of innovation potential, $\bar{x}_a(t)$. Before continuing to the estimation, it remains to develop an explicit model of $\bar{x}_a(t)$ and how it may shift over time. This sub-model is presented in the next section.

3.1 Locating Shifts in Life-Cycle Productivity

In this section we will add parametric structure to the definition of x_i and, by extension, $\bar{x}_a$. In particular, we will embed in equation (2) a sub-model of innovation potential that allows for specific and testable shifts in the life-cycle productivity of innovators.

In choosing an appropriate modeling strategy, it is helpful to first consider the existing empirical literature on creative careers, which suggests the following general pattern (e.g., Lehman 1953, Bloom 1985, Simonton 1991). First, the life-cycle begins with a period of full-time training in which there is no substantive creative output. Second, there is a rapid rise in output to a peak in the late 30's or early 40's. Third, innovative output declines slowly through later years. While laboratory experiments do suggest that creative thinking becomes more difficult with age (e.g. Reese et al, 2001), the decline in innovative output at later ages may largely be due to declining effort, which a range of sociological, psychological, institutional, and economic theories have been variously proposed to explain (see Simonton 1996 for a review).

To illustrate these life-cycle patterns, Figure 4 presents data for the creative output of 31 long-lived individuals, all of whom were born in the mid to late 19th century. Eleven of the individuals are inventors who hold significant numbers of patents, and their innovation output is defined by patent frequency as a function of age. The remaining individuals are eminent American scientists across a variety of disciplines; their innovation potential is defined by publication frequency.[4] The average innovation potential is presented in the figure along with a smoothed estimate using an Epanechnikov kernel with a bandwidth of 3 years. These data match closely with the general pattern noted above.

Given this pattern, consider the following simple model. First, we assume that innovators start their lives with a period of education during which they do not innovate. Let the (stochastic) length of education required for an individual i be e_i.[5] Additionally, we will define $g(a_i; z_i)$ as the individual's innovation potential if fully educated, where z_i is some (stochastic) measure of talent, effort, health, and any other factor that influences innovative ability. The innovation potential of

[4] See the Data Appendix for further details.

[5] Two reasons we might expect a distribution in the age at which innovators arrive at the knowledge frontier are (i) different fields contain different types or amounts of knowledge, and (ii) individuals vary in the speed at which they educate themselves within a specific field.

individual i as a function of their age is then,

$$x_i = I(a_i \geq e_i)g(a_i; z_i)$$

where $I(a_i \geq e_i)$ is an indicator function equal to 1 if $a_i \geq e_i$ and 0 otherwise.

With this specific description of individual innovation potential, we can employ a law of large numbers to write the cohort average innovation potential as,

$$\bar{x}_a(t) \xrightarrow{p} E[I(a_i \geq e_i)g(a_i; z_i)]$$

Assuming additionally that e_i and z_i are independent[6] this expectation simplifies to,

$$\bar{x}_a(t) \xrightarrow{p} \Pr(a_i \geq e_i)E[g(a_i; z_i)] \tag{3}$$

To estimate $\Pr(a_i \geq e_i)$, we will assume first that e_i is distributed logistically within cohorts,

$$\Pr(a_i \geq e_i) = \frac{1}{1 + e^{-(a_i - \mu(t))/\omega(t)}} \tag{4}$$

where $\mu(t) = E[e_i]$ and $\omega(t)$ is a variance parameter. A logistic specification seems reasonable as it is parametrically simple, flexible, and captures the "S" shape we see leading up to the peak of Figure 4.[7] Figure 5 presents a graph to clarify this logistic specification and the meaning of the parameters.

To estimate the expectation of $g(a_i; z_i)$ I will assume a second logistic curve with parameters θ and ρ,

[6] In general, it might be more reasonable to assume that the distributions of e_i and z_i are not independent: high ability (high z_i) individuals may be more efficient in their education (low e_i). While this possibility is not considered formally in this paper, the data set used here considers great inventors, for whom we might take the view that z_i is essentially constant.

[7] Another natural option is a normal distribution, a variation that has no substantive effect on the results.

$$E[g(a_i; z_i)] = 1 - \frac{1}{1 + e^{-(a_i - \theta(t))/\rho(t)}} \tag{5}$$

This reverse "S" curve appears reasonable, as it can capture the initially slow decline in output in later middle age, followed by the more rapid decline and then tailing off of output into old age, as seen in Figure 4 and documented in other literature noted above.

With equation (3) and its sub-components (4) and (5), we now have a model for innovation potential over the life-cycle. We can estimate this model to determine how the propensity to produce great achievements in knowledge changes with age. By articulating specific underlying models for both the "front end" of the life-cycle, (4), and the "back end" of the life-cycle, (5), we can ask not only whether innovation potential has been shifting over time but, more specifically, whether any shifts are coming from the early years of life, the late years of life, or both.

A motivational question in this paper is whether $\mu(t)$, the mean age at which the active career begins, is changing over time. Shifts in this mean over time can be generally modeled by a polynomial expansion,

$$\mu(t) = \mu_0 + \mu_1 t + \mu_2 t^2 + ... \tag{6}$$

Shifts in the standard deviation of the can be modeled similarly. The main estimation presented in Section 4 will allow for a linear trend in $\mu(t)$ and a fixed variance parameter, ω; more general specifications will also be considered as robustness checks.

As with the beginning of the innovative career, we can further allow for shifts in innovation potential at the end of the career,

$$\theta(t) = \theta_0 + \theta_1 t + \theta_2 t^2 + ... \tag{7}$$

For example, as noted above, shifts that increasingly favor experience over raw ability may increase later life innovation potential. Meanwhile, improved health technology may lead to clearer thinking and/or increased physical stamina, while, alternatively, rising incomes could encourage earlier retirement and a decline in average innovation potential among older innovators. In the estimation I will constrain $E[g(30; z_i)] > 0.9$ to ensure that (5) and movements in it are describing effects later in life, which will make the results more transparent to interpret. This strategy will help us to substantially limit theories for the increased age at great innovation over the 20th Century.

Taken together, equations (2), (3), and the distributional assumptions in (4), (5), (6) and (7) produce a stochastic model that integrates demographic effects with a model of knowledge accumulation.[8] We can now attempt to determine what is driving the upward age trend among inventors.

4 Estimation and Results

This section proceeds in three parts. It first discusses the data used to estimate $p_a(t)$. It then turns to estimates of the sources of the inventors' age trend, presenting the paper's central results, and concludes by considering the robustness of the estimations.

4.1 Population Data

The great innovators come from many different countries and are therefore drawn from populations with differing age distributions. Data on these age distributions are difficult to find for many countries, particularly over the time-frame of the entire 20th century. For this reason, the estimation will focus on the American subset of great innovators. The American innovators show a similar trend in mean age at great achievement as the larger group and provide a significant number of

[8] The log-likelihood function is:

$$\sum_i \log \left(\frac{p_{a_i}(t_i) \frac{1}{1+e^{-(a_i-\mu(t))/\omega(t)}} \left(1 - \frac{1}{1+e^{-(a_i-\theta(t))/\rho(t)}}\right)}{\sum_{\{a_i \in A\}} p_{a_i}(t_i) \frac{1}{1+e^{-(a_i-\mu(t))/\omega(t)}} \left(1 - \frac{1}{1+e^{-(a_i-\theta(t))/\rho(t)}}\right)} \right)$$

observations on their own.[9]

Figure 6 shows the population age density, $p_a(t)$, of the United States for three selected census years. The densities are calculated from large micro-samples of the U.S. census. With these micro-samples, it is possible to determine not only the age distribution for (i) the entire national population, but also the distribution for subgroups of (ii) active workers and (iii) professional scientists and engineers. The scientist and engineer data are appealing as they may capture a closer approximation of the relevant age distribution of innovators. However, the sample sizes are small in early census years, and the occupational codes in the census are not entirely consistent across time, raising concerns that shifts in the age distribution for this sub-group may partly be an artifact of shifting classifications. The maximum likelihood model will be estimated using each of these population data sets. As we will see, the estimates are quite insensitive to the choice of population. See the Data Appendix for further discussion of these census data and the construction of the science and engineers sub-sample.

4.2 Central results

Table 3 presents the maximum likelihood estimates for innovation potential. Specifications (1) through (3) present estimates of innovation potential assuming it is stationary over the 20th Century. Specifications (4) through (6) present the main estimates, which allow for linear trends in life-cycle innovation potential over time.

There are two striking results. First there is a large and statistically significant shift in life-cycle innovator productivity over time, even when controlling for an aging population. Second, the shift in innovation potential is felt entirely at the beginning of the life cycle. In particular, we see that the mean age at which innovators begin making active contributions has increased by about 8 years over the course of the 20th Century, rising from a mean age of about 23 in 1900 to approximately

[9] There are 294 American-born great innovators. The trend in age at great achievement is 8.24 years/century with a standard error of 2.58 years/century.

31 in the year 2000. These results are robust to the choice of population data.

Meanwhile, there is no shift in innovation potential in middle age or beyond. The estimates show virtually zero movement, although robustness checks (see below) suggest non-zero but still highly insignificant movements in other specifications. Interestingly, this stationarity implies that demographic effects have driven the rising density of innovators beyond middle age that was seen in Figure 2.

Figure 7 compares the estimated life-cycle curves for the year 1900 and the year 2000, using specification (6). We see that the peak ability to produce great achievements in knowledge came around age 30 in 1900 but shifted to nearly age 40 by the end of the century. An interesting aspect of this graph is the suggestion that total lifetime innovation potential has declined. Other things equal, if individuals delay the start of their innovative careers without increasing their productivity at older ages, then their life-cycle output will fall. In particular, the area under the 2000 curve in Figure 7 is 30% smaller than it was in 1900. This point will be further discussed in Section 5.

Finally, note that the estimated life-cycle innovation potential follows the general pattern noted in the literature on creative careers. We see a period of essentially no innovation potential at very young ages, followed by a steep rise and then a more gradual tailing off. We also see a substantially steeper decline in the capacity for great innovations than we see in Figure 4, which considered more ordinary innovations, unweighted for quality. This suggests that while innovators often remain active into later ages, great achievements remain a younger person's game.

4.3 Robustness

In this section I consider a number of robustness checks. Columns (1) through (3) of Table 4 allow for additional linear trends in the variance parameters of the logistic distributions. We see that, regardless of the population data used, the only statistically significant movement in innovation potential continues to come at the beginning of the life cycle, in the mean age at which the career begins. The point estimates suggest that the variance may tighten in the initial part of the life cycle, although this is far from significant. Many other functional forms have also been examined, including those that do not assume logistic distributions. Regardless of the functional form, I find that the significant trends appear only in the beginning of the life-cycle, showing an increasing age at which innovators begin innovating.[10]

As a second robustness check, columns (4) and (5) re-estimate the model using all unique observations of great inventors worldwide. While the population data used covers only the U.S., one might consider the population patterns in the U.S. to be reasonable approximations for other developed countries where most great inventors live. The empirical estimates using the worldwide data agree substantially with the pattern estimated for U.S. inventors alone. The trend coefficient is slightly lower, similar to what we saw in Table 1 when we do not control for country fixed effects, and the standard errors smaller, as expected given the increased sample size.

Lastly, Table 5 considers the U.S. Nobel Prize and U.S. great inventor datasets separately. While the Nobel Prize is in principle given for distinct achievements, we might be concerned that other criteria affect the selection, and that these criteria have shifted over time to favor older innovators. For example, an increasing bias towards lifetime achievement would have this effect. Furthermore, the Nobel Prize is retrospective and requires award winners to be alive; therefore, increased longevity beyond the achievement age could also increasingly favor the old.

These possible selection concerns regarding Nobel Prizes are unlikely to be important for several reasons. The great inventor data set, which simply lists the great technological achievements in each given year, appears more immune to these kinds of selection biases and yet, on every dimension, produces similar results. As seen in Figure 1, the Nobel Prize and great inventors data sets have

[10] The logistic functions are appealing for the reasons discussed in Section 3.1. The estimation methodology can, however, also employ polynomial or piecewise linear functions to describe sequential pieces of the life cycle. These specifications produce similar results: an increase, prior to middle age, in the age at which innovators begin producing great ideas.

extremely similar age distributions. Second, these two groups show extremely similar mean trends, as detailed in Table 1. Third, Table 5 shows that the structural trends are similar for both groups when they are estimated independently; the coefficients for the great inventors are the same as for the whole, and the standard errors rise slightly as would be expected given the smaller sample size. These common patterns suggest common forces rather than idiosyncratic selection effects. Finally, the results of Tables 3, 4, and 5 show shifts in innovation potential at the *beginning* of the life-cycle and not at the end, which is not consistent with selection stories based on longevity or increasing favoritism for lifetime achievement.

5 Interpretations

This paper has presented a series of increasingly specific facts. The age at great innovation has trended upwards by approximately 6 years over the course of the 20th Century. This trend is not simply due to an aging population but reflects a substantial change in the life-cycle productivity of innovators. Specifically, the age at which the young begin their innovative careers has risen by approximately 8 years. Meanwhile, there is no compensatory increase in creative potential beyond early middle age.

This rising delay at the beginning of the innovator's life cycle has also been echoed in other research. A study of patenting behavior since 1975 found that the age at first patent among the general innovator population has shown a similar upward trend - at a rate of about 6 to 7 years per century if extrapolated (Jones, 2005). That age trend holds across widely different fields and institutional environments, including both government and corporate research. Taking the facts together, we see similar trends among both the greatest minds and ordinary inventors. We appear to be seeing a general phenomenon.

The increasing delay at the beginning of the innovator's life cycle may be an important element in understanding aggregate inventive output. Other things equal, the shorter the period that innovators spend innovating, the less their output as individuals. If innovation is central to technological progress, then forces that reduce the length of active innovative careers will reduce the rate of technological progress, working against growth in standard models. This effect will be particularly strong if innovators do their best work when they are young. In fact, aggregate data patterns, much debated in the growth literature, have noted long-standing declines in the per-capita output of inventors, both in terms of patent counts and productivity growth (Machlup 1962; Evenson, 1991; Jones 1995a; Kortum, 1997). Simple calculations suggest that the typical R&D worker contributes approximately 30% as much to aggregate productivity gains today as she did at the opening of the 20th Century.[11] This paper provides micro-evidence that explains part of that trend. Other things equal, the estimates of Section 4 indicate a 30% drop in the lifetime innovation potential over the century, or nearly half of the overall decline in individual research productivity.[12]

The maximum likelihood estimates focus interpretations on effects limited to the young. Explanations must confront not a general aging effect but a specific delay at the beginning of the life-cycle.[13] One natural and intriguing possibility is that innovators are choosing more education

[11]Combining Machlup's data on growth in knowledge producing occupations for 1900-1959 (Machlup 1962, Table X-4) with similar NSF data for 1959-1999 (National Science Foundation, 2005), we see that the total number of knowledge-producing workers in the United States has increased by a factor of approximately 19. Meanwhile, the U.S. per-capita income growth rate, which proxies for productivity growth over the long-run, suggests a 6-fold increase in productivity levels (based on a steady growth rate of 1.8%; see Jones 1995b). The average rate at which individual R&D workers contribute to productivity growth is $\dot{A}/L_R = gA/L_R$, where A is aggregate productivity, g is the productivity growth rate, and L_R is the aggregate number of R&D workers. The average contribution of the individual R&D worker in the year 2000 is then a fraction $\left(A^{2000}/A^{1900}\right)/(L_R^{2000}/L_R^{1900}) = 6/19$ (32%) of what it was in 1900.

[12]This paper estimates the relative innovation potential across age groups, so that forces that enhance or reduce the impact of all innovators, regardless of age, are not captured. Other influences, on top of delays at the beginning of the life-cycle, may therefore help to explain further portions of the declining trend in the average contributions of innovators. Suggested mechanisms include innovation exhaustion or "fishing out" stories (e.g. Evenson, 1991; Kortum, 1997), as well as narrowing expertise and innovative capacity as an endgenous response to an increased educational burden. Jones (2005) provides theory and empirical support for this latter mechanism.

[13]Theories that focus on productivity in the later life-cycle, such as improved health effects, find little support. Theories that suggest delays in innovation at both young and old ages will also have trouble explaining the specific empirical patterns we see. For example, research on creative careers in the arts (Galenson & Weinberg 2001; Galenson

over time, and the balance of this section will explore this interpretation in detail. In the first part, I will examine trends in the age at Ph.D. of Nobel Prize winners. I will then employ a simple formalization of educational choice to analyze why innovators would delay their careers, even if the delay results in less time for innovation.

5.1 The Age at Ph.D.

Given the increasing delay at the beginning of the life-cycle, an obvious question is whether this delay is reflected in longer periods of formal education. Indeed, several studies have documented upward trends in educational attainment among the general population of scientists.[14] While the Ph.D. is an institution that only approximately captures the end of the training phase and the beginning of the primary research phase, I will briefly consider the basic facts.

For 92% of the Nobel Prize winners, it was possible to determine the age and location for the highest degree. In 96% of these cases, the highest degree was a doctorate, in 2% a Master's degree, and in the remaining cases the type of degree was unclear. I analyze trends in the age by running the following regression:

$$a_i^D = \alpha + \beta t_i^D + \gamma X_f + \varepsilon_i \tag{8}$$

where a_i^D is the age of individual i at the time of their highest degree, t_i^D is the year of the highest degree, and X_f are fixed effects for the country of the degree and the field of the ultimate achievement.

2004a; Galenson 2004b) has suggested a useful distinction between "conceptual" innovation and "experimental" innovation, where the former favors the young and the latter favors the old – often the very old. However, these important ideas are not wholly satisfactory here because an increasing experimental bias would presumably be felt to a large degree at older ages.

[14] For example, the age at which individuals complete their doctorates rose generally across all major fields in a study of the 1967-1986 period, with the increase explained by longer periods in the doctoral program (National Research Council, 1990). The duration of doctorates as well as the frequency and duration of post-doctorates has been rising across the life-sciences since the 1960s (Tilghman et al, 1998). A study of electrical engineering over the course of the 20th century details a long-standing upward trend in educational attainment, from an intitial propensity for bachelor degrees as the educational capstone to a world where Ph.D.'s are common (Terman, 1998).

The results are presented in Table 6. We see that Nobel Prize winners are completing their formal education at substantially older ages today than they were a century ago. There is an upward age trend of approximately 4 years per century, and the trend is robust across specifications. Interestingly, while this trend is large, it accounts for only about half of the delay seen in the maximum likelihood estimates, although it is within those estimates' confidence intervals. Institutional variations in Ph.D. requirements complicate further interpretations. For example, the country fixed effects in column (4) are jointly significant with a p-value of less than .0001. This suggests that variations in degree requirements differ across countries; institutional variations over time are then likely as well. For example, the rise of post-doctorates and/or "on-the-job training" may explain part of the upward trend in ways not captured by charting ages at Ph.D..[15]

5.2 The Educational Decision

This paper begins by noting an intriguing tradeoff innovators confront at the beginning of their lives. Innovators build on earlier knowledge, but knowledge acquisition delays the active production of new ideas. More generally, innovators must compare the return to active production with the return to further training, whatever the type of benefits that training brings. In this final section, I will analyze the educational decisions of innovators in a simple model to both clarify and limit theories for the empirical patterns uncovered in this paper.

A reasonable specification, especially for highly motivated innovators, is that they choose their educational attainment to maximize their lifetime research contribution. In particular, their choice problem is:[16]

[15] "On-the-job training" can be substantial. For example, Nobel prize-winning physicist Philip Anderson describes his early years at Bell Labs, after his Ph.D., as a period where "Lannier taught me many fundamental techniques" and "I learned chrystallography and solid state physics...and most of all the Bell mode of close experiment" (Lundqvist, 1992).

[16] More generally, this objective function captures cases where utility is defined to maximize fame or income, so long as fame and income are monotonic functions of lifetime innovative output.

$$\max_E \int_E^T f(E;k)g(a)da$$

where $f(E;k)$ represents the value of education to their innovative output, and $g(a) > 0$ represents the individual's natural ability as a function of age. Individuals spend some number of years, E, focused on education during which time they do not innovate, followed by a career of innovation until they die at time T. The amount of education will influence their ultimate productivity, where $\partial f/\partial E > 0$ and $f(0) = 0$. The parameter k represents the state of knowledge or other environmental conditions that influence the value of education.

The first order condition for this problem is:

$$\frac{\partial f(E;k)}{\partial E} \int_E^T g(a)da = f(E;k)g(E) \tag{9}$$

which clarifies the central tradeoff. Greater education brings a benefit: the incremental effect on innovative output, $\partial f/\partial E$, weighted by the innovative ability that remains over your lifetime. But it also brings an opportunity cost, $f(E;k)g(E)$, the current innovation potential foregone.

Figure 8 presents a graph for analyzing the optimization condition (9), with the terms usefully rearranged. In particular, I label the upward sloping curve the "longevity curve" and the downward sloping curve the "educational return curve".[17]

Consider the educational return curve first. One apparently intuitive idea is that innovators will seek more education if education becomes "more valuable". However, increases in the value of education can – surprisingly – have no effect or even the opposite effect. Increasing the value of education may increase the marginal return, but it will also increase the inframarginal value

[17] The longevity curve has some finite value at age 0 and approaches ∞ as $E \to T$, since $\lim_{E\to T} \int_E^T g(a)da = 0$. The educational return curve has some finite value at age T and approaches ∞ as $E \to 0$, since $f(0) = 0$ and therefore $\lim_{E\to 0} f'(E)/f(E)$ is unbounded. Hence there is an interior solution to this problem. I have drawn these curves as monotonic in E, so that the curves have a single crossing property, but this is not necessarily the case under general functional specifications.

of what is already learned – raising the opportunity cost of further education. For example, if f increases by a fixed multiple, then the educational return curve does not shift in Figure 8 and innovators' optimal educational attainment will not change. Alternatively, if the value of education, f, increases by a constant, then the optimal amount of education can actually decline.[18]

A second idea is that knowledge becomes more difficult to acquire. To assess this possibility, consider a specific, learning-based model. First, knowledge acquisition begins with fundamental skills that are not field specific, such as basic mathematics training. These skills are necessary foundations for more advanced learning, but do not on their own raise the future innovator's immediate innovative potential to any appreciable degree. For example, the typical twelve-year-old has mastered a number of essential concepts, upon which future training will build, but the twelve-year-old still has no appreciable capacity to innovate. The required acquisition of preliminary knowledge means that f remains low through much of one's training. At some point, education begins to bring the individual close to the research frontier. Here, f begins to rise substantially.

In this investment process, increases in the stock of preliminary knowledge cause the marginal benefit of education to remain low for longer periods at the beginning of the life-cycle. In particular, a rightward translation of f by a fixed amount will cause the "educational return curve" in Figure 8 to shift by an equivalent amount and, therefore, E^* to rise.[19] Increasing educational attainment in a field is then an optimal response to a growing set of foundational knowledge. Useful knowledge for research simply takes longer to acquire – there is an increasing distance to the knowledge frontier.

Finally, increases in life expectancy, T, can also produce endogenous increases in educational attainment. Increasing longevity raises the marginal benefit of education. In Figure 8, the "longevity" curve shifts to the right, resulting in higher E^*. This effect is likely small however,

[18] Formally, in the first case, write $f(E;k) = k\widetilde{f}(E)$. Then $(\partial f/\partial E)/f$ is insensitive to the value of k. In the second case, write $f(E;k) = \widetilde{f}(E) + k$. Then $(\partial f/\partial E)/f$ is decreasing in the value of k at all values of E.

[19] Formally, let $f(E;k) = \widetilde{f}(E-k)$ and define $E_1 = E + k$. Then $\widetilde{f}'(E)/\widetilde{f}(E) = \widetilde{f}'(E_1-k)/\widetilde{f}(E_1-k)$, so a translation of $\widetilde{f}(E)$ by an amount k causes an equivalent translation in the educational return curve.

because raw ability and/or effort are low at the end of life. Adding low-productivity years at the end of life, especially years beyond typical retirement ages, will do little to compensate for shorter active innovation periods at younger ages.[20]

Other possible explanations can be considered within or outside this optimization framework.[21] However, the interesting result from this theoretical analysis is to suggest that education-based theories for the empirical patterns in this paper must follow particular forms. General shifts in the value of education may substantially impact the overall lifetime productivity of innovators and yet, to the extent that *time* opportunity costs are definitive, have no effect on educational attainment. In particular, the evolution of innovative opportunities, a subject much debated in the growth and innovation literatures, appears orthogonal in a basic analysis to the trends seen in this paper.[22] The drop in lifetime innovation potential seen in Figure 7 may be either exacerbated or compensated for by general shifts in creative opportunities.

Meanwhile, the empirical patterns are consistent with an increase in the amount of foundational knowledge passed from one generation to the next. In this theory, a natural by-product of innovation is an increasing educational burden imposed on successive generations of innovators.[23] Such an accumulation of foundational knowledge runs against the prominent ideas of Thomas Kuhn, who saw scientific progress as a process of revolutionary paradigm shifts (Kuhn, 1962). And here

[20] If innovators are impatient, then the effect of increased longevity is reduced further: with a 5% discount rate, a 30-year-old would discount innovative output at age 70 by 90%. Lastly, if we take the view that $f(E)$ is essentially a step function, with foundational knowledge acquisition providing no return until you get close to the frontier, then the educational return function may be very steep in the vicinity of the equilibrium, so that E^* will react little to shifts in the longevity curve.

[21] For example, institutional changes that force students to spend more time in teaching or other duties might delay active innovative careers; in that sense the f curve may shift for institutional reasons. Alternatively, if establishing a reputation is a prerequisite to doing grant-based research, and research has become more grant-based (expensive) over time, then extensions of formal education may serve as a signalling mechanism - a different interpretation of the educational return curve. This idea is perhaps more compelling, although (unreported) results show a substantial age increase at Ph.D. and great achievement among the economists alone, and these prize winners have had little need for large grants.

[22] Popular ideas include a "fishing-out" mechanism, where good ideas become increasingly hard to find (e.g. Kortum 1997), as well stories for increasing returns, where the knowledge of one generation increases the creative potential of future generations (e.g. Weitzman 1998).

[23] See Jones (2004) for a range of other evidence consistent with an increasing burden of knowledge.

an interesting irony appears. Scientists, disinclined to believe that their current views are fundamentally wrong, tend to see the Kuhnian view as a pessimistic description of scientific progress. But once we consider the educational burden that accumulations of knowledge can impose – increasingly limiting the life-cycle contributions of innovators – the revolutionary view may be the more optimistic one.

6 Conclusions

Great minds produce their greatest insights at substantially older ages today than they did a century ago. This upward age trend is not due simply to an aging population, but comes from a substantial decline in the innovative output of younger innovators. Meanwhile, there is no compensatory expansion of innovative output at later ages. Innovators are the engines of technological change and, other things equal, the less time an innovator spends successfully innovating, the less her lifetime output. The estimates point to a 30% decline in life-cycle innovation potential over the 20th Century.

Innovators build directly on earlier knowledge and spend significant portions of their early years in education. This human capital acquisition provides a possible explanation for the trends we see. In a simple optimization framework, the accumulation of knowledge – a rising distance to the frontier – can explain increased educational attainment. Whether or not this particular theory is correct, the economics literature has focused little on the human capital investments of innovators. Given that innovators spend some of their youngest and potentially brightest years undertaking educational investments, understanding the tradeoffs at the beginning of the life-cycle may be first-order for understanding the ultimate output of these individuals. Certainly, great innovation is less and less the provenance of the young.

7 Data Appendix

This appendix describes the data sources used in the paper, providing both reference material and some underlying details of the methodology used in data collection.

7.1 Data on great innovators

A wealth of biographical information is available for Nobel Prize winners. The most useful immediate source is the official website of the Nobel Foundation, nobelprize.org. This website provides lists and written biographies of all winners, and was used to obtain dates and locations of birth, the field of the prize, the year and location of the highest educational degree, and the year(s) in which the prize-winning research was performed. Altogether, I was able to determine dates of birth for 547 of the 564 Nobel Prize winners (97.0%), and the period of key research for all but 3 of these. The period of key research was usually straightforward to ascertain, although for 30% of the prizewinners it spans multiple years, in which case the apparent initial and final years of the key research were collected. In cases where the Nobel Foundation's web-site did not accurately identify the year or period of key research, other sources were consulted. The primary printed materials used were:

Schlessinger, B. and Schlessinger, J. *The Who's Who of Nobel Prize Winners*, 1901-1995. Oryx Press, Phoenix AZ 1996.

which was cross-referenced with,

Daintith, J. and Gjertsen, D. *The Grolier Library of Science Biographies.* Vols. 1-10. Grolier Educational, Danbury CT 1996.

Debus, A.G. ed. *World Who's Who in Science: A Biographical Dictionary of Notable Scientists from Antiquity to the Present.* Marquis Who's Who Inc., Chicago 1968.

McMurray, E.J., Kosek, J.K., and Valade, R.M. *Notable Twentieth-Century Scientists.* Vols. 1-4. Gale Research, Detroit 1995.

Williams, T.I. ed. *Biographical Dictionary of Scientists.* John Wiley and Sons, New York 1974.

Where necessary, the estimations in the text use the middle year of the research to define the age at great achievement, although results using either the first or last year of the key research are extremely similar in general and, in particular, nearly identical in the size of the trends and their statistical significance.

Data on great inventors were collected from two technological almanacs that provide, by year, a list of notable technological advances. These almanacs typically provide the date and location of birth of the innovator responsible, providing a set of 286 inventors in the 20th Century. The almanacs used were,

Bunch, B. and Hellemans, A. *The Timetables of Technology.* Simon and Schuster, New York 1993.

Ochoa, G. and Corey, M. *The Timeline Book of Science.* Ballantine Books, New York 1995.

In the event that the almanacs did not provide dates or location of birth, the four general biographical resources listed above were consulted, in addition to two further biographical resources:

Bachman, T.M. ed. *Who's Who of Science and Engineering: Millennial Edition.* Marquis Who's Who, Wilmette, IL 1999.

Inventors' Hall of Fame. Website: http://www.invent.org/. Akron, OH.

The field of research is given according to set categories in both the Bunch et al and Ochoa et al sources. I condense the categorizations across these sources into nine fields: Communication, Computers & Electronics, Energy, Food & Agriculture, Materials, Medicine, Tools & Devices, Transportation, and Other. These categorizations define the field fixed effects in the econometric specification (1), but the results are not sensitive to specific categorizations.

7.2 Data on great athletes

For the Most Valuable Player Award in baseball, dates of birth and dates of the award were taken electronically from the *Lahman Baseball Database*, version 5.1, which was released on January 24, 2004. This database contains an enormous set of data for baseball players from 1871 through the present. It can be found at http:/baseball1.com/.

For world record breakers in track & field, dates of birth, dates of record breaking event, and nationalities were taking from the following source:

Lawson, G. *World Record Breakers in Track and Field Athletics*, Human Kinetics, Champaign, IL 1997.

7.3 Data on population age distribution

One and five percent micro-samples of the U.S. census are available electronically through IPUMS, the Integrated Public Use Microdata Series, which is maintained by the University of Minnesota:

Steven Ruggles, Matthew Sobek, Trent Alexander, Catherine A. Fitch, Ronald Goeken, Patricia Kelly Hall, Miriam King, and Chad Ronnander. *Integrated Public Use Microdata Series:* Version 3.0 [Machine-readable database]. Minneapolis, MN: Minnesota Population Center [producer and distributor], 2004.

The smallest sample used was for the 1900 census, whose micro-sample provided data on approximately 100,000 individuals. The largest sample used was for the 2000 census, whose micro-sample provided data on approximately 2.8 million individuals. Existing census research available on the website (www.ipums.umn.edu/usa/chapter3/chapter3.html) indicates that these micro-samples provide accurate estimates of the population at large with regard to age. Population data for years in between decennial census years were determined by linear interpolation. Data for the year 1930 are not available, requiring interpolation for years between 1920 and 1940.

The subgroup of active workers is defined by having active labor force participation (LABFORCE=1) in the IPUMS data. The IPUMS attempts to recode census data to allow comparisons over time on a common basis, even if the census questions asked are not entirely consistent. From 1940, LABFORCE=1 for any individual who is actively working or seeking work. In 1900, the variable requires that individuals report any profession and have worked in the last 12 months. In 1910-1920, it includes those who report any "gainful occupation". This subsample is still large, with a minimum of 40,000 observations in 1900 and 1.3 million observations in 2000.

The subgroup of professional scientists and engineers requires further construction. The IPUMS uses the variable OCC1950 to define occupations across census years according to a common set of categories. I then take a subsample of these occupation codes that include the following relevant descriptions: professors and instructors in all subjects except social sciences (codes 012-026); engineers (codes 041-049); and natural scientists (codes 061-069). These data have two potential difficulties: first, the samples are substantially smaller, with only 56 observations in 1900 and 353 in 1910, rising to approximately 25,000 in the year 2000; second, the occupation is not defined until it is begun, in which case those still in school are not included. To create reasonable population estimates I first smooth these population data with an Epanechnikov kernel and a bandwidth of 2 years. Second, I impute the number of innovators still in training (those aged 15-29) based on the number of employed workers ten years later who are ten years older. The results are not sensitive to particular kernel bandwidths or age imputation schemes. In results not reported, I have also considered a broader set of all "professional, technical" workers (codes 001-099), which gives similar results.

7.4 Data on innovation potential

The data in Figure 4 considers the life-cycle output of prominent inventors and scientists. Patent data for 9 of the 11 patentees used in estimating life-cycle innovation potential were drawn from the following source,

Kraeuter, D.W. *Radio and Television Pioneers: A Patent Bibliography.* Scarecrow Press, Metuchen NJ 1992.

Patent data for the remaining 2 patentees, as well as data on the publication frequencies for 20 eminent American scientists, were drawn from multiple volumes of the following series,

National Academy of Sciences, *Biographical memoirs*, Washington, D.C.

The year in which a patent was granted was used to determine the age of the patentee at the time of innovation. The year of publication was used for the scientists. For the scientists, publications of any kind were included and equally weighted, whether they be journal articles, books, or lighter articles or editorials. All 31 persons lived into old age. The two youngest at their death were 67. Seventy percent lived beyond the age of 79, most well into their 80's. Kernel estimations of the patentees alone or of the scientists alone produce estimates that are less smooth but similar in their larger features.

References

[1] Bloom, B. "Generalizations about Talent Development," in *Developing Talent in Young People*, New York: Ballentine Books, 1985.

[2] Ericsson, K.A. and Lehmann, A. C. "Expert and Exceptional Performance: Evidence of Maximal Adaptation to Task Constraints," *Annual Review of Psychology*, 1996, 47, 273-305.

[3] Evenson, Robert E. "Patent Data by Industry: Evidence for Invention Potential Exhaustion?" *Technology and Productivity: The Challenge for Economic Policy*, 1991, Paris: OECD, 233-248

[4] Jones, Benjamin F. "The Burden of Knowledge and the Death of the Renaissance Man: Is Innovation Getting Harder?" NBER Working Paper #(TBD), 2005.

[5] Jones, Charles I. "R&D-Based Models of Economic Growth," *Journal of Political Economy*, 1995a, 103, 759-784.

[6] ——. "Time Series Tests of Endogenous Growth Models," *Quarterly Journal of Economics*, 1995b, 110, 495-525.

[7] Galenson, David W. "A Portrait of the Artist as a Young or Old Innovator: Measuring the Careers of Modern Novelists," NBER Working Paper #10213, 2004a.

[8] ——. "A Portrait of the Artist as a Very Young or Very Old Innovator: Creativity at the Extremes of the Lifecycle," NBER Working Paper #10515, 2004b.

[9] Galenson, David W., and Weinberg, Bruce A. "Creating Modern Art: The Changing Careers of Painters in France from Impressionism to Cubism," *American Economic Review*, September 2001, 91 (4), 1063-1071.

[10] Kortum, Samuel S. "Research, Patenting, and Technological Change," *Econometrica*, November 1997, 65, 1389-1419.

[11] Kuhn, Thomas S. *The Structure of Scientific Revolutions*, Chicago: University of Chicago Press, 1962.

[12] Lehman, Harvey C. *Age and Achievement*, Princeton, NJ: Princeton University Press,1953.

[13] Lundqvist, Stig, ed. *Nobel Lectures, Physics 1971-1980*, Singapore: World Scientific Publishing Co., 1992.

[14] Machlup, Fritz. *The Production and Distribution of Knowledge in the United States*, Princeton, NJ: Princeton University Press, 1962.

[15] National Research Council, *On Time to the Doctorate: A Study of the Lengthening Time to Completion for Doctorates in Science and Engineering*, Washington, DC: National Academy Press, 1990.

[16] National Science Foundation, *Industrial Research and Development Information System*, Table H-19, www.nsf.gov/sbe/srs/iris/start.cfm, 2005.

[17] Reese, H,W. Lee, L.J., Cohen, S. H., and Puckett, J.M. "Effects of Intellectual Variables, Age, and Gender on Divergent Thinking in Adulthood", *International Journal of Behavioral Development*, November 2001, 25 (6), 491-500.

[18] Simonton, Dean K. *Scientific Genius: A Psychology of Science*, Cambridge: Cambridge University Press, 1988.

[19] ——. "Age and Outstanding Achievement: What Do We Know After a Century of Research?" *Psychological Bulletin*, September 1988, 104, 251-267.

[20] ——. "Career Landmarks in Science: Individual Differences and Interdisciplinary Contrasts," *Developmental Psychology*, January 1991, 27, 119-130.

[21] ——. "Creativity," in *The Encyclopedia of Gerontology*, San Diego: Academic Press,1996.

[22] Terman, F.E. "A Brief History of Electrical Engineering Education", *Proceedings of the IEEE*, August 1998, 86 (8), 1792-1800.

[23] Tilghman, Shirley (chair) et al. *Trends in the Early Careers of Life Sciences*, Washington, DC: National Academy Press, 1998.

[24] Weitzman, Martin L. "Recombinant Growth," *Quarterly Journal of Economics*, May 1998, 113, 331-360.

[25] Zuckerman, Harriet and Merton, Robert. "Age, Aging, and Age Structure in Science," in Merton, Robert, *The Sociology of Science*, Chicago, IL: University of Chicago Press, 1973, 497-559.

Table 1: Age Trends among Great Innovators

	Dependent Variable: Age at Great Achievement					
	Nobel Prize Winners			Great Inventors		
	(1)	(2)	(3)	(4)	(5)	(6)
Year of Great Achievement (in 100's)	5.83*** (1.37)	6.34*** (1.36)	7.79*** (1.54)	4.86** (2.31)	6.60** (2.58)	8.18** (3.29)
Field Fixed Effects	No	Yes	Yes	No	Yes	Yes
Country of Birth Fixed Effects	No	No	Yes	No	No	Yes
Number of observations	544	544	544	286	286	248
Time span	1873-1998	1873-1998	1873-1998	1900-1991	1900-1991	1900-1988
Average age	38.6	38.6	38.6	39.0	39.0	38.9
R^2	0.032	0.068	0.189	0.016	0.098	0.173

Notes: Coefficient on year of great achievement gives age trend in years per century. Standard errors are given in parentheses. Field fixed effects for Nobel Prizes comprise four categories: Physics, Chemistry, Medicine, and Economics. Field fixed effects for great inventors comprise nine categories: Communication, Electronics and Computers, Energy, Food and Agriculture, Materials, Medicine, Tools and Devices, Transportation, and Other.

** Indicates significance at a 95% confidence level.

*** Indicates significance at a 99% confidence level.

Table 2: Age Trends among Great Athletes

	Dependent Variable: Age at Great Achievement			
	Baseball MVP	Track & Field World Record Breakers		
	(1)	(4)	(5)	(6)
Year of Great Achievement (in 100's)	-0.01 (1.18)	1.04 (1.19)	1.22 (1.14)	-0.89 (1.33)
Field Fixed Effects	--	No	Yes	Yes
Nationality Fixed Effects	--	No	No	Yes
Number of observations	160	328	328	224
Time span	1911-1999	1908-1996	1908-1996	1908-1996
Average age	29.0	25.0	25.0	24.8
R^2	0.004	0.003	0.247	0.348

Notes: Coefficient on year of great achievement gives age trend in years per century. Standard errors are given in parentheses. Errors are clustered by individual. Field fixed effects for Track & Field World Record Breakers comprise 10 categories: the 100m, 1500m, Mile, Javelin, and High Jump for Men, and the first four of these for Women.
** Indicates significance at a 95% confidence level.
*** Indicates significance at a 99% confidence level.

Table 3: Maximum Likelihood Estimation of Life-Cycle Innovation Potential

		(1)	(2)	(3)	(4)	(5)	(6)
Early Life Cycle Logistic Curve	μ_0 Initial Mean, in Years	28.1 (1.05)	27.9 (1.20)	28.0 (1.02)	24.0 (2.05)	23.3 (2.30)	23.4 (1.95)
	μ_1 Trend, in Years/Century	--	--	--	7.76 (3.22) [.016]	8.29 (3.49) [.018]	8.32 (2.71) [.002]
	ω Variance Parameter	2.53 (0.35)	2.65 (4.55)	2.55 (0.41)	2.40 (0.37)	2.47 (0.43)	2.38 (0.40)
Later Life Cycle Logistic Curve	θ_0 Initial Mean, in Years	45.6 (3.48)	46.6 (2.58)	50.1 (3.89)	45.5 (2.36)	46.6 (2.21)	50.0 (3.90)
	θ_1 Trend, in Years/Century	--	--	--	-0.00e-03 (8.63e-03)	-0.00e-03 (6.70e-03)	0.14e-03 (6.88e-03)
	ρ Variance Parameter	7.09 (1.16)	7.54 (9.75)	9.15 (1.82)	7.05 (1.10)	7.54 (0.91)	9.12 (1.60)
Data	Population	Entire U.S. Population	All Active Workers	Scientists and Engineers	Entire U.S. Population	All Active Workers	Scientists and Engineers
	Number of Great Achievement Observations	294	294	294	294	294	294
Log Likelihood		-1053.8	-1055.8	-1059.7	-1050.9	-1053.0	-1056.7

Notes: All estimates are maximum likelihood. Standard errors are given in parentheses and calculated using the inverse of the information matrix. P-values for the trend in the early life-cycle are given in square brackets.

Table 4: Maximum Likelihood Estimation: Further Specifications

		(1)	(2)	(3)	(4)	(5)
Early Life Cycle Logistic Curve	μ_0 Initial Mean, in Years	24.8 (1.94)	24.6 (1.84)	24.5 (2.36)	25.8 (1.07)	25.5 (1.19)
	μ_1 Trend, in Years/Century	6.32 (3.27) [.053]	6.05 (2.94) [.039]	6.47 (3.95) [.101]	5.32 (1.97) [.007]	5.89 (2.12) [.005]
	ω_0 Initial Variance Parameter	2.86 (0.83)	3.42 (1.01)	3.08 (1.14)	2.21 (0.20)	2.02 (4.18)
	ω_1 Trend, Variance Years/Century	-0.88 (1.42)	-1.74 (1.60)	-1.29 (1.80)	--	0.43 (7.35)
Later Life Cycle Logistic Curve	θ_0 Initial Mean, in Years	45.0 (4.49)	45.3 (3.22)	48.5 (4.88)	46.9 (3.70)	46.8 (3.79)
	θ_1 Trend, in Years/Century	0.99 (4.06)	2.29 (3.28)	2.79 (6.79)	-0.71 (5.17)	-0.93 (2.53)
	ρ_0 Initial Variance Parameter	6.81 (2.06)	6.97 (1.47)	8.42 (2.13)	7.38 (7.40)	7.67 (0.73)
	ρ_1 Trend, Variance Years/Century	0.45 (1.77)	1.04 (1.46)	1.27 (3.07)	--	-0.42 (1.09)
Data	Population	Entire U.S. Population	All Active Workers	Scientists and Engineers	Entire U.S. Population	Entire U.S. Population
	Nationality	U.S. Born	U.S. Born	U.S. Born	All	All
	Number of Great Achievement Observations	294	294	294	738	738
Log Likelihood		-1050.8	-1052.6	-1056.4	-2641.2	-2641.1

Notes: All estimates are maximum likelihood. Standard errors are given in parentheses and calculated using the inverse of the information matrix. P-values for the trend in the early life-cycle are given in square brackets. Specifications (4) and (5) use U.S. population data but all great achievements worldwide in the 20th Century.

Table 5: Maximum Likelihood Estimation: U.S. Nobel Prizes vs. U.S. Great Inventors

		(1)	(2)	(3)	(4)	(5)	(6)
Early Life Cycle Logistic Curve	μ_0 Initial Mean, in Years	22.6 (2.84)	23.2 (3.59)	23.4 (3.32)	22.8 (2.84)	21.6 (4.22)	20.8 (5.72)
	μ_1 Trend, in Years/Century	8.91 (5.23) [.088]	8.64 (5.84) [.139]	7.85 (5.18) [.130]	10.29 (4.56) [.024]	11.53 (4.46) [.010]	12.82 (6.67) [.055]
	ω Variance Parameter	2.43 (0.73)	2.83 (0.79)	2.64 (0.81)	2.30 (0.46)	2.32 (0.90)	2.23 (0.68)
Later Life Cycle Logistic Curve	θ_0 Initial Mean, in Years	53.7 (8.16)	50.8 (13.1)	61.5 (14.0)	43.7 (3.54)	44.4 (7.94)	46.8 (10.1)
	θ_1 Trend, in Years/Century	-4.61 (12.1)	-1.8 (17.7)	-7.20 (22.4)	0.00e-03 (1.28e-03)	0.00e-03 (2.20e-02)	0.09e-03 (7.40e-02)
	ρ Variance Parameter	6.64 (2.72)	8.64 (3.01)	8.43 (6.87)	6.24 (0.99)	6.55 (2.23)	7.65 (3.72)
Data	Population	Entire U.S. Population	All Active Workers	Scientists and Engineers	Entire U.S. Population	All Active Workers	Scientists and Engineers
	Source for Great Achievement (US Born Only)	Technology Almanacs	Technology Almanacs	Technology Almanacs	Nobel Prize	Nobel Prize	Nobel Prize
	Number of Great Achievement Observations	127	127	127	181	181	181
Log Likelihood		-463.6	-463.8	-466.6	-633.6	-635.4	-635.5

Notes: All estimates are maximum likelihood. Standard errors are given in parentheses and calculated using the inverse of the information matrix. P-values for the trend in the early life-cycle are given in square brackets.

Table 6: Age Trends at Highest Degree among Nobel Prize Winners

	Dependent Variable: Age at Highest Degree				
	(1)	(2)	(3)	(4)	(5)
Year of Highest Degree (in 100's)	4.11*** (0.61)	3.85*** (0.62)	3.86*** (0.62)	4.39*** (0.65)	3.22*** (1.22)
Field Fixed Effects	No	No	Yes	Yes	Yes
Country of Degree Fixed Effects	No	No	No	Yes	--
Data	All	Doctorate Only	All	All	U.S. Degree
Number of observations	505	484	505	505	213
Time span	1858-1990	1858-1990	1858-1990	1858-1990	1888-1990
Average age	26.5	26.6	26.5	26.5	26.6
R^2	.084	.075	.096	0.283	.060

Notes: Coefficient on year of highest degree gives age trend in years per century. Standard errors are given in parentheses. Field fixed effects for Nobel Prizes comprise four categories: Physics, Chemistry, Medicine, and Economics.
** Indicates significance at a 95% confidence level.
*** Indicates significance at a 99% confidence level.

Figure 1: The Age Distribution of Great Innovation

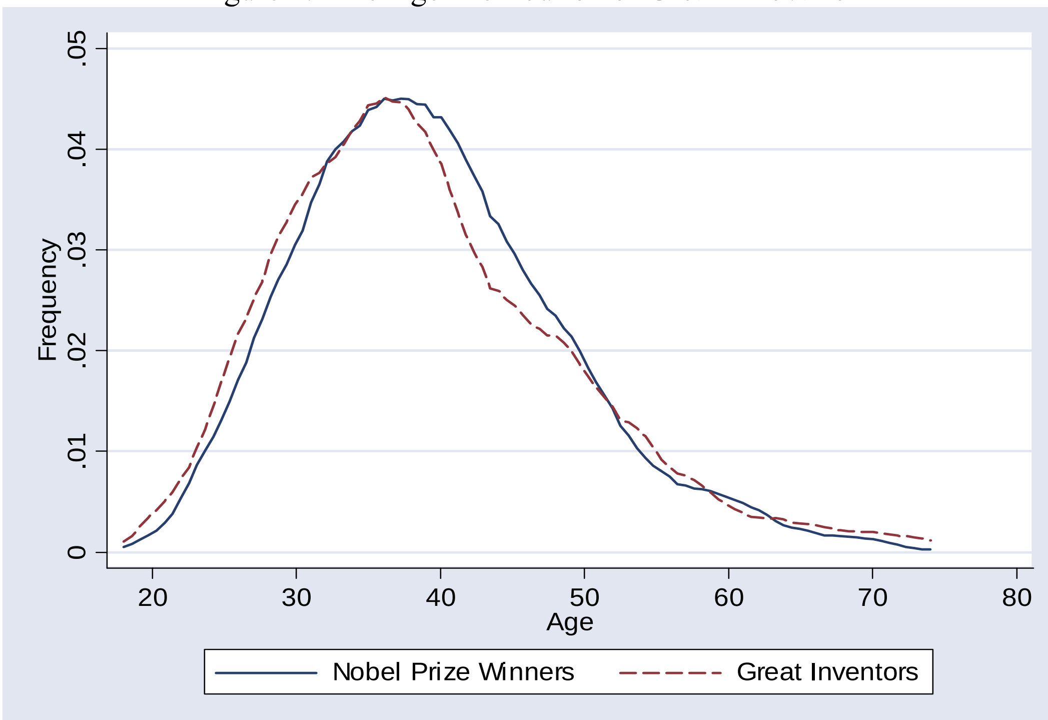

Figure 2: Shifts in the Age Distribution of Great Innovation

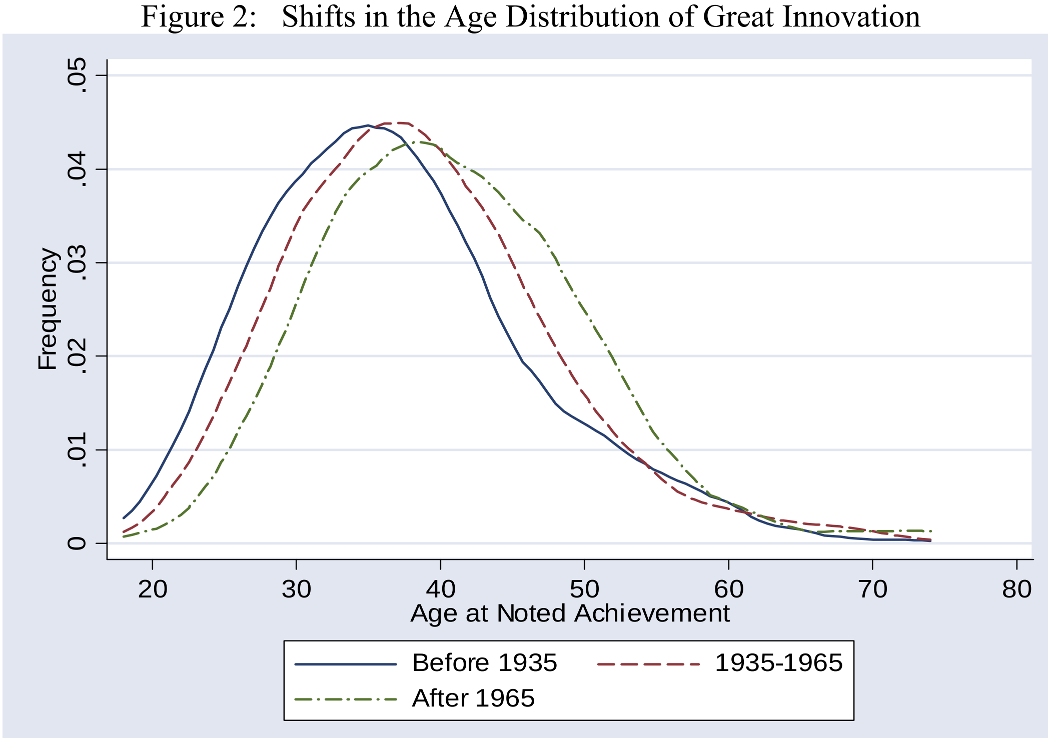

Figure 3: The Age Distribution of Baseball's Most Valuable Player

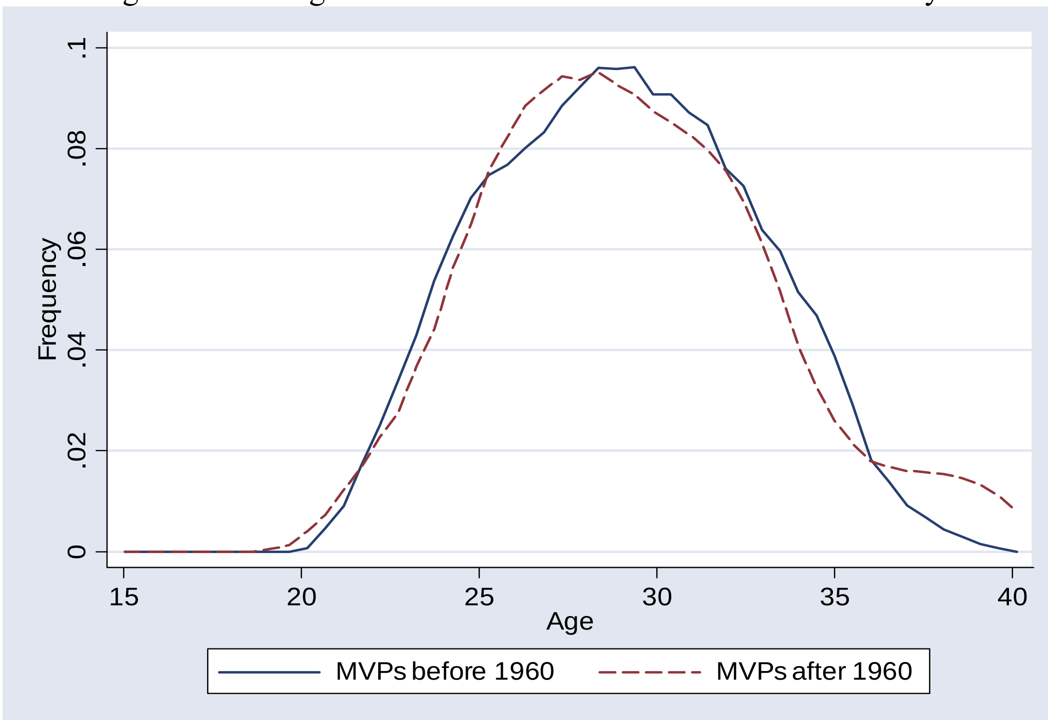

Figure 4: Innovative Output over the Life-Cycle for 31 Eminent Innovators

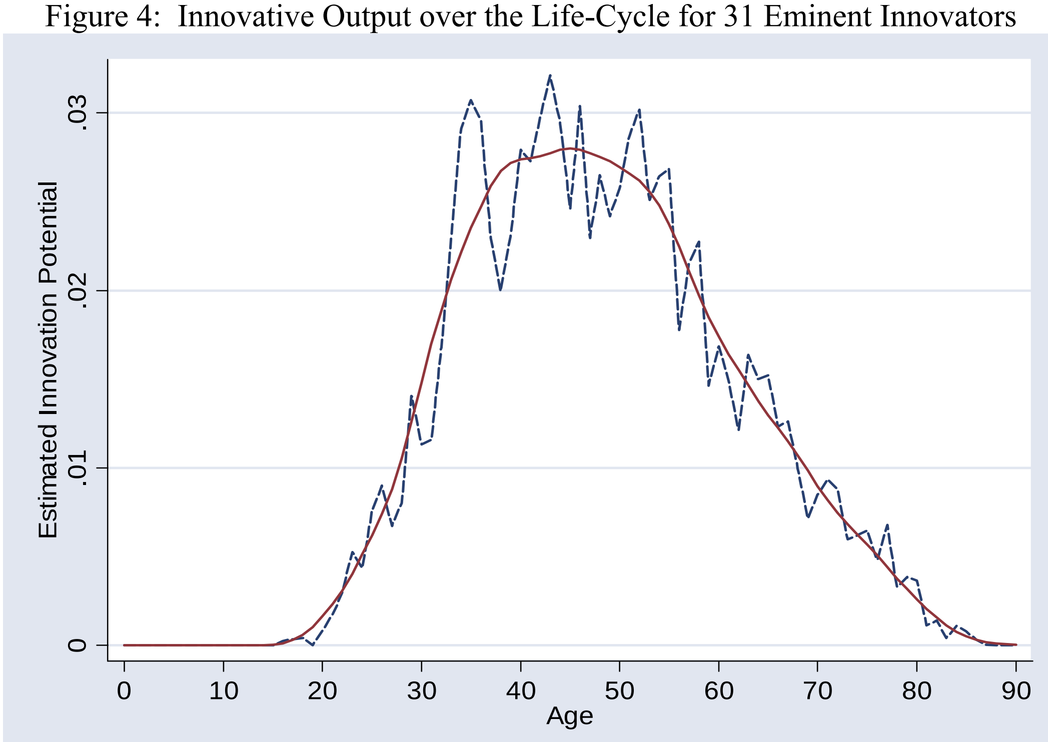

Figure 5: Model of Innovation Potential over the Life-Cycle

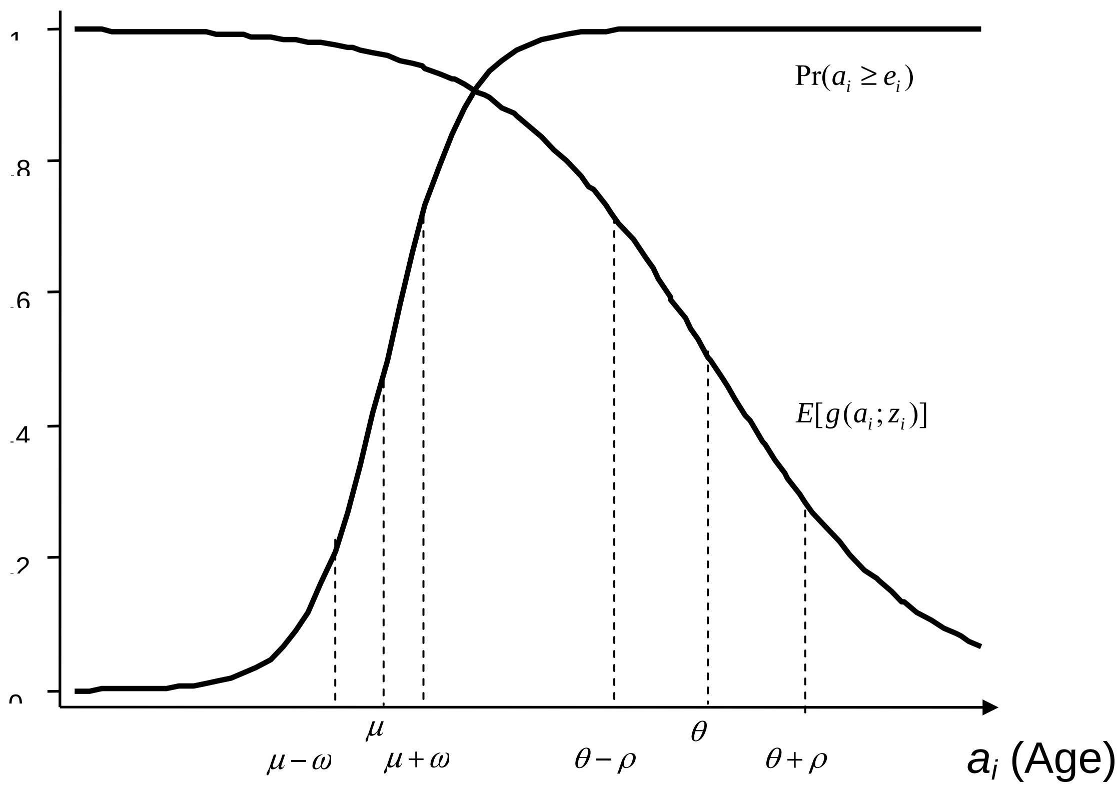

Figure 6: The United States Population Distribution in Two Selected Years

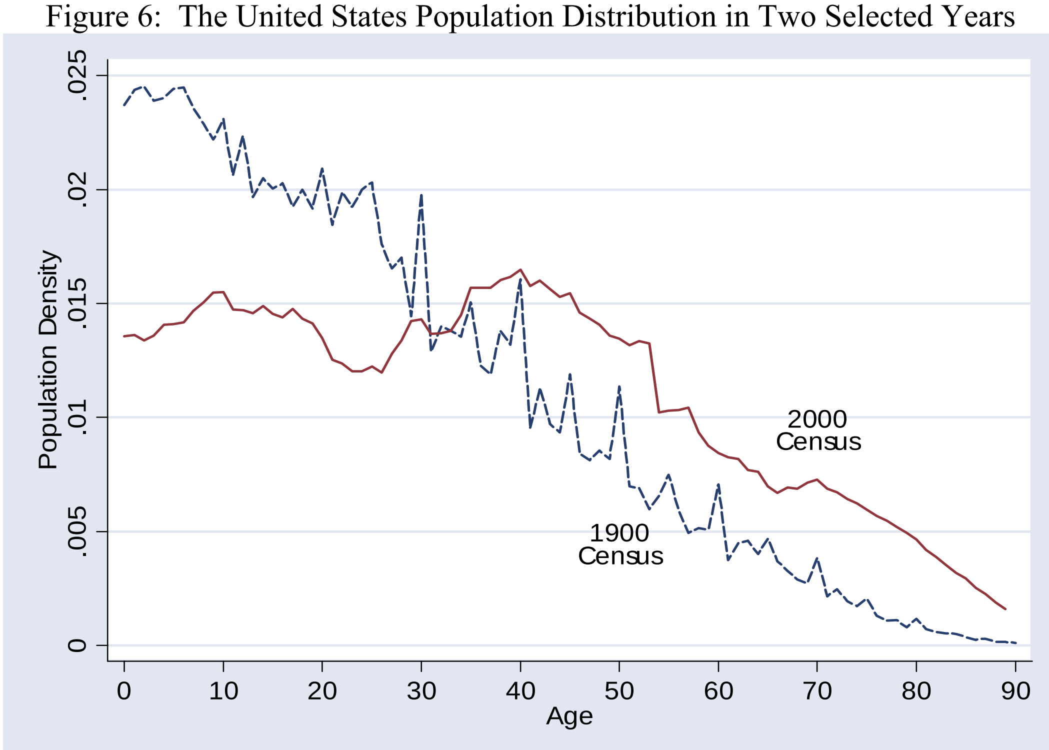

Figure 7: Maximum Likelihood Estimates for the Potential to Produce Great Innovations as a Function of Age

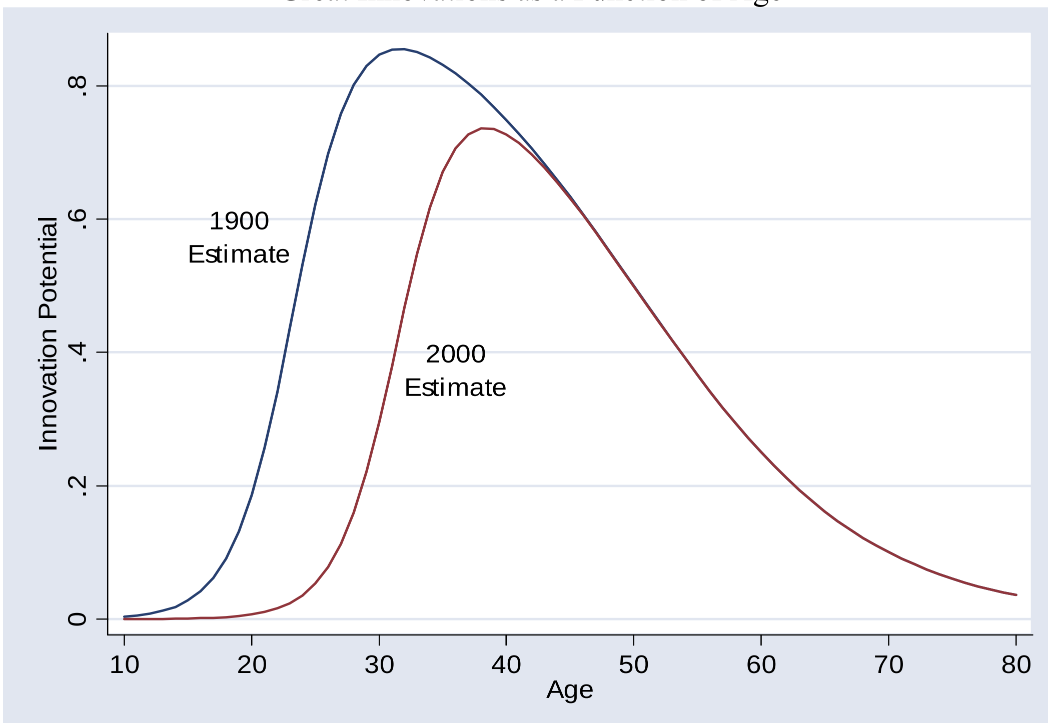

Figure 8: The Equilibrium Choice of Educational Attainment