The Review *of* Economics *and* Statistics

Vol. XCII | February 2010 | Number 1

AGE AND GREAT INVENTION

Benjamin F. Jones*

Abstract—Great achievements in knowledge are produced by older innovators today than they were a century ago. Nobel Prize winners and great inventors have become especially unproductive at younger ages. Meanwhile, the early life cycle decline is not offset by increased productivity beyond middle age. The early life cycle dynamics are closely related to age when the PhD was received, and I discuss a theory where knowledge accumulation across generations leads innovators to seek more education over time. More generally, the narrowing innovative life cycle reduces, other things equal, aggregate creative output. This productivity drop is particularly acute if innovators' raw ability is greatest when young.

> 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

I. 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 tend to support the idea that innovative activity is greater at younger ages, although great achievement before the age of thirty is not typical. Rather, a researcher's output tends to rise steeply in the twenties and thirties, peak in the late thirties or early forties, 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 further it is by standing on ye sholders 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.[2] Certainly, innovation would be a very difficult enterprise if every generation had to reinvent the wheel.

These two observations suggest an intriguing trade-off. If innovators are especially productive when young but education is an important preliminary input to innovation, then the opportunity cost of the time spent in education may be significant. Moreover, how innovators approach this trade-off may change as the economy evolves. For example, accumulations of knowledge across generations may create increasing educational demands, so that expanding the time costs of education delays 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 twentieth century occurred at later and later ages. The mean age at great achievement for both Nobel Prize winners and great technological inventors rose by about six years over the course of the twentieth century. This aging phenomenon appears to be substantially driven by declining innovative output in the early life cycle. Moreover, the early life cycle effects appear to be substantially explained by increases in training.

Section II presents the main fact: a substantial increase in the age at great invention. Several hypotheses for the trend are then introduced. In one type of hypothesis, the life cycle productivity of innovators may have shifted. For example, increasing educational attainment may delay the onset of active innovative careers. Alternatively, innovator productivity may increase at more advanced ages due to improved health, effort, or an increased role for experience. In another type of hypothesis, the upward age trend in the data could simply reflect underlying demographic shifts. 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 twentieth century. Put another way, if

Received for publication April 13, 2007. Revision accepted for publication May 13, 2008.

* Kellogg School of Management and NBER.

I thank Daron Acemoglu, Amy Finkelstein, Abhijit Banerjee, Leemore Dafny, James Dana, Ben Olken, Mike Mazzeo, Joel Mokyr, Sendhil Mullainathan, Scott Stern, Peter Temin, Bruce Weinberg, and two anonymous referees for helpful comments and suggestions. Alexander Karlan, Gowoon Lee, and Atanas Stoyanov provided excellent research assistance. All errors are my own.

[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 & Lehmann, 1996, for a review).

[2] Dirac built on Heisenberg's uncertainty principle and Hamiltonian mechanics, while Einstein's early insights built on the work of Planck and Maxwell.

people lived shorter lives in the past, then innovators in the past will also appear younger.

Section III tests these competing explanations and locates any specific shifts in life cycle productivity. 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. The estimates suggest that, on average, the great minds of the twentieth century typically became research active at age 23 at the start of the twentieth century, but only at age 31 at the end—an upward trend of eight years. Meanwhile, there has been no compensating shift in the productivity of innovators beyond middle age.

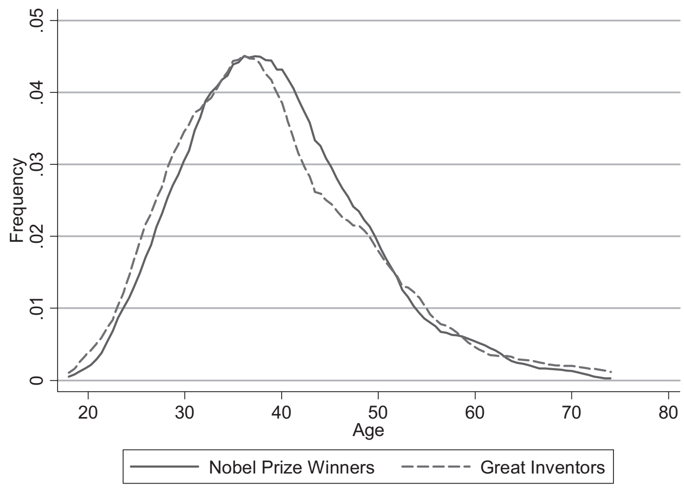
Section IV presents additional analysis to further understand the delayed start to the career. I first show that PhD age increases substantially over the twentieth century. I next harness World Wars I and II as natural experiments, testing the idea that training is a prerequisite for innovation and showing that interruptions to training must be "made up" after the war. Next, I investigate cross-field, cross-time variation and show that variations in PhD age typically predict variations in the age-invention relationship. Collectively these analyses suggest that training plays a key role in explaining the age-invention patterns.

Section V clarifies interpretations of the empirical patterns and considers their implications. I present a simple theory to explore the relationship between human capital investments and life cycle productivity and show how accumulations of knowledge within fields can explain the set of facts. Further evidence from ordinary inventions underscores this perspective and also shows that the aging phenomenon extends broadly across the innovator population. Section V closes by detailing implications of the empirical patterns for core issues in economic growth and the history of science. I show specifically how contractions in the life cycle of innovation can help explain the decline in innovative output per researcher seen over the twentieth century. Section VI concludes.

II. Age and Great Achievement

This section presents benchmark facts about the age of individuals at the time of their great achievements in knowledge. Two types of data were collected. The first set considers research that leads to the Nobel Prize in Physics, Chemistry, Medicine, and Economics. The second set considers great technological achievements as presented in almanacs of the history of technology. Recipients of the Nobel Prize 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 and agriculture, materials, and tools and 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 the Nobel Prize, the year of achievement was determined by consulting various biographical resources. The data appendix to this paper describes the data collection and sources in further detail.

Figure 1.—Age Distribution of Great Innovation



Note: Data are pooled across time.

As a first look at the data, figure 1 presents innovators' ages at their great innovation, considering all twentieth-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 30s (42%), but a substantial amount also came in the 40s (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 upward by five or six 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 eight years over the course of the twentieth century. The strengthening effect of the controls on the trend suggests a compo-

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	5.83***	6.34***	7.79***	4.86**	6.60**	8.18**
(in 100s)	(1.37)	(1.36)	(1.54)	(2.31)	(2.58)	(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

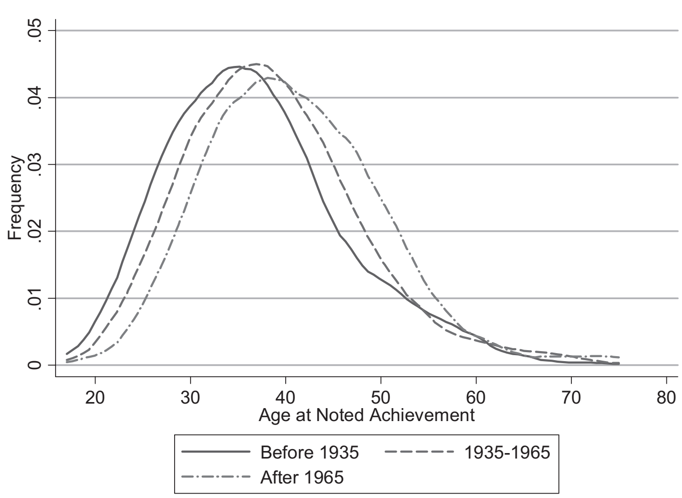
Notes: All specifications consider trends in the age at great achievement. Columns 1–3 consider Nobel Prize winners, and columns 4–6 consider great inventors listed in technological almanacs, as detailed in the text and data appendix. Field and country-of-birth fixed effects are included or excluded as indicated. The coefficient on year of great achievement gives the 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. Results are similar when allowing for separate country-of birth-fixed effects by field, or when using country of highest degree fixed effects instead of country of birth.
**Significance at a 95% confidence level. ***Significance at a 99% confidence level.

sitional shift in great innovation toward 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 twentieth century into three chronological periods: 1900–1935, 1935–1965, and 1965 to the present. This figure combines all unique individuals in the Nobel Prize and great inventors data sets. We observe a general shift of the age distribution, with a distinct drop in the presence of those in their 20s and an increased presence of those in later middle age.

One obvious hypothesis for this outward age shift is a shift in the life cycle productivity of great minds. 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, for example, due to improved health or an increased role for experience.

FIGURE 2.—SHIFTS IN THE AGE DISTRIBUTION OF GREAT INNOVATION



Note: The figure pools the Nobel Prize winner and great inventor data.

But we must be careful in interpreting the distributional shifts we see. An alternative 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 is 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 twentieth century.

The following section develops a formal econometric model to identify specific shifts in innovation potential, controlling for demographic effects. We 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 innovators' 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?

III. Life Cycle Productivity

This section presents an econometric model to define the probability that witnessed innovations are produced by innovators at particular ages. Empirical analysis follows, using this model to determine sources of the upward trend in the age of great achievement.

Given a great innovation, the probability that this innovation was produced by an individual of age a will depend on two things. First, it will depend on the relative innovation potential of innovators in different cohorts. For example, according to Dirac, a physicist below the age of 30 has more innovative potential on average than one who is older. Second, it will depend on the density of innovators of various ages. If a population is full of 50-year-old researchers but has very few 20-year-old researchers, then the likelihood a particular innovation came from a 20 year old is low, even if young innovators have good ideas.

Formally, consider a population $N(t)$ at time t. Given a witnessed innovation, the probability the innovation was produced by an individual i at time t is defined by

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

where $x_i(t)$ represents the innovation potential of person i at time t. Innovation potential measures the relative innovative strength of an individual.[3]

For estimation, consider the model in terms of cohorts of equally aged individuals. Define the set of cohorts as $A(t)$, where $a \subset A(t)$ represents a cohort with age a. Let the set of individuals in this cohort be $N_a(t) \subset N(t)$, and let the number of individuals in such a set be defined as $|N_a(t)|$. The probability a witnessed innovation is produced by an individual of age a is

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

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

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

Intuitively, the probability an innovation comes from a person of age a is just the relative innovation potential across cohorts weighted by the population age density or, equivalently, the population age density weighted by the innovation potential. Shifts in the age distribution of great innovation, as seen in figure 2, are thus seen to be driven by shifts in either the population age distribution, $p_a(t)$, or the average innovation potential of various age groups, $\bar{x}_a(t)$.[4]

Equation (2) is the central vehicle for the maximum likelihood estimation to follow. 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 submodel is presented in the next section.

A. A Model of Life Cycle Productivity

In choosing an appropriate modeling strategy for life cycle innovation potential, it is helpful to first consider the empirical literature on creative careers, which suggests the following pattern (e.g., Lehman, 1953; Bloom, 1985; Simonton, 1991; Stephan & Levin, 1993). 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, following and S-curve, to a peak in the late 30s or early 40s. Finally, innovative output declines slowly through later years, following a declining S-curve.[5]

Given this pattern, consider a simple model,

$$\bar{x}_a = L_1(a) L_2(a), \tag{3}$$

where $L_1(a)$ captures early life cycle effects and $L_2(a)$ captures late life cycle effects. This model can be given a specific basis that draws naturally on the life cycle considerations above. In particular, let an innovator i start life with a stochastic duration of education, e_i, during which she does not innovate. Additionally, 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, or other factor that influences innovative ability. The innovation potential of individual i as a function of her 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.

Employing a law of large numbers, we can write the cohort average innovation potential as $\bar{x}_a \xrightarrow{p} E[I(a_i \geq e_i) g(a_i; z_i)]$. Assuming additionally that e_i and z_i are independent, this expectation simplifies to $\bar{x}_a \xrightarrow{p} L_1(a) L_2(a)$, where $L_1(a) = \Pr(a_i \geq e_i)$ captures the average year in which the career starts and $L_2(a) = E[g(a_i; z_i)]$ captures the path of life cycle productivity once educated.

To estimate $L_1(a)$, we assume first that e_i is distributed logistically within cohorts,

$$L_1(a) = \frac{1}{1 + e^{-(a-\mu)/\omega}}, \tag{4}$$

where $\mu = E[e_i]$ is the average age at which the career begins and ω is a variance parameter. A logistic specification seems reasonable as it is parametrically simple and flexible, and captures the S shape one sees in early life cycle output.

[3] 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.

[4] Several other useful points can now be made. First, the stochastic process represented in equation (2) can produce innovators with a large variance in age, as demonstrated in figure 1. Second, any presumption that the innovators' upward age trends are driven mechanically by increasing life expectancy may be misleading if the innovation potential, $\bar{x}_a(t)$, of those in their later years is low—if only because people retire. 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.

[5] While laboratory experiments do suggest that creative thinking becomes more difficult with age (Reese et al., 2001), the decline in innovative output at later ages may largely be due to declining effort, which a range of theories have been proposed to explain (see Simonton, 1996, for a review).

FIGURE 3.—MODEL OF INNOVATION POTENTIAL OVER THE LIFE CYCLE

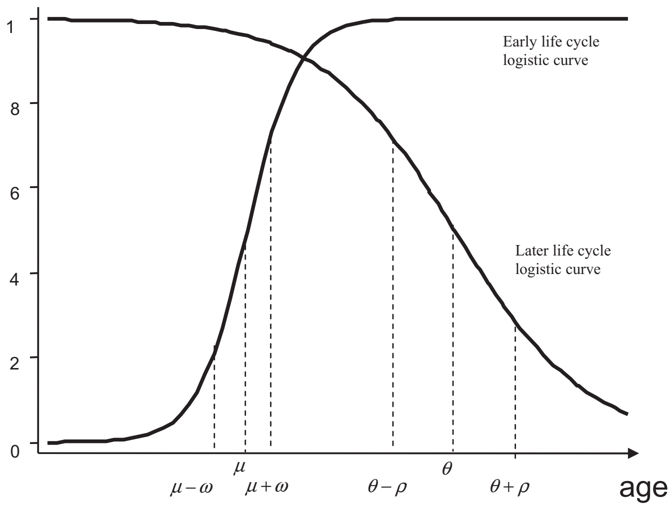


Figure 3 presents a graph to clarify this logistic specification and the meaning of the parameters.

To estimate $L_2(a)$, we assume a second logistic curve with parameters θ and ρ,

$$L_2(a) = 1 - \frac{1}{1 + e^{-(a-\theta)/\rho}}. \tag{5}$$

With equation (3) and its subcomponents (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. Moreover, by articulating specific underlying models for both the front end of the life cycle, (4), and its back end, (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 μ, 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,

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

Shifts in the variance parameter can be modeled similarly. The main estimation below will allow 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,

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

For example, shifts that increasingly favor experience over raw ability may increase later-life-innovation potential. Meanwhile, improved health technology may lead to clearer thinking or increased physical stamina, or both, while, alternatively, rising incomes could encourage earlier retirement and a decline in average innovation potential among older innovators. In the estimation I will constrain $L_2(30) > 0.9$ 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 twentieth century.

Taken together, equation (2) and the submodel of innovation potential given by equations (3) to (7) integrate demographic effects with life cycle productivity considerations. The following section discusses the data used to estimate $p_a(t)$. We then present the central results.

B. 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 twentieth 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 observations on their own.[6]

The population age densities are calculated from large microsamples of the U.S. census. With these microsamples, one can determine the age distribution for the entire national population, as well as for subgroups of active workers and professional scientists and engineers. The scientist and engineer data 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 change somewhat across time, raising concerns that shifts in the age distribution for this subgroup 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 subsample.)

C. Results

Table 2 presents the maximum likelihood estimates for shifts in innovation potential. Columns 1 through 3 allow for linear trends in the mean parameters, which we take as the main specification. There are two striking results. First, there has been a large shift in life cycle innovation potential, 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

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

Table 2.—Maximum Likelihood Estimation of Life Cycle Innovation Potential

		(1)	(2)	(3)	(4)	(5)	(6)
Early life cycle logistic curve	μ_0	24.0	23.3	23.4	22.6	22.8	25.8
	Initial mean, in years	(2.05)	(2.30)	(1.95)	(2.84)	(2.84)	(1.07)
	μ_1	7.76	8.29	8.32	8.91	10.29	5.32
	trend, in years/century	(3.22)	(3.49)	(2.71)	(5.23)	(4.56)	(1.97)
		[.016]	[.018]	[.002]	[.088]	[.024]	[.007]
	ω	2.40	2.47	2.38	2.43	2.30	2.21
	Variance parameter	(0.37)	(0.43)	(0.40)	(0.73)	(0.46)	(0.20)
Later life cycle logistic curve	θ_0	45.5	46.6	50.0	53.7	43.7	46.9
	initial mean, in years	(2.36)	(2.21)	(3.90	(8.16)	(3.54)	(3.70)
	θ_1	−0.00e-03	−0.00e-03	0.14e-03	−4.61	0.00e-03	−0.71
	Trend, in years/century	(8.63e-03)	(6.70e-03)	(6.88e-03)	(12.1)	(1.28e-03)	(5.17)
	ρ	7.05	7.54	9.12	6.64	6.24	7.38
	Variance parameter	(1.10)	(0.91)	(1.60)	(2.72)	(0.99)	(7.40)
Data	Population	U.S. population	Active workers	Scientists and engineers	U.S. population	U.S. population	U.S. population
	Inventor type	All	All	All	Technology almanacs	Nobel Prize	All
	Nationality	U.S. born	U.S. born	U.S. born	U.S. born	U.S. born	All
	Number of invention observations	294	294	294	127	181	738
Log likelihood		−1,050.9	−1,053.0	−1,056.7	−463.6	−633.6	−2,641.2

Notes: This table presents maximum likelihood estimates for the econometric model of life cycle innovation potential, as derived in section III and described visually in figure 3. These estimates allow for trends in the mean parameters for the early and late life cycle logistic curves. Table 3 additionally allows for trends in the variance parameters. 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 mean are given in brackets.

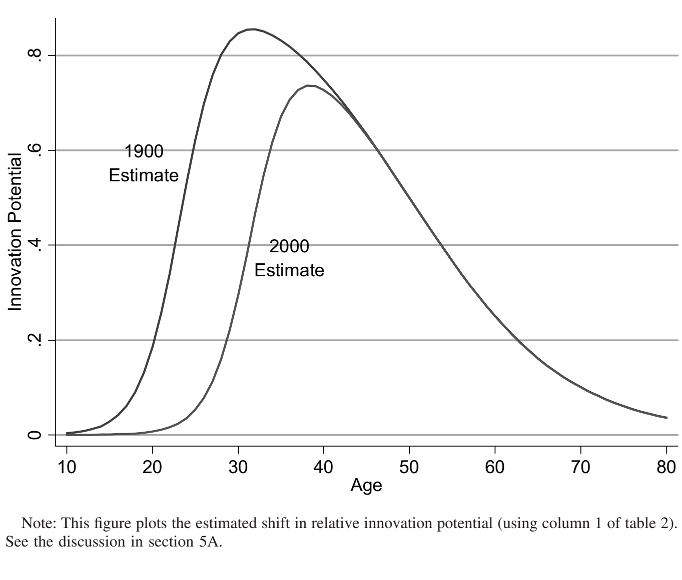
which innovators begin making active contributions increased by about eight years over the course of the twentieth century, rising from a mean age of about 23 in 1900 to approximately 31 in the year 2000. These results are robust to the choice of population data. Meanwhile, there is little shift in innovation potential in middle age or beyond. The estimates tend to show most modest and in all cases highly insignificant movements.

Figure 4 compares the estimated life cycle curves for the year 1900 and the year 2000, using specification (3). 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 other things equal, lifetime innovation potential has declined. This point will be further discussed in section V.

One may also now decompose the trends of section II into demographic and productivity components. Holding population distribution fixed using 1950 data, the productivity shift in table 2 implies an approximately five-year increase in the mean age of great achievement. Meanwhile, holding innovation potential fixed at its 1950 estimate, the aging population suggests an approximately three-year increase in mean age. Hence productivity shifts account for about 60% of the eight-year age trend seen in section II, while the aging population captures 40%.[7]

These basic results are robust to a number of alternative specifications. The specifications in table 3 allow additional linear trends in the variance parameters of the logistic distributions. We see that 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. Columns 4 and 5 of table 2 further consider the U.S. Nobel Prize and U.S. great inventor data sets separately, showing similar results in each case. 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.[8] Possible selection concerns regarding Nobel Prizes are unlikely to be important here, however, mainly because the great inventor data set, which simply lists the great technological achievements in each given year, ap-

Figure 4.—Maximum Likelihood Estimates for the Potential to Produce Great Innovations as a Function of Age



Note: This figure plots the estimated shift in relative innovation potential (using column 1 of table 2). See the discussion in section 5A.

[7] Moreover, much of the demographic effect comes in the first two-thirds of the century; more recently, the baby boom significantly attenuated the aging of the working population and even reversed it in the 1970s and 1980s.

[8] For example, an increasing bias toward lifetime achievement could have this effect.

TABLE 3.—MAXIMUM LIKELIHOOD ESTIMATION: FURTHER SPECIFICATIONS

		(1)	(2)	(3)	(4)
Early life cycle logistic curve	μ_0 Initial mean, in years	24.8 (1.94)	24.6 (1.84)	24.5 (2.36)	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]	5.89 (2.12) [.005]
	ω_0 Initial variance parameter	2.86 (0.83)	3.42 (1.01)	3.08 (1.14)	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)	46.8 (3.79)
	θ_1 Trend, in years/century	0.99 (4.06)	2.29 (3.28)	2.79 (6.79)	−0.93 (2.53)
	ρ_0 Initial variance parameter	6.81 (2.06)	6.97 (1.47)	8.42 (2.13)	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
	Nationality	U.S. born	U.S. born	U.S. born	All
	Number of invention observations	294	294	294	738
Log likelihood		−1,050.8	−1,052.6	−1,056.4	−2,641.1

Note: This table presents maximum likelihood estimates for the econometric model of life cycle innovation potential, as derived in section III and described visually in figure 3. These estimates allow for trends in mean and variance parameters for both the early and late life cycle logistic curves. Table 2 allows for trends in the mean parameters only, providing a more parsimonious model. 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 mean are given in brackets.

pears more immune to these kinds of selection biases and yet has produced very similar results.[9]

IV. Inside the Early Life Cycle

Age at great invention has trended upward by approximately six years over the course of the twentieth century. This trend is not simply due to an aging population but reflects a substantial change in the life cycle productivity of innovators. Furthermore, the maximum likelihood estimates focus interpretations on effects limited to the young. Explanations must confront not a general aging effect but a specific, substantial delay at the beginning of the life cycle.[10]

A natural and intriguing hypothesis for the rising delay in the early life cycle is the possibility that training time has increased. A viewpoint that emphasizes training would build on two claims: that training is an important preliminary input to the innovative career and that variations in training duration can help explain the age-invention relationship.

This section focuses on the early life cycle and the role of training to open up the black box of age and invention. I undertake three analyses. The first analysis looks directly at evidence from PhD age and shows that PhD age increases substantially over the twentieth century. The second analysis harnesses world wars, as exogenous interruptions to the young career, to test the basic idea that training is an important preliminary input to innovation. I show that while the world wars do not explain the twentieth century's age trend, they do indicate the unavoidable nature of training: lost years of training appear to be made up after the war. The final analysis explores cross-field, cross-time variation. I show that variations in training duration predict variations in age at great invention, and I close by discussing a perspective in which shifts in foundational knowledge explain major training and achievement age patterns within fields and over time.

[9] The Nobel Prize and great inventors data sets have extremely similar age distributions (figure 1) and extremely similar mean trends (table 1). Table 2 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 2 and 3 show shifts in innovation potential at the beginning of the life cycle, and not at the end, which is inconsistent with selection stories based on longevity or increasing favoritism for lifetime achievement.

[10] 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, 2004a, 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.

TABLE 4.—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 100s)	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

Note: All specifications consider trends in the age at highest degree among Nobel Prize winners. Columns 1, 3, and 4 consider the full sample for whom degree age can be obtained. Column 2 considers the subsample for whom the highest degree was a PhD (as opposed to a master's or bachelor's degree). Column 5 considers only those where the highest degree came in the United States. Field and country of degree fixed effects are included or excluded as indicated. The coefficient on year of highest degree gives the 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. Results are similar when allowing for separate country of degree effects by field or when using country of birth fixed effects instead of country of degree.

**Significance at a 95% confidence level. ***Significance at a 99% confidence level.

A. *Age at Highest Degree*

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.[11] While the PhD is an institution that only approximately captures the end of the training phase and the beginning of the primary research phase, it is also the most obvious delimiter between these phases. I consider here the basic trend.

For 93% 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. I analyze trends in the age by running the following regression,

$$a_i^D = \alpha + \beta t_i^D + \gamma X_f + \varepsilon_i, \qquad (8)$$

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.

The results are presented in table 4. We see that Nobel Prize winners complete their formal education at substantially older ages today than they did a century ago. There is an upward age trend of approximately four years per century, and the trend is robust across specifications. This result suggests that training duration may be intimately related to the drop in innovative output in the early life cycle.[12] As we will see, patterns in the age of highest degree also inform substantially more detailed variation in the data.

B. *War Interruptions*

World Wars I and II created interruptions in many careers. For certain cohorts, the interruption of the war was felt largely in the training phase. This provides a natural experiment to investigate the role of training in the early life cycle.

In particular, in every year, there are people who have completed their undergraduate degree but not their graduate degree. Every person is at some point between degrees, so by drawing any year at random, we draw a sample of people with similar innate characteristics on average. We can then ask whether individuals who happened to be between degrees at the outset of world war, in 1914 or 1939, as opposed to being between degrees in other years, experienced unusual delays in completing their training and in their ensuing innovative careers. For 68% of the Nobel Prize winners, it was possible to collect the year of the undergraduate degree and hence identify individuals who are between degrees. I run regressions of the form

$$y_i = \alpha + \gamma_1 WW1 + \gamma_2 WW2 + \beta t_i + \gamma X_f + \varepsilon_i, \qquad (9)$$

where y_i is the outcome variable of interest: the age at highest degree, the number of years between the undergraduate and graduate degree, or the age at great achievement. The variables $WW1$ and $WW2$ are dummies equal to 1 if the individual happened to be between degrees at the outset of the indicated war. The control t_i captures background trends in the dependent variables, and X_f includes field and country

[11] Indeed, several studies have documented upward trends in educational attainment among the general population of scientists. For example, the age at which individuals complete their doctorates rose generally across all major fields in a study of the 1967–1986 period (National Research Council, 1990). The duration of study for the doctorate as well as the frequency and duration of postdoctorates has risen across the life sciences since the 1960s (Tilghman et al., 1998). A study of electrical engineering over the course of the twentieth century details a long-standing upward trend in educational attainment, from an initial propensity for bachelor degrees as the educational capstone to a world where PhDs are common (Terman, 1998).

[12] Interestingly, while the increase in training age is large, it accounts for only half the shift seen in the maximum likelihood estimates, although it is within those estimates' confidence intervals. Institutional variations in PhD requirements may complicate further interpretations. For example, the country fixed effects in (8) 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. The well-known rise of postdoctorates (e.g., Tilghman et al., 1998) or increased "on-the-job training," or both, could suggest further extension of the training phase in ways not captured by charting ages at PhD.

TABLE 5.—WAR INTERRUPTIONS

	PhD Age		Lag between Degrees		Achievement Age	
	(1)	(2)	(3)	(4)	(5)	(6)
World War II	1.941***	2.034***	2.835***	2.797***	2.286**	2.727**
	(0.558)	(0.522)	(0.544)	(0.517)	(1.148)	(1.187)
World War I	2.339**	1.911*	2.790**	2.194**	0.374	0.335
	(1.149)	(1.104)	(1.125)	(1.096)	(2.723)	(2.830)
Year of doctorate	0.037***	0.042***	0.021***	0.027***	—	—
	(0.006)	(0.006)	(0.008)	(0.008)		
Year of great achievement	—	—	—	—	0.068***	0.079***
					(0.014)	(0.016)
Missing educational observation	−0.634*	−1.098***	—	—	2.111**	0.452
	(0.366)	(0.403)			(0.834)	(0.985)
Field fixed effects	No	Yes	No	Yes	No	Yes
Country of birth fixed effects	No	Yes	No	Yes	No	Yes
Number of observations	508	508	348	348	544	544
R^2	0.12	0.37	0.10	0.36	0.05	0.20

Note: This table considers the effect of world war on PhD age (columns 1 and 2), the time lag between undergraduate and graduate degrees (columns 3 and 4), and the age at great achievement (columns 5 and 6). Results are OLS with standard errors in parentheses. WW2 and WW1 are dummies equal to 1 for individuals who happened to be between their undergraduate and graduate degrees at the outset of the indicated war (1939 for World War II and 1914 for World War I).
*Significance at a 90% confidence level. **Significance at a 95% confidence level. ***Significance at a 99% confidence level.

fixed effects and dummies for cases where educational data are not observed.[13]

The results are presented in table 5. We see that both world wars resulted in a two-year increase in the age at PhD for individuals who happened to be caught between degrees. Related, there is a two- to three-year increase in the number of years between the undergraduate and graduate degrees. Both suggest that interruptions to training must be made up. These early life cycle delays can be further associated with increased age at great achievement—two to three years for World War II, although there is little effect for World War I. Given that innovation potential remains high through middle age, it is not clear that a small number of early life cycle interruptions will show an increase in the mean age of great innovation.[14] One expects, more precisely, a decline in innovative potential at younger ages. In fact, one can show that experiencing either war when aged 20 to 25 implies a substantially reduced probability of producing innovations when aged 25 to 30, even though the wars were over.[15]

In short, great minds do not magically arrive at high innovation potential at a certain age; rather, their behavior in their early life cycle informs their ensuing innovative output. Interruptions during the training phase create delays to their education and their achievements, suggesting that training is an important preliminary input to the innovative career.[16]

C. *Knowledge Accumulation and Revolution*

We can go further in understanding any training-achievement nexus by harnessing additional variation in the data. Figure 5 plots the evolution of age at great invention separately for the four Nobel subfields (dark lines, left axis). These are nonparametric regressions, so that the patterns are seen without imposing a functional form. The age at PhD is separately plotted (gray lines, right axis), as are 95% confidence intervals.[17]

We see that PhD and achievement age tend to follow remarkably similar dynamics within fields.[18] The shared dynamics are most apparent in the hard sciences—physics, chemistry, and medicine—and less so in economics, although this case is obscured by outliers.[19] Most strikingly, both achievement and PhD age in physics experienced a unique decline in the early twentieth century. This unusual feature, beyond reinforcing the relationship between training and achievement age, may also inform basic theories for the underlying dynamics and differences across fields.

[13] One can also consider country-specific war entry dates. The major potential change would be to the United States, which entered World War I only in 1917 and World War II only in 1941 and encompassed 40% of the relevant sample of individuals who were between degrees at the outset of war. In practice, such an adjustment has little effect on the sample and strengthens the results slightly.

[14] The weak World War I result may reflect low power, since only ten individuals were caught between degrees in 1914 (there were 66 observations in the World War II case).

[15] These results are available from the author on request.

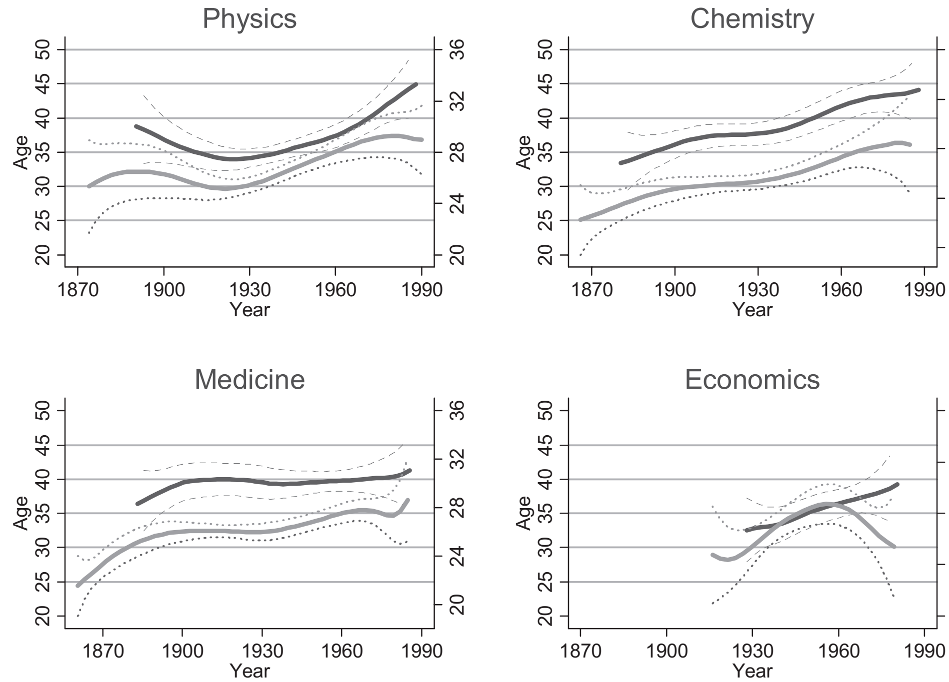
[16] Note also that the world wars do not drive the overall twentieth-century trends. While wars interrupted certain careers, table 5 shows that the background trends in achievement age and PhD age are essentially unchanged.

[17] These plots are the results of Fan regressions with a quartic kernel, 25% bandwidth, and bootstrapped standard errors (Fan, 1992; Deaton, 1997).

[18] It can also be demonstrated in regression format that age at highest degree is an economically and statistically significant predictor of achievement age (results available from the author on request).

[19] The movement in mean PhD age for economics is less predictive, although the confidence intervals are wide. There are only 53 observations for economics, which limits the inference, and the sharp rise to 1960 is driven largely by a few significant outliers (Allais, Coase, and Stone, who received their PhDs at ages 38, 41, and 44, respectively, long after their first academic appointments and years of successful research). These outliers mask a continuing decline in the completion of formal education at young ages. Prior to 1950, 32% of eventual Nobel Prize winners in Economics completed their highest degree prior to age 25. After 1950, only 7% completed their highest degree by age 25. Similar, large declines in PhD completion by the very young are also seen in physics, chemistry, and medicine.

FIGURE 5.—AGES AT PHD AND ACHIEVEMENT OVER TIME, BY FIELD



Note: This figure shows, by Nobel field, nonparametric Fan regressions estimating shifts over time in the mean age at great invention (dark lines, left axes) and mean age at PhD (gray lines, right axes).

An intriguing side effect of innovation is the possibility that new ideas impose an increasing educational burden on future innovators. If the set of foundational ideas expands, training time may expand, making innovators less productive in their early life cycle. Progress in science need not require expansions in foundational ideas, however. New ideas sometimes serve not as extensions or refinements but rather as revolutions, leading to contractions in the knowledge space that ease training requirements. Whether scientific progress is fundamentally cumulative or revolutionary in nature is an empirical question—and one much debated by historians of science. Thomas Kuhn (1962) distinguished between periods of "normal" science (accumulation) and periods of "paradigm" shifts (revolution), with early-twentieth-century physics as his quintessential example of the latter.

Jones and Weinberg (2008) consider the U-shaped age relationship in physics through the lens of accumulation and revolution. That analysis details a remarkable coincidence between the quantum mechanics revolution and the unusual behavior in the age data. Many historians chart the specific period from 1900 to 1927 as the time when the entire worldview of physics changed (Kuhn, 1962; Jammer, 1966; Galison, Gordin, & Kaiser, 2001).[20] From a training point of view, physicists found themselves in the early 1900s wrestling with a new, limited set of empirical puzzles and the failure of existing theory, allowing young minds to achieve the research frontier relatively easily. The firm establishment of quantum mechanics in the late 1920s then led back toward normal science, a long period of accumulation and refinement that has continued since. In the data, the ages among physicists during the early twentieth century reach a minimum in the late 1920s, coincident with the establishment of quantum mechanics, and have risen since then. This age pattern is not seen in other fields, which did not experience such revolution.

In sum, training dynamics appear to usefully predict variations in the age-invention relationship. Digging deeper, the distinction between knowledge accumulation and revolution can provide a useful lens through which to view major training and age-invention patterns in the data, an interpretation that is further explored below.

V. Interpretations and Implications

The evidence in this paper points to early life cycle effects and a particular role for training in understanding the age-invention relationship. In this section, I explore more formal reasoning for innovators' training decisions. I clarify the potential explanatory power of several hypotheses and close by considering implications of the patterns uncovered in this paper.

Consider the following simple model. An innovator is born with no knowledge but is endowed with time. Innovators can invest in training, but knowledge acquisition

[20] Starting with Planck's idea in 1900 that radiation comes in discrete energy packets (quanta), the work of many others followed, breaking down classical physics and reaching firm footing only with the advent of a consistent "quantum mechanics," developed by Schrödinger, Heisenberg, and others in the mid- to late 1920s.

delays the active production of new ideas. Innovators compare the return to active production with the return to further training, whatever the benefits that training brings.

A reasonable specification, especially for highly motivated innovators, is that educational attainment is chosen to maximize one's lifetime research contribution. In particular, the choice problem is[21]

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

where $f(E)$ represents the value of education to innovative output and $g(a) > 0$ represents other sources of the individual's innovation potential (e.g., natural ability, health) 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 influences their ultimate productivity, where $f'(E) > 0$ and $f(0) = 0$.

The first-order condition for this problem is[22]

$$f'(E) \int_E^T g(a) da = f(E)\, g(E), \tag{10}$$

which clarifies the central trade-off. Greater education brings a benefit: the incremental effect on innovative output, $f'(E)$, weighted by the innovative ability that remains over a lifetime. But it also brings an opportunity cost, $f(E)\, g(E)$—the current innovation potential forgone.

Consider the implications of expansions in foundational knowledge. If knowledge accumulates across generations, then the training curve, $f(E)$, can shift rightward, causing E^* to rise. This is a natural outcome if there is a set of preliminary skills one must master to reach the frontier in a field, and the set of these skills expands. For example, imagine that innovators need to know A, B, C to attack D, the frontier. Following success at D, however, ensuing cohorts may need to know A, B, C, D to attack the new frontier. In its simplest form, one imagines $f(E)$ as a step function with a step up after training for a particular number of years, with this number increasing when knowledge accumulates.

We may also consider the influence of increased life expectancy (T). Longer life expectancy, by increasing $\int_E^T g(a)da$, can lead endogenously to increased training time by providing a longer period over which to reap the gains of further training. However, there are several reasons to believe such a mechanism may not provide an adequate explanation for the empirical patterns. First, mean life expectancy at age 10 was already above 60 in 1900, while sections II and III show that innovation potential is modest above age 60, so that adding years of life beyond this age would have at most mild effects on the optimization.[23] Related, even modest discounting would substantially limit the effect of gains felt from age 35 beyond the end of training on the marginal training decision. Finally, common life expectancy changes cannot explain the cross-field and cross-time variation explored in section IV, such as the unique behavior of physics.

Another set of explanations involves institutional or sociological effects on the training curve, $f(E)$. A plausible story might involve signaling. If establishing a reputation is a prerequisite for grant-based research, and research has become more grant based (expensive) over time, then extensions of formal training or apprenticeship may serve to signal reputation—a different interpretation of the educational return. While this force may be operating, it also suffers as a general explanation; for example, figure 5 shows a substantial age increase at great achievement among the economists alone, and these prize winners have had little need for large grants.

Other forces might also be considered within this framework, none of which are mutually exclusive and many may be operating.[24] At the same time, reasoning about shifts in foundational knowledge can provide a parsimonious explanation for the major patterns in this paper. A difficulty in directly establishing this thesis is the difficulty in directly measuring the stock of foundational knowledge—the "distance to the frontier." But we can go further by considering additional, indirect evidence. To see this, we further articulate the training decision.

If we imagine that the innovators must reach the frontier to do innovative work, the innovator still has a decision over the breadth of expertise along the frontier. To fix ideas simply, let

$$E = bD,$$

where D measures distance to the frontier and b measures the innovator's breadth along the frontier. Total training time is then determined in part by the capacity for specialization. For example, the chemist may choose to come a frontier expert in the synthesis of metal alloys (narrower b),

[21] More generally, this objective function captures the maximization of fame or income, so long as fame and income are monotonic functions of lifetime innovative output. One can incorporate time discounting in the $g(a)$ function.

[22] This maximization problem has an interior solution, which is seen by noting that $\int_E^T f(E)\, g(a)da = 0$ if $E = 0$ or $E = T$, and that $\int_E^T f(E)\, g(a)da > 0$ if $E \in (0, T)$.

[23] The life expectancy data are for white males in the United States (Department of Health and Human Services, 2004). The life expectancy of innovators, who have several advantages, would likely be higher still. In fact, the age at death for great innovators in the sample who were born between 1900 and 1910 averages 80.

[24] As a final example, one may imagine declining educational efficiency, shifting the training curve rightward. Such a decline in training efficiency would have to be very large to explain the estimated eight-year delay, from age 23 to 31, in achieving high innovation potential, and one might imagine that educational efficiency increases with technology, rather than decreases, suggesting some skepticism for this particular point of view. The biographies of Nobel Prize winners suggest a degree of focus that is not commensurate with slow or undirected training.

or both alloys and organics (wider b).[25] This specialization margin presents a useful empirical application. Inferences about knowledge accumulation, D, based only on investigations of training time, E, may be clouded by other possible forces, as discussed above. But we might infer knowledge accumulation more definitively by observing E in combination with some measure of breadth of expertise, b. That is, $D = E/b$. If people spend longer in training (more E) and yet come out the other end more specialized (less b), then the distance to the frontier has increased.

This reasoning is explored in Jones (2009), which studies "ordinary" inventors, looking at all U.S. patents in the 1975–2000 period. There are two key results. First, the age at first patent is rising at a rate of six years per century. Age at first patent provides an outcome-based measure to delimit the training and research phases. Remarkably this estimate for the extending training period is extremely similar to the trends noted in this paper. Second, proxy measures for specialization show increased specialization across the full range of technological fields. One proxy measure is research collaboration in patenting, measured as team size, which is increasing at over 10% per decade.[26] A more direct measure of specialization considers the probability that an individual switches technological areas between consecutive patents. Jones (2005) shows that the probability of switching technological areas is substantially declining with time. These analyses indicate that training time, E, is rising, while measures of breadth, b, are simultaneously declining. It is then difficult to escape the conclusion that the distance to the knowledge frontier is rising. This evidence also acts to confirm, with "ordinary" invention, the rising age pattern found among the great minds.

A. *Implications*

Shifts in life cycle research productivity can have diverse implications, from the efficient targeting of grants to the design of tenure processes and the timing of child rearing. Here I emphasize two aggregate implications, for core issues in economic growth and scientific progress, that are suggested by the particular life cycle shifts identified in this paper.

First, other things equal, the shorter the period that innovators spend innovating, the lower their output is 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. 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 R&D workers in terms of both patent counts and productivity growth (Machlup, 1962; Evenson, 1991; Jones 1995a; Kortum, 1997). Simple calculations from aggregate data 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 twentieth century.[27] This paper provides microevidence that can explain part of that trend. Other things equal, the estimates of section III indicate a 30% drop in the lifetime innovation potential over the century, or nearly half of the overall decline in individual research productivity.[28]

Second, the facts in this paper can also inform basic debates about the nature of scientific progress. A core question in the history of science is whether scientific progress happens primarily through accumulation and refinement of ideas or through radical, "Kuhnian" revolution. These debates are traditionally enjoined through historical argument, such as Kuhn's seminal analysis of physics. The age data in this paper can provide, alternatively, a data-driven test. If the aging phenomena detailed in this paper suggest accumulations of knowledge, as discussed above, then Kuhnian revolutions appear rare.

VI. Conclusion

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 inno-

[25] The capacity for specialization will likely be imperfect. For example, the chemist may specialize more or less on certain types of synthesis but, regardless, must understand theories of valence and molecular structure. The knowledge space is thus a mix of common foundational ideas and more specialized ideas. When the stock of knowledge grows, we might imagine that innovators respond on both dimensions, partly by increasing their training time and partly by increasing specialization (though the argument in the text is empirical and does not require this assumption).

[26] General upward trends in collaboration are also found in journal publications (e.g., Wuchty, Jones, & Uzzi, 2007).

[27] 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 sixfold 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 A/L_R, or 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 $(A^{2000}/A^{1900})/(L_R^{2000}/L_R^{1900}) = 6/19$ (32%) of what it was in 1900.

[28] 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 endogenous response to an increased educational burden. Jones (2005) provides theory and empirical support for this latter mechanism.

vating, the less her lifetime output. The estimates point to a 30% decline in life cycle innovation potential over the twentieth century.

This paper further explores the role of training in understanding the early life cycle dynamics, investigating evidence from world wars, age at PhD, and cross-time variation in training and achievement ages. These analyses point toward the training phase as a key explanation for the trends we see. Yet 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 trade-offs 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.

REFERENCES

Bloom, Benjamin S., "Generalizations about Talent Development," in Benjamin S. Bloom (Ed.), *Developing Talent in Young People* (New York: Ballantine Books, 1985).

Deaton, Angus, *The Analysis of Household Surveys* (Baltimore, MD: Johns Hopkins University Press, 1997).

Ericsson, Karl A., and Andreas C. Lehmann, "Expert and Exceptional Performance: Evidence of Maximal Adaptation to Task Constraints," *Annual Review of Psychology* 47 (1996), 273–305.

Evenson, Robert E., "Patent Data by Industry: Evidence for Invention Potential Exhaustion?" Yale Economic Growth Center discussion paper no. 620 (January 1991).

Fan, Jianqing, "Design-Adaptive Nonparametric Regression," *Journal of the American Statistical Association* 87 (1992), 998–1004.

Galenson, David W., "A Portrait of the Artist as a Young or Old Innovator: Measuring the Careers of Modern Novelists," NBER working paper no. 10213 (2004a).

——— "A Portrait of the Artist as a Very Young or Very Old Innovator: Creativity at the Extremes of the Lifecycle," NBER working paper no. 10515 (2004b).

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

Galison, Peter, Michael Gordin, and David Kaiser (Eds.) *The History of Modern Physical Science in the Twentieth Century, Quantum Histories,* Vol. 4 (New York: Routledge, 2001).

Jammer, Max, *The Conceptual Development of Quantum Mechanics* (New York: McGraw-Hill, 1966).

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

Jones, Benjamin F., and Bruce Weinberg, "Age and Scientific Creativity," Northwestern University mimeograph (2008).

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

——— "Time Series Tests of Endogenous Growth Models," *Quarterly Journal of Economics* 110 (1995b), 495–525.

Kortum, Samuel S., "Research, Patenting, and Technological Change," *Econometrica* 65 (1997), 1389–1419.

Kuhn, Thomas S., *The Structure of Scientific Revolutions* (Chicago: University of Chicago Press, 1962).

Lehman, Harvey C., *Age and Achievement* (Princeton, NJ: Princeton University Press, 1953).

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

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).

National Science Foundation, *Industrial Research and Development Information System,* table H-19, www.nsf.gov/sbe/srs/iris/start.cfm (2005).

Reese, Hayne W., Liang-Jei Lee, Stanley H. Cohen, and James M. Puckett, "Effects of Intellectual Variables, Age, and Gender on Divergent Thinking in Adulthood," *International Journal of Behavioral Development* 25:6 (2001), 491–500.

Simonton, Dean K., *Scientific Genius. A Psychology of Science* (Cambridge: Cambridge University Press, 1988).

——— "Age and Outstanding Achievement: What Do We Know after a Century of Research?" *Psychological Bulletin* 104 (1988), 251–267.

——— "Career Landmarks in Science: Individual Differences and Interdisciplinary Contrasts," *Developmental Psychology* 27 (1991), 119–130.

——— "Creativity," in *The Encyclopedia of Gerontology* (San Diego, CA: Academic Press, 1996).

Stephan, Paula, and Sharon Levin, "Age and the Nobel Prize Revisited," *Scientometrics* 28 (1993), 387–399.

Terman, Frederick E., "A Brief History of Electrical Engineering Education," *Proceedings of the IEEE* 86:8 (1998), 1792–1800.

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

U.S. Department of Health and Human Services (2004, November 10). *National Vital Statistics Report.*

Wuchty, Stefan, Benjamin F. Jones, and Brian Uzzi, "The Increasing Dominance of Teams in the Production of Knowledge," *Science* 316 (2007), 1036–1039.

Zuckerman, Harriet, and Robert Merton, "Age, Aging, and Age Structure in Science" (pp. 497–559), in Robert Merton (Ed.), *The Sociology of Science* (Chicago: University of Chicago Press, 1973).

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.

A. Data on Great Innovators

The official Web site of the Nobel Foundation, nobelprize.org, provides 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 all 547 Nobel Prize winners between 1901 and 2003 and the period of key research for all but three of these. In practice, the data identify a single year of great achievement—the year of success—for 75% of the cases. For the remainder, the Nobel citation appears to encompass multiple subcontributions, in which cases early and late dates of achievement were collected. In these cases, the estimations in the text use the middle year 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. The year or period of key research was usually straightforward to ascertain through the Nobel Foundations biographies, but in cases where these did not accurately identify the year or period of key research, other sources were consulted. The primary printed materials used were:

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

Daintith, J., and D. Gjertsen, *The Grolier Library of Science Biographies,* vols. 1–10 (Danbury, CT: Grolier Educational, 1996).

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

McMurray, E. J., J. K. Kosek, and R. M. Valade, *Notable Twentieth-Century Scientists,* vols. 1–4 (Detroit: Gale Research, 1995).

Williams, T. I. (Ed.) *Biographical Dictionary of Scientists* (New York: Wiley, 1974).

Data on great inventors were collected from two technological almanacs, which indicate a list of notable technological advances in each year and typically provide the birth date and birthplace of the innovator responsible. Together the almanacs provided a set of 286 inventors in the twentieth century. These inventors include Reginald Fessenden (first audio radio transmission), Leo Baekeland (Bakelite, a breakthrough plastic), Georges Jean Marie Darrieus (vertical axis wind turbine), Frank Whittle (jet engine), Virginia Apgar (newborn health rating), and Steve Wozniak (early personal computer). The almanacs consulted were:

Bunch, B., and A. Hellemans, *The Timetables of Technology* (New York: Simon and Schuster, 1993).
Ochoa, G., and M. Corey, *The Timeline Book of Science* (New York: Ballantine Books, 1995).

Fields of research are given in both sources. I condense their categorizations into nine fields: communication, computers and electronics, energy, food and agriculture, materials, medicine, tools and 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.

B. Data on Population Age Distribution

One percent and 5% microsamples of the U.S. census are available electronically through IPUMS, the Integrated Public Use Microdata Series, which is maintained by the University of Minnesota:

Ruggles Steven, et al., *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 microsample provided data on approximately 100,000 individuals. The largest sample used was for the 2000 census, whose microsample provided data on approximately 2.8 million individuals. Existing census research available on the Web site (www.ipums.umn.edu/usa/chapter3/chapter3.html) indicates that these microsamples 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 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 past twelve 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 two years. Second, I impute the number of innovators still in training (those aged 15–29) based on the number of employed workers 10 years later who are 10 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.