Age Dynamics in Scientific Creativity

Benjamin F. Jones[1,3] and Bruce A. Weinberg[2,3]

1. Kellogg School of Management, Northwestern University, Evanston IL 50208 USA
2. Department of Economics, Ohio State University, Columbus OH 43210 USA
3. National Bureau of Economic Research, Cambridge, MA 02138 USA

August 2011

Data on Nobel Laureates show that the age-creativity relationship varies substantially more over time than across fields. The age dynamics within fields closely mirror field-specific shifts in (i) training patterns and (ii) the prevalence of theoretical contributions. These dynamics are especially pronounced in physics and coincide with the emergence of quantum mechanics. Taken together, these findings demonstrate fundamental shifts in the life-cycle of research productivity, inform theories of the age-creativity relationship, and provide observable predictors for the age at which great achievements are made.

Introduction

At what age do scientists tend to produce great ideas? Focusing on great scientific achievements of the 20th century, this paper shows that the age-creativity relationship demonstrates much greater variation over time than across fields. Moreover, field-specific dynamics in the age-creativity relationship are closely associated with variation in other field-specific characteristics, including the prevalence of theoretical contributions, educational duration, and citation patterns. These dynamics are especially pronounced in physics during the 1920s and 1930s, when quantum mechanics was developing. Thus, while the iconic image of the young, great mind making critical breakthroughs was a good description of physics at that time, it turns out to be a poor descriptor of age-creativity patterns more generally or even of physics today, where the mean age of Nobel Prize winning achievements since 1980 is 48.

This paper makes two new contributions to research on the age-creativity relationship. First, existing work, dating from the 19th century and spanning multiple disciplines, has emphasized differences in when creativity peaks across various fields assuming that each field shows a fixed age-creativity pattern (1-8). By contrast, this paper demonstrates that such cross-field age differences are small compared to cross-time differences within fields. Moreover, the field-specific dynamics are large enough that the traditional ordering of fields by the age at which breakthrough contributions are made is unstable. Second, this paper shows that these age dynamics are closely associated with several observable metrics. This analysis draws together two strands of work on the age-creativity relationship, which have only been studied separately in prior work, including our own: how the training requirements related to acquiring foundational knowledge may explain the age at which scientific careers begin (9-10) and the distinction between conceptual and experimental work in explaining creative peaks across the life-cycle (11-12). While we do not identify causal mechanisms, we show that measures drawn from this prior work, in addition to a new measure for foundational knowledge based on backwards citations ages, all move in a striking and intuitive way with shifts in the tendency for scientific contributions by the young. These collective dynamics are especially pronounced in physics during the early 20th century.

Our primary analysis focuses on the complete set of Nobel Prizes given between 1900 and 2008 in physics, chemistry, and medicine. Through extensive historical and biographical analysis, we determined the years (and hence ages) at which each Nobel-prize winner produced their prize-winning work, providing a data set of 525 prizewinners, with 182 in physics, 153 in chemistry, and 190 in medicine. Alternative data sources are also considered below. (Data-collection methods, raw data, and summary statistics are described in detail in the Supporting Information.)

Results

The image of the young, great mind making critical breakthroughs is iconic in the hard sciences. Moreover, traditional analyses of Nobelists show that, on average, physicists have made important contributions at earlier ages than chemists or medical scientists (5, 7-8). Our first results reconsider this evidence, studying differences in the age of peak creativity between fields and changes over time within fields.

Fig. 1, Panel A, presents the mean age at which Nobel laureates did their Prize-winning work, showing field averages for the whole sample period and separately for an early period (Prize-winning work before 1905) and late period (Prize-winning work after 1985). Two key patterns emerge. First, shifts in the mean age over time are large. As summarized in Panel B, the mean age of Prize-winning work has increased by 7.4 years ($p<.05$) in medicine, 10.2 years ($p<.01$) in chemistry and 13.4 years ($p<.01$) in physics. These magnitudes are much larger than the cross-field differences, the largest of which in the whole sample is 3.0 years ($p<.01$) between chemistry and physics. Second, the traditional cross-field comparisons are highly unstable. As summarized in Panel C, the rank ordering of fields from youngest to oldest interchanges. Physics, for example, has the oldest mean age at great achievement in the late period, even though it is the youngest field over the period as a whole. A variance decomposition further demonstrates the relative importance of shifts over time. Not surprisingly given the wide range of ages at which individuals make important contributions, most of the variation cannot be explained by field or year effects. Nevertheless, static, cross-field differences only account for 2.48% of the overall variance in ages while within field dynamics account for 12.33%, or 5.0 times the variance explained by the cross-field differences focused on in the literature.[1]

Figure 2 (physics), Figure 3 (chemistry) and Figure 4 (medicine) detail the field-specific age-dynamics and examine factors relating to these shifts. Panel A in each figure shows the percentage of great achievements produced by ages 30 and 40. The predicted values and indicated 95% confidence intervals are given by logistic fractional polynomial regressions (13).[2] Letting $\textit{Achievement Age}_i$ and $\textit{Year}_i$ denote the age and year in which laureate *i* made his or her great achievement, the (second-degree) regressions are from models of the form

$$\Pr\left[\textit{Achievement Age}_i < 30 \middle| \textit{Year}_i, x_i\right] = \left(1 + \exp\left\{x_i\beta_0 + \textit{Year}_i^{\,p_1}\beta_1 + \textit{Year}_i^{\,p_2}\beta_2\right\}\right)^{-1},$$

where $p_j \in \{-2,-1,-1/2,0,1,2,3\}$, for $j=1,2$. The estimation procedure searches over values of p_j to obtain the best fit to the data. This approach smoothes the data in the way that a polynomial regression does but, by searching over functional forms, it provides a more flexible relationship between $Year_i$ and the dependent variable than a polynomial regression.

In the early years, great achievement at young ages is common in all three fields. Prior to 1905, 69% of chemists; 63% of medical scientists, and 60% of physicists did their Prize-winning work before age 40, while Prize-winning work done before age 30 accounted for approximately 20% of cases. The ensuing 100 years exhibit large dynamics, with two key features. First, in all three fields, great achievement before age 30 becomes increasingly rare, converging towards 0% of cases by the end of the century. This shift away from the very young also extends to higher age thresholds in physics and chemistry. In physics, great achievement by age 40 occurs in only 19% of cases by the year 2000, less than one third its rate in 1900. In chemistry, great achievement by age 40 converges toward 0% by 2000, while it accounted for 66% of cases in 1900. Supporting Figures 1A and 1B present the underlying data and additional non-parametric estimates that show similar patterns. Supporting Figures 2 and 3, together with associated analysis in the supporting materials drawing on (10), show that underlying demographic shifts do not explain the dynamics in early life-cycle innovation.

Beyond the general aging pattern over the 20th century as a whole, a notable and exceptional dynamic appears in the initial increase in the frequency of young achievement in physics. The share of physics laureates who did their prize-winning work at young ages rises sharply, peaking at 31% in 1923 (with a 95% confidence interval of 1905-1942) for work done by age 30 and peaking at 78% in 1934 (with a 95% confidence interval of 1924-1944) for work done by age 40, before declining over the remainder of the century. The shift in physics stands out from both chemistry, where young achievement declines more consistently, and medicine, which shows a decline in achievement before age 30 but otherwise no substantial trends. The lower average age for physics over the whole period (Fig. 1) thus arises from the temporarily increased incidence of contributions by young physicists during the first third of the 20th century but does not represent a stable feature of physics. Supporting Figure 4 demonstrates similar age shifts in physics when using sources other than the Nobel Prize, and Supporting Figure 5 shows that the physics pattern is robust to controlling for region of birth.

The substantial shift toward youthful achievements in early 20th century physics occurs in a similar period as the development of quantum mechanics. Historians of physics often identify 1900-1927 as the key period in this development, starting with Planck's introduction of quanta in 1900 and continuing through the formulation of consistent theoretical foundations in 1925-1927 (14-16). Figure 2 Panel A shows that the probabilities of great achievement by ages 30 and 40 peak during this period.

Werner Heisenberg, who developed his matrix mechanics in 1925 at age 23 and his uncertainty principle two years later, may provide a useful window into early career contributions during this period. Strikingly, he did not seem particularly young for an important physicist at the time – Pauli and Dirac made contemporary, prize-winning contributions at ages 25 and 26. In the previous ten years, the majority of Nobel Prizes in physics had been given to individuals for work done by their early 30s, and Dirac and Einstein suggested that by age thirty a physicist was effectively dead (17-18). Yet as Panel A of Figs. 2-4 demonstrate, one cannot make similar claims about chemistry or medicine at that time, nor about physics today.

Heisenberg's example points toward two features of this period that may illuminate the age dynamics: the prevalence of abstract / deductive work and the obsolescence of existing knowledge. One line of age-creativity research has emphasized that abstract / deductive contributions tend to come at earlier ages than inductive contributions, which draw more heavily on accumulated knowledge (11-12). Kuhn points to the role of theoretical contributions like Heisenberg's in this episode (14). Thus, there may have been a shift toward theoretical work, which tends to be abstract and deductive, in physics at this time. A second line of age-creativity research has emphasized that the expansion of foundational knowledge in a field may increase training requirements, making contributions at younger ages more difficult (9-10). From this perspective, age dynamics might be associated with changes in a field's foundational knowledge, which may typically expand with time but may also contract in a case where new knowledge devalues old knowledge. Heisenberg, for example, nearly failed his Ph.D. exams at age 21 because he knew little of classical electromagnetism (19); his contributions in the subsequent four years suggest that training in classical physics may have become less salient.

To examine the importance of theoretical work, we classified all Nobel-Prize winning achievements according to whether the work had an important theoretical component (see Supporting Information for methods and data). To examine the relationship between the age dynamics and shifts in training, we identified the age at which each Nobel laureates received his or her highest degree (a doctorate in 98% of cases).

Figure 5 shows how the mean age at laureates' great achievements is jointly related to the theoretical versus empirical content of their contribution and their age at high degree. This figure summarizes these relationships using an ordinary least squares regression with a linear term in age at high degree and a categorical variable for a theoretical great achievement. Theorists make their great achievements 4.434 (SE=.936) years earlier than empiricists on average, and a 1 year increase in the laureate's age at highest degree is associated with a .304 (SE=.106) year increase in the average age of the laureate's great achievement. Supporting Table 3 presents regression estimates for a range of specifications (including the one in Figure 5, reported in Panel B), demonstrating that (i) training duration and (ii) theoretical research are independent, robust, and powerful predictors for the age at which great scientific achievements are made.

Given that the nature of a laureate's work and the length of the laureate's training are strong independent predictors for the age at prize-winning contributions, we turn to how they co-move over time with the age at prize-winning contributions. Fractional polynomial estimates of the dynamics for theoretical contributions are presented in Panel B of Figs. 2-4. Supporting Figure 1C presents the underlying data and additional non-parametric estimates that show similar patterns. In physics, the prevalence of theoretical contributions is hump-shaped over the 20th century (Fig. 2B), demonstrating a striking association with the age dynamics (Fig. 2A). The probability a contribution is theoretical peaks at 46% in 1933 (with a 95% confidence interval of 1925-1942).[3] The dynamics in theoretical contributions in chemistry (Fig. 3B) and medicine (Fig. 4B) also resemble their respective age dynamics, although the temporal shifts in these fields are less precisely estimated.

Panel C in Figs. 2-4 presents fractional polynomial estimates of the shares of Nobelists who received their highest degree by age 25. The training patterns also mirror the achievement-age patterns closely. While the majority of Nobel laureates received their degrees by age 25 in the early 20th century, all three fields show substantial declines in this tendency, with physics and chemistry converging towards 0% of cases by the end of the century. Furthermore, in physics, the tendency to receive a Ph.D. at young ages follows the same inverted U-shape. The dynamics in chemistry and medicine, while less precisely estimated, match closely with the dynamics in the propensity for great achievement by age 40.

To further examine shifts in foundational knowledge, we develop a novel measure based on backward citations to prior work. Specifically, we take the top 100 most-cited papers published annually over the 20th century in each of the 3 Nobel fields and in an “other” category (comprising all other science and engineering fields). The data are drawn from the Thomson Reuters-ISI’s Century of Science (covering 1900-1955) and its Web of Science data (covering 1955 to the present). We measure the mean age of each of these highly cited papers’ backwards citations. To eliminate background trends in citation age dynamics, we study the difference of the mean age of backward citations in each of the 3 Nobel fields from the mean age in the “other” category and normalize this difference by the standard deviation in the other category. Note that the ISI data are independent of the Nobel Prize data and allow more precise estimation of the dynamics due to greater sample size.

We use this citation age measure – i.e. the temporal distance to prior work – to examine knowledge dynamics, where a tendency to cite older papers suggests that top research draws on longer-established knowledge and a tendency to cite recent papers suggests that top research primarily draws on recent work.[4] Panel D of Figs. 2-4 presents fractional polynomial estimates of the evolution in backward citation age for physics, chemistry, and medicine. Supporting Figure 1D presents the underlying data and additional non-parametric estimates that show similar patterns. Again, the citation age dynamics match closely the achievement age dynamics. In physics, the tendency toward recent citations peaks in 1935 (with a 95% confidence interval of 1930-1940), which is close to the peak in youthful achievement.[5] Citation age dynamics in chemistry and medicine also reflect age-creativity patterns in those fields. Overall, the dynamics in achievement age appear similar to the citation age dynamics.

Summary/Conclusions

This paper demonstrates that the frequency of great achievement at young ages is more a function of time than field. The analysis further shows strong, independent associations between age dynamics within fields and both the prevalence of theoretical work and measures of the stock of foundational knowledge. Further work is needed to assess causal mechanisms underlying these empirical relationships and consider alternative forces, possibly emanating from the norms and institutions of science or the scale of the scientific enterprise (20-21). Notably, the dynamics in age at great achievement, prevalence of theory, Ph.D. age, and mean citation age are especially pronounced in physics and are coincident with the development quantum mechanics, which Kuhn placed at the center of his analysis of scientific revolutions (12, 22). The findings thus may provide candidate, quantitative markers to help identify such revolutionary events, providing an intriguing direction for future research.

---

[1] The share of the variance explained by cross-field differences is the R^2 of a regression of age at great achievement on field dummy variables. The share of the variance explained by within-field dynamics is the R^2 of a regression of age at great achievement on field-specific fractional polynomial regressions.

[2] Here and below, we estimate 2nd degree logistic fractional polynomial regressions for physics and medicine and a 3rd degree logistic fractional polynomial model for chemistry, which shows more complicated dynamics.

[3] A portion of the shift toward younger ages in the early part of the 20th century is reflected in a shift toward more theoretical work. At the same time, Supporting Figure 6 shows that the shift toward the young in early 20th century physics and the ensuing aging phenomenon persist when looking within theorists and within empiricists.

[4] Although related, our measure differs from a citation half-life insofar as half-lives measure durability using forward citations, whereas our measure captures reliance on previous work using backward citations. Our measure is also distinct from conventional citation metrics for research performance in that it measures the amount of foundational knowledge in a field at a point in time as opposed to identifying important papers or researchers (e.g. the H-index). See Supporting Information for citations methods and data details.

[5] We further study whether the temporary reduction in citation age in physics was driven by new scholars entering the field, or whether existing scholars also started citing more recent work. The latter phenomenon would suggest that our findings describe general changes in the knowledge space itself, not simply changes in which physicists were active. Supporting Table 4 presents the results of regressions that predict citation age over time while controlling for individual researcher fixed effects, thus eliminating changes in the composition of physicists and focusing on variation in the citation tendencies within individual careers. The regression results show that the humped-shaped phenomenon in Fig. 2 Panel D persists when looking at changes over time within individuals' careers. The tendency to cite recent papers peaks in 1920 in physics, suggesting a contraction in foundational knowledge and a turn of existing researchers towards the new frontier.

Notes

1. Beard, George M. (1874, 1979). “Legal Responsibility in Old Age: Based on Researches into the Relation of Age to Work.” reprinted in *The “Fixed Period” Controversy: Prelude to Ageism*, Gerald J. Gruman, Ed. New York: Arno Press.
2. Zuckerman, Harriet and Robert Merton, "Age, Aging, and Age Structure in Science," in Robert Merton (Ed.) *The Sociology of Science* (Chicago, IL: University of Chicago Press, 1973), 497-559.
3. Simonton, Dean K. *Scientific Genius: A Psychology of Science* (Cambridge: Cambridge University Press, 1988a).
4. _____. "Career Landmarks in Science: Individual Differences and Interdisciplinary Contrasts," *Developmental Psychology* 27 (January 1991), 119-130.
5. Stephan, Paula G. and Sharon E. Levin, "Age and the Nobel Prize Revisited", *Scientometrics* 28 (November 1993), 387-399.
6. Levin, Sharon E. and Paula G. Stephan. 1991. “Research Productivity over the Life Cycle: Evidence from Academic Scientists.” *American Economic Review* 81 (No. 1, March): 114-32.
7. Lehman, Harvey C., *Age and Achievement* (Princeton, NJ: Princeton University Press, 1953).
8. Simonton, Dean K., "Age and Outstanding Achievement: What Do We Know After a Century of Research?" *Psychological Bulletin* 104 (September 1988b), 251-267.
9. Jones, Benjamin F. 2009 "The Burden of Knowledge and the Death of the Renaissance Man: Is Innovation Getting Harder?" *Review of Economic Studies* 76:1 (January), 283-317.
10. Jones, Benjamin F. 2010. "Age and Great Invention," *Review of Economics and Statistics* 99:1 (February), 1-14.

11. Galenson, David W., and Bruce A. Weinberg. 2001. "Creating Modern Art: The Changing Careers of Painters in France from Impressionism to Cubism," *American Economic Review* 91:4 (September), 1063-1071.
12. Weinberg, Bruce A. and David W. Galenson. 2008. “Creative Careers: Age and Creativity among Nobel Laureate Economists.” Working Paper.
13. Royston, P. and D. G. Altman, “Regression using fractional polynomials of continuous covariates: parsimonious parametric modeling,” *Applied Statistics* 43 (1994), 429-467.
14. Kuhn, Thomas S., *The Structure of Scientific Revolutions* (Chicago: University of Chicago Press, 1962).
15. Jammer, Max, *The Conceptual Development of Quantum Mechanics* (New York: McGraw-Hill, 1966).
16. 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).
17. Brodetsky, S. 1942. “Newton: Scientist and Man.” *Nature* 150 (December 19): 698-9.
18. Jungk, Robert. 1958. *Brighter than a Thousand Suns: A Personal History of the Atomic Scientists*. Translated by James Cleugh. New York: Harcourt, Brace.
19. Cassidy, David C. 1992. *Uncertainty: The Life and Times of Werner Heisenberg*. New York: W. H. Freeman and Co.
20. Dasgupta, Partha, and Paul A. David. 1994. "Toward a New Economics of Science," *Research Policy* 23: 5 (September), 487-521.
21. Goodstein, David. 1994. “The Big Crunch.” August 8, 2011 <http://www.its.caltech.edu/~dg/crunch_art.html>.
22. Stove, David. 1982. *Popper and After: Four Modern Irrationalists*, Pergamon Press.

Acknowledgements: We thank David Kaiser, Dean Keith Simonton, and Scott Stern for helpful comments and Sharon Levin and Paula Stephan for Nobel Prize data. The John Templeton Foundation, the National Institutes of Health, the National Science Foundation, and National Bureau of Economic Research provided financial support.

	Physics	Chemistry	Medicine
Whole Period	37.2 (0.7)	40.2 (0.7)	39.9 (0.6)
Early Period	36.9 (2.0)	36.1 (1.6)	37.6 (1.5)
Late Period	50.3 (3.2)	46.3 (2.5)	45.0 (3.0)

A Mean Age of Great Achievement

	Physics	Chemistry	Medicine
Age Difference Over Time	13.4*** (3.8)	10.2*** (3.0)	7.4** (3.3)
	Ch - Ph	**Ch - Me**	**Me – Ph**
Age Difference Across Fields	3.0*** (0.9)	0.3 (0.9)	2.7*** (0.9)

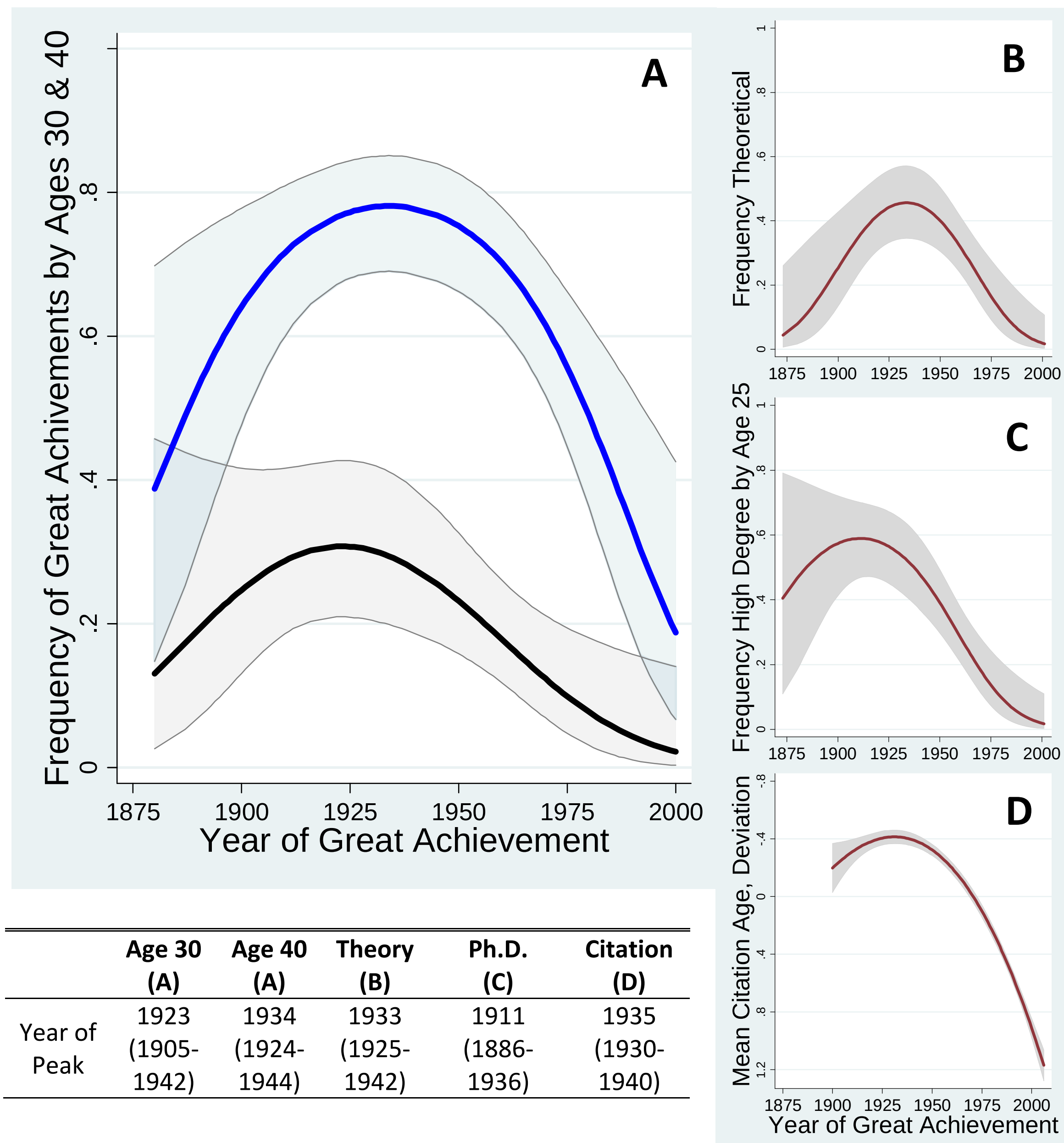
B Age Differences Over Time and Across Fields

	Physics	Chemistry	Medicine
Whole Period	1	3	2
Early Period	2	1	3
Late Period	3	2	1

C Age Rank Instability

Figure 1: Age of Great Achievement over Time and by Field
Panel A presents the mean age at which Nobel laureates produced their prize-winning work in physics, chemistry, and medicine across all years (All) and in the early period (through 1905) and late period (from 1985). Panel B (top row) presents the change in mean age of great achievement between the early and late periods within each field. For each pair of fields, Panel B (bottom row) shows the difference between fields in the mean age over all years. Panel C presents the rank ordering of the fields by mean age for the whole period, early period, and late period (1 indicates lowest mean age, 3 indicates highest mean age). Standard errors are given in parentheses. ** indicates significance at 5%; *** indicates significance at 1%.

PHYSICS



	Age 30 (A)	Age 40 (A)	Theory (B)	Ph.D. (C)	Citation (D)
Year of Peak	1923 (1905-1942)	1934 (1924-1944)	1933 (1925-1942)	1911 (1886-1936)	1935 (1930-1940)

Figure 2: Dynamics in Physics. The figure presents dynamics in age and associated factors for Nobel Prize winning achievements in Physics. Panel A presents the evolution over time in the frequency of prize-winning contributions by ages 30 (lower) and 40 (upper). Shaded regions indicate 95% confidence intervals. Panels B through D present related dynamics: the frequency of prize-winning achievements with an important theoretical component (Panel B); the frequency of Nobel laureates who received their Ph.D. by age 25 (Panel C); and the backward citation age for the top papers in each field (Panel D, inverted y-axis), as defined in the text. The table summarizes the year when each measure peaks, with 95% confidence intervals, further indicating similar dynamics across the measures.

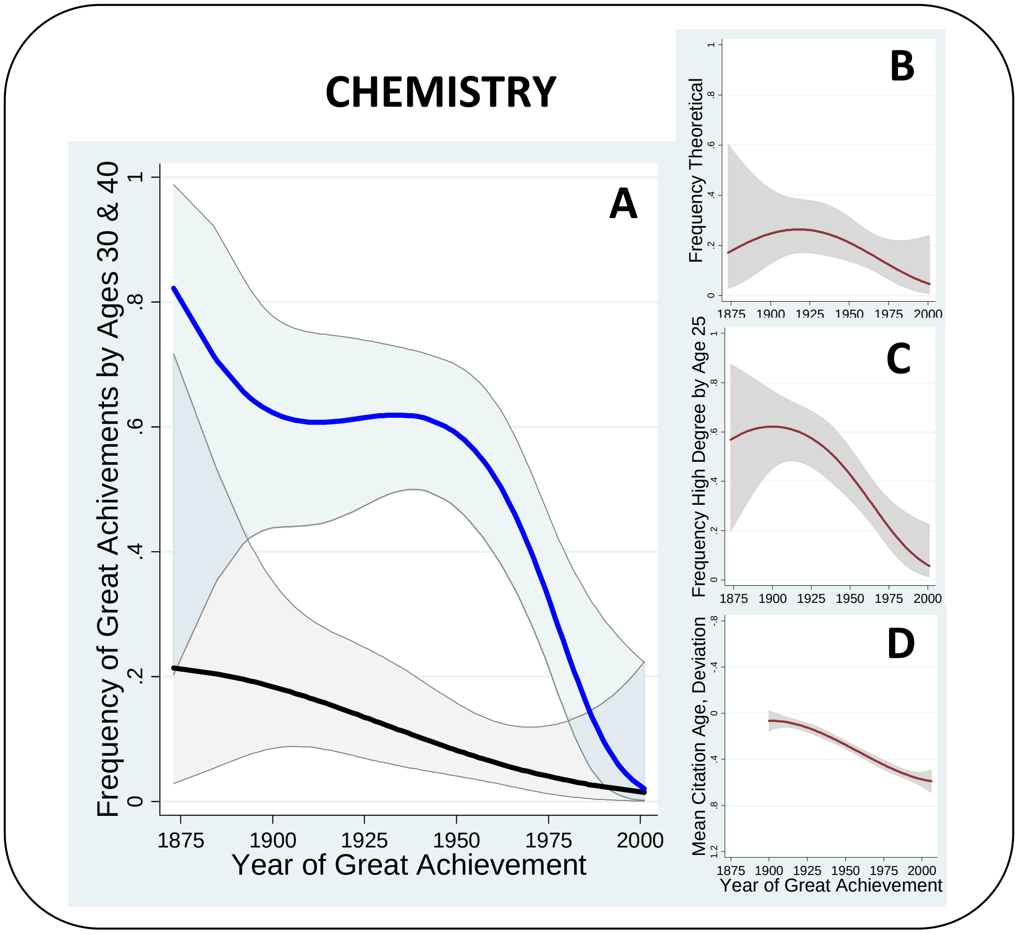


Figure 3: Dynamics in Chemistry. Age and associated factors for Nobel Prize winning achievements in Chemistry. See notes for Figure 2.

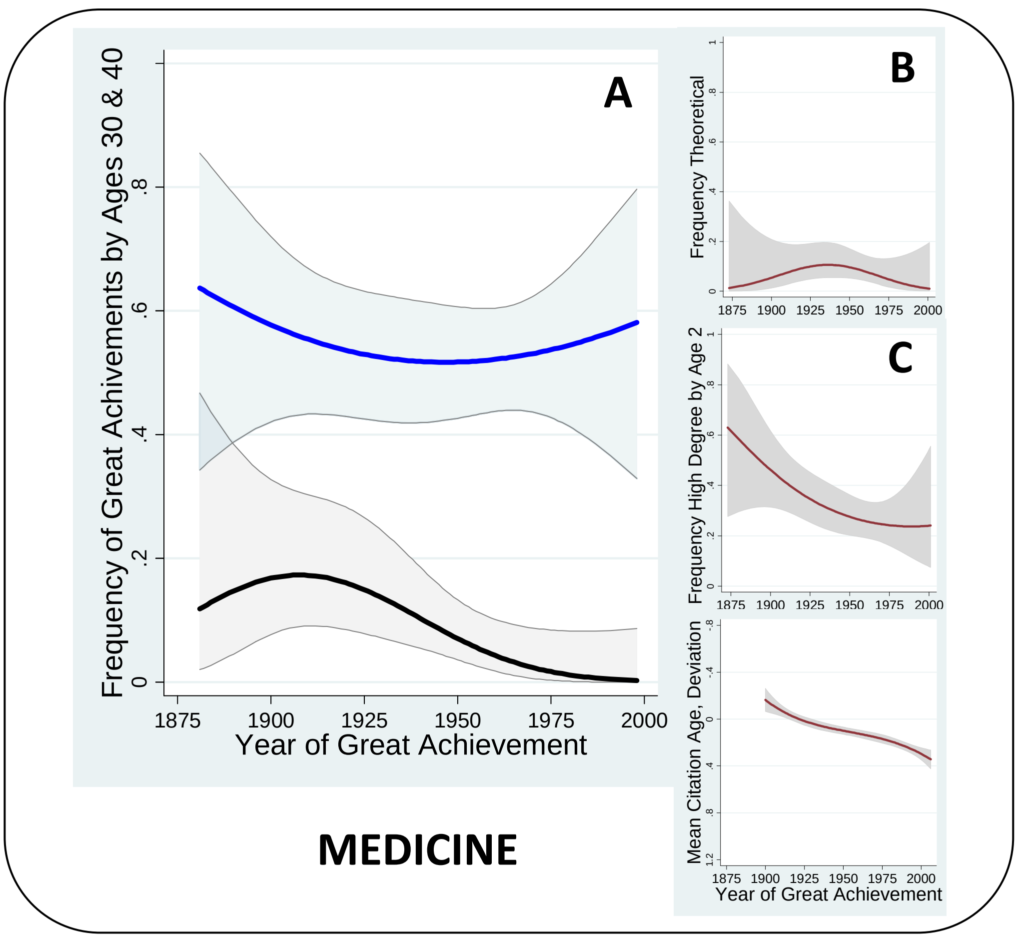


Figure 4: **Dynamics in Medicine.** Age and associated factors for Nobel Prize winning achievements in Medicine. See notes for Figure 2.

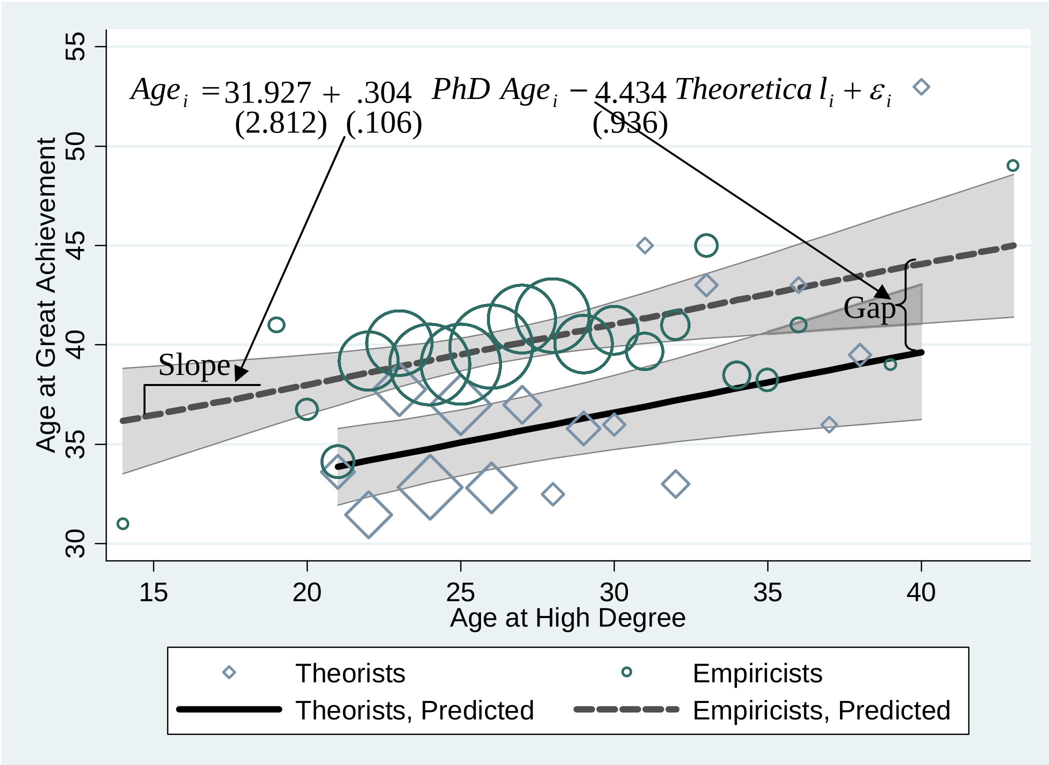


Figure 5. Predictors of Age at Great Achievement
The figure shows how the age at which a laureate produces prize-winning work is related to the laureate's age at highest degree and whether the great achievement had a theoretical component. Each square (circle) represents the average achievement age for the laureates who received their high degree at a given age and the prize for theoretical (empirical) work. They are scaled in proportion to the number of laureates in that cell. Supporting Table 3 reports regressions of achievement age on the nature of work (theoretical versus empirical) and age at high degree for a range of specifications. The regression and 95% confidence intervals are based on the specification shown in column 3 of panel