Why the US science and engineering workforce is aging rapidly

David M. Blau[a,b,1] and Bruce A. Weinberg[a,b,c]

[a]Department of Economics, Ohio State University, Columbus, OH 43210; [b]Institute of Labor Economics (IZA), 53113 Bonn, Germany; and [c]National Bureau of Economic Research, Cambridge, MA 02138

Edited by Shirley M. Tilghman, Princeton University, Princeton, NJ, and approved February 14, 2017 (received for review August 12, 2016)

The science and engineering workforce has aged rapidly in recent years, both in absolute terms and relative to the workforce as a whole. This is a potential concern if the large number of older scientists crowds out younger scientists, making it difficult for them to establish independent careers. In addition, scientists are believed to be most creative earlier in their careers, so the aging of the workforce may slow the pace of scientific progress. We develop and simulate a demographic model, which shows that a substantial majority of recent aging is a result of the aging of the large baby boom cohort of scientists. However, changes in behavior have also played a significant role, in particular, a decline in the retirement rate of older scientists, induced in part by the elimination of mandatory retirement in universities in 1994. Furthermore, the age distribution of the scientific workforce is still adjusting. Current retirement rates and other determinants of employment in science imply a steady-state mean age 2.3 y higher than the 2008 level of 48.6.

science of science | demography | aging | retirement | innovation

The US science and engineering workforce is aging rapidly. This is a potential problem for two reasons: (*i*) older scientists may not retire at a fast enough rate to free up positions for younger researchers to establish independent careers (1–4), and (*ii*) scientific creativity is thought to peak at a relatively young age (5–9), although the evidence is in fact somewhat mixed. The aging of the scientific workforce has been called a crisis (10). Policy proposals have focused on directing more research support to new and early-stage investigators to maintain the quantity and quality of scientific research and the sustainability of the scientific workforce (11, 12). However, we are not aware of rigorous analyses of the causes of the aging of the scientific workforce, and the implications of current trends for the long-run age distribution of scientists.

This article develops and simulates a demographic model of the scientific workforce to (*i*) determine the causes of the recent aging trend and (*ii*) predict the long-run effects of these factors on the age distribution. First, we show that "demographic momentum" in the form of the aging of the large baby boom cohort has driven much of the recent rapid aging of the scientific workforce, and will continue to do so for the next two decades as the later cohorts of the baby boom pass through their 60s and early 70s. However, sharp declines since 1993 in the rate at which scientists retire from employment can account for 8% of the increase in the mean age of scientists. The decline in retirement was most likely triggered by the elimination of mandatory retirement at universities in 1994. We also find that the aging of the workforce as a whole (due to lower fertility) accounts for 13% of the increase in the mean age of the scientific workforce. Second, we show that the scientific workforce was very far from its implied steady-state age distribution when our analysis begins in 1993 (4.9 y younger on average). Strikingly, the scientific workforce remains far from steady state even as of 2008—current entry, exit, and transition rates imply that the mean age of the scientific workforce will increase by another 2.3 y from that level.

The main source of data on US scientists and engineers is the restricted-use 1993–2010 Survey of Doctorate Recipients (SDR) of the National Science Foundation (NSF), a typically biannual longitudinal sample survey of the population with a research doctorate in science, engineering, or health, earned in the United States (https://www.nsf.gov/statistics/srvydoctoratework/). We use detailed information on age, field of degree, job tenure, previous employment, occupation, and sector of employment on about 73,000 scientists aged 76 or less, across all fields (we refer to this population as "scientists," although we include people with engineering, health, and social science degrees and all sectors of employment). We supplement the SDR with census data from the Current Population Survey (CPS), the 1980 and 1990 US Decennial Censuses, and the 2000–2013 American Community Surveys. The census data provide information on trends in the age distribution of the US workforce as a whole (defined as individuals who work at least 13 wk per y and 15 h per wk), and they also help fill two gaps in coverage of the SDR: scientific workers in the United States who obtained a PhD abroad, and pre-1993 data. Complete details are provided in *SI Appendix, section 1*. (Consent was obtained from SDR participants by NSF. Our work was approved by Ohio State University and National Bureau of Economic Research's institutional review boards.)

Descriptive Results

Fig. 1*A* shows the age distribution of the scientific workforce in 1993 and 2010, the first and last years of SDR data available to us. The aging of the workforce is evident, with a significant decline in the share of scientists aged 35–53 and a significant increase in the share older than 53. The workforce as a whole is also aging, as shown in Fig. 1*A*. In 1993, the scientific workforce was disproportionately concentrated at ages 30–56 compared with the workforce as a whole, which had substantial mass at younger ages. (The

Significance

The science and engineering workforce has aged rapidly, both absolutely and relative to the workforce, which is a concern if the large number of older scientists crowds out younger scientists. Moreover, scientists are believed to be most creative earlier in their careers, so the aging of the workforce may slow the pace of scientific progress. We study the causes of this aging, showing that a substantial majority is a result of the aging of the large baby boom cohort of scientists, but the elimination of mandatory retirement in universities in 1994 was also an important factor. Strikingly, current patterns imply a steady-state mean age 2.3 y higher than the 2008 level of 48.6.

Author contributions: D.M.B. and B.A.W. designed research, performed research, analyzed data, and wrote the paper.

Conflict of interest statement: The authors declare B.A.W. was supported under P01 AG039347 directly by the National Bureau of Economic Research and on a subcontract to Ohio State University.

This article is a PNAS Direct Submission.

Data deposition: The data are restricted use data from the Survey of Doctorate Recipients, which are available from the National Science Foundation subject to their application process. We will make our code available upon request.

[1]To whom correspondence should be addressed. Email: blau.12@osu.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1611748114/-/DCSupplemental.

ECONOMIC SCIENCES

Fig. 1. The age distribution of the US scientific workforce and the US workforce as a whole. *A* shows the age distribution of scientists [calculated from the Survey of Doctorate Recipients (SDR)] and the US workforce [calculated from the Current Population Survey (CPS)] for 1993 and 2010. *B* shows the share of scientists in the US workforce by age in 1993 and 2010. *C* plots trends in the mean age of scientists and the US workforce as well as the share of scientists and the US workforce age 55 and over.

SDR includes only people with PhDs, so our definition of the scientific workforce excludes graduate students. The age distribution estimates for the 20s and early 30s should be interpreted with this in mind.) In 1993, the distributions at ages 57+ were nearly identical. By 2010, the share of scientific workers aged 55+ was much larger than the corresponding share for all workers. Thus, scientists in 2010 were employed at older ages to a much greater extent than the workforce as a whole, in contrast to 1993. Fig. 1*B* shows the share of scientific workers in the workforce by age (the ratio of scientists to the workforce as a whole from Fig. 1*A*). In 2010, the share of scientists increases from 0.27% at age 50 to over 0.8% at age 71, in clear contrast to 1993, when the share of scientific workers peaked at 0.4% and there is no strong age pattern. Comparing the scientific workforce to the highly educated workforce as a whole (reported in *SI Appendix*, Fig. S1) shows that scientists are aging less relative to the more educated workforce than relative to the workforce as a whole.

Fig. 1*C* illustrates trends in two summary statistics of the age distribution, the mean age and the share 55 and older. The average age of the scientific workforce increased from 45.1 to 48.6 between 1993 and 2010. The mean age of the workforce as a whole increased at a slightly slower pace, from 42.2 to 45.4. There was a larger divergence in the share aged 55 and above. In 1993, the shares were 0.18 for scientific workers, and 0.15 for all workers. By 2010, the share of all workers aged 55+ increased to 0.23, whereas the share of scientific workers rose to 0.33. Thus, the aging of the scientific workforce has been especially concentrated at older ages and was considerably more rapid than that of the workforce as a whole. *SI Appendix*, Figs. S2–S4 present age distributions for individual years, and breakdowns by field and different segments of the science workforce. They show that the aging pattern is pervasive across fields—biomedicine, where aging has received considerable attention, is not exceptional; and computer and information science is initially younger than the other fields but ages considerably more rapidly. Not surprisingly, scientists whose primary or secondary activity is research tend to be somewhat younger. Scientists in academia tend to be slightly younger than those outside, but the trends are similar. *SI Appendix*, Fig. S5 shows that the share of academics has declined slightly from 42% to 38% between 1993 and 2008.

The three main determinants of the age distribution of the scientific workforce, beyond the part that can be explained by aggregate demographic trends, such as declining fertility and mortality rates and the large size of the baby boom cohort, are as follows: (*i*) the proportion of the population that completes a PhD in a science discipline and the distribution of age at completion; (*ii*) entry to the US scientific workforce by immigrants, both those who obtain a PhD in the United States and scientists who obtained a PhD abroad (scientists who obtained a PhD abroad are excluded here because they are not covered by the SDR, but included in the analysis of census data below); and (*iii*) the rate of exit from the scientific workforce. We begin by documenting trends in these factors before turning (in the next section) to a formal model that quantifies their impact.

Fig. 2 shows the biannual hazard rate of exit from the scientific workforce to nonemployment in 1993 and 2008. The hazard rate is the probability of exiting the work force conditional on not having previously exited, measured empirically by the proportion of individuals employed as scientists in a particular year who were not employed 2 y later. We refer to it as the hazard rate of retirement, because most exits to nonemployment are in fact self-reported as being due to retirement. (We report biannual hazard rates because

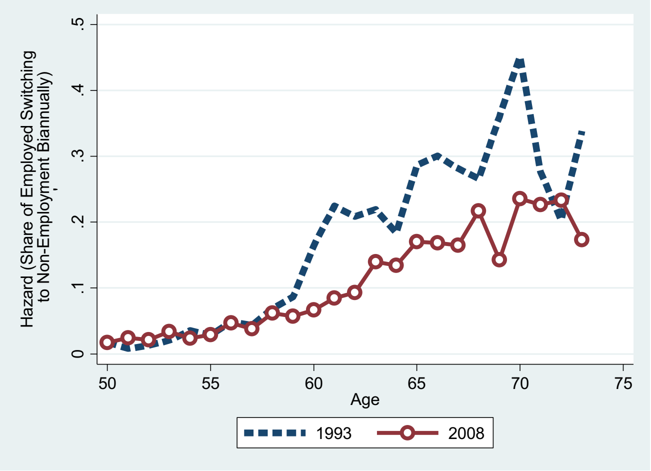


Fig. 2. Biannual transition rates from science employment to nonemployment, 1993 and 2008. The figure shows the share of science doctorates employed in science in 1993 (dashed blue) and 2008 (red circles), who are not employed as of the next survey.

the reference periods for the 1993 and 1995 SDR surveys and the 2008 and 2010 surveys were exactly 2 y apart. *SI Appendix, section 3* describes how we deal with cases in which the surveys were more than 2 y apart.) In 1993, the shape of the retirement hazard was similar to, but lower than the typical age pattern of retirement (*SI Appendix*, Fig. S6), with a substantial increase in the exit rate between ages 60 and 62, a jump at age 65, and a very large spike at age 70. The most recent data show a much slower and more gradual increase in the exit hazard rate, and no major spikes. In particular, the large spike at age 70 in 1993 completely disappeared by 2008. This change is consistent with the end of mandatory retirement at age 70 in universities in 1994 (due to eliminating an exemption for universities to the 1986 Age Discrimination Act), which caused a substantial reduction in the rate of retirement of university faculty (13). *SI Appendix*, Fig. S6 *B* and *C* show that the spike in the retirement hazard at age 70 in 1993 is much larger among scientists in academia than among those outside of academia, and the spike declined by a much larger amount in academia after 1993. As demonstrated below, this change in retirement behavior has had a substantial influence on the age distribution of the scientific workforce.

Fig. 3*A* shows the trend in the scientific PhD completion rate, expressed as a share of births 30 y earlier for illustrative purposes. (These results are not sensitive to using births in a 3-y span on either side of the 30-y base. We are able to extend back to the 1960s by using data on year of PhD completion among scientists up to age 76.) (*SI Appendix, section 2* and Figs. S7 and S8 provide additional data on science graduates.) From 1985 to 2008, the rate of science doctorate completion doubled from about 0.005 to 0.010 as a share of lagged births. All other things being equal, this would tend to reduce the age of the scientific workforce. However, as shown in Fig. 3*B*, the average age at completion of a science PhD in the United States increased from 30 in the 1970s to around 33 in 1993. This clearly contributed to the aging of the scientific workforce in the 1970s, 1980s, and early 1990s, but in the period we study (1993–2008), the mean age at PhD completion actually declined slightly. Fig. 3*C* focuses on the change over the period we study, from 1993 to 2008, in the number of science PhDs awarded by age. During this period, there was a shift toward PhDs awarded at ages below 36 and away from the late 30s and 40s. Hence, changes in the distribution of age at PhD during this period will not be able to explain aging of the scientific workforce.

The share of foreign recipients among new science doctorates awarded in the United States grew rapidly, from 10% to 15% in the 1960s, to more than 40% in recent years, as illustrated in Fig. 3*D*. In the absence of foreign PhDs, growth in the number of new science PhDs would likely have been much lower. However, it turns out that this would not have affected the rate of aging of the scientific workforce, because foreign-born and native-born US PhD recipients have very similar employment patterns (see below).

Modeling Changes in the Age Distribution of Scientific Workers

Drawing on standard demographic simulation methods, this section outlines a formal stock-flow model of entry and exit from the scientific workforce by age, which we use to numerically analyze changes in the age distribution of scientists. As illustrated in Fig. 4 and detailed in *SI Appendix, section 3*, the model allows for entry to the scientific workforce from (*i*) US natives obtaining a PhD in the United States and (*ii*) nonnatives obtaining a PhD in the United States (in subsequent analysis described below, we also incorporate entry by nonnatives obtaining PhDs abroad). Doctorate recipients then flow between being employed in the United States in science, being employed in the United States outside of science, and being out of the labor force. *SI Appendix, sections 4 and 5* discuss the entry and exit rates symbolized by the arrows in Fig. 4, and *SI Appendix*, Figs. S9–S11 illustrate their trends.

We use the model to analyze the change in the age distribution of scientists between 1993 and 2008. (We stop in 2008 because the 2010 SDR does not contain data on all PhDs awarded in 2009.) Simulating the model generates a predicted 2008 age distribution conditional on the observed 1993 age distribution as a function of (*i*) the observed set of survey-year-and-age–specific hazard rates for employment transitions, (*ii*) the observed year-and-age–specific PhD completion rate, (*iii*) observed fertility and mortality rates by year, and (*iv*) the observed year-and-age–specific share of foreign-born US PhD recipients. We use the model to explore explanations for the aging of the scientific workforce by conducting

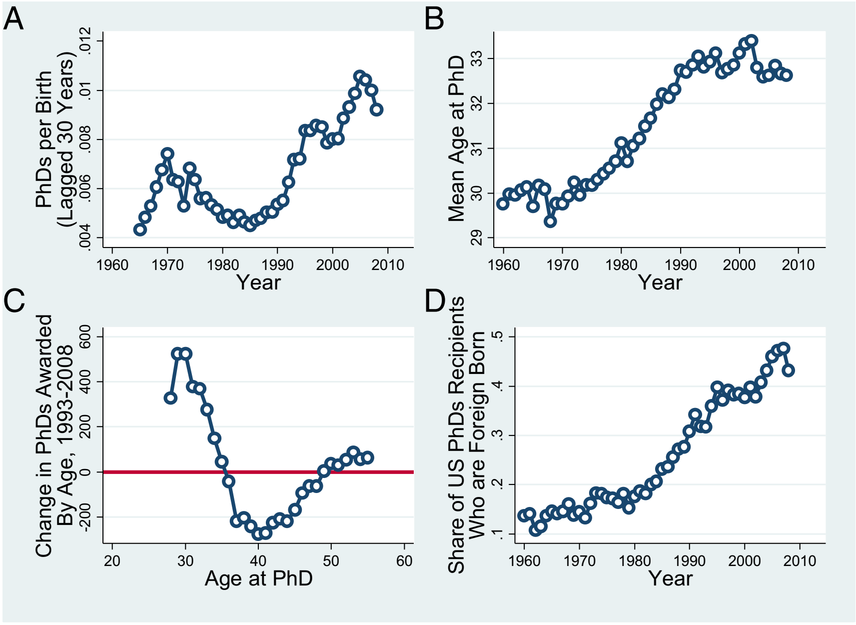


Fig. 3. Trends in US science doctorates. *A* shows the number of PhDs granted (from the SDR) per birth 30 y earlier (from Vital Statistics of the United States). *B* displays the mean age at PhD completion calculated from the SDR. *C* shows the change between 1993 and 2008 in the number of PhDs awarded by age, from the SDR. *D* reports the share of PhDs awarded in the United States to nonnatives.

ECONOMIC SCIENCES

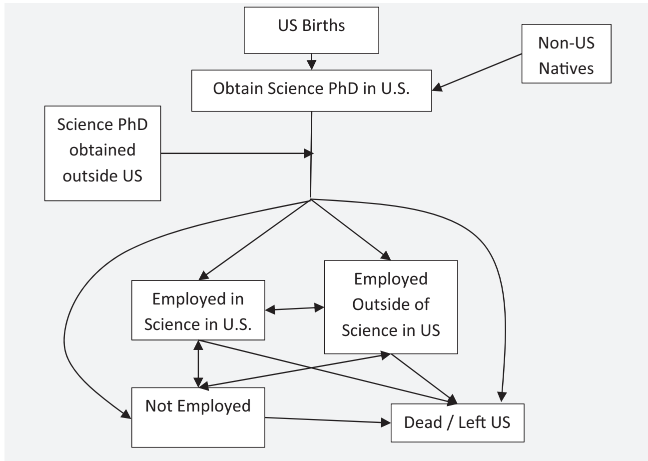


Fig. 4. Model schematic. The figure illustrates the model in *SI Appendix, section 3*, with boxes showing states and arrows showing transitions. Note that the SDR does not include people who leave the United States immediately after completing the PhD, even if they later return to the United States.

counterfactual simulations of the 2008 age distribution in which each transition rate is set to its 1993 age-specific value, one at a time and in combination. The only exception is that we use the 1960 birth rate as the counterfactual, because the birth rate affects the workforce with a long lag—scientists aged 27–75 in 1993 were born from 1918 to 1966. We use 1960 as a representative high-fertility year that was part of the baby boom.

Fig. 5*A* shows the observed age distributions in 1993 (dashed blue) and 2008 (red circles), the 2008 distribution predicted by the model based on the observed hazard, PhD completion, and birth and death rates for each year (green triangles), and the counterfactual predicted 2008 distribution, using the same model, but holding all transition rates fixed at their 1993 values (yellow diamonds). (*SI Appendix*, Fig. S12 shows analogous results for the share of scientists in the workforce by age.) As shown above, there was a substantial change in the age distribution from 1993 (dashed blue) to 2008 (red circles). The model slightly overpredicts the decline in the shares until the early 40s, underpredicts the decline in the 40s and 50s, and is quite accurate at later ages. Holding all transition rates fixed at their 1993 values, the model predicts smaller changes at all but the very youngest ages (yellow diamonds) compared with the predicted change, but the differences are very small at ages 48–63. Small random measurement errors in the hazard and other transition rates can cumulate over a 15-y period, because the model is nonlinear. Under the circumstances, the model is remarkably accurate.

To summarize the results, we compute the fraction of the observed change at each age that can be accounted for by the model [(predicted 2008 − observed 1993)/(observed 2008 − observed 1993)] and report the median of these age-specific changes. We also report in *SI Appendix*, Table S1 the fraction of the observed change in mean age and the share aged 55 and older that can be explained by the model. Fig. 5*B* and *SI Appendix*, Table S1 show that the median age-specific share of the change explained by the model is 95.5%. *SI Appendix*, Table S1 shows that the model explains 97.3% of the change in the share ages 55 and above, and overexplains the change in the mean age: 120.3%. The next row of Fig. 5*B* shows that with all transition rates held fixed at their 1993 values, the explanatory power of the model is about 15 percentage points lower for the median of the age-specific shares (and 10–14 percentage points lower for the mean age and share aged 55 and above; *SI Appendix*, Table S1). Thus, the changes in transition rates over time can account for a relatively small but not negligible part of the predicted change in the age distribution.

The remaining rows of Fig. 5*B*, in which one rate at a time is held constant, indicate that the change in the hazard rate of retirement is the most powerful explanatory factor (aside from change in the birth rate, which has a population-wide effect, not specific to scientists). If all transition rates had evolved as observed

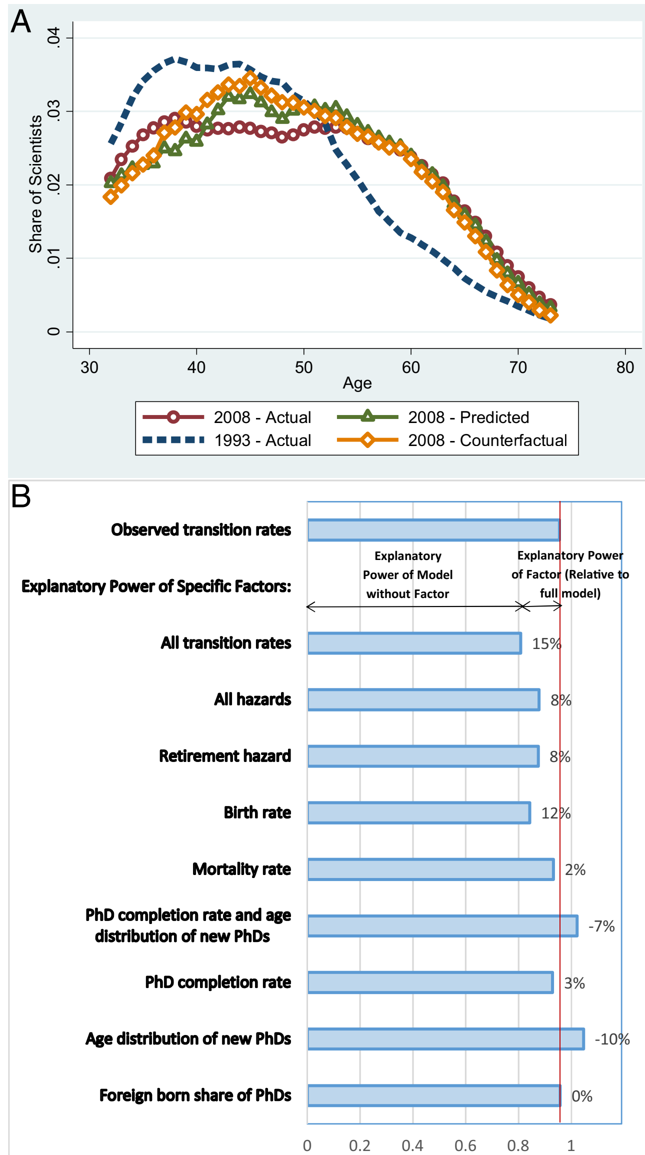


Fig. 5. Changes in the age distribution of the science workforce and explanations. *A* shows the actual age distributions in 1993 (dashed blue) and 2008 (red circles), the predicted 2008 distribution derived from the model in *SI Appendix, section 3* using the observed hazard, PhD completion, and birth and death rates for each year (green triangles), and the counterfactual predicted 2008 distribution, using the same model but holding all transition rates fixed at their 1993 values (yellow diamonds). To facilitate interpretation, we show 5-y moving averages by age. The first row of *B* shows that, on average, 95% of the age-specific changes in the distribution can be accounted for by changes in observed transition rates. The remaining rows quantify the importance of each factor by using the observed 2008 rates except for the one indicated, for which the 1993 rate is used. For example, holding the retirement hazard constant at its 1993 age-specific level reduces the fraction explained from 0.95 to 0.87, for a loss of 0.08 in the fraction explained, or 8% as a share of the 0.95 explained by the model.

except for the retirement hazard rate, and it had remained constant at the 1993 level, the model could account for 87.5% of the observed change in the median age-specific fraction explained, instead of 95.5%. So the sharp decline in the retirement hazard alone can explain 8.4% of the predicted change in the age structure: [(0.955 − 0.875)/0.955 = 8.4%]. The decline in the birth rate is also important because it reduces the growth rate of the scientific workforce: it can explain 12% of the predicted change. Declining mortality had relatively little impact on the age distribution because mortality has little impact on the age structure at low levels of fertility (14). The decline in the age at PhD completion since 1993 was in the wrong direction to explain aging of the scientific workforce. The slight decline in the rate of production of science PhDs since 1993 had little impact on the age distribution. The other hazard rates account for only small changes in the age distribution of scientists. Strikingly, the large increase in the share of US science PhDs awarded to foreign-born individuals has had virtually no impact on the age distribution because foreign-born and native science PhD recipients behave very similarly (i.e., have very similar hazard rates; see ref. 15 for analysis of foreign-born US-trained scientists). *SI Appendix*, *section 6* and Figs. S13–S17 show the large increase in women among science and engineering graduates and extend the model to allow differences in transition rates by gender and by nativity. As detailed in *SI Appendix*, *section 6*, these extensions had virtually no impact on the results.

The Long-Run Age Distribution of the Scientific Workforce

We use the model to investigate the implications of current transition rates for the long-run age distribution of the scientific workforce. This is important because the age distribution can take many years to reach the steady state implied by a given set of transition rates. Thus, the current age distribution may not be a good indicator of the future. We also show how the steady-state age distribution is affected by alternative values of the transition rates.

In this part of the analysis, we are able to extend the model to incorporate immigration of scientists who received a science PhD abroad (or received a US PhD, left the United States, and then returned), which is not measured in the SDR, using data from the American Community Survey (ACS). We combine data from the ACS for the years 2000–2013 to compute the average annual number of recently arrived immigrant scientists. Our methods and results are presented in *SI Appendix*, *section 7* and Figs. S18–S20.

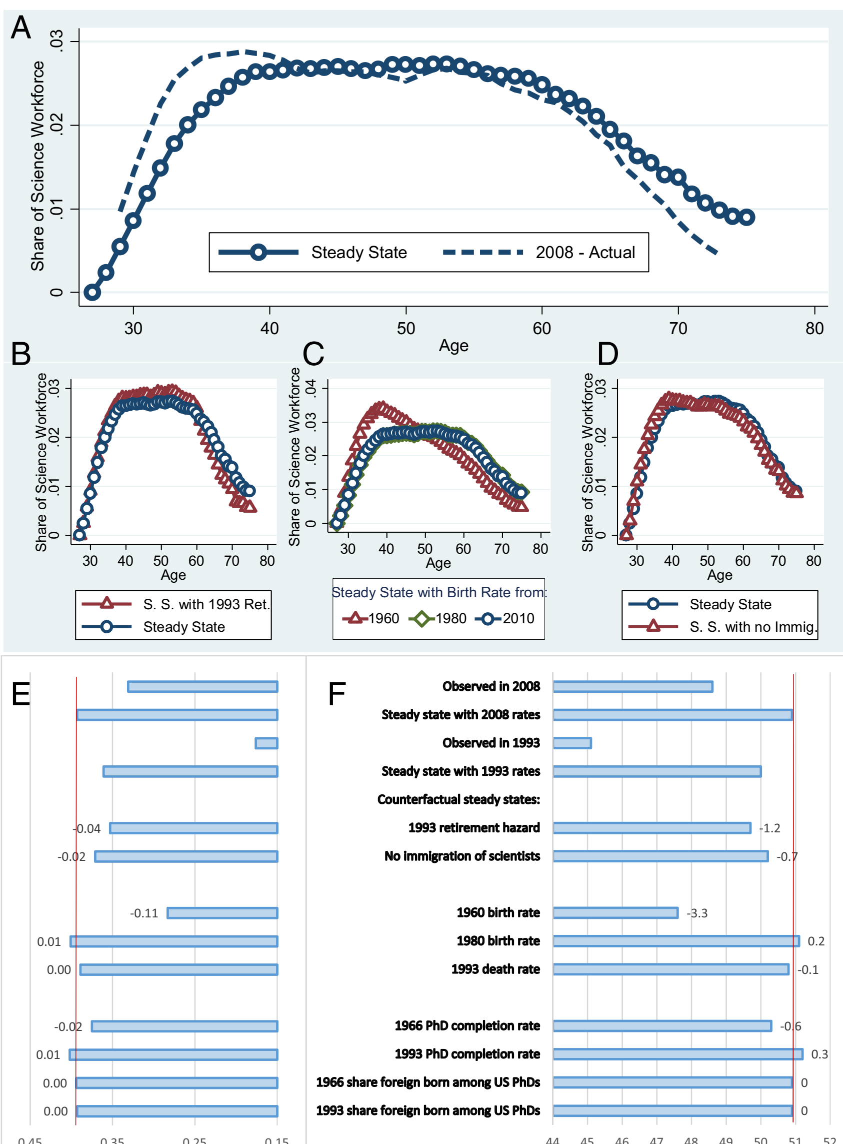


Fig. 6. Age distribution of science workforce: Actual, steady state, and counterfactuals. *A* shows the actual 2008 age distribution (from the SDR, dashed) and implied steady state (circles). *B* shows the steady state from *A* (blue circles) and the counterfactual steady state assuming the 1993 retirement hazard rates (red triangles). *C* shows the steady state from *A* (blue circles) and the counterfactual assuming the 1960 birth rate (red triangles) and 1980 birth rate (green diamonds, just behind the blue circles). *D* shows the steady state from *A* (blue circles) and the counterfactual steady state assuming no immigration of science workers (red triangles). The top four rows of *E* and *F* provide basic descriptive statistics on the share of the science workforce over age 55 (*E*) and the mean age (*F*). The remainder of *E* and *F* report the effect on the implied steady state of the change illustrated. The row labeled "1993 retirement hazard" uses observed 2008 transition rates except for the retirement hazard, which is set to its 1993 level. The other counterfactuals have a similar interpretation. The 2008 transition rates are based on 1-y transitions observed between the 2008 and 2010 survey waves. The 1993 transition rates are based on 1-y transitions observed between the 1993 and 1995 survey waves.

SI Appendix, Fig. S18 shows the rapid growth in the share of immigrants in the US scientific workforce, especially at ages 45–54 (*SI Appendix*, Fig. S19). *SI Appendix*, Fig. S20 shows that the number of recent arrivals (less than or equal to 1 y in the United States at the time of the survey) drops sharply from 1,500 at age 31 to 500 at age 40 and 250 by age 50. Summing over ages, the total number of scientists trained abroad immigrating to the United States has been around 17,000 per year on average since 2000. This compares to about 35,000 new science doctorates produced in the United States per year, including those earned by nonnatives who remain in the United States (*SI Appendix*, Fig. S8). (We have no data on the transition rates of foreign-trained scientists, so we assume they have the same rates as those of US-trained scientists.)

We simulate the steady-state distribution by letting the model run with the 1993 transition rates until the age distribution converges. Strikingly, the steady-state mean age implied by the 1993 transition rates is 4.9 y greater than the observed age in 1993, indicating that one reason for the aging from 1993 to 2008 was that the 1993 age distribution was very far from the steady state.

Fig. 6*A* reveals that the 2008 transition rates imply a substantially older scientific workforce than that observed. Summary statistics in the top part of Fig. 6 *E* and *F*, the top part of *SI Appendix*, Table S2, and *SI Appendix*, Fig. S21 indicate that, if 2008 transition rates persist, in the long run the mean age of the scientific workforce will increase by 2.3 y, from 48.6 in 2008 to 50.9 (Fig. 6*F*), and the fraction aged 55 and older will increase by 6.2 percentage points, from 0.331 to 0.393 (Fig. 6*E*). Thus, despite the already rapid aging of the scientific workforce from 1993 to 2008 (compare the first and third rows), the 2008 transition rates imply a substantially older age distribution. The transition takes about 40 y.

We resimulate the model to examine how alternative values of the transition rates affect the steady-state age distribution. Fig. 6 *E* and *F*, and *SI Appendix*, Table S2 and Fig. S22 indicate that changing the retirement hazard from the 2008 level to the 1993 level would imply a mean age of the steady-state age distribution of 49.7, 1.2 y lower than the 50.9 y implied by the 2008 hazard (Fig. 6*E*) and a 0.040 smaller share of the science workforce over age 55. Fig. 6 *B–D* illustrates the implied steady-state distribution (red triangles for a variety of counterfactuals). Fig. 6 *E* and *F* and *SI Appendix*, Table S2 show that the total fertility rate of 3.76 in 1960, compared with 1.85 in 1980 and 1.93 in 2010, would imply a mean steady-state age of 47.6 y and a share of scientists over 55 of 0.283, both of which are lower than the observed 2008 levels.

We simulate the impact of immigration of scientists who obtained a PhD abroad by comparing a hypothetical scenario of zero immigration with the observed immigration level shown in *SI Appendix*, Fig. S20. Fig. 6 and *SI Appendix*, Table S2 and Fig. S23 show that, in the absence of any immigration, the steady-state mean age would be 50.2, 0.7 y younger than in the steady state implied by current immigration, and the share age 55 and above would be 0.371, or 0.022 lower. This rather surprising finding is a consequence of the older average age of entry to the scientific workforce by scientists trained outside the United States compared with US-trained scientists.

The other rows of Fig. 6 *E* and *F* and *SI Appendix*, Table S2 indicate that changes in mortality, the PhD completion rate, and the share of foreign-born PhDs will have little impact on the steady-state age distribution of the scientific workforce.

Conclusions

Our major findings are that (*i*) the scientific workforce has aged rapidly in recent years relative to the workforce as a whole; (*ii*) the main causes have been a decline in the retirement rate of older scientists, which occurred after the elimination of mandatory retirement in universities, and a convergence to the steady-state distribution as the baby boom cohort has aged; and (*iii*) current trends imply a further substantial increase in the age of the scientific workforce in coming years. Although we have taken entry and transition rates as given, if instead one assumes that the size of the scientific workforce is largely fixed, then these factors may further crowd out young scientists. However, this "lump of labor" hypothesis has been tested and rejected in many contexts (16). The implications of these findings depend on whether and how rapidly scientific productivity declines with age, and whether the life cycle pattern of scientific productivity will change in response to the aging of the scientific workforce. If scientific productivity is much lower at older ages, and if this is mainly due to inherent physiological factors, then the aging of the scientific workforce will have an adverse impact on scientific productivity in the United States.

We acknowledge limitations in our study resulting from the fact that scientists without a US science doctorate, including physician scientists who do not have a PhD, and scientists trained abroad are excluded. We also acknowledge that our simulation model is mechanical and does not account for possible behavioral responses to the changing age distribution of the scientific workforce in domains such as whether to obtain a science doctorate, whether to become part of the scientific workforce, and whether to focus on research.

ACKNOWLEDGMENTS. We thank Wei Cheng and Wei Yang Tham for outstanding research work; seminar participants at the Institute of Labor Economics (IZA) and the Berlin Applied Micro Seminar for helpful comments; John Ham, Donna Ginther, participants at the Innovation in an Aging Society Meetings, for comments; and the NIH Advisory Committee to the Director Biomedical Research Workforce Working Group for inspiration. We both gratefully acknowledge the support from the NIH (National Institute on Aging, Office of Behavioral and Social Sciences Research, and NSF's Science of Science and Innovation Policy Program via Grant P01 AG039347). B.A.W. also thanks the NSF (Grants 1064220, 1348691, and 1535399) and the Kauffman and Sloan Foundations. Any opinions, findings, and conclusions or recommendations in this material are those of the authors and do not necessarily reflect the views of our funders.

1. Teitelbaum MS (2008) Research funding. Structural disequilibria in biomedical research. *Science* 321(5889):644–645.
2. Teitelbaum MS (2014) *Falling Behind: Boom, Bust, and the Global Race for Scientific Talent* (Princeton Univ Press, Princeton).
3. Alberts B, Kirschner MW, Tilghman S, Varmus H (2014) Rescuing US biomedical research from its systemic flaws. *Proc Natl Acad Sci USA* 111(16):5773–5777.
4. Heggeness ML, Carter-Johnson F, Schaffer WT, Rockey SJ (2016) Policy implications of aging in the NIH-funded workforce. *Cell Stem Cell* 19(1):15–18.
5. Lehman HC (1953) *Age and Achievement* (Princeton Univ Press, Princeton).
6. Simonton DK (1990) *Scientific Genius: A Psychology of Science* (Cambridge Univ Press, Cambridge, UK).
7. Levin SG, Stephan PE (1991) Research productivity over the life cycle: Evidence for academic scientists. *Am Econ Rev* 81(1):114–132.
8. Jones BF (2009) The burden of knowledge and the "death of the renaissance man": Is innovation getting harder? *Rev Econ Stud* 76(October):283–317.
9. Jones BF, Weinberg BA (2011) Age dynamics in scientific creativity. *Proc Natl Acad Sci USA* 108(47):18910–18914.
10. Kaiser J (2008) The graying of NIH research. *Science* 322(5903):848–849.
11. National Institutes of Health (2016) *New and Early Stage Investigator Policies*. Available at https://grants.nih.gov/policy/new_investigators/index.htm. Accessed January 21, 2017.
12. White JC, Rush M, Schaffer WT (2009) *Workforce Modeling for the National Institutes of Health (NIH)* (National Institutes of Health, Bethesda, MD).
13. Ashenfelter O, Card D (2002) Did the elimination of mandatory retirement affect faculty retirement? *Am Econ Rev* 92(4):957–980.
14. Lee RD (1994) The formal demography of population, aging, and the economic life cycle. *Demography of Aging*, eds Martin LG, Preston SH (National Academy Press, Washington, DC), pp 8–49.
15. Grogger J, Hanson GH (2015) Attracting talent: Location choices of foreign-born PhDs in the United States. *J Labor Econ* 33(3, Part 2):S5–S38.
16. Gruber J, Wise DA (2010) Social Security and Retirement Programs Around the World: The Relationship to Youth Unemployment (Univ of Chicago Press, Chicago).

Supporting Information

Why the United States Science and Engineering Workforce Is Aging Rapidly

David Blau and Bruce Weinberg, the Ohio State University

January 16, 2017

1. Data

The main source of data on US scientists is the Survey of Doctorate Recipients (SDR), a longitudinal sample survey of the population with a research doctorate in science, engineering, or health, earned in the US (http://www.nsf.gov/statistics/srvydoctoratework/). The sampling frame is the Survey of Earned Doctorates, an annual census of individuals receiving a research doctorate (http://www.nsf.gov/statistics/srvydoctorates/). Once in the sample, individuals are surveyed repeatedly until age 76. The sample is refreshed with new doctorate recipients at each wave. The SDR is sponsored by the National Science Foundation, and is usually administered every two years, with an occasional two and a half year gap. We use restricted-access microdata from nine survey waves conducted from 1993 through 2010, with detailed information on age, field of degree, job tenure, previous employment, occupation, and sector of employment.[1] The data contain observations on about 73,000 scientists aged 76 or less, with an average of four observations per sample member. We supplement the SDR with data from the Census Bureau.[2]

We define scientific workers as individuals with a research doctorate who are currently employed and work in a scientific occupation. Scientific occupations include the life sciences (biology, medical science, etc.), health-related occupations[3], physical sciences (chemistry, physics, astronomy, and geology), engineering, computer science and mathematics, and social science (economics, psychology, etc.). Scientific workers are employed in universities, hospitals, national laboratories, for-profit and not-for-profit corporations, and federal and state government

[1] Earlier waves are available but a redesign for the 1993 wave makes it difficult to use them in a trend analysis.

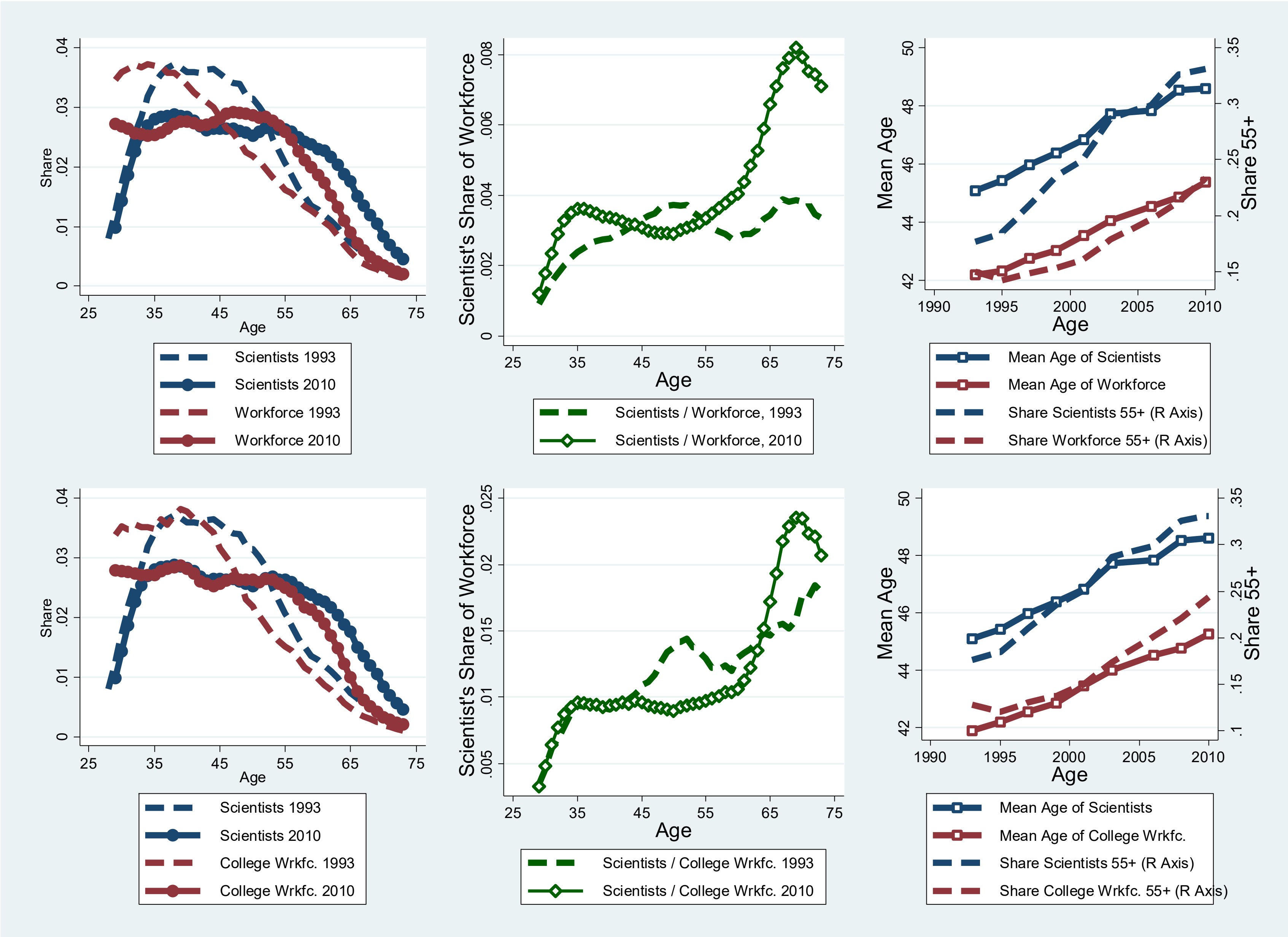
[2] The Census data were downloaded from IPUMS (https://usa.ipums.org/usa/).

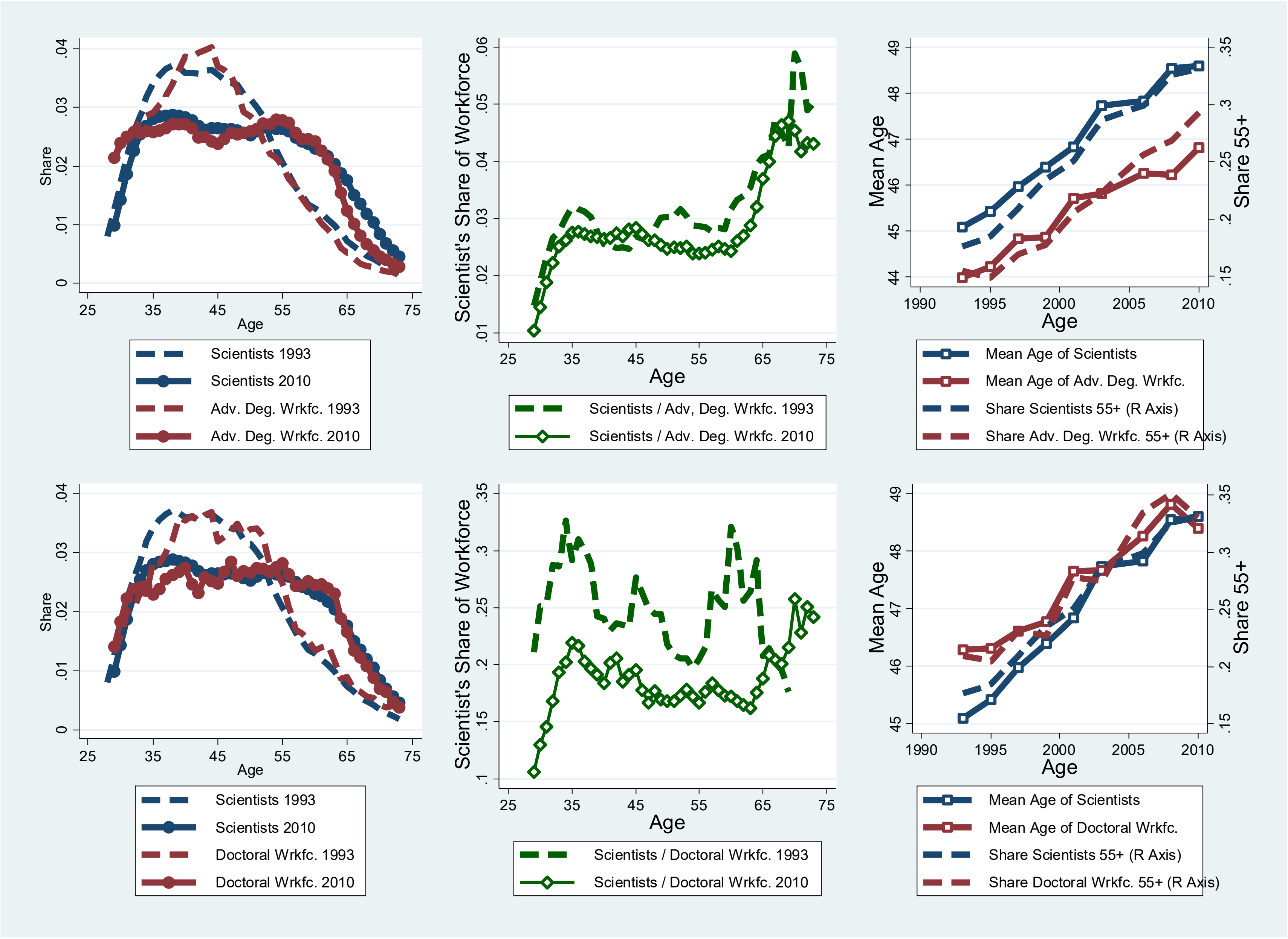
[3] Health-related occupations include those physicians and other diagnosing and treating practitioners, nurses, pharmacists, and others who have a research doctorate.

agencies. 85% of the individuals we classify as scientists report that they were engaged in a research-related activity (basic, applied, development, programming, and design of equipment, structures, models, and processes). Over three-quarters (76%) reported that one of these activities is their primary or secondary activity, based on hours of effort.[4]

We use data from the American Community Survey (ACS) for information on scientists who obtained a PhD outside of the United States and then migrate to the US. We assume somewhat arbitrarily that if an individual with a PhD arrived in the US at age 32 or older, the PhD was completed abroad. However, the sample sizes are too small to produce meaningful results by year and age at arrival. Instead, we combine data from the ACS for the years 2000-2013 to compute the average annual *number* of recently arrived immigrant scientists. We cannot compute the *rate* of arrival because we do not have data on the population of scientists who obtained the PhD outside the US.

[4] We experimented with alternative definitions of the scientific workforce, based on whether research was the primary or secondary activity, and the results were very similar. We also tried a specification in which no distinction was made between scientists and non-scientists – all SDR respondents were treated as scientists regardless of their reported activity and occupation. This yielded virtually identical results.


Share
Age
Scientists 1993
Scientists 2010
Workforce 1993
Workforce 2010
Scientist's Share of Workforce
Age
Scientists / Workforce, 1993
Scientists / Workforce, 2010
Mean Age
Share 55+
Age
Mean Age of Scientists
Mean Age of Workforce
Share Scientists 55+ (R Axis)
Share Workforce 55+ (R Axis)
Share
Age
Scientists 1993
Scientists 2010
College Wrkfc. 1993
College Wrkfc. 2010
Scientist's Share of Workforce
Age
Scientists / College Wrkfc. 1993
Scientists / College Wrkfc. 2010
Mean Age
Share 55+
Age
Mean Age of Scientists
Mean Age of College Wrkfc.
Share Scientists 55+ (R Axis)
Share College Wrkfc. 55+ (R Axis)



Share
Age
Scientists 1993
Scientists 2010
Adv. Deg. Wrkfc. 1993
Adv. Deg. Wrkfc. 2010
Scientist's Share of Workforce
Age
Scientists / Adv, Deg. Wrkfc. 1993
Scientists / Adv. Deg. Wrkfc. 2010
Mean Age
Share 55+
Age
Mean Age of Scientists
Mean Age of Adv. Deg. Wrkfc.
Share Scientists 55+ (R Axis)
Share Adv. Deg. Wrkfc. 55+ (R Axis)
Share
Age
Scientists 1993
Scientists 2010
Doctoral Wrkfc. 1993
Doctoral Wrkfc. 2010
Scientist's Share of Workforce
Age
Scientists / Doctoral Wrkfc. 1993
Scientists / Doctoral Wrkfc. 2010
Mean Age
Share 55+
Age
Mean Age of Scientists
Mean Age of Doctoral Wrkfc.
Share Scientists 55+ (R Axis)
Share Doctoral Wrkfc. 55+ (R Axis)

Figure SI1. The age distribution of the US scientific workforce and the US workforce, by level of education. The left panels show the age distribution of scientists (calculated from the Survey of Doctorate Recipients (SDR)) and the US workforce (calculated from the Current Population Survey (CPS)) for 1993 and 2010. The center panels show the share of scientists in the US workforce by age in 1993 and 2010. The right panel plots trends in the mean age of scientists and the US workforce as well as the share of scientists and the US workforce age 55 and over. The top panel includes the entire workforce (regardless of education). The second panel down focuses on the college workforce, defined as those with a 4 year college degree (or 16 years of school) or more. The third panel down focuses on workers with advanced degrees, defined as those with a masters or equivalent (or 17 years of school) or more. The last panel focuses on workers with doctoral degrees, defined as those with a doctorate or equivalent (or 20 years of school) or more. The figures show that the scientific workforce is aging overall and relative to the various populations. The aging is particularly pronounced when looking at the share of the workforce that is 55 and older. The aging of the scientific workforce is smallest relative to people with advanced degrees, but clear relative to the other groups, including people with doctorates.

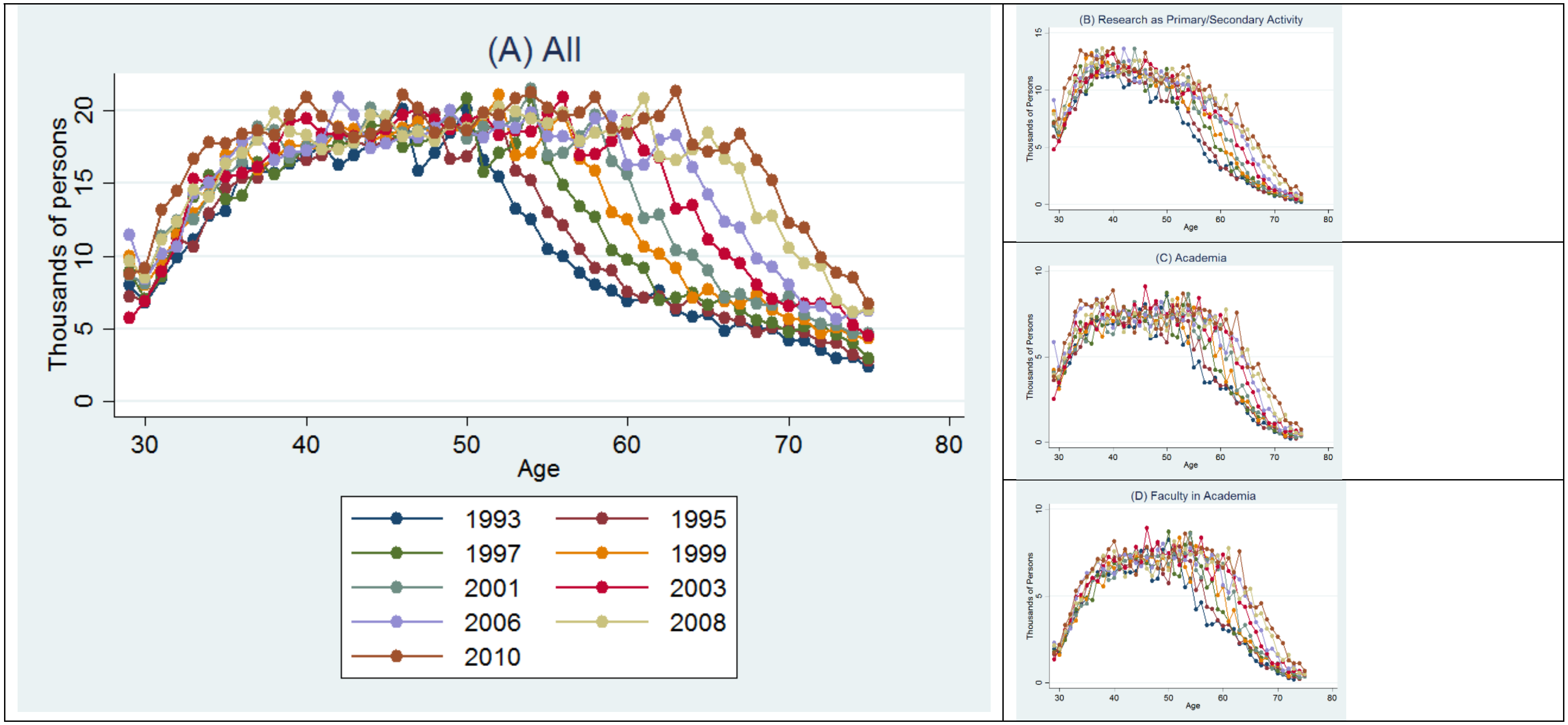


Figure SI.2. Age Distribution by Year. Panel A plots the size of the US research doctorate workforce in thousands, by age and year from 1993 to 2010. There are few research doctorates in their 20s, with a rapid increase in numbers around age 30. The counts continue to increase until the early 40s when they plateau. There is little discernable age trend from the early 40s to age 50. The age at which counts increase are relatively constant over time and while the counts begin to fan out in the 30s and remain spread out in the 40s, they do not exhibit clear trends. By contrast the data after age 50 show a remarkably strong pattern. The series for 1993 begins to turn down around age 50. That for 1995 begins to turn down two years later, 1997 is delayed by two more years. Each successive curve runs remarkably parallel to the series before it. Thus the aging of the workforce can be seen in the downturn in counts occurring at ever later ages. An important implication of the lack of consistent changes in the plateau and ever later downturns is that the size of the workforce is growing consistently over time. Estimates for people whose primary and secondary work activity is research (in Panel B) peak earlier, indicating that people move away from research positions as they age.

Academics as a whole (shown in Panel C) are similar to A but the curves for the later years remain high until later ages. Panel D shows the distribution for non-postdocs in academia. Because many people are in postdocs before age 40, this figure increases more rapidly until age 40. Despite these intuitive differences, the four samples show a remarkable and consistent aging pattern.

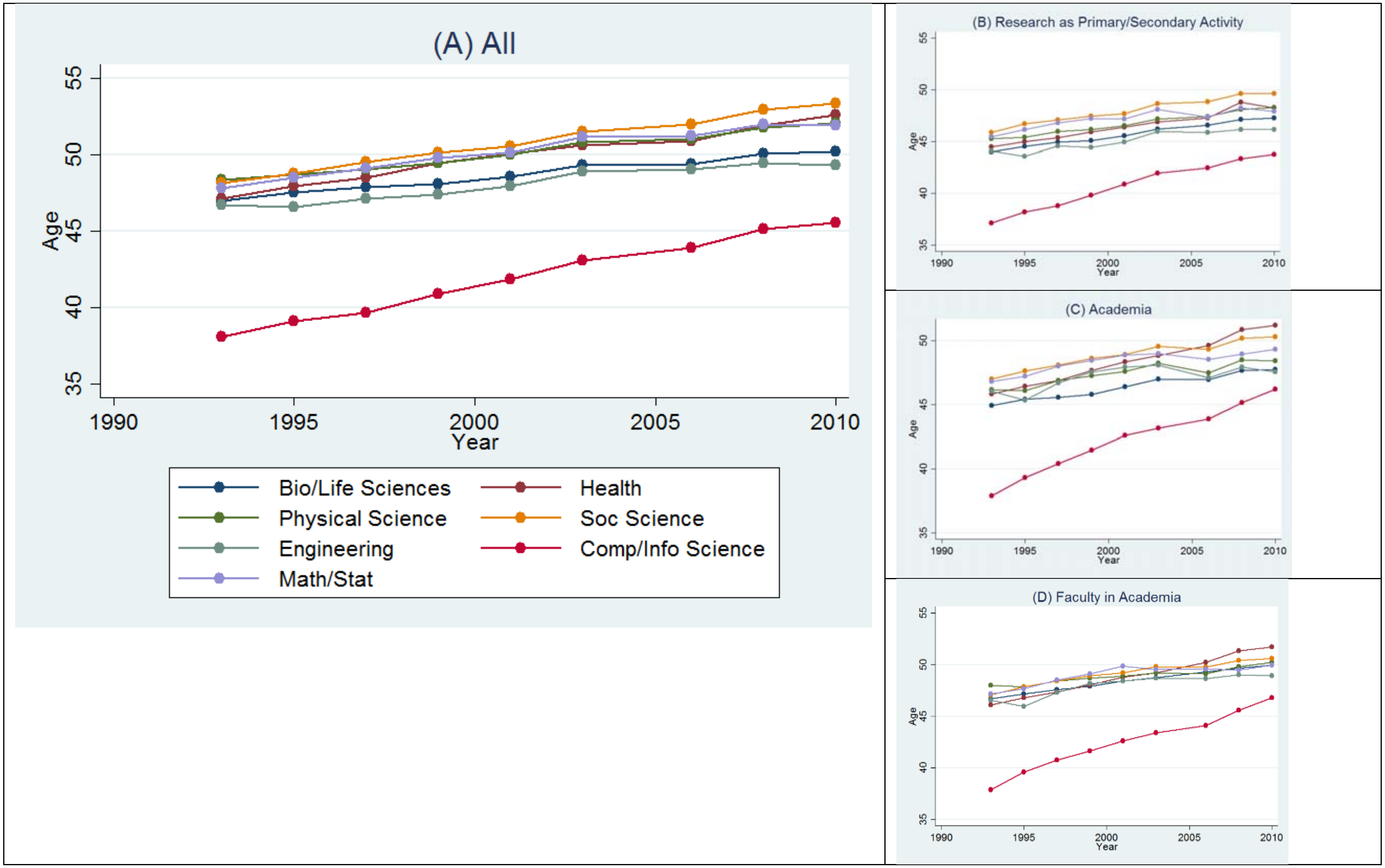


Figure SI.3. Mean Ages by Field. Panel A reports the mean age of the U.S. research doctorate workforce by field over time. The estimates show that the aging pattern is common across fields. The mean age in most of the fields is between 46 and 49 in 1993 and increases to between 49 and 54 in 2010. Computer and Information Sciences are the notable exception, with a mean age of only 38 in 1993. At the same time, Computer and Information Sciences age more rapidly than any of the other fields (by 7 years) to 2010. Within the other cluster of fields, the social sciences are the oldest while Engineering is the youngest. The aging of the biomedical research workforce, including the ages of principal investigators funded by the National Institutes of Health (NIH) has received considerable attention. It is therefore noteworthy that while Health shows the greatest increase in age among the non-Computer and Information Sciences cluster, the Biological and Life Sciences is among the youngest and does not age more rapidly than the other fields. As indicated by Panel B, scientists primarily or secondarily engaged in research are somewhat younger than the average scientist, indicating a tendency to move away from research at older ages. Academics (in panel C) are noticeably older than people whose primary or secondary activity is research, but the patterns are broadly similar to those in Panel A. Scientists employed in academia are only slightly older than those employed outside of academia (mean age of 45.9 in academia versus 45.7 outside). Perhaps not surprisingly, the largest difference between academia and non-academia is for engineering, where the age of academics is considerably older than their non-academic counterparts. Panel D further restricts the sample of academics to scientists who are not in postdocs. This restriction naturally increases the average age, with among the largest effects in biomedicine in recent years. This reflects the large size of the postdoc population in biomedicine. However, even after postdocs are excluded, the Biological and Life Sciences are not exceptionally old. (We have explored a range of methods to identify postdocs (1).

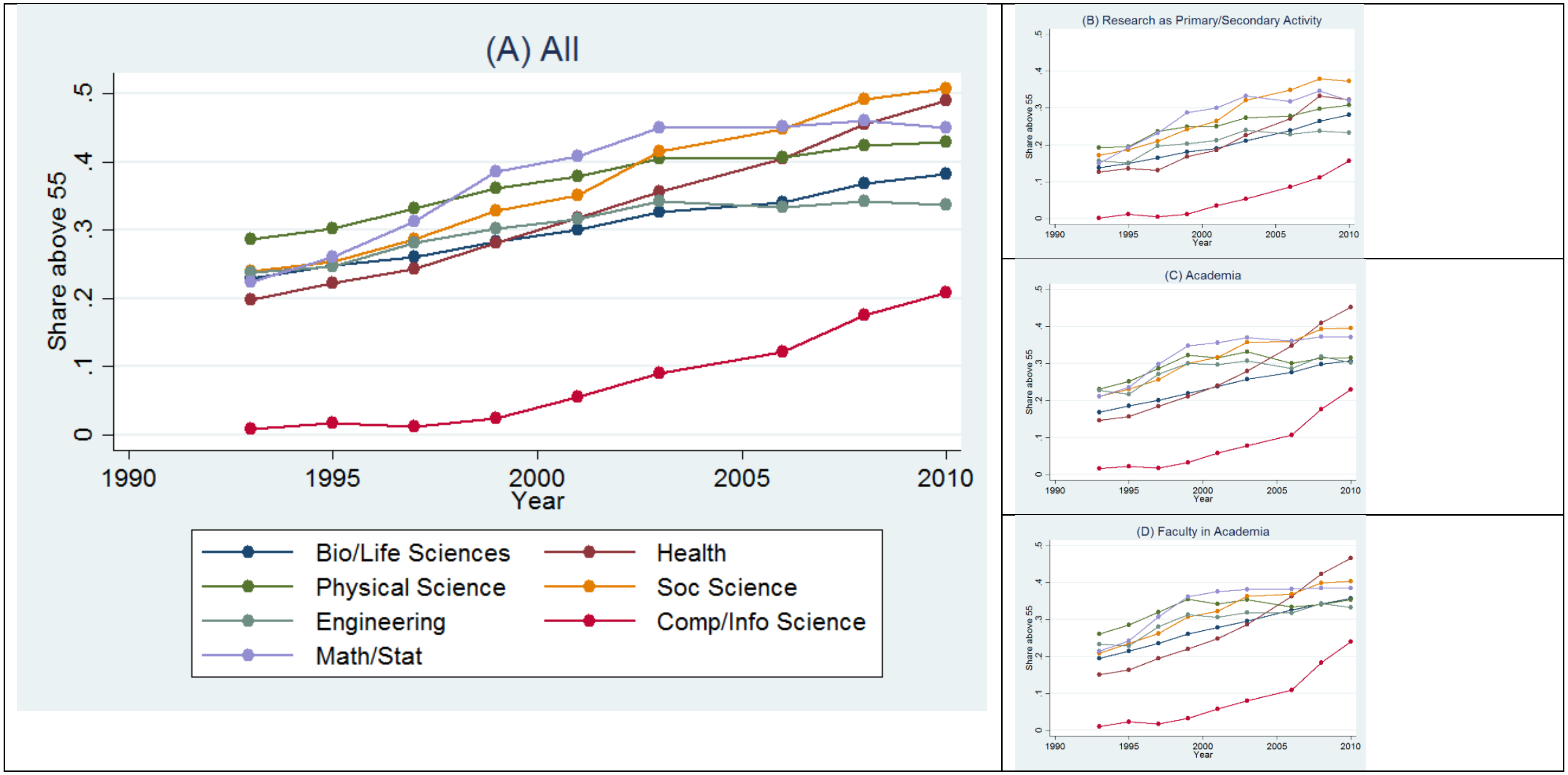


Figure SI 4. Share of Field above 55. The figure shows the share of the STEM workforce by field and position over age 55. Panel A reports the share of the U.S. research doctorate workforce over 55 by field over time. Like Figure SI.3, the estimates show that the aging pattern is common across fields, with Computer and Information Sciences initially being younger, but aging more quickly. Panel B shows that people primarily or secondarily engaged in research are somewhat younger than the average research doctorate. Academics (in panel C) are noticeably more likely to be over 55 than those whose primary or secondary activity is research. Academics are somewhat more likely to be over 55 (23.2%) than those employed outside of academia (21.3) (not shown). Again, the patterns are broadly similar to those in Panel A. Panel D further restricts the sample of academics to people who are not in postdocs.

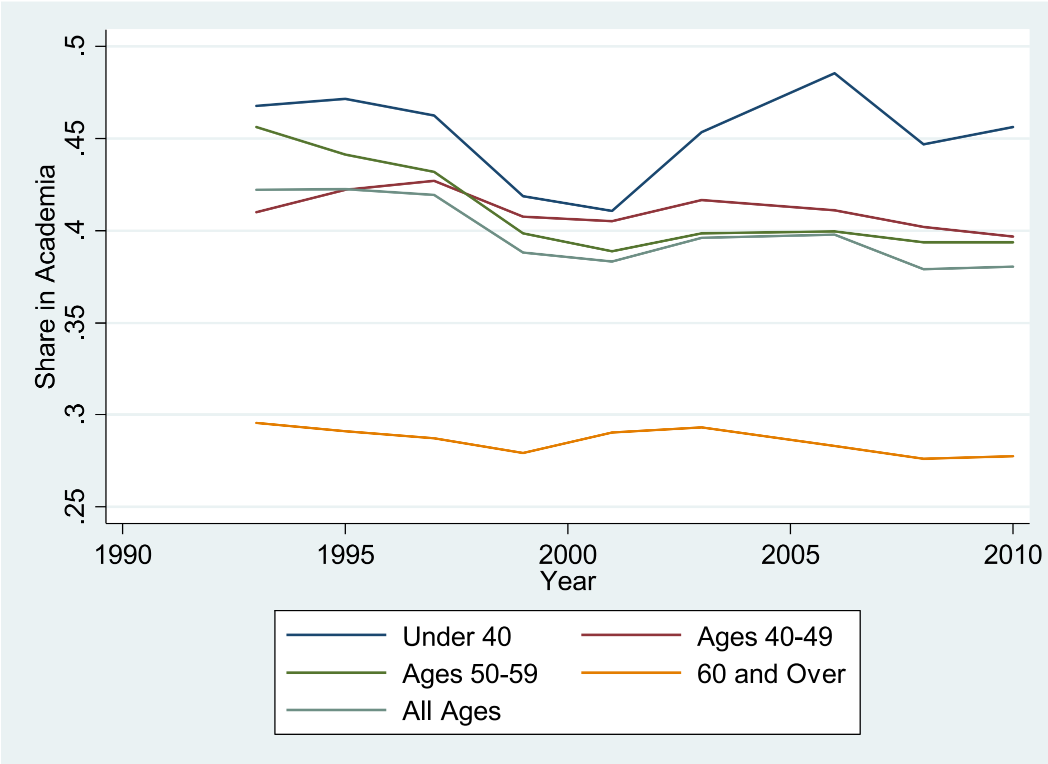


Figure SI.5. Share of scientists in academia. The figure shows that the share of scientists who are in academia declines with age and has been essentially flat within age categories. The share of scientists in academia (across all age groups) has declined because the scientific workforce has aged and older doctorates are less likely to be in academia than younger ones.

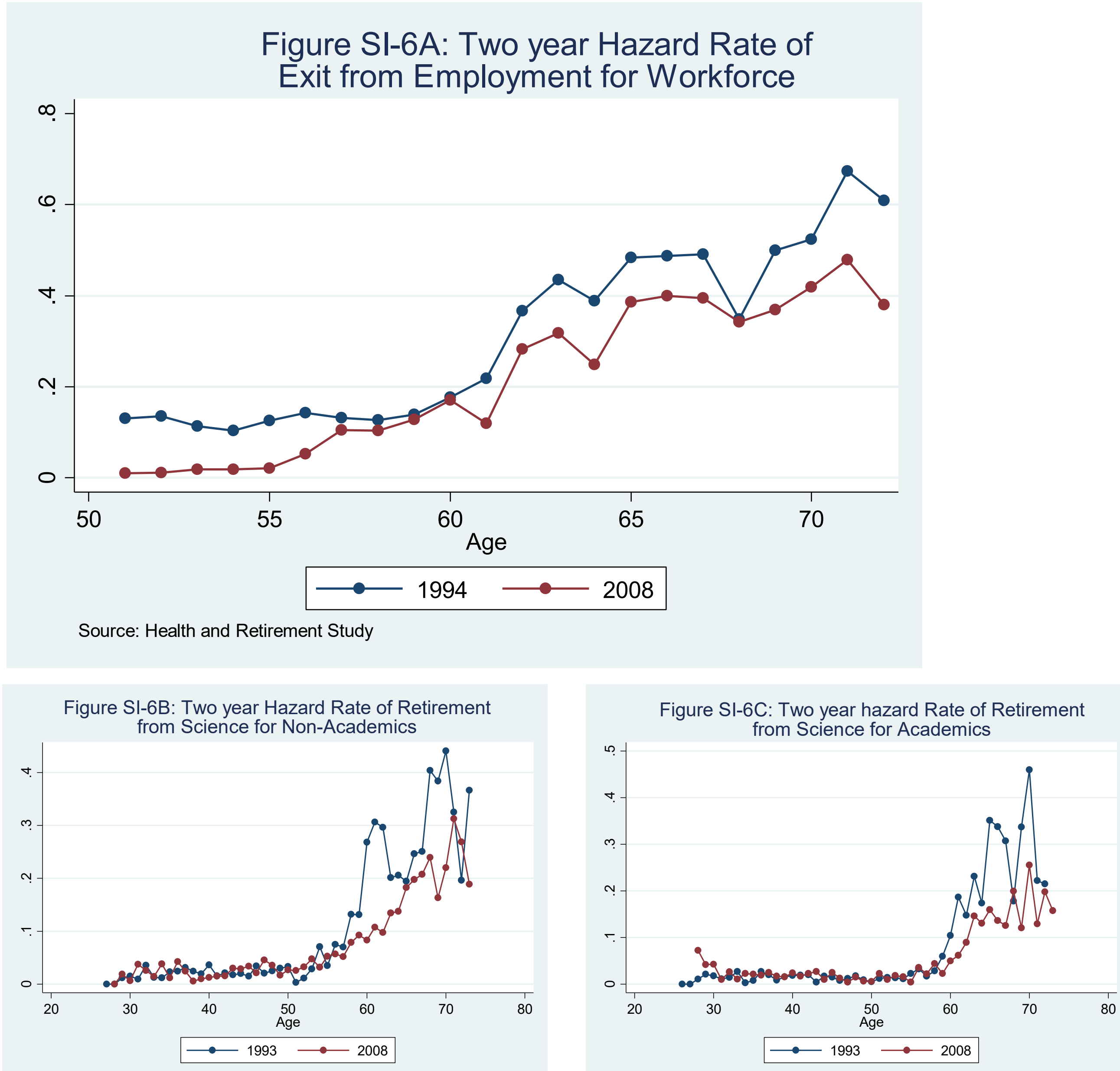


Figure SI.6. Panel A shows biannual transition rates from employment to non-employment for the US workforce, 1994 and 2008. The figure shows the share of employed workers 1994 (blue) and 2008 (red), who are not employed two years later. Panels B and C shows biannual transition rates from employment to non-employment for scientists employed outside of academia (Panel B) and in academia (Panel C) for 1993 (blue) and 2008 (red). These are calculated as the share of science doctorates employed in science in the 1993 (2008) survey, who were not employed as of the 1995 (2010) survey. The spike in the retirement hazard at age 70 in 1993 and the decline is more pronounced among scientists in academia than among those outside of academia.

2. Trends in Completion of a PhD in Science

We break the scientific PhD completion rate into three components: completion of a bachelor's degree, the share of science majors among bachelor's degree recipients, and the share of undergraduate science majors that completes a PhD in science. Figure SI-7 illustrates trends in these variables, going back to the 1960s.[5] The 4-year college completion rate as a share of births 22 years earlier has grown dramatically, from less than 20% in the 1960s to more than 40% since 2005. The share of bachelor's degrees earned in a scientific field trended down from 35% in the 1960's to 31% in 1990, and has fluctuated in a narrow range around 32% since the mid 1990's. The rate at which science majors who graduated with a bachelor's degree in year *t*-6 completed a science PhD by year *t* dropped from over 10% in the early 1970's to less than 6% in 1980.[6] It then rose back up to 8% by 1995, and has fluctuated between 6 and 8% since then. If we look at the 1993 to 2010 period covered by the SDR, there is no trend. Thus science is no more popular today than in the 1990s, but it is not much less popular. The dramatic increase in college completion has resulted in steady growth in the *number* of new science PhDs produced in the US, as shown in Figure SI-8, from around 12,000 in the 1960s to 35,000 recently.

---

[5] The data described in this paragraph are from the NCSES web site (https://nces.ed.gov/datatools/index.asp?DataToolSectionID=4), and are derived from the Survey of Earned Doctorates (SED) and the Survey of College Graduates (SCG).

[6] We use a six year lag to approximate the average time to completion of a PhD. The trend is very similar using other lags.

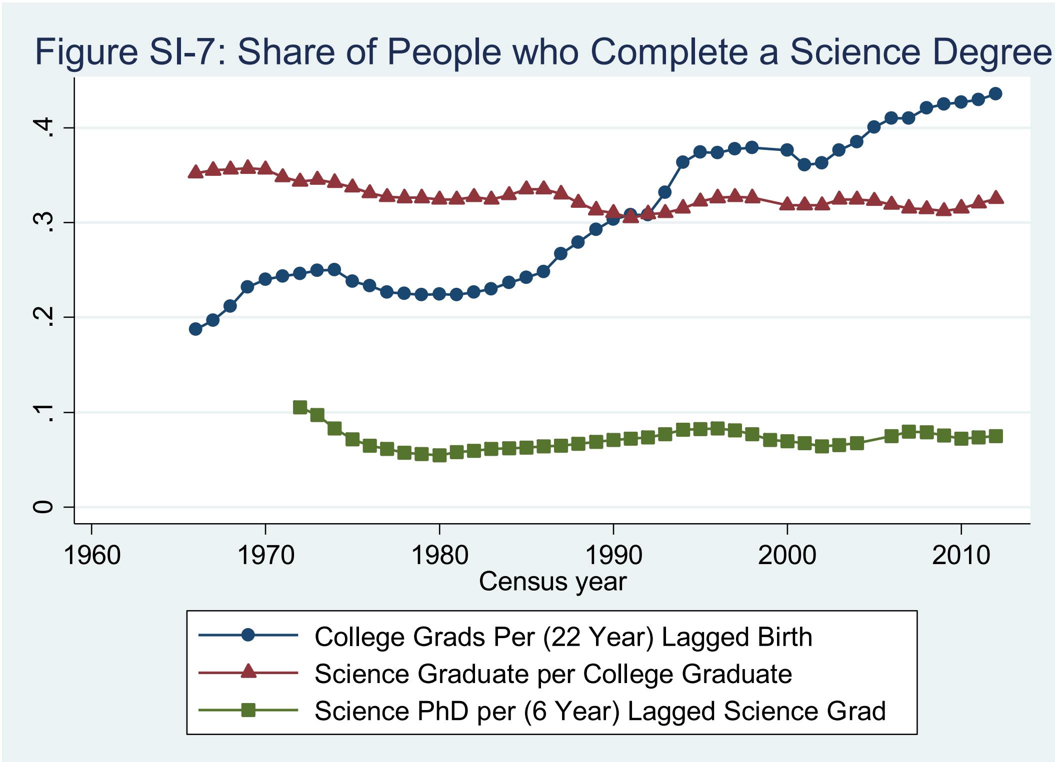


Figure SI-7: Share of People who Complete a Science Degree
0
.1
.2
.3
.4
1960
1970
1980
1990
2000
2010
Census year
College Grads Per (22 Year) Lagged Birth
Science Graduate per College Graduate
Science PhD per (6 Year) Lagged Science Grad

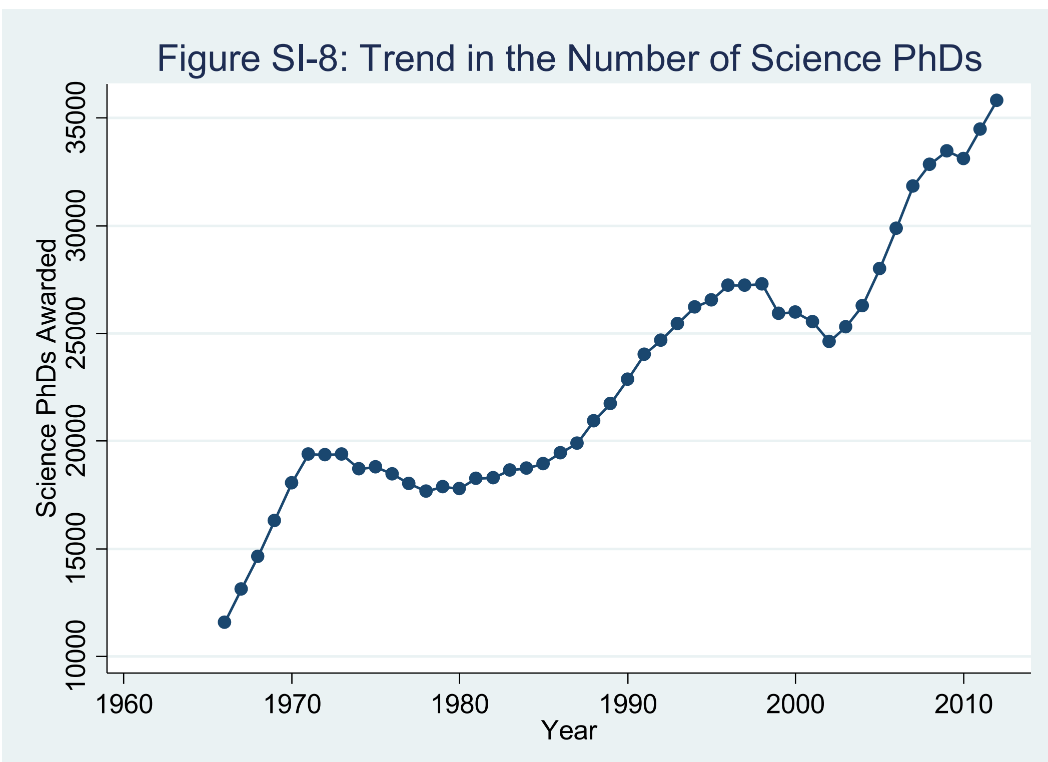


Figure SI-8: Trend in the Number of Science PhDs
Science PhDs Awarded
10000
15000
20000
25000
30000
35000
1960
1970
1980
1990
2000
2010
Year

3. Model

Here, we analyze the change in the age distribution of scientists between 1993 and 2008.[7] We use a simulation model to generate a predicted 2008 age distribution of the scientific workforce conditional on the observed 1993 distribution and the observed set of survey-wave-and-age-specific transition rates among the employment states of (1) scientific worker, (2) non-scientific worker, and (3) not employed.[8] The population is individuals who earned a PhD in a science field in the US. The model incorporates hazard rates for movements among the three states, as well as birth rates, death rates, PhD completion rates, and the share of foreign-born among US science PhDs (for brevity we refer to these factors collectively as transition rates). We use the model to explore explanations for the aging of the scientific workforce by conducting counterfactual simulations in which each transition rate is held fixed at its 1993 age-specific values.

Define N_{at}^{S} as the number of individuals with a PhD in science who are employed as scientific workers of age a in year t, N_{at}^{N} as the number of individuals with a PhD in science who are employed outside of science, and N_{at}^{O} as the number of individuals with a PhD in science who are out of the labor force. Define h_{at}^{ij} as the biannual hazard rate for a transition from state *i* to state *j*, *i*, *j* = *S*, *N*, *O*, between years *t* and *t*-2. The age-specific transition equation for the scientific workforce from *t*-2 to *t* is

$$N_{at}^{S} = N_{a-2,t-2}^{S}(1 - h_{at}^{sn} - h_{at}^{so} - m_{sat}) + N_{a-2,t-2}^{N}h_{at}^{ns} + N_{a-2,t-2}^{O}h_{at}^{os} + (B_{t-a}g_{a-1,t-1} + B_{t-a-1}g_{a-2,t-2})(e_{at} + \alpha_t f_{at})$$

Here, m_{sat} is the biannual age-a mortality rate of scientists, B_{t-a} is the number of births *t*-*a* years ago, g_{at} is the proportion of individuals born in period *t*-*a* who obtain a science PhD in period *t* at age a, e_{at} is the proportion of US science PhD graduates who obtained a PhD at age a-1 in year *t*-1 or age a-2 in year *t*-2 that are in the scientific workforce at *t*, f_{at} is the corresponding proportion for foreign students, and α_t is the ratio of foreign to native US students

[7] We use 2008 instead of 2010, the last year of the SDR available to us, because the 2010 SDR does not contain complete data on all PhDs awarded in 2009.

[8] We do not distinguish scientists employed in academia from those employed in industry because the differences in the age distribution are not large, nor are there large time trends – roughly 52% of scientists are in academia compared to 48% outside. See figure S.5.

completing a US scientific PhD program. h_{at}^{so} is the biannual hazard rate of exiting the scientific workforce to out of the labor force, i.e. the retirement hazard. Note that we have dropped the intermediate step of obtaining a Bachelor's degree in science because we do not have information on the number of students completing such a degree by age in a given year. We allow for PhDs obtained at ages 27-57, since there is quite a dispersed distribution of the age at PhD completion. We have annual data on new PhDs, so we aggregate these data to a biannual basis, as indicated in the equation. There are similar equations for transitions into non-science employment and out of the labor force.

The share of science PhDs earned by women has grown from 10% in 1970 to more than 40% in 2010 (see Figure SI-13) and women have longer life expectancy. We modify the model to allow for sex-specific mortality and PhD completion rates as follows. Let the r superscript denote gender, r = m for males and f for females.

$$N_{at}^{Sr} = N_{a-2,t-2}^{Sr}(1 - h_{at}^{sn} - h_{at}^{so} - m_{sat}^{r}) + N_{a-2,t-2}^{Nr}h_{at}^{ns} + N_{a-2,t-2}^{Or}h_{at}^{os} + (B_{t-a}g_{a-1,t-1}^{r} + B_{t-a-1}g_{a-2,t-2}^{r})(e_{at} + \alpha_t f_{at})$$

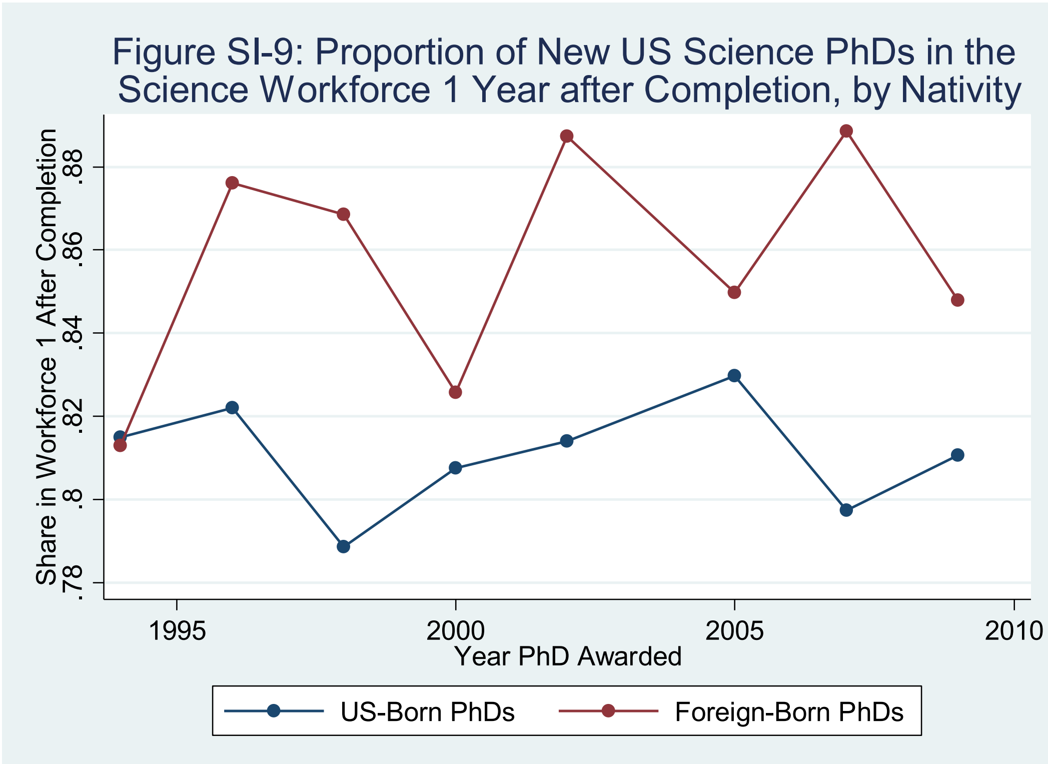
There are similar transition equations for non-science and not employed, in each case differentiated by gender. Note that we allow mortality and the PhD completion rate to differ by gender, but we assume that the hazard rates, the initial entry rates into science (e and f) and the share foreign born are the same. We relax this restriction in Section SI.6.

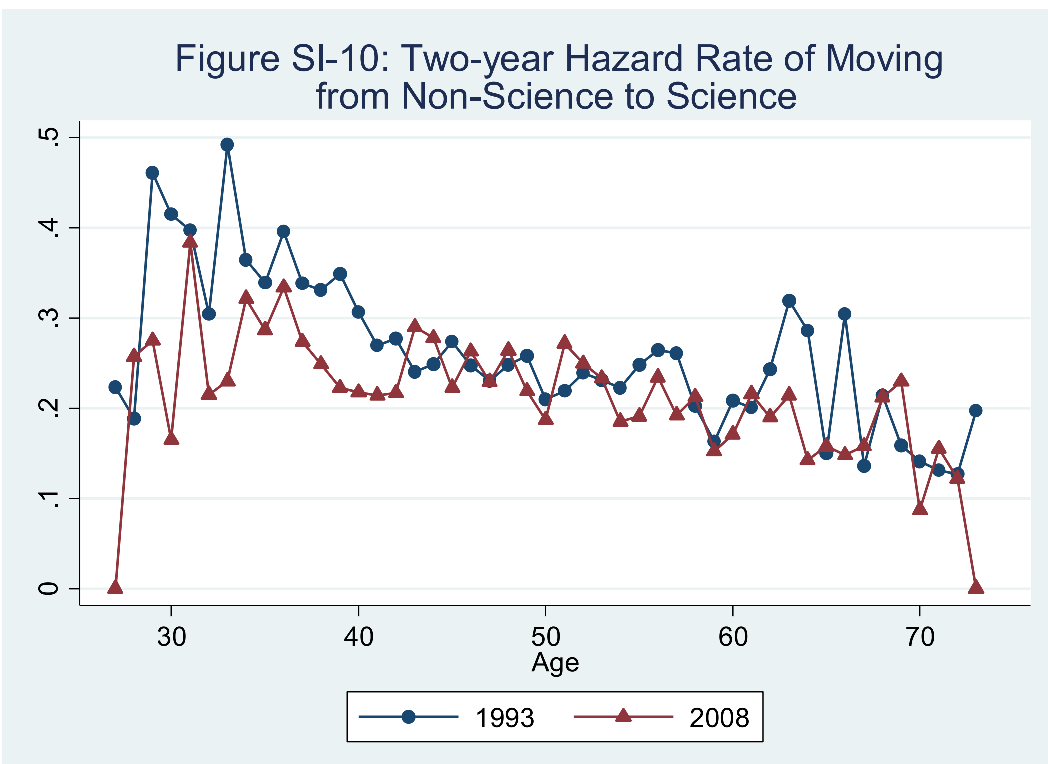
As noted in the main text, we use biannual hazard rates because the SDR has been conducted biannually, with one triannual exception. The precise length of the intervals between the surveys is not always exactly two (or three) years. This is not a problem because our goal is to explain the observed changes in the age-specific stock of scientists from wave to wave, and the hazard rates for transitions from wave *w* to wave *w*+1 should correspond to the actual interval between the periods. We do not attempt to adjust the biannual mortality and PhD completion rates for differing intervals between survey waves.

4. Entry to the Scientific Workforce

Figure SI-9 shows the trend in the share of individuals who obtained a science PhD in year *t* who are part of the US scientific workforce in year *t*+1, by nativity. Roughly 80% of natives who obtain a US science PhD enter the scientific workforce quickly, within one year. Foreign-born students who obtained a PhD in the US and who stayed in the US are on average about five percentage points *more* likely than US-born PhDs to be in the US scientific workforce one year after the PhD, indicating the important role of foreign born recipients of US PhDs. Figure SI-10 shows the biannual hazard rate of entry to the scientific workforce from a non-science occupation, conditional on having a science PhD.[9] There is considerable mobility of individuals with a science PhD into the scientific workforce, with evidence of a modest increase between 1993 and 2009 at younger ages.

[9] The hazard of entry is the proportion of all individuals with a US science PhD working in a non-science field in year *t* who had entered the scientific workforce by year *t*+2.

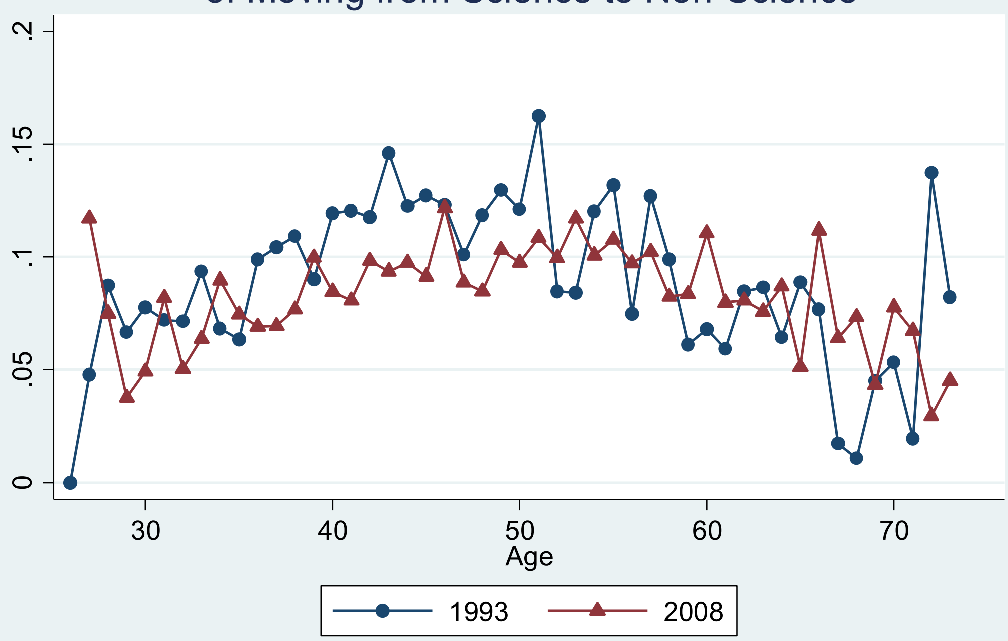




5. Exit from the Scientific Workforce

Figure 2 in the text shows the biannual hazard rate of exit from the scientific workforce to non-employment in 1993 and 2008. This includes exits for any reason, but we will refer to it as the hazard rate of retirement, since most exits to non-employment are in fact self-reported as being due to retirement. There is very little exit from the labor force until around age 58. In 1993, the retirement hazard at older ages looked much like the typical pattern of retirement, with large jumps at certain ages. The hazard rate of exit from the scientific workforce decreased substantially after age 60 between 1993 and 2008. The biannual hazard of retirement at age 70 declined by almost half from 0.45 in 1993 to 0.23 in 2008. The large spike at age 70 in 1993 had disappeared by 2008. There is evidence that this is due to the termination in 1994 of a special exemption from the 1986 Age Discrimination Act that allowed universities to impose mandatory retirement at age 70 (2). Figure SI-11 shows the hazard rate of exiting the scientific workforce to a non-science occupation in 1993 and 2008. The exit rate from science to non-science dropped by a modest amount between these years, mainly at younger ages. A lower rate of exit at younger ages would tend to increase the share of older workers.

Figure SI-11: Two-year Hazard Rate of Moving from Science to Non-Science



The third source of exit from the scientific workforce is death. Life tables from the Centers for Disease Control indicate that the annual risk of death at age 65, conditional on being alive at age

64, was .021 for men and .012 for women in 1993, and had declined to .015 and .009, respectively, by 2010. There is a strong mortality gradient by education, so the mortality decline among scientists may have been different than for the population as a whole. It is possible to estimate mortality rates by level of education using micro data from the National Health Interview Survey Linked Mortality File (3). The highest education category in their analysis is college graduate. Using their results, it is possible to compute the ratio of mortality of college graduates to mortality of the population as a whole for several age groups. In the 2000-2006 period, the ratios are 0.69 at 45-54, 0.55 at 55-64, and 0.51 at 65-74. We use these figures in the analysis below.

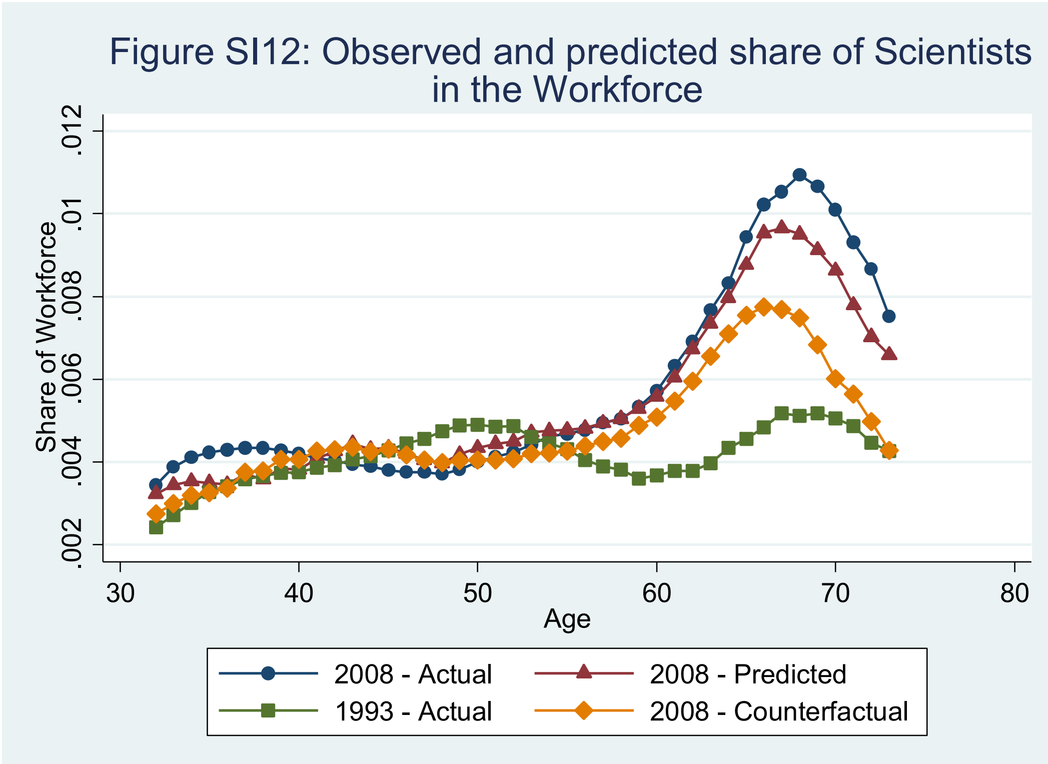


Figure SI-12 shows the age distribution of the share of scientific workers in the workforce in 1993 (green squares) and 2008 (blue circles). It also shows the 2008 distribution predicted by the model based on the observed hazard, PhD completion, and birth and death rates for each year (red triangles), and the counterfactual predicted 2008 distribution, using the same model, but holding all transition rates fixed at their 1993 values (yellow diamonds).

	Median of the age-specific fraction explained	Fraction of the change in mean age explained	Fraction of the change in the share age 55 and above explained
Observed transition rates	0.955	1.203	0.973
Holding all transition rates at 1993 levels	0.808	1.079	0.835
Fraction explained using observed transition rates except holding the following at 1993 values, one at a time:			
Retirement hazard	0.875	1.119	0.897
All hazards	0.878	1.053	0.874
Birth rate	0.842	0.963	0.832
Mortality rate	0.932	1.185	0.957
PhD completion rate and age distribution of new PhDs	1.022	1.440	1.059
PhD completion rate	0.929	1.310	0.959
Age distribution of new PhDs	1.046	1.348	1.072
Foreign born share of PhDs	0.958	1.208	0.977

Table SI.1. Fraction of the change in the age distribution of the scientific workforce from 1993 to 2008 explained by the simulation model. Fraction explained = (predicted 2008 – observed 1993) / (observed 2008 – observed 1993). The first row reports the fraction explained using the observed transition rates for each year (with averages of adjacent years for non-survey years). The second row shows the fraction explained if all transition rates had remained at their 1993 values. The remaining rows quantify the importance of each factor by using the observed 2008 rates except for the one indicated, for which the 1993 rate is used. For example, holding the retirement hazard constant at its 1993 level reduces the fraction explained from 0.955 to 0.875, for a loss of 0.08 in the fraction explained.

6. Model extensions

We extend the model in two directions in order to determine whether other factors might help account for the aging of the scientific workforce. First, if male and female scientists have different transition rates, in addition to the different mortality and PhD completion rates already incorporated in the model, the change in gender composition of the scientific workforce, shown in Figure SI-13, could affect the change in the overall age distribution. Figures SI-14 and SI-15 depict the biannual hazard rate of exit from science to non-science by gender in 1993 and 2008. The data are somewhat noisy (even using a five-year moving average) but show somewhat higher transitions out of scientific careers for women in the 40s, when many women have young children in the household. Otherwise the gender differences are modest. Figures SI-16 and SI-17 show the retirement hazards by gender. In both 1993 and 2008, women retired from science at a *slower* rate than men. As the female share of scientists increased, this would have tended to cause aging of the scientific workforce. However in quantitative terms, the gender difference is small compared to the rapid decline in the retirement hazard for both men and women. The quantitative results from a version of the model that allows hazard rates to differ by gender indicate that the change in gender composition of new PhDs can explain less than 2% of the observed aging of the scientific workforce, and none of the change in the age distribution of the fraction of the scientific workforce.

Figure SI-13: Share of Science PhDs earned by Females

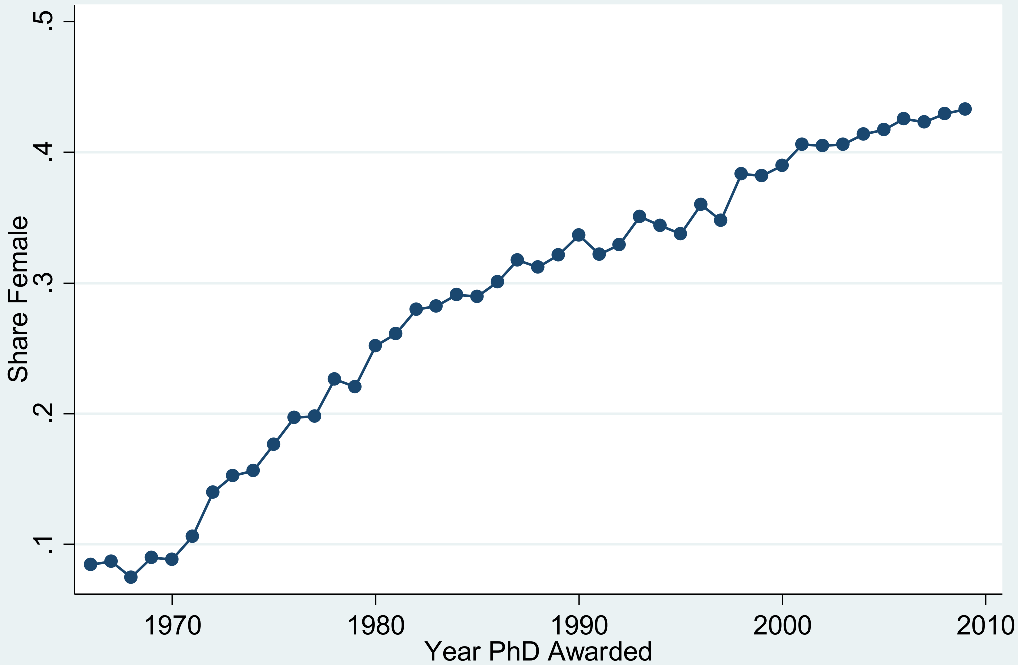
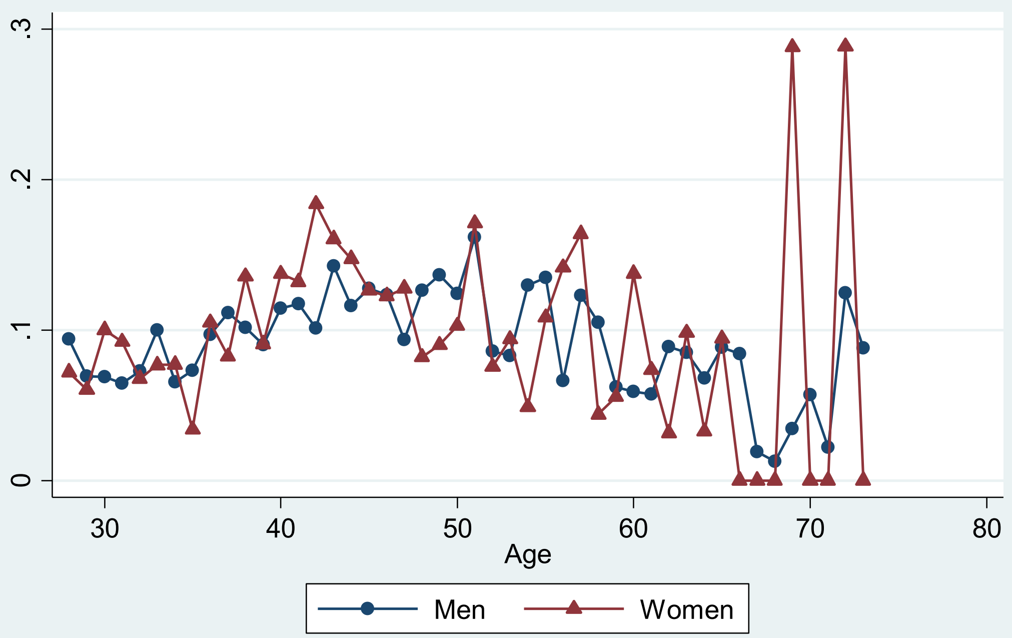


Figure SI-14: Two-year Hazard Rate of moving from science to Non-science by gender, 1993



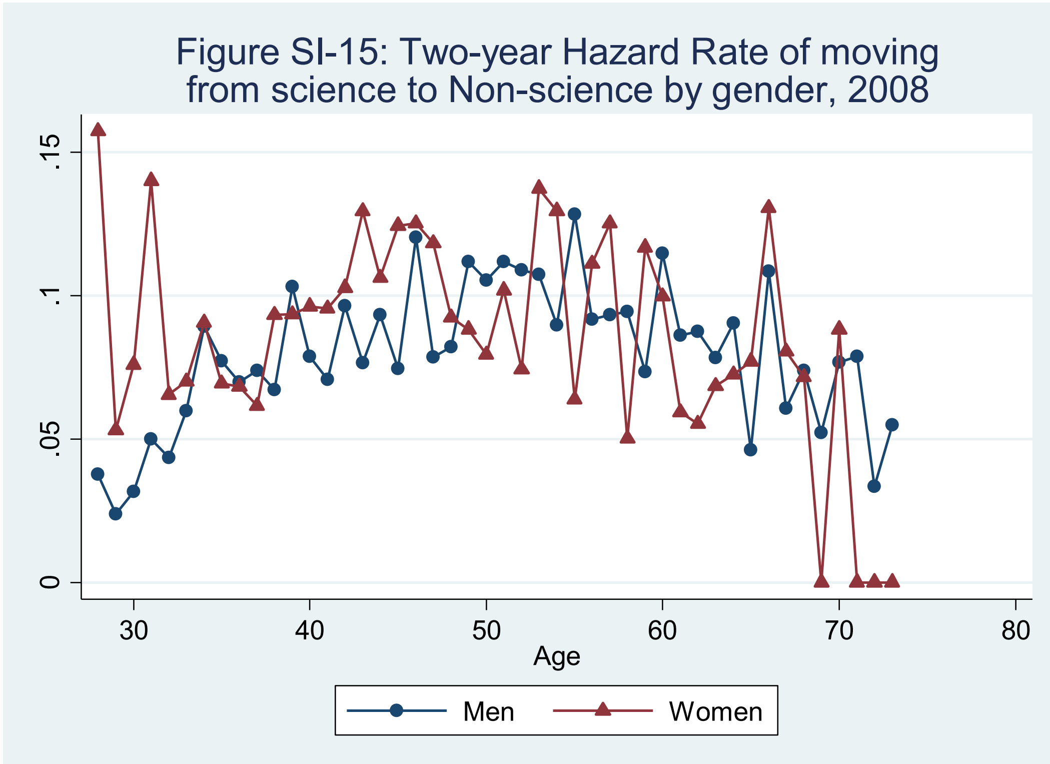


Figure SI-15: Two-year Hazard Rate of moving
from science to Non-science by gender, 2008
0
.05
.1
.15
30
40
50
60
70
80
Age
Men
Women

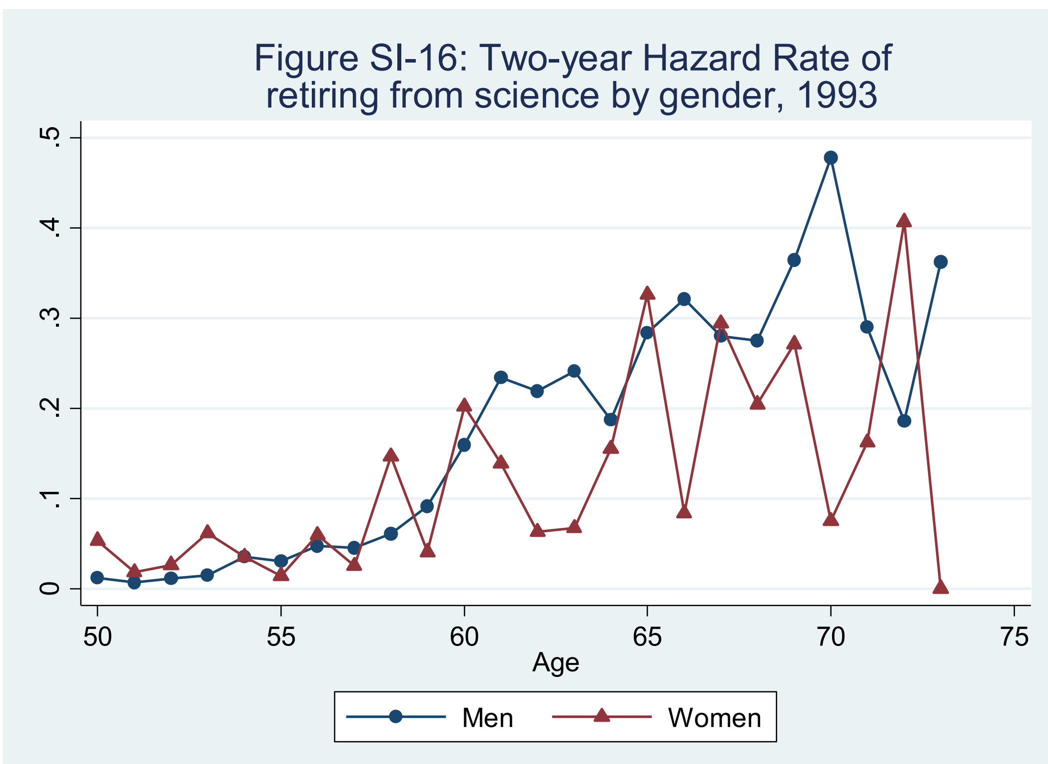
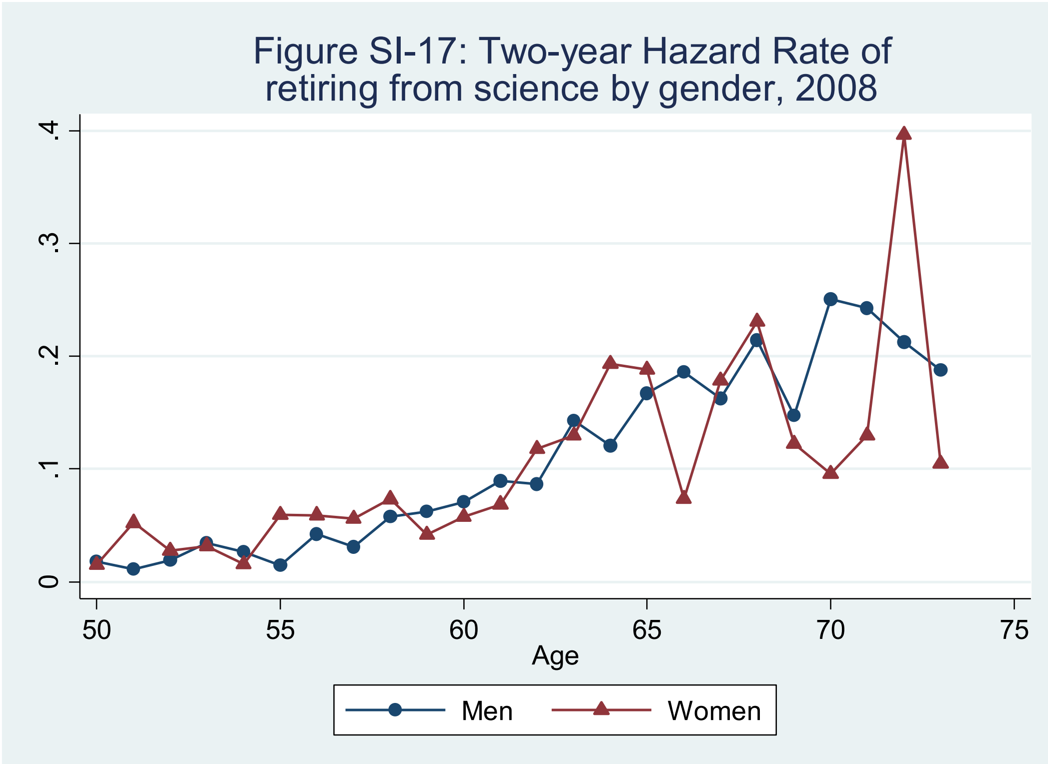


Figure SI-16: Two-year Hazard Rate of
retiring from science by gender, 1993
0
.1
.2
.3
.4
.5
50
55
60
65
70
75
Age
Men
Women



The second extension is to allow for differences in behavior of native and foreign-born scientists who earned a PhD in the US (recall that the SDR does not include scientists who immigrated to the US after earning a PhD abroad). If transition behavior differs by nativity, the increasing share of foreign-born PhDs could have affected the age distribution of the scientific workforce. However, the hazard rates for natives and non-natives who obtained the PhD in the US are quite similar. A quantitative analysis allowing hazard rates to differ by nativity was not possible because the sample of non-natives was too small.

7. Immigration

We rely here on the ACS/Census, which is the only source for information on immigrants who obtained a PhD outside the US (the sample size of immigrant scientists in the CPS is too small to be useful). The ACS/Census data do not contain information on the age at which the PhD was obtained. We assume somewhat arbitrarily that if an individual with a PhD arrived in the US at age 32 or above, the PhD was completed abroad. However, even in the ACS and Census, the sample sizes are too small to produce meaningful results disaggregated by year and age at arrival. Instead, we show trends in the overall share of immigrants in the US scientific workforce, including those who obtained a PhD in the US. An immigrant is defined as an individual who was not born in the US.

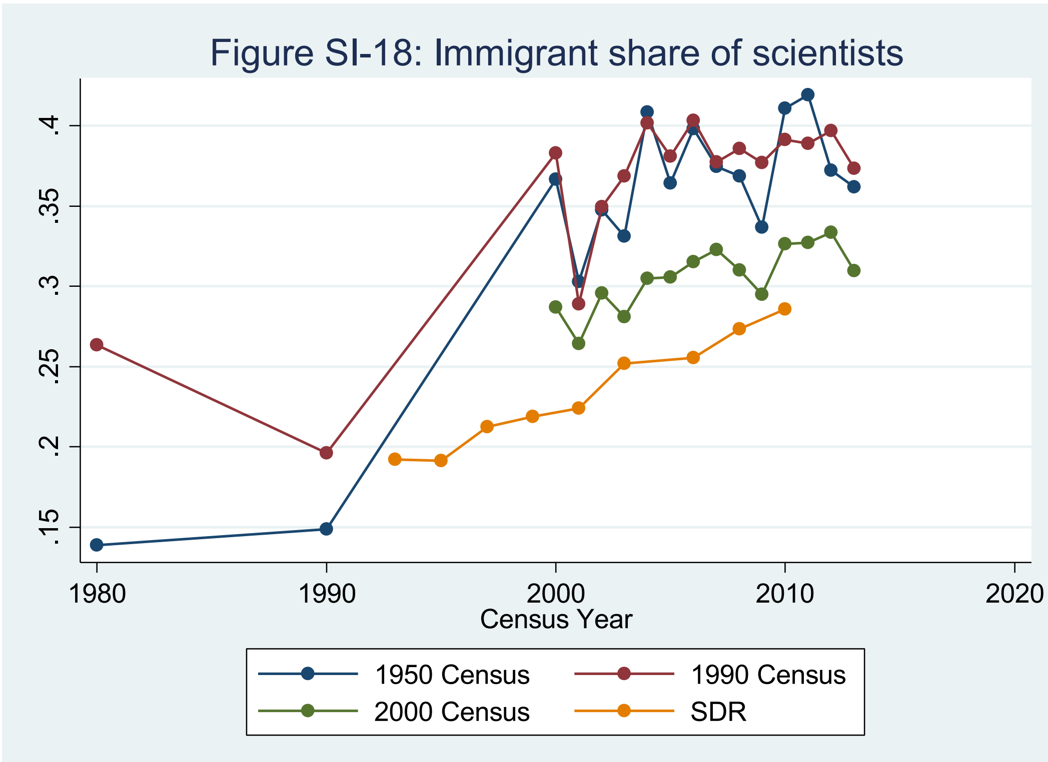


Figure SI-18 shows trends in the immigrant share of scientists using three alternative Census occupational coding schemes: one based on the 1950 occupational classification, a second based on the 1990 classification, and a third based on the classification used since 2000. The IPUMS provides recoded occupations for all years based on the 1950 and 1990 classifications, but the more detailed codes used in the 2000 classification cannot be used to recode earlier years.[10] For comparison, the proportion foreign born in the SDR is also shown. The share of immigrant scientists differs substantially in 1980 for the 1950 and 1990 classifications, differs moderately in 1990, and hardly differs at all since 2000. Both the 1950 and 1990 classifications show an immigrant share 5-10 percentage points higher than the measure based on the 2000 classification system. The latter in turn is about 5 percentage points higher than the share based on the SDR. This is expected, since the ACS/Census includes all immigrant scientists, while the SDR includes only those who obtained a PhD in the US. Despite uncertainty about the level of

[10] The 2000 classification system changed in a few minor ways since 2000.

immigration, all of the series show an upward trend, more or less parallel since the 1990s. The immigrant share of the scientific workforce was between 30% and 40% in 2010.

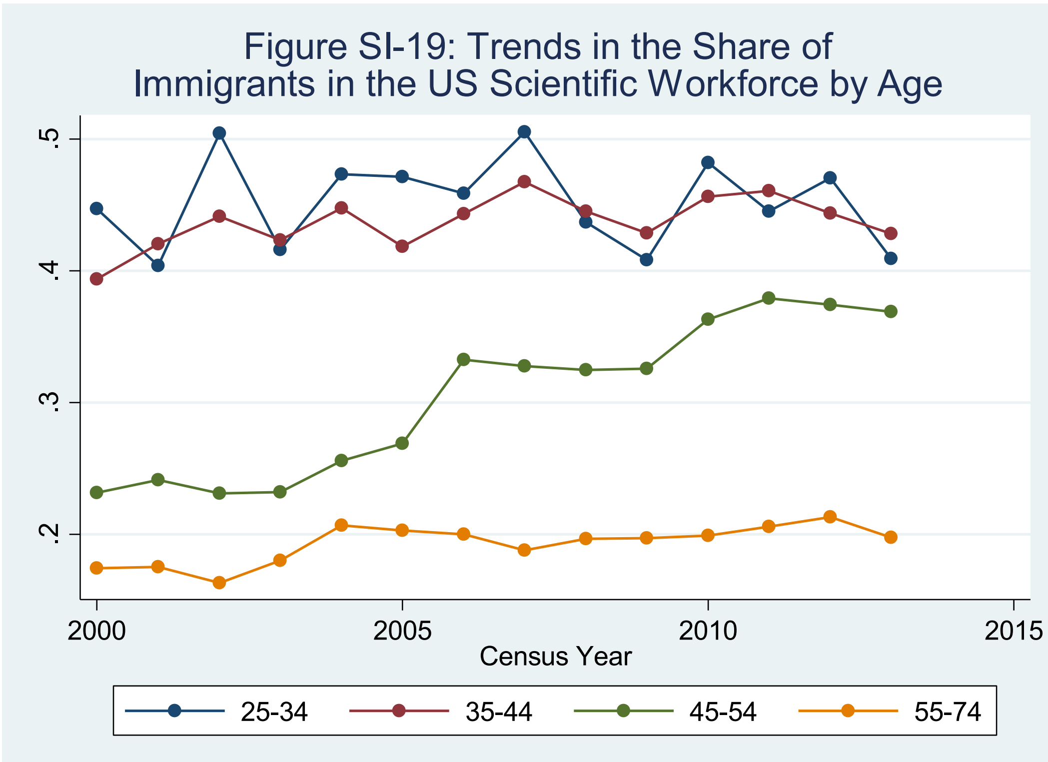


Figure SI-19 based on the 2000 classification system shows that the immigrant share is highest at younger ages, but the share at younger ages has not been increasing in recent years. In contrast, there was a sharp upward trend in the immigrant share at ages 45-54, and a modest upward trend at ages 55-64.

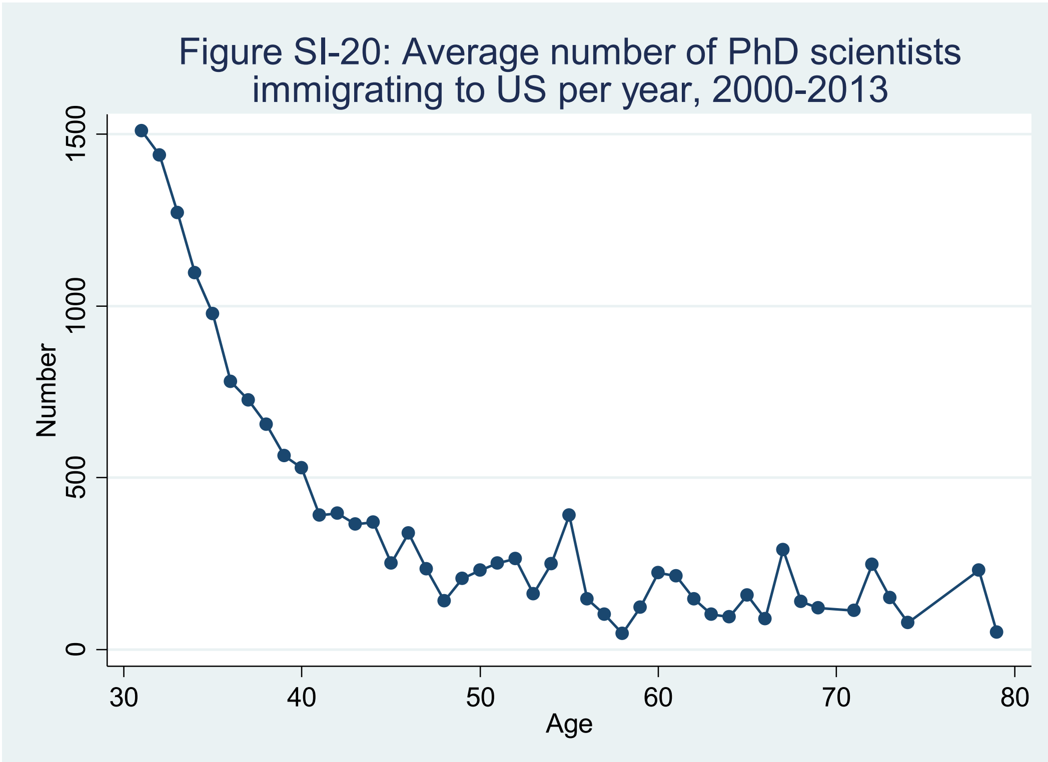


Figure SI-20 combines data from the ACS for the years 2000-2013 to illustrate the number of *new* immigrant scientists (arrived in the year of or the year before the survey) by age at arrival. The sample size is too small to do this separately by year. The number drops sharply from 1,500 at age 31 to 500 at age 40 and 250 by age 50. Summing over age, the data indicate that the total number of scientists trained abroad immigrating to the US was around 17,000 per year since 2000. This compares to about 35,000 new PhDs produced in the US per year (see Figure SI-8 above), including those earned by non-natives.

A note of caution is warranted here. The ACS/Census data are hardly ideal for measuring immigration of scientists by age, but they are the only data available. Differences in the share of immigrants as a function of occupation coding scheme, illustrated in Figure SI-18 above, and the small sample sizes available for measuring trends in immigration by age suggest considerable uncertainty about the level and age distribution of immigration of scientists.

	Mean age	Fraction aged 55+
Observed in 2008	48.6	0.331
Steady state with 2008 transition rates	50.9	0.393
Observed in 1993	45.1	0.176
Steady state with 1993 transition rates	50.0	0.361
Counterfactuals:		
1993 retirement hazard	49.7	0.353
1960 birth rate	47.6	0.283
1980 birth rate	51.1	0.401
No immigration of scientists	50.2	0.371
1993 death rate	50.8	0.389
1966 PhD completion rate	50.3	0.375
1993 PhD completion rate	51.2	0.402
1966 share of foreign born among new US PhDs	50.9	0.394
1993 share of foreign born among new US PhDs	50.9	0.393

Table SI.2: Steady state age distribution implications of alternative values of transition rates. 2008 transition rates are based on one-year transitions estimated from biannual transition rates observed between the 2008 and 2010 survey waves. 1993 transition rates are based on one-year transitions estimated from biannual transition rates observed between the 1993 and 1995 survey waves. The annual rates are estimated from the biannual rates under the assumption that the annual rates are equal between survey years. If b is the biannual rate and a is the annual rate, then b = a + (1-a)a. This equation is solved for a to obtain measures of the annual rates. The row labelled "1993 retirement hazard" uses observed 2008 transition rates except for the retirement hazard, which is set to its 1993 level. The other counterfactuals have a similar interpretation.

8. Additional steady state model simulation results

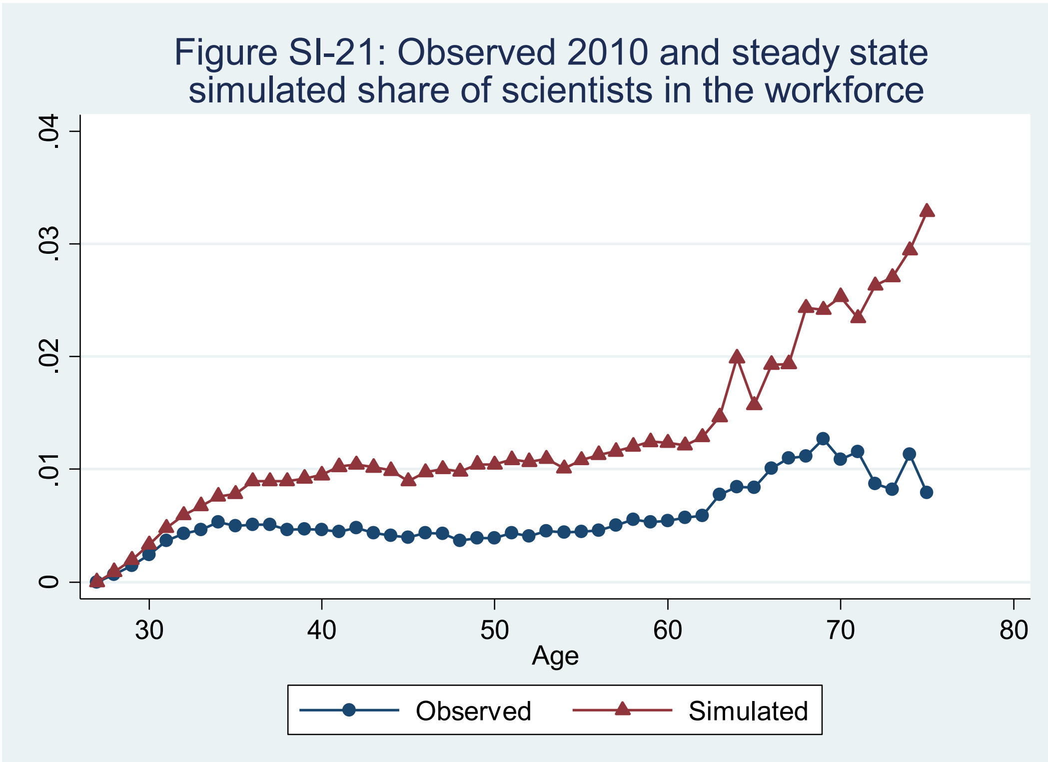


Figure SI-21 shows the observed 2010 share of scientists in the workforce by age, along with the simulated steady state share using 2008 transition rates. As noted in the text, the workforce as a whole is aging but the scientific workforce is aging much more rapidly, resulting in an increasing share at all ages, but especially beyond age 65.

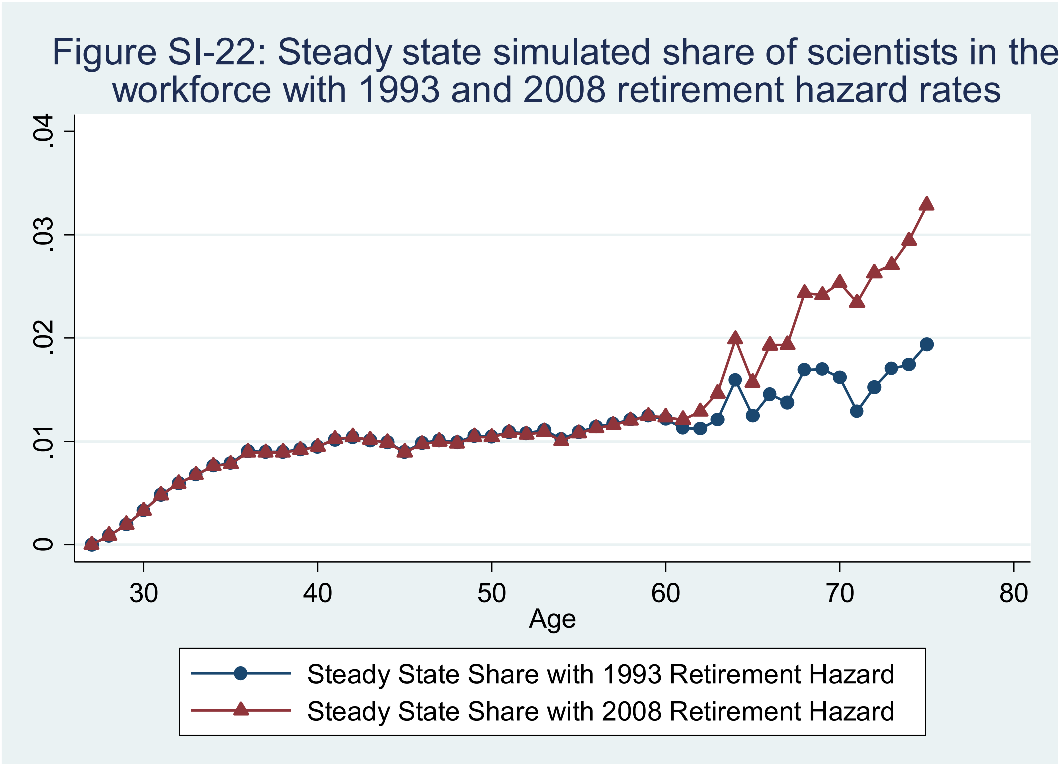


Figure SI-22 shows the steady state share of scientists in the workforce by age using 1993 and 2008 retirement hazard rates. The pre-mandatory retirement (1993) hazard rate implies a substantially lower share of scientists in the workforce at older ages.

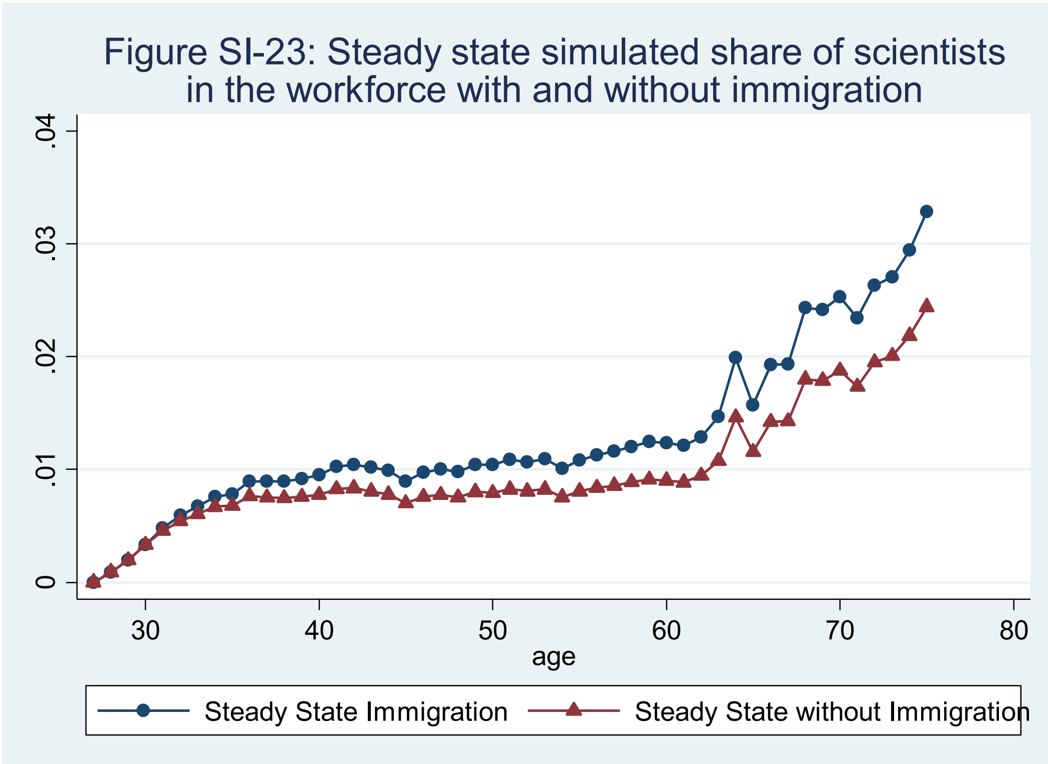


Figure SI.23 shows the steady state share of scientists in the workforce by age using current immigration rates of scientists and assuming no immigration of scientists. In the absence of immigration, share of scientists in the workforce would be lower, with the largest decline at older ages.

References

(1) Ginther D, Kahn S. (2013) Postdoctoral Training and the Career Outcomes of Biomedical PhDs. University of Kansas. Working Paper.

(2) Ashenfelter O, Card D (2002) Did the elimination of mandatory retirement affect faculty retirement? *Am Econ Rev* 92(4):957–980.

(3) Montez JK, Hummer RA, Hayward MD, Woo H, Rogers RG. (2011) Trends in the Educational Gradient of U.S. Adult Mortality from 1986 to 2006 by Race, Gender, and Age Group. *Research on Aging* 33(2): 145–171.