Although blaming the looming insolvency of developed country social security systems on aging populations is popular, critics often fail to recognize that many countries have benefit structures that add to cost by discouraging work. In those countries, people who work longer receive lower lifetime benefits. To avoid being penalized, they take early retirement. Early retirement increases expenditures by increasing the number of retirees. It also reduces the tax revenues generated by people in the labor force. As a result, reforms that increase the early retirement age, or impose reductions in retirement benefits for those who retire earlier, could reduce overall program costs in some cases by 20 to 50 percent.

In Social Security Programs and Retirement Around the World: Fiscal Implications, Introduction and Summary (NBER Working Paper No. 11290), NBER Research Associates Jonathan Gruber and David Wise summarize the evidence underlying these conclusions. They report on the findings of a large group of economists from 12 different countries. The group used a common analytical structure to model how features of national Social Security systems affect retirement behavior. They then used these models to analyze the fiscal impact of several types of reforms, considering the effects on both benefits paid out and taxes collected from older workers.

The first reform considered is increasing by three years the age at which individuals become entitled to their Social Security benefits. Despite lengthening life spans and better health in old age, the proportion of men who are officially out of the labor force between ages 55 and 65 has increased substantially in the industrialized countries since 1960. It now ranges from 0.7 in Belgium to about 0.2 in Japan. Increasing the retirement age is a natural way to increase labor force participation, and to improve Social Security’s fiscal position as well. Raising retirement ages in existing social security systems by three years would generate savings of over 40 percent in the United Kingdom, about 30 percent in the United States, and slightly over 15 percent in Italy.

The second reform that Gruber and Wise consider is making benefits “actuarially fair” with respect to early retirement, so that individuals who retire early get lower benefits and those who retire later get higher benefits. This is currently the approach taken by several nations, such as the United States, but it is not common in European nations. The modeling that underlies each nation’s analysis shows that these financial incentives have very significant effects on the retirement decisions of workers. As a result, in those European nations moving to actuarially fair pensions, by increasing the amount of time people choose to work and pay taxes and decreasing the amount of time they collect benefits, the costs will be lower. Instituting actuarially fair pensions would reduce costs by about 40 percent in Germany and by 10 to 20 percent in Belgium, Denmark, Japan, the Netherlands, and Italy.

Finally, the teams of authors consider a “common reform” with an early retirement age of 60, a normal retirement age of 65, and actuarially fair benefits. The effect of the common reform illustrates how retirement age and benefits structure affect fiscal balance. In countries with relatively parsimonious pension systems or relatively high early retirement ages — Canada, the United States, and the United Kingdom — the common reform will raise costs by more than 40 percent. In countries with early retirement and benefit structures that penalize people who work longer — Germany, France, Italy, Spain, and the Netherlands — the common reform will generate savings of more than 40 percent of base costs.

— Linda Gorman
Great Inventions Come Later in Life

Innovative thinkers are innovating later than they used to. While conventional wisdom holds that creative thinkers do their best work when they are young, a study by NBER researcher Benjamin Jones shows that over the past century the average age at which individuals produce notable inventions and ideas has increased steadily.

In Age and Great Invention (NBER Working Paper No. 11359), Jones considers data on Nobel Prize winners in Physics, Chemistry, Medicine, and Economics over the past 100 years, and on outstanding technological innovations over the same period. For comparative purposes, Jones also considers the ages of track and field record-setters and ball players who have received Most Valuable Player awards.

The data on the innovators reveal three initial characteristics. First, there is large variation in age: 42 percent of innovations came about when their creators were in their 30s, while 40 percent occurred when the inventors were in their 40s, and 14 percent appeared when the inventors were over 50. Second, there were no great achievements produced by innovators before the age of 19, and only 7 percent were produced by innovators at or before the age of 26 (Einstein's age when he performed his prize winning work). Third, the age distributions for the Nobel Prize winners and the technologists are nearly identical.

The most striking finding, however, is that the age distribution shifts over time, with the mean age of great achievement rising by five or six years per century. This parallels another study showing a similar upward trend in the age of persons receiving their first patents (NBER Working Paper No. 11360). Jones finds the trends among great innovators are significant and robust, even after controlling for nationality and field of study. Indeed, these controls strengthen the age trend, causing it to rise to about eight years over the course of the twentieth century. This suggests a compositional shift in great innovation towards fields and countries that favor the young.

One possible explanation for this age shift is a decline in the productivity of younger innovators in favor of older innovators. It may well be that the younger innovators are devoting themselves to an increasing amount of education and training. Or, it may be that the productivity of older innovators is increasing in relative terms simply because innovators are living longer. If we accept that raw ability declines over the life cycle while experience increases, then the shift in the distribution may indicate the rising importance of experience over ability. Alternatively, improved health care may spell increased ability and effort at later ages.

However, Jones urges caution in interpreting these distributional shifts. The shifts, he suggests, may reflect a simple demographic effect. If the population of innovators is getting older, then the older innovators will be more likely to produce substantial innovations even if the relationship between age and innovative potential is fixed. That is to say, the greater the ratio of 50-year-old innovators to 25-year-old innovators, the more likely the Nobel Prize-winning work or the groundbreaking technological development is to come from one of the 50-year-olds. Such demographic effects may be important, because age expectancy and the average age of the population have risen substantially throughout the twentieth century.

By subjecting these various hypotheses to econometric analysis, Jones concludes that the upward trend for productive innovators does not merely reflect the aging population, but in fact is a result of a substantial decline in the innovative output of younger individuals.

Jones notes that, unlike athletes, who do not require increased training demands over time, innovators appear to spend increasingly significant portions of their early years in education—a kind of human capital acquisition that might well explain the age trends in his study. Because the rules and requirements of their fields of endeavor remain fixed, athletes are not obligated to increase their human capital; accordingly, the data show no distributional shift in the ages of top athletes over the years. But thinkers must increasingly invest in acquiring intellectual capital, and the accumulation of knowledge—the rising distance to the frontier—can explain increased educational attainment.

Jones notes that economists have not focused much on the human capital investments of innovators. Because innovators customarily devote their youngest and perhaps brightest years to acquiring their education, understanding the tradeoffs at the beginning of the life-cycle may be of primary importance for understanding the ultimate output of these individuals—and for understanding why great innovation is steadily declining among younger thinkers.

The Value of Health and Longevity

During the twentieth century, life expectancy at birth for a representative American increased by roughly thirty years. In 1900, nearly 18 percent of males born in the United States died before their first birthday—today, it isn’t until age 62 that cumulative mortality reaches that level. This remarkable increase in longevity reflects progress against a
variety of afflictions and diseases, driving reductions in mortality at all ages. It illustrates a substantial, but unmeasured, increase in social welfare attributable to improvements in health.

Rising longevity, and health improvements more generally, are aspects of economic progress. Valuation of these gains is important for two reasons. First, traditional measures of economic growth and welfare, based on national income accounts, make no attempt to account for this source of rising living standards. Therefore, they underestimate improvements in well-being. Second, public expenditures account for a large portion of both medical research and the provision of medical care. Efficient decisions require a framework for measuring the value of treatment and of research-based medical progress.

In The Value of Health and Longevity (NBER Working Paper No. 11405), authors Kevin Murphy and Robert Topel develop and apply an economic framework for valuing improvements in health and longevity, based on individuals’ willingness to pay. They then use that framework to estimate the economic gains from declining mortality in the United States over the twentieth century, and to value the prospective gains that could be obtained from further progress against major diseases.

Murphy and Topel find that reductions in mortality from 1970 to 2000 had an economic value to the 2000 U.S. population of about $3.2 trillion per year. Over the longer term, the cumulative longevity gains during the twentieth century were worth about $1.3 million per person. Valued at the date they occurred, the production of longevity-related “health capital” would raise estimates of per capita output in the United States from 10 to 50 percent, depending on the time period in question.

The authors distinguish between two types of health improvements — those that extend life by reducing mortality and those that raise the quality of life. Life extension is valued because the utility from goods and leisure accrues over a longer period, and improvements in the quality of life raise the utility from given amounts of goods and leisure. This study provides a framework for calculating the economic value of a life-year, the value of remaining life, and changes in these values when health improves.

Based on a lifecycle model of consumption and survival, the authors show that the social value of improvements in health is greater: 1) the larger is the population; 2) the higher are average lifetime incomes; 3) the greater is the existing level of health; and 4) the closer are the ages of the population to the age of onset of disease. These factors point to an increasing valuation of health improvements over the past several decades and into the future. As the U.S. population grows, as lifetime incomes grow, as health levels improve, and as the baby boom generation approaches the primary ages of disease-related death, the social value of improvements in health will continue to rise.

The authors also find that improvements in health tend to be complementary. For example, improvements in life expectancy (from any source) raise willingness to pay for further health improvements by increasing the value of remaining life. This means that advances against one disease, say heart disease, raise the value of progress against other age-related ailments, more valuable than gains in longevity — and for these reasons as well they are likely to be conservative. The authors show that these values will increase in the future because of economic growth and, more interestingly, because health itself continues to improve.

Large as they are, these values may be offset by the costs of developing and implementing improvements in health. Current public and private spending on health-related research is a tiny fraction of what is available, but such investments may not be worthwhile if the costs of implementing new technologies are large.

An analysis of the social value of improvements in health is a first step toward evaluating the social returns to medical research and health-augmenting innovations. Improvements in health and longevity are partially determined by society’s stock of medical knowledge, for which basic medical research is a key input. The United States invests over $50 billion annually in medical research, about 40 percent of which is federally funded, accounting for 25 percent of government research and development outlays. The $27 billion federal expenditure for health-related research in FY 2003 represented a real-dollar doubling over 1993 outlays. The authors suggest that the returns to basic research may be quite large, so that substantially greater expenditures may be worthwhile. Using the authors’ estimate that a single percent reduction in cancer mortality would be worth about $500 billion, then a “war on cancer” that would spend an additional $100 billion on cancer research and treatment would be worthwhile if it has a 1-in-5 chance of reducing mortality by one percent.

Prospectively, even modest progress against mortality-causing diseases, such as cancer and heart disease, would have enormous social value. A single percent reduction in mortality from cancer or heart disease would be worth nearly $500 billion to current and future Americans. These estimates ignore the value of health advances to individuals in other countries, so they likely underestimate the aggregate social value of possible innovations. They also ignore corresponding improvements in the quality of life — which evidence suggests may be even more valuable than gains in longevity.

“A single percent reduction in mortality from cancer or heart disease would be worth nearly $500 billion to current and future Americans.”

— Les Picker
Better schools and school desegregation tended to raise the earnings of southern-born African-American men, but not all of that progress can be attributed to the Supreme Court’s 1954 decision in *Brown v. Board of Education*. The public profile of that landmark ruling overshadows the slow, long-term process that raised the quality of schooling available to southern black children. In *Evaluating the Role of *Brown* v. *Board of Education* in School Equalization, Desegregation, and the Income of African-Americans* (NBER Working Paper No. 11394), co-authors Orley Ashenfelter, William Collins, and Albert Yoon study the labor market implications of, first, providing more equal resources for black schools in the South and, later, bringing about school desegregation.

In *Plessy v. Ferguson* (1896), the Supreme Court affirmed the right of southern states to enforce racial segregation “for the promotion of the public good.” During the first third of the twentieth century, wide racial disparities in basic measures of school inputs, such as the length of the school year and students per teacher, were the norm in the South. In Alabama in 1910, for example, the average school year for white students was more than 30 days longer than for blacks, and there were approximately 12 more black students per teacher than white students per teacher. These disparities began to narrow twenty years before the *Brown* decision. As legal pressures mounted, southern state and local governments took the “equal” part of “separate but equal” more seriously.

The authors ask: If black workers who were born in the South in the 1920s and 1930s had attended schools with the same measurable characteristics (for example, length of school year and students per teacher) as white schools, how much higher might their income have been later in life? To answer this question, the authors use individual-level data from the 1970 census to estimate the average labor market returns to school quality for southern-born black men. The results suggest that southern-born black men from the 1920s birth cohorts would have earned 6 to 9 percent more than they actually did in 1970 if they had gone to “equal” schools. For southern-born black men from the 1930s birth cohorts, income would have been 2 to 5 percent higher. The relatively small difference indicates that this later birth cohort attended schools that were fairly similar to white schools in terms of school-year length and students per teacher. The link between school quality and income is forged largely by the connection between school quality and years of educational attainment — southern-born black men who went to better schools completed more grades and therefore earned higher incomes. Thus, the disparities in school resources were a significant factor in determining the earnings gap between southern-born blacks and whites, but a large earnings gap remains after accounting for differences in school quality.

Ashenfelter, Collins, and Yoon argue that the education level of black parents is an important consideration in this long-run context. Their analysis suggests that parental education had a strong influence on children’s educational attainment and subsequent earnings. In this sense, discrimination in the allocation of school resources in one generation tended to spill over to the education and earnings of the subsequent generation of African American workers.

In the second part of the paper, the authors examine the impact of post-1964 southern school desegregation. On the eve of the *Brown* decision in 1954, the Southern Education Reporting Service found that essentially no black children attended school with white children in public schools in the Deep South and that very few black children in Border States did. In 1956, one poll found that only 14 percent of Southern whites thought that black and white students should attend the same school, and that poll included whites in the Border States and Washington, D.C. With that public opinion in the background, southern states and local governments exercised a variety of legal tactics to forestall any meaningful school integration. The Civil Rights Movement was making inroads through the courts, but even five years after the infamous 1957 standoff in Little Rock, Arkansas, only 1 percent of southern black students attended school with whites. Starting in the mid-1960s, however, school desegregation in the South proceeded rapidly as a combined result of the Civil Rights Act of 1964, the Elementary and Secondary Education Act of 1965, and a series of federal court orders.

“The earnings gap between southern-born black men and non-southern-born black men in the same birth cohort narrowed by about 10 percent in the post-desegregation group.”

Ashenfelter, Collins, and Yoon use individual-level data from the 1990 census to evaluate the impact of this sudden and widespread pattern of desegregation on the earnings of southern-born black males. Essentially, the authors compare the earnings of southern-born black men who would have completed their schooling under the segregated regime with the earnings of those who followed behind them in school by a few years or more and therefore, most likely, would have attended desegregated schools. Controlling for several factors, they find that the earnings gap between southern-born black men and non-southern-born black men in the same birth cohort narrowed by about 10 percent in the post-desegregation group. This finding is consistent with “an economically significant, positive effect on blacks’ income and high school completion rates.” The authors stress that the pattern is suggestive, not conclusive, and they encourage further scrutiny of the hypothesis.

— David R. Francis
The Mexican Workforce in the United States

The population of Mexican-born persons residing in the United States has increased at an unprecedented rate in recent decades. This increase can be attributed to both legal and illegal immigration. During the entire decade of the 1950s, only about 300,000 legal Mexican immigrants entered the United States, making up 12 percent of the immigrant flow. In the 1990s, 2.2 million Mexicans entered the United States legally, making up almost 25 percent of the legal flow, according to the U.S. Immigration and Naturalization Service.

In addition, there were seven million illegal aliens residing in the United States as of January 2000, with 4.8 million (68 percent) being of Mexican origin. As a result of the increase in the number of legal and illegal Mexican immigrants, nearly 9.2 million Mexican-born persons resided in the United States in 2000, comprising about 29.5 percent of the foreign-born population.

In The Evolution of the Mexican-Born Workforce in the United States (NBER Working Paper No. 11281), NBER Research Associates George Borjas and Lawrence Katz use data from 1900 through 2000 to document the evolution of the Mexican-born workforce in the U.S. labor market. While it is well known that there has been a rapid rise in Mexican immigration to the United States in recent years, they find that the share of Mexican immigrants in the U.S. workforce declined steadily after the 1920s before beginning to rise again in the 1960s. It was not until the 1970s that the relative number of Mexican immigrants in the U.S. workforce was back to the 1920s level.

Analyzing the economic performance of these immigrants throughout the twentieth century, the authors find that Mexican immigrants have much less education than either native-born workers or non-Mexican immigrants. These differences in what economists call “human capital” account for nearly three-quarters of the very large wage disadvantage suffered by Mexican immigrants in recent decades.

“Wage convergence has been weaker on average for Mexican immigrants than for other immigrant groups.”

While the earnings of non-Mexican immigrants converge to approximate those of their native-born counterparts as the immigrants accumulate work experience in the U.S. labor market, the authors find that this wage convergence has been weaker on average for Mexican immigrants than for other immigrant groups. Although native-born workers of Mexican ancestry have levels of human capital and earnings that far exceed those of Mexican immigrants, the economic performance of these native-born workers lags behind that of native workers who are not of Mexican ancestry. Much of the wage gap between the two groups of native-born workers can be explained by the large difference in educational attainment between the two groups.

The authors also find that the large Mexican influx in recent decades has contributed to the widening of the U.S. wage structure by adversely affecting the earnings of less-educated native workers and improving the earnings of college graduates. These wage effects have, in turn, lowered the prices of non-traded goods and services that are low-skill labor intensive.

There is little evidence that the influx of Mexican-born workers into the United States is slowing down as we enter a new century, and there is also little evidence that the skill composition of the Mexican immigrants is changing from what it has been in the past. The continued migration of Mexican workers into the United States, and the inevitable rapid growth of the group of native-born workers of Mexican ancestry, suggest that the economic consequences of this migration influx are only beginning to be felt.

— Les Picker

Is China Mercantilist?

Strong evidence of China’s emergence as a global economic powerhouse are these twin facts: a large foreign exchange reserve that China is holding, especially in dollar-denominated assets, and a large amount of foreign direct investment (FDI) going into China that rivals FDI into the United States. A popular (and politically charged) explanation for these facts runs as follows: China’s rapid rise in the foreign exchange reserve is a consequence of its mercantilist policy, exporting like mad by relying on a deliberately undervalued currency, cheap labor, and foreign investors, particularly those from the United States.

But in The Chinese Approach to Capital Inflows: Patterns and Possible Explanations (NBER Working Paper No. 11306), authors Eswar Prasad and Shang-Jin Wei suggest that the reasons behind China’s increased foreign exchange reserve and its success at attracting FDI — as opposed to more volatile financial and equity markets — are too complex for this kind of simple theory. They argue that the mercantilist explanation is an “intriguing story, but the facts do not support it.”

To start with, they note that more than 87 percent of the acceleration in the increase in China’s foreign reserve holding from the period 1988–2000 to the period 2001–4 can be explained by a surge in non-FDI type of capital inflows.
(sometimes called “hot money”), including a dramatic reversal of capital flight. Only 13 percent of the increment can be attributed to an increase in the current account surplus and an acceleration of the inward FDI.

Furthermore, they note that China’s FDI, which in 2004 was $61 billion, comes from countries that are running a current account surplus with China rather than those with a deficit. The main contributors are based in advanced Asian economies such as Japan, Korea, and Singapore. Europe and the United States combined account for, at most, about 30 percent of China’s FDI.

Prasad and Wei also view the focus on an undervalued currency as off the mark. As recently as 1997 and 1998, China chose not to devalue its currency even though such a move would have aided exports. And in the 1980s and 1990s, China’s currency was more likely to be overvalued than undervalued.

“Further research will be needed to disentangle the competing explanations for [the rise in FDI in China], but there is little evidence that mercantilist stories are the right answer,” they write.

Prasad and Wei nonetheless believe it is important to understand more about China’s success in “tilting” the flow of money into China toward FDI, “especially as China continues its integration into world financial markets and becomes more exposed to the vagaries of these markets.” As it now stands, the fact that so much foreign money has been of the FDI variety has helped shield China from the kind of jolts recently administered to other Asian economies where a higher proportion of investment was indirect — such as bank lending or stock portfolios — and tended to be withdrawn at the first hint of trouble. “It is not just the degree of financial opening but the composition of capital inflow that determines the quality of a developing country’s experience with globalization,” the authors state. “In particular, FDI appears to be less subject to sharp reversals than other types of inflows, particularly bank lending.”

Prasad and Wei observe that during the Asian financial crisis of the 1990s, “FDI inflows were only marginally affected” while other forms of investment showed “sharp increases in outflows” and took two or three years to recover. According to their analysis, China’s capital controls along with incentives offered to foreign investors “appear to have played a big part in encouraging FDI inflows.” For example, foreign firms investing in China don’t have to pay corporate income tax on their first two years of profits and in subsequent years pay only half the corporate income tax rate of Chinese companies. Overall, Prasad and Wei find that China probably “offers more incentives to attract FDI than most countries in the world.”

But every theory, they note, seems to have certain flaws. For example, they point to a seemingly plausible argument linking China’s upsurge in FDI to government actions that deprived private firms of capital in favor of state-owned firms. According to this theory, FDI increased as private Chinese firms aggressively used pro-FDI policies to secure the investments they needed to expand. Prasad and Wei believe this explanation could account for some of the FDI flowing into China in the 1980s. But it appears flawed in explaining the more recent surge, as Chinese banks have become “increasingly willing” to make loans to private firms.

Similarly, the notion that China’s specific incentives for FDI have provided the spark has its problems, since China also erects many barriers to investment. “The story is not that straightforward since one would expect a counteracting effect from factors such as weak governance, legal restrictions on investment by foreigners, and poor legal infrastructure and property rights,” they note.

— Matthew Davis