NEW PAPERS ON HEALTH

By Victor R. Fuchs
NEW PAPERS ON HEALTH

The Robert Wood Johnson Foundation, helping Americans lead healthier lives and get the care they need. www.rwjf.org
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>iii</td>
</tr>
<tr>
<td>How and Why U.S. Health Care Differs From That in Other OECD Countries</td>
<td>1</td>
</tr>
<tr>
<td>The Gross Domestic Product and Health Care Spending</td>
<td>6</td>
</tr>
<tr>
<td>Major Trends in the U.S. Health Economy since 1950</td>
<td>9</td>
</tr>
<tr>
<td>The Physician’s Role in Controlling Costs</td>
<td>14</td>
</tr>
<tr>
<td>If Accountable Care Organizations Are the Answer, Who Should Create Them?</td>
<td>17</td>
</tr>
<tr>
<td>Shortening Medical Training by 30%</td>
<td>20</td>
</tr>
<tr>
<td>Specialization in Medicine. How Much Is Appropriate?</td>
<td>23</td>
</tr>
<tr>
<td>The New Demographic Transition: Most Gains in Life Expectancy Now Realized Late in Life†</td>
<td>26</td>
</tr>
<tr>
<td>Geographic and Racial Variation in Premature Mortality in the U.S.: Analyzing the Disparities</td>
<td>38</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>52</td>
</tr>
</tbody>
</table>
Introduction

This booklet contains nine pieces published by me and co-authors in 2012 and the first half of 2013. Unlike the two previous booklets which dealt almost exclusively with health care reform, this one covers a wide range of subjects including medical care per se, medical education, and demographic issues. Health economics provides a unifying perspective, but each subject is developed to be accessible to health professionals, policymakers, and interested lay readers.

The first paper addresses a long-standing question – how and why does the U.S. health care system differ from those of the other OECD countries? This difference is manifest not only in much higher U.S. expenditures per capita but also in a different mix of expenditures. The average OECD country has more physicians and hospital beds per capita delivering more “bread and butter” medical care. The U.S. has proportionately more specialists and sub-specialists, delivering much more hi-tech medicine such as MRI scans, and has more stand-by capacity providing easier access to hi-tech diagnostic and therapeutic interventions. The U.S. also provides more amenities such as space and privacy for hospitalized patients.

Life expectancy is higher in the other OECD countries, but that is probably due more to non-medical than medical factors. For example, the poverty rate in the U.S. is double that in the average OECD country, and poverty is a good predictor of earlier death within and across countries.

The paper speculates about why the U.S. has such a different health care system and proposes three, not mutually exclusive reasons. In the U.S. there is greater distrust of government, more resistance to redistribution, and a political system that permits special interests that benefit from the existing health system to block major changes.

“The Gross Domestic Product and Health Care Spending,” the next and most recent paper, addresses the question why health care expenditures have grown so slowly in recent years. From 2007 to 2011, the average annual rate of change of national health expenditures (NHE) per capita adjusted for general inflation was 1.35%; the average over the last 60 years was 4.4% per annum. Some observers point to the recession as an explanation; real GDP per capita actually declined between 2007 and 2011 at -0.58% per annum. The President’s Council of Economic Advisers, however, concluded that the recession was not a major factor relative to improved efficiency in hospitals and physician groups, payment reforms, and early responses to the Affordable Care Act.

My paper looks at the relationship between annual changes in real GDP and changes in real NHE since 1950 and finds a robust relationship most of the time. It seems premature to dismiss the worst recession since the 1930s as the major explanation for the recent slowdown in health care spending. The one big exception to the correlation between changes in GDP and changes in NHE occurred in the middle of the 1990s when health expenditures grew at an unusually slow rate while the GDP was increasing rapidly. This was precisely the period when managed care became widespread. Subsequently, there was a “backlash” to managed care from physicians and patients, accompanied by a large surge in health care spending.

Despite the “backlash” the paper finds what might be a lasting effect of managed care on expenditures. From 1950 to 1995 real health expenditures per capita grew at 4.7% per annum; from 1995 to 2011, the rate was 3.1% per annum. Slower growth of real GDP per capita from 2.1% per annum prior to 1995 to 1.4% after 1995 could explain part of the change, but the “gap” between growth of GDP and NHE fell from 2.6% (4.7-2.1) to 1.7% (3.1-1.4).

The paper concludes that the rate of growth of health care spending in the past was, with the exception of the 1990s, substantially related to the growth of GDP. There has, however, been a decrease in the gap between the two series.
since the mid-1990s, possibly as a result of some persistent aspects of managed care.

“Major Trends in the U.S. Health Economy since 1950,” the next paper, was written at the invitation of the New England Journal of Medicine to help celebrate its 200th anniversary of publication. Without doubt the most important trend in health care during that period was the increase in expenditures as a share of the gross domestic product from 4.6% in 1950 to over 17% in 2010. This increase has posed a particularly large problem for public policy because the government’s share of the health care bill increased from one-quarter to almost one-half and the federal share of government spending grew from one-half to three-quarters. This has led many observers to conclude that the U.S. government doesn’t have a general spending problem, it has a health care problem. A second important trend is the decline of out-of-pocket share of payment from 56% in 1950 to 14%, replaced partly by government but also by the growth of private sector insurance.

Within the health care sector, the most important trend has been a decline in the relative importance of days in hospital – one half of what it was 40 years ago on a per capita basis. The paper also notes an increase in the proportion of physicians who are women, are specialists or sub-specialists, are hospital based, and use electronic medical records. A small recent trend with potential to dramatically change the organization and delivery of care is the replacement of the current, fragmented system with competition among large accountable care organizations (ACOs) that serve defined populations for risk-adjusted per capita annual payments.

“The Physician’s Role in Controlling Costs” was another invited paper, given at a joint meeting of British and American cardiologists in San Francisco, March 2013. Here, I try to explain to American cardiologists why they along with other physicians will be under pressure to control costs despite the spectacular advances in life expectancy attributable to declines in diseases of the heart. Cardiologists, like other physicians, will be expected to pay more attention to the value of many prescriptions, tests, and referrals for surgery that they generate. Some interventions are clearly worth their cost, but others may not be. To practice cost-effective medicine, physicians will need more information about the value of alternative interventions, will need access to an infrastructure of technology and personnel appropriate for efficiently treating their patients, and must face incentives that will motivate them to take advantage of that information and infrastructure.

Controlling the cost of care is widely seen as imperative for the fiscal health of the federal government. A Congressional Budget Office assessment of 10 major Medicare demonstration projects indicated that none of the programs using fee-for-service payment reduced federal spending. Much discussion and a few demonstration projects have begun to focus on accountable care organizations (ACOs) that provide care to large, self-defined populations (through prospective annual enrollment) for a risk-adjusted capitation payment.

“If Accountable Care Organizations Are the Answer,” written with Leonard Schaeffer, asks who will create ACOs. Employers? Physicians? Hospitals? We examine the pros and cons for each of these possibilities and conclude that for a wide variety of reasons none of them are likely to create enough ACOs within a reasonable time frame to cover a significant portion of the population. Our conclusion is that only large health insurance plans, both not-for-profit and for-profit, have the operational infrastructure, access to capital, and depth of professional management required to create and manage ACOs on a large scale reasonably quickly.

The movement to reform the health sector has focused primarily on changing funding, organization, and delivery of care. Universal coverage achieved through taxes or mandates-cum-subsidies, multi-specialist groups integrated with hospitals replacing small-scale fragmented physician practices, capitation payment, and managed competition among groups have been the major goals of many reformers.

I have commented frequently on various aspects of reform, but, with one exception until recently have not addressed medical education. This now seems to me to be a serious omission. Surely the goals of health care reform would have more chance of success if the training of physicians was not still embedded in a structure that was created to serve a medical care system now regarded as dysfunctional. Two commentaries in this collection deal with different aspects of reforming medical education.

“Shortening Medical Training by 30%,” written with Ezekiel Emanuel, M.D., is concerned primarily with the length of training. The present rigid system typically requires a minimum of 14 years from high school graduation until certification as a sub-specialist. We believe that this could be shortened by 30% through relatively modest changes at each stage of medical education: pre-med, medical school, residency, and sub-specialty training.

“Specialization in Medicine: How Much Is Appropriate?” the other commentary on medical education, written with Allan S. Detsky, MD, PhD and Stephen R. Gauthier, discusses the rapid growth of specialization and sub-specialization in the U.S., the advantages and disadvantages, and the appropriate role for governments regarding specialists. In 1960 in the U.S. there were only 18 specialty boards and just a handful of sub-
specialties, but by 2011 there were 158 specialties and subspecialties. No other country has as many as 100.

The last two papers focus on demographic issues rather than medical care, *per se*. “The New Demographic Transition: Most Gains in Life Expectancy Now Realized Late in Life,” written with Karen Eggleston, calls attention to the current negative impact of increases in life expectancy on labor force participation and GDP per capita. This is in direct contrast to the effect of increases in life expectancy during the first demographic transition in the 19th and early 20th centuries. Around 1900 in the U.S. and other rapidly developing countries, most of the gains in life expectancy resulted from declining mortality of infants, children, and young adults. Less than 20 percent of the additional years of life were realized after age 65. Now more than 75 percent are realized after age 65. This paper covers the years 1900-2007 for the United States and 16 other developed countries. We focus on life expectancy at birth and at age 65, the percent living until 65, and the share of additional years of life lived after 65. For the U.S., we also calculate effects of the changes in age distribution on expected labor force participation as a share of the population. The paper concludes with a brief discussion of the long-run implications of this new demographic transition.

Life expectancy has been the primary measure for describing and analyzing health across populations and its utility is beyond question. In “Geographic and Racial Variation in Premature Mortality in the U.S.: Analyzing the Disparities,” written with Mark Cullen, MD, and Clint Cummins, PhD, however, we focus on a different measure of the health of a population – the percentage of a cohort that survives to age 70 ($S_{70}$). We believe that this is a more sensitive measure for describing and analyzing geographic and racial disparities and can be more accurately calculated for small populations.

Most of the paper concentrates on 4 race-sex groups in 510 U.S. counties, describing the large disparities across and within the counties and showing their close relationship to a group of 22 socio-economic variables. For white males, the $R^2$ for a regression of $S_{70}$ on the 22 socio-economic variables is 0.86. Among other notable results, the average differential within counties of 17 percentage points in $S_{70}$ between black and white males shrinks to 2.4 percentage points when white values for the socio-economic variables are substituted for black values in the black estimating equation.

References
1. Fuchs, VR. Health Care Reform. SIEPR, January 2010.
2. Fuchs, VR. More Health Care Reform. SIEPR, January 2012

Victor R. Fuchs
Henry J. Kaiser Jr. Professor Emeritus
Senior Fellow, Stanford Institute for Economic Policy Research (SIEPR)
Stanford University
How and Why U.S. Health Care Differs From That in Other OECD Countries

By Victor R. Fuchs, Ph.D.

United States health care, often hailed as “the best health care system in the world,” is also faulted for being too costly, leaving many millions of individuals uninsured, and having avoidable lapses in quality. Criticism often draws on comparisons with other countries of the Organization for Economic Co-operation and Development (OECD). This Viewpoint also makes such comparisons, over a broad range of variables, and reaches one inescapable conclusion—US health care is very different from health care in other countries. Potential reasons for the differences are discussed, leading to the conclusion that future efforts to control cost, provide universal coverage, and improve health outcomes will have to consider the United States’ particular history, values, and political system.

US vs OECD: Health Expenditures and Outcomes

Compared with the average OECD country, US health care expenditures differ in 3 important ways.1 First, as a percentage of gross domestic product, US expenditures are twice as high. Second, the US share of health expenditures funded by government is much lower, 46% vs 75%. Third, the mix of services provided (technology intense vs more basic care) is very different (eTable, available at http://www.jama.com).

The larger role of government in health in OECD countries and the difference in mix of services are the main proximate explanations for the higher level of spending in the United States. Because funding in most OECD countries is usually through a tax-supported system, administrative costs are usually much lower than in the United States, with its fragmented sources of funding and payment. Also, the OECD countries use the concentration of funding to negotiate aggressively with drug companies and physicians and to control investment in hospitals and equipment. The United States could try to use the buying power of Medicare in a similar way, but legislation and political pressure prevent such an approach. The OECD countries provide more physicians and more acute care hospital beds, whereas the United States provides much more high-tech services, such as magnetic resonance imaging (MRI) scans and mammograms, proportionately more specialists, more amenities (privacy and space in hospitals), and more standby capacity as evident in a higher ratio of MRI scanners available to MRI scans performed. The greater number of physician visits and hospital days in OECD countries does not result in higher spending because of differences in services provided during a visit or a hospital day. In general, the United States has an expensive mix, whereas the OECD countries have an inexpensive one.

The effect of these differences in mix and total expenditures on health outcomes is uncertain. Measured by life expectancy, the OECD countries do slightly better than the United States, but firm conclusions are elusive because life expectancy depends on many factors in addition to medical care. For instance, the percentage of population in poverty is much higher in the United States than in the OECD countries (17% vs 9%), and poverty is a predictor of early death. Health is probably distributed less equally in the United States than in the OECD countries because the United States has more individuals without insurance and greater income inequality.

Why the Differences?

Three basic differences between the United States and most other OECD countries might explain why health policy differs. First, US individuals appear more distrustful of government, a distrust that has deep historical roots. It

---

was an armed rebellion against the government of King George III that led to the founding of the United States. It was Thomas Jefferson, a principal founding father, who said, “That government is best which governs least.” The initial antigovernment sentiment has received recurrent “booster shots” from waves of immigrants who came to the United States seeking freedom. Their willingness to risk life in a new land was frequently fueled by negative experience with government in their home country, a government that oppressed them because of their political beliefs, religion, ethnicity, or social class. Medicare and Medicaid appear to be an exception to distrust of government, but these programs provide insurance for populations that were not and could not be served by private insurance. A Pew public opinion survey of a representative sample of US individuals about their attitude toward elected officials showed more than twice as many negative as positive views.2

Closely related to the weaker support for government action in the United States is a reluctance to achieve more equal outcomes for the population through redistributive public policy. Although US individuals have always rejected European-style class distinctions that required deference and subservience,3 the declaration that “all men are created equal” did not carry any suggestion of equality of outcomes, such as in income or health. The income tax is less progressive in the United States than in most OECD countries, and the redistributive effect is augmented in the OECD countries by more egalitarian transfers of money and services. In response to a Pew survey,4 4 of 5 US individuals agreed that “everyone has it in their own power to succeed.” Only 1 in 5 agreed that “success in life is pretty much determined by forces outside our control.” Whether this view reflects reality is another matter. It is attitude and beliefs that shape voting behavior.4

Heterogeneity of the US population tends to strengthen resistance to redistribution. Diversity of race, religion, ethnic origin, and sometimes language contribute to a weaker sense of empathy for less fortunate members of society, whose identity may differ greatly from one’s own. In more homogeneous nations, such empathy is more likely to be experienced and acted upon. Weak support for redistribution at the national level in the United States stands in sharp contrast with redistribution within self-defined more homogeneous groups (for example, Mormon Relief Societies, Jewish homes for the aged in almost every major city, and the founding of Baptist, Catholic, Lutherans, Methodists, and other sectarian hospitals).

The third, and probably most important, difference between the United States and most OECD countries is the political system. Many observers attribute US failure to enact comprehensive health care reform to the opposition of “special interests,” such as pharmaceutical and device manufacturers, insurance companies, physicians (especially those in high-income specialties), and hospitals. But all countries have special interests; only in the United States have they been particularly successful in blocking comprehensive reform. This success can be explained in part by noting that the US political system is different from the parliamentary systems of most OECD countries in ways that make special interests more effective. Some of these differences are built into the US Constitution, including the checks and balances provided by 2 separate houses of Congress with their powerful committees, plus an independent executive branch with veto power. Some differences have evolved over time, such as expensive primary election battles, long election campaigns, and the Senate filibuster. Thus, the US system provides many “choke points” for special interests to block or reshape legislation. Also, in recent years, contributions from special interest groups significantly influence who runs for office, who gets elected, and how elected officials vote.

Lessons for Future Reform

President Obama’s Affordable Care Act (ACA), if fully implemented, would involve significant redistribution with tens of millions of poor and sick persons obtaining health insurance paid for by others. If the ACA is pared back, there will be less redistribution and tens of millions of persons would not have coverage, and the more difficult task of controlling health costs would remain. This review suggests a strategy for obtaining further reform.

First, government’s role should be limited to what is necessary, not just desirable. Efficiency and equity in financing require a dedicated tax to fund basic care for all.5 Second, provision of basic coverage for all should not require equality

---


for obtaining additional coverage. As in Australia, Israel, the Netherlands, and Switzerland, individuals should be free to purchase more than basic care. Third, reform should have features that would appeal to some special interests, or to some elements within each special interest group (for example, some physicians or some health plans). Comprehensive health care reform in the United States is necessary to avoid a financial disaster, but enactment of such reform will require attention to US history, values, and politics. Author Affiliations: Stanford Institute for Economic Policy Research, Stanford University, Stanford, California. Corresponding Author: Victor R. Fuchs, PhD, Stanford Institute for Economic Policy Research, Stanford University, 366 Galvez, Stanford, CA 94305-6015 (vfuchs@stanford.edu).

Conflict of Interest Disclosures: The author has completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Dr Fuchs reported that his institution received grants from the Blue Shield of California Foundation and the Robert Wood Johnson Foundation, and that he received royalties for a book from World Scientific Publishing.

Online-Only Material: The eTable and Author Audio Interview are available at http://www.jama.com.

Additional Contributions: The financial support from the Robert Wood Johnson Foundation and Blue Shield of California Foundation is gratefully acknowledged. Rossannah Reeves, BA (Stanford University Institute for Economic Policy Research, Stanford, California), provided excellent research assistance by tracking down the data and putting it into tabular form. Ms Reeves received no compensation for her contribution outside of her employment.
Supplementary Online Content

Fuchs VR. How and why US health care differs from that in other OECD countries. JAMA. doi:10.1001/jama.2012.125458, This supplementary material has been provided by the authors to give readers additional information about their work.

eTable. United States vs Other OECD Countries, 2008

<table>
<thead>
<tr>
<th>Part A</th>
<th>United States</th>
<th>OECD²</th>
<th>Range (Number of Countries)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health expenditures per capita (US $2008 ppp)</td>
<td>7,760</td>
<td>2,998</td>
<td>5,246 to 882 (29)</td>
</tr>
<tr>
<td>Health expenditures as percent of GDP</td>
<td>16.4</td>
<td>8.8</td>
<td>11.1 to 5.8 (28)</td>
</tr>
<tr>
<td>Government health expenditures as percent of total health expenditures</td>
<td>46.0</td>
<td>75.0</td>
<td>84.7 to 46.9 (27)</td>
</tr>
<tr>
<td>Practicing physicians (doctors) per 1,000 population</td>
<td>2.4</td>
<td>3.1</td>
<td>4.6 to 1.9 (22)</td>
</tr>
<tr>
<td>Acute care hospital beds per 1,000 population</td>
<td>2.6</td>
<td>3.3</td>
<td>8.2 to 1.7 (31)</td>
</tr>
<tr>
<td>Number of MRI scans per 100,000 population</td>
<td>9340</td>
<td>3806</td>
<td>9730 to 600 (15)</td>
</tr>
<tr>
<td>Number of MRI machines per 100,000 population (2007)</td>
<td>2.59</td>
<td>0.67</td>
<td>1.93 to 0.15 (23)</td>
</tr>
<tr>
<td>Mammograms per 1,000 women</td>
<td>40.2</td>
<td>14.4</td>
<td>47.4 to 3.2 (20)</td>
</tr>
<tr>
<td>Caesarean sections per 1,000 live births</td>
<td>323</td>
<td>246</td>
<td>439 to 143 (28)</td>
</tr>
<tr>
<td>Life expectancy at birth-males</td>
<td>75.5</td>
<td>77.6</td>
<td>79.0 to 68.6 (33)</td>
</tr>
<tr>
<td>Life expectancy at birth-females</td>
<td>80.5</td>
<td>82.4</td>
<td>83.2 to 75.8 (33)</td>
</tr>
<tr>
<td>Life expectancy at age 65-males</td>
<td>17.3</td>
<td>17.5</td>
<td>18.9 to 13.6 (33)</td>
</tr>
<tr>
<td>Life expectancy at age 65-females</td>
<td>20.0</td>
<td>20.7</td>
<td>21.6 to 15.8 (33)</td>
</tr>
<tr>
<td>Infant mortality</td>
<td>6.5</td>
<td>3.7</td>
<td>15.2 to 1.8 (32)</td>
</tr>
<tr>
<td>Infant mortality US whites</td>
<td>5.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Daily smokers as percent of 15+populationmales</td>
<td>17.9</td>
<td>25.9</td>
<td>46.3 to 13.1 (19)</td>
</tr>
<tr>
<td>Daily smokers as percent of 15+populationfemales</td>
<td>15.1</td>
<td>17.6</td>
<td>22.3 to 7.4 (19)</td>
</tr>
<tr>
<td>Percent obese (age 15+) (&gt;30 kg/m²) – males (2009)</td>
<td>32.2</td>
<td>16.0</td>
<td>26.0 to 3.6 (33)</td>
</tr>
<tr>
<td>Percent obese (age 15+) (&gt;30 kg/m²) – females (2009)</td>
<td>35.5</td>
<td>16.1</td>
<td>4.5 to 3.5 (33)</td>
</tr>
<tr>
<td>Percent overweight (age 15+) (&gt;25 kg/m²) – males (2009)</td>
<td>40.1</td>
<td>42.0</td>
<td>47.9 to 26.1 (33)</td>
</tr>
<tr>
<td>Percent overweight (age 15+) (&gt;25 kg/m²) – females (2009)</td>
<td>28.6</td>
<td>29.5</td>
<td>37.4 to 17.3 (33)</td>
</tr>
<tr>
<td>Percent of population in poverty (late 2000s)</td>
<td>17</td>
<td>9</td>
<td>21 to 5 (33)</td>
</tr>
</tbody>
</table>

1 Except where noted.
2 Excluding United States.
### Part B<sup>3</sup>  U.S. relative to OECD countries

<table>
<thead>
<tr>
<th></th>
<th>U.S. relative to OECD countries</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent of population without health insurance</td>
<td>Higher</td>
</tr>
<tr>
<td>Inequality across income levels in access to medical care</td>
<td>Higher</td>
</tr>
<tr>
<td>Open-ended insurance as percent of total</td>
<td>Higher</td>
</tr>
<tr>
<td>Non-government acute care hospital beds as percent of total</td>
<td>Higher</td>
</tr>
<tr>
<td>Physician fees (adjusted for cost of education)</td>
<td>Higher</td>
</tr>
<tr>
<td>Specialists as percent of all physicians</td>
<td>Higher</td>
</tr>
<tr>
<td>Ability to see a specialist in less than 4 weeks</td>
<td>Higher</td>
</tr>
<tr>
<td>Standby capacity of high-tech equipment and personnel</td>
<td>Higher</td>
</tr>
<tr>
<td>Drug prices</td>
<td>Higher</td>
</tr>
<tr>
<td>Administrative costs</td>
<td>Higher</td>
</tr>
<tr>
<td>Amenities (in hospitals and clinics)</td>
<td>Higher</td>
</tr>
</tbody>
</table>

<sup>3</sup> Data available for only a small number of countries or other problems in obtaining a precise comparison between the U.S. and other OECD countries.

Countries in the OECD: Australia, Austria, Belgium, Canada, Chile, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Luxembourg, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, United Kingdom, United States.

The Gross Domestic Product and Health Care Spending

By Victor R. Fuchs, Ph.D.

How much will the United States spend on health care during the next decade or two? The answer matters greatly to physicians, federal and state governments, businesses, and the general public. The answer will determine the type and extent of care that physicians can provide to their patients, as well as the amount of physicians’ take-home pay. It will also determine how much everyone else can consume or invest in other goods and services. Unfortunately, forecasting health care spending is extremely difficult. Future spending depends in part on developments within the health care sector and in part on developments in the economy as a whole. The former include changes in the prevalence of health problems such as obesity, infectious diseases, and dementia, as well as changes in medical technology such as new drugs, imaging devices, and surgical procedures. The economy as a whole includes variables such as the unemployment rate, trends in average wages, and prices of securities and housing.

The 2013 Economic Report of the President takes an optimistic view of future national health care expenditures, which is based on the slowdown in the rate of growth of those expenditures in recent years. Like most commentators, the report notes that one possible explanation is the recent recession, but it argues that this was not a major factor relative to improved efficiency in hospitals and physician groups, payment reforms, and early responses to the Affordable Care Act. If the United States is entering a new era of modest growth in health care spending, the current pressure for radical changes in funding, modes of payment, organization, and delivery of care would abate. On the other hand, if the current slowdown is primarily attributable to the most severe recession since the 1930s, or to one-time changes that are not relevant to future trends, then rapid growth in health care expenditures is likely to return when the economy becomes more robust. In that case, the heavy lifting to control cost growth remains to be done.

An examination of data from the past 60 years for the economy as a whole and for health care expenditures indicates that there has been a robust relationship between the two. It seems premature to dismiss the sluggish economy as the major explanation for the spending slowdown of recent years. In the line graph, the economy is represented by the gross domestic product (GDP), which is the total value of all goods and services produced in a given year or its equivalent, the total income received by all contributors to production (e.g., labor, management, and capital). The GDP and national health care expenditures are adjusted for population growth and general inflation. Between 1950 and 2011, real GDP per capita grew at an average of 2.0% per year, while real national health care expenditures per capita grew at 4.4% per year. The gap between the two rates of growth — 2.4% per year — resulted in the share of the GDP related to health care spending increasing from 4.4% in 1950 to 17.9% in 2011. Most experts believe that a gap of close to this magnitude over many future years would have catastrophic consequences for the federal government and the U.S. economy.


In order to observe whether fluctuations in national health care expenditures are related to fluctuations in the GDP, annual data for each series are smoothed with a 5-year moving average (to increase reliability), and the GDP value for each year is increased by 2.4% (the average gap) to facilitate visual comparison of short term movements in the two series. The correlation is not perfect, but over a period of 60 years, most sharp increases (and decreases) in the GDP have been accompanied by similar movements in health care expenditures. Note the long acceleration in both series in the 1960s, the slowdown around 1980, the subsequent acceleration in the late 1980s, and the recent sharp deceleration when both national health care expenditures and GDP rates of growth fell by more than 2.0% annually in just a few years.

The one big exception to the correlation is the mid-1990s, when growth of real national health care expenditures per capita was below 3% per year even though real GDP per capita was accelerating. This was precisely the period during which managed care became widespread. Prior to the 1990s, most insured patients could choose freely among providers, physicians were paid on a fee-for-service basis, and their decisions were rarely questioned by insurers. Under managed care, insurance companies selectively contracted with hospitals and physicians, fees and prices were negotiated in advance, physician decisions were subject to outside review, patients faced financial penalties if they obtained care “out of plan,” and providers sometimes shared in the insurance risk. A backlash from patients and providers followed, accompanied by a large increase in health care spending.

The spread of managed care in the 1990s, however, seems to have had an effect on long-term trends in expenditures as well as on short-term changes. Between 1950 and 1995, real health care expenditures per capita grew at an average annual rate of 4.7%, while real GDP per capita grew at 2.1%. Between 1995 and 2011, the average rates were 3.1% for real health care expenditures per capita and 1.4% for real GDP per capita. Thus, the average gap fell from 2.6% in the pre-1995 period to 1.7% in the post-1995 period. Resumption of the 60-year gap of 2.4% per year until 2040 would result in health care’s absorbing 30% of the GDP, as compared with the current 18%. Continuation of a 1.7% gap until 2040 would result in health care’s absorbing 26% of the GDP, a level that would still pose problems for the economy and especially the federal budget.

Some observers place great emphasis on the particularly slow growth of national health care expenditures in 2010 and 2011. How useful is the experience of growth over a period of 2 years in predicting the growth rate over the next 20 years? The answer seems to be not at all. The scatterplot shows 2-year growth rates on the horizontal axis and the corresponding subsequent 20-year growth rates on the vertical axis. The period covered is 1950 to 1991 (the last year for which we have 20 subsequent years of observations). The correlation between the 2-year and 20-year rates is actually negative, −0.22, but not statistically significant.

When speculating about future growth of health care spending, it is also important to note that some of the reasons for the slow growth in the past 2 years, such as the switch from brand-name drugs to generics and the reductions in hospital readmissions, are one-time gains, not alterations in such determinants of
long-term growth as new medical technology and the aging of the population.

In conclusion, the rate of growth of national health care expenditures in the past appears to have been substantially related to the growth of the GDP. There has been some slowing of the growth of health care spending relative to the GDP, but it began not just a few years ago, but in the 1990s, for reasons that remain to be determined. One possibility is that the movement to managed care in the 1990s resulted in long-term slowing of health care spending, an effect temporarily obscured by the increase in spending during the backlash.

Disclosure forms provided by the author are available with the full text of this article at NEJM.org.

From the Departments of Economics and Health Research and Policy, Stanford University, Stanford, CA.
Rapid advances in medical science and technology, substantial gains in health outcomes attributable to medical care, and budget-busting increases in health care expenditures fueled by private and public insurance have marked the past six decades of health care in the United States. As the country struggles to emerge from a multiyear financial and economic crisis, policymakers and the public have increasingly homed in on those skyrocketing health care expenditures. What lessons can be drawn from the evolution, since 1950, in the sources of payment and objects of expenditures in the health care arena?

Health Expenditures

The rapid growth of health expenditures is one of the most important economic trends in the United States in the post–World War II era. It has implications for the financial viability of federal and state governments and has resulted in stagnation of wages in most industries. In 1950, health expenditures accounted for only 4.6% of the gross domestic product (GDP). In 2009, they accounted for more than 17%, a larger share than all manufacturing, or wholesale and retail trade, or finance and insurance, or the combination of agriculture, mining, and construction. According to public finance experts such as Alan Blinder and Alice Rivlin, control of health care expenditures is the greatest fiscal policy challenge facing the United States.

From 1950 through 2009, there was an almost continuous increase in annual real per capita health expenditures, with the exception of one 2-year pause in the mid-1990s, when the effect of managed care was at its peak1 (see line graph). The absolute rate of growth has been increasing over time, as evidenced by the concave shape of the curve in the graph. The relative rate of increase was greater between 1950 and 1980 than between 1980 and 2009 — 4.6% versus 4.1% per year — primarily because of the introduction of Medicare and Medicaid in 1965.

Unfortunately, the slight slowing in the rate of growth of health expenditures since 1980 was accompanied by even greater slowing in the growth of the GDP (per capita adjusted for inflation), from 2.6% per year in 1950–1980 to 1.8% per year in 1980–2009. Thus, the gap between the rate of growth of health expenditures and that of GDP increased from 2.0% to 2.3% per year between the two periods. Most experts believe that such a gap is not sustainable over the long term, because health expenditures would cut too drastically into the availability of other essential goods and services.

The most important explanation for the increase in real per capita health expenditures is the availability of new medical technology2 and the increased specialization that accompanies it. Between 1974 and 2010 alone, the number of U.S. patents for pharmaceutical and surgical innovations increased by a factor of six. Second in importance is the spread of public and private health insurance, which diminishes the effect of health care prices on demand.3 There is a positive-feedback loop between new technology and the spread of health insurance: new technology stimulates the demand for insurance, and the spread of insurance stimulates the demand for new technology.4 Finally, a small portion of the increase, typically 0.1 or 0.2 percentage points per year, is due to the increased health care needs of the aging population.5

---

2 Pauly MV. Competition and new technology. Health Aff (Millwood) 2005;24:1523-35.
It’s not possible to estimate how much of the increase in expenditures reflects higher health care prices and how much reflects greater quantities of care, because the content of a day in the hospital or a visit to a physician keeps changing. No doubt some of the increase in expenditures reflects an increase in the quantity of medical care, if quantity is adjusted for improvements in the quality of care.

**Sources of Payment**

The sources of payment for medical care have changed significantly since 1950 (see table). The most important trends have been a decline in out-of-pocket payment and a rise in third-party payment (both private and public), an increase in government’s share of payment and a decrease in the private share, and an increase in the federal government’s share as compared with that of state and local governments.

Third-party payment has grown partly because of expensive interventions that expose individuals to large financial risk and partly because employers’ contributions to employee health insurance are not considered part of employees’ taxable income. Since World War II, there has been a large increase in the number of workers who must pay income tax and an even greater increase in the number who must pay payroll taxes. These increases have made tax exempt employer-based health insurance more attractive. A shift from individual to group insurance has also contributed to the spread of coverage by reducing marketing and administrative costs and, thanks to compulsory participation within firms, limiting the risk of adverse selection for insurance companies.

The growth of government’s share, and especially the federal share, can be explained by the public’s desire to cover more of the public with insurance and private insurers’ difficulty in providing coverage for the elderly and the poor. Federal legislation also substantially extended public coverage for children.

**Objects of Expenditures**

Throughout the period since 1950, health expenditures have gone primarily to hospitals, physicians, and drugs. Moreover, the rate of growth of expenditures in each of these

---

* The percentage of payments by the federal government was calculated on the basis of National Health Care Expenditure data. Data are from the Department of Health and Human Services and the U.S. Census Bureau.
categories between 1950 and 2009 has been fairly close to the rate of growth of total health expenditures (see bar graph). Drug expenditures may appear to have grown more slowly, but that's probably due to a data mismatch: the 1950 figure includes sundries, whereas the 1980 and 2009 figures are for prescription drugs only. Such stability in the share of these categories is remarkable, given the great changes that have occurred in medical technologies, sources of payment, and health policy since 1950. As a rule of thumb, the ratio 3:2:1 does a fairly good job of describing the relative importance (in dollar terms) of hospitals, physicians, and drugs. The “other” expenditures are divided among many categories, the most important of which are public administration and the net cost (premiums minus benefits paid) of private health insurance, nursing homes, and dental services.

There have been periods in the past 60 years when individual categories accounted for greater or lesser proportions of expenditures. Spending for hospital care and physicians received a boost between 1950 and 1980 from the introduction of Medicare and Medicaid. Spending for drugs accelerated sharply after 1980 following the introduction of a host of new products for treating heart diseases, mental illness, gastrointestinal disorders, and cancer and a large increase in private and public insurance coverage for drugs.

The ability of hospitals to maintain their high share is particularly noteworthy, because between 1950 and 2009 the industry had several large shocks. Psychiatric hospitals virtually emptied out. Admission rates to acute care hospitals (“community” hospitals) dropped precipitously after 1970, as did the average length of stay. As a result, the average daily census, adjusted for population growth, has decreased by almost 50% over the past four decades. Hospitals have maintained and increased their revenues in part through more intensive treatment of inpatients. Despite shorter stays, the cost per case (in 2009 dollars) jumped from $6,600 in 1997 to $9,200 in 2009. Hospitals’ total incomes were also preserved through expansion of outpatient services, including same-day surgery, magnetic resonance imaging and computed tomography, and outpatient clinics for diagnosing and treating cancer, heart disease, and other illnesses.

Community Hospitals

Community hospitals (including academic medical centers), the recipients of the largest share of health expenditures, have seen dramatic shifts in demand for and supply of inpatient care since 1950. During the first three decades of this period, the number of inpatient days per 1000 population increased by more than a third, driven by Medicare and Medicaid, the spread of employer based insurance, and lax utilization controls by public and private payers (see Table 1 in the Supplementary Appendix, available with the full text of this article at NEJM.org). A slight decline in the average length of stay was more than offset by a 50% increase in the number of admissions per 1000 population. The industry’s 31% increase in the number of beds per 1000 population, abetted by consultants’ predictions of ever-growing demand, proved to be an expensive mistake. In the late 1960s and early 1970s, there was mounting evidence that many hospital admissions were ill-advised and that lengths of stay for many patients were overly long (see the Supplementary Appendix).

Between 1980 and 2009, the number of inpatient days per 1000 population fell by almost half, with declines in admissions and average length of stay contributing almost equally. The decline in length of stay was particularly spectacular in some major categories of patients. For example, stays for uncomplicated myocardial infarction dropped from 3 weeks to 3 days; for uncomplicated vaginal delivery, from 1 week to 1 day; and for herniorrhaphy, from 6 days to same-day surgery. The average decrease among all patients, however, was smaller than those for individual causes of admission, because the average severity of patients’ conditions on admission increased. The hospital industry responded to the drop in demand by closing some hospitals (net decrease of 18%) and closing off some beds as unavailable, but even so, the average occupancy rate fell by 10 percentage points to the inefficient level of 65.5%.

Physicians

The number of active physicians in the United States increased by a factor of approximately four between 1950 and 2009 (see Table 2 in the Supplementary Appendix). As the...
population grew, the number of active physicians per 1000 population increased from 1.41 to 2.73, an annual growth rate of 1.1%. That figure may overstate the growth of physicians' availability, however, since the number of hours the average physician worked probably decreased appreciably between 1950 and 2009. Major trends in the physician supply that had important implications for the health economy were large increases in the percentages of female physicians, specialist physicians, and hospital based physicians.

Because women, even professional women, still bear a disproportionate share of domestic responsibilities, female physicians tend to differ from their male peers in preferences regarding annual hours of work, night coverage, self-employment, specialty choice, and other aspects of practice.

The increase in the proportion of physicians who are specialists and subspecialists has resulted in a considerable increase in the number of years the average physician spends in training, although a restructuring of medical education could change that. There has been a large increase in the number of specialists and an even larger increase in the number of specialties and subspecialties, from a few dozen 50 years ago to more than 150 now.

The shift away from office based practice, along with possible changes in payment systems, may portend a time when most medical care will be delivered by teams of physicians and other health care providers (e.g., nurse practitioners and physician assistants) working in accountable care organizations.

Changes in Organization and Delivery

An important recent trend affecting hospitals and physicians is a sharp division between physicians who treat outpatients and others, called hospitalists, who treat only inpatients. The number of hospitalists has grown rapidly, from no more than 1000 15 years ago to 7000 10 years ago to approximately 30,000 in 2011, according to physician-economist David Meltzer of the University of Chicago. Hospitalists are said to improve both the efficiency of care (mostly through reducing lengths of stay) and its quality. Though primary care physicians initially resisted this change in professional responsibilities, many now prefer the new system because they perceive that hospital visits were not an efficient use of their time.

Another trend attracting wide attention is the use of electronic medical records (EMRs) in physicians' offices. Opinions vary regarding the effects of EMRs on the efficiency and quality of care. I believe a well-organized health care system can benefit substantially from EMRs, but the fragmented nonsystem of U.S. medical care is not likely to derive enough benefit to justify the cost.

During this period, another change that affected hospitals and physicians was the development of managed care. Until about 1990, most insured patients could choose freely among providers, physicians' decisions were not subject to frequent questions by insurers, and payment was typically fee for service. The rapid growth of health care expenditures in the late 1980s, combined with sluggish growth of the GDP, fueled a demand for change. In the 1990s, insurers selectively contracted with providers, fees and prices were negotiated in advance, physicians' decisions became subject to insurance-company review, and patients faced financial penalties for obtaining out-of-plan care. The effect on health care expenditures was dramatic: growth rates fell to 2% per year by the mid-1990s. At the same time, GDP growth accelerated to about 3% per year. Both physicians and patients, however, grew increasingly critical of managed care. Physicians complained about a squeeze on their incomes and interference with their autonomy. Patients resented restrictions on their choice of providers and worried that cuts in spending would necessarily result in a poorer quality of care. The complaint by physicians and patients that health outcomes were adversely affected by managed care, fueled by many anecdotes, has not been supported by systematic evidence.

The term “managed care” still carries negative connotations for many observers, but as long as concern about cost is strong, it's difficult to imagine a widespread call for unmanaged care. Stakeholders will disagree about who should do the managing, about the relative roles of regulation and competition, and what form competition should take. Perhaps the most important future trend, too nascent to quantify, let alone evaluate, is the replacement of the current system of organization and delivery with competition among large accountable care organizations serving defined populations for risk-adjusted per capita annual payments.

Past and Future

The six decades since 1950 have been remarkable for the U.S. health economy in many ways, especially the extraordinary increase in health care expenditures. Future historians may, with some irony, refer to this period as a golden age for U.S. medicine because health care's share of the GDP quadrupled from 4.6% in 1950 to more than 17% in 2009; in most peer countries, the share is 9 to 11%. Other noteworthy trends in the health economy have been the spread of private and public health insurance to the point where almost 90% of

---

6 Fuchs VR. Alan Gregg Lecture: The structure of medical education — it's time for a change. Presented at the Annual Meeting of the American Association of Medical Colleges, Denver, November 6, 2011.
the total bill for care is paid by third parties; the increased role of the federal government in funding health care; the decline in inpatient use of hospitals (fewer admissions and shorter stays) and the expansion of hospital outpatient services; the shift in the physician workforce toward more women, more specialists, and more hospital-based physicians; and the deluge of new medical technologies confronting clinicians with a menu of 6000 drugs and 4000 procedures to choose from.

It is difficult to see how the health sector can continue to expand rapidly at the expense of the rest of the economy, but every past prediction of a sustained slowing of the growth of health expenditures has been proved wrong. Rapid growth may continue as a result of political gridlock regarding the form that curbs on expenditures should take. There is no public consensus about how much care should be provided for the poor and sick or how it should be done. Similarly, there's no public consensus regarding efforts to increase the efficiency of care. A rational approach to the financing, organization, and delivery of care seems politically impossible. However, the observation by de Tocqueville that in the United States “events can move from the impossible to the inevitable without ever stopping at the probable” may prove to be prescient.

Disclosure forms provided by the author are available with the full text of this article at NEJM.org.
The Physician’s Role in Controlling Costs
By Victor R. Fuchs, Ph.D.

The decline in age-adjusted mortality from diseases of the heart and stroke is the greatest medical success since antibiotics. Figure 1 shows that since 1965, age-adjusted mortality in the U.S. from diseases of the heart has fallen approximately by two-thirds and from cerebrovascular diseases by almost 80 percent. Once these causes of death are subtracted from total mortality, there has been a small decline since 1965 and no decline in U.S. mortality since 1980. The medical profession as whole likes to take credit for extending life expectancy, but most of the credit goes to the physicians who treat heart disease and stroke.

Despite this success, mention of health care doesn’t always bring cheers from the American public. Why is that? Probably the most important reason is the ever increasing cost of care. Figure 2 shows indexes of expenditures (1974=100) by source of payment for care and GDP minus health expenditures, adjusted for population growth and general inflation. Two major conclusions emerge from Figure 2. First, spending on health care, regardless of the source of payment, has grown much more rapidly than the rest of the economy (GDP minus health expenditures). The result has been traumatic for each of the three sources of payment. In the private sector, wages for many workers have been stagnant for three decades because productivity increases have mostly gone to pay for higher health insurance premiums. State and local governments have been covering their higher health costs by slashing expenditures for education, skipping needed maintenance of roads, bridges, and tunnels, and skimping on contributions to employee pension funds. The federal government has mostly dealt with its problem of higher health care costs by borrowing from the Chinese, the Japanese, and other lenders.

The second important takeaway from Figure 2 is the special squeeze of health care on the federal government. That squeeze is at the heart of current political and economic controversies over the budget deficit and mounting government debt. Peter Orszag, Citigroup Vice Chairman and former Director of CBO and OMB, has written, “Healthcare costs are the core long-term fiscal challenge facing the U.S....President Obama is right. We do not have a spending problem—we have a healthcare problem.”

1 Financial Times, 22 January 2013.
The pressure on costs has already begun, and as it gets stronger, it is understandable that cardiologists, the main authors of the medical success story might feel they are being punished unfairly. They are not. Indeed, punishment is not the issue. Cardiologists will just be caught up in the general effort to control costs.

If cardiologists get special attention, it will be for the same reason that Willie Sutton is supposed to have given for always robbing banks – “that’s where the money is.” Spending on circulatory system conditions is by far the largest category of health care expenditures, almost one dollar in five\(^2\). The next highest category, “mental disorders,” accounts for only about half as much spending as circulatory system conditions. Total health care spending in the U.S. this year will approach 3 trillion dollars. That’s over $9,000 per person – man, woman, and child. The share going to the circulatory system category will be about 500 billion dollars, putting it roughly on a par with defense expenditures, or what Americans spend on education at all levels.

What to do? And what role can and should physicians play in controlling costs?

There is an identity equation in economics that tells us that expenditures must be equal to the product of three terms: First, the quantity of output \((Q_0)\), that is, visits, tests, prescriptions, hospitalizations, and procedures. Second, the reciprocal of productivity \((Q_1/Q_0)\): how much input of labor, equipment, facilities, and supplies is required for a given level of output. Third is the prices of inputs \((P_1)\), including wages and salaries, physicians’ incomes, drug and equipment prices, and the like.

\[
E = (Q_0)(Q_1/Q_0)(P_1)
\]

Physicians are limited in what they can do about input prices. If they agreed on more standardization, they could get some reductions on the price of drugs, devices, and supplies. They can contribute to improving productivity through standardization, restructuring of work processes, but some productivity is beyond their control. The biggest potential impact of physicians is on the first term, the quantity of output. There can be no doubt that physicians play a major role in determining the frequency of visits, tests, and procedures. They write the prescriptions. It’s physicians who admit patients to hospitals and who tell them when should leave.

What is the appropriate amount of output? Or as a physician might say, what is the optimal quantity of care? As a practical matter, this is often a very difficult question to answer because the necessary data are not known. But lack of data aside, it is a problematic question because the optimal quantity can be defined from a social perspective or a medical perspective, and the two will rarely be the same.

Figure 3 illustrates this dilemma with two simplified diagrams. Version A is based on total cost and total benefit, and version B makes the same points with marginal cost and marginal benefit (which are the first derivative of the total cost and total benefit curves). As quantity increases for an individual patient or a population of patients, cost increases as do potential benefits, but the latter usually at a decreasing rate. For example, as the indication for joint replacement is broadened, the marginal benefit tends to decrease, and might reach a point where the risks outweigh the benefits. Marginal costs and marginal benefit are determined by the slope of the total cost or total benefit curve at any particular quantity of output.

---

What is the optimal quantity of care? From the social point of view it is the amount of output where the marginal cost equals the marginal benefit. It is $Q_1$ in Figure 2. Anything less than that amount would yield additional benefit greater than the additional cost; anything more than that would add more to cost than to benefit. Society as a whole is best off when marginal benefits and marginal costs are equal for every kind of output. If there wasn't this equality everywhere, total benefit to society could be increased by moving resources to that kind of output where marginal benefit is greater than marginal cost and away from output where the reverse is true.

The medical optimum, however, is to increase care until any addition yields no benefit. It is $Q_2$. An insured patient might say, “Do everything that could possibly help regardless of cost.” But no nation can give all patients all the care that might do them some good. Costs must be controlled.

Who should have the major role in controlling cost? Patients? Regulators? Physicians? Some observers advise focusing on patients because health insurance leads patients to want “too much” care. If patients had “more skin in the game” as the saying goes, expenditures would be less. A small amount of patient co-pay at the point of service might be helpful, but the rate of co-pay should vary with the benefit of the service, e.g., preventive care at zero co-pay and some services of unproven benefit at very high co-pay.

At bottom, however, reliance on patients to have a major role in controlling costs is questionable because patients’ knowledge about costs and especially benefits is often very limited. Also, when ill or facing serious illness for a loved one is not a good time to make rational decisions about care. The prospect of a barely conscious patient choosing between two or three stents and between bare metal or coated seems preposterous. Moreover, health care expenditures are concentrated among a small percentage of patients with very high utilization. Five percent of patients account for half of all expenditures in a year, and they are often those with least capacity to assess costs versus benefits. This precludes having patients play a major role in controlling costs.

Control by regulators also seems unreliable. They often lack the necessary expertise, and are often subject to political pressures from administrators and legislators, who have other objectives.

Of the three possibilities, it seems to me that physicians are likely to be better informed than patients or regulators and could potentially do the best job of providing cost-effective care.

But they need 3 “I’s” to do so. Without information physicians can’t be expected to provide the appropriate amount of care. This information about the costs and especially the benefits of alternative interventions must come from research and must be available to physicians in a timely, user-friendly way. Physicians also need access to the technologies and personnel, i.e., infrastructure, most appropriate to dealing with any particular problem. It is one thing to know that a team approach to caring for a patient with diabetes is preferred, but without access to such a team, the information is of little value. Finally, the actual practice of cost-effective care depends on physicians having the right incentives. The wrong incentives will not lead to appropriate care, even if information and infrastructure are available. Many health policy experts question whether the prevailing system in the U.S. of fee-for-service payment provides the right incentives for practicing cost-effective care.

A final question is whether cost-effective care, that is the social optimum, is consistent with medical ethics. The new Physician Charter, as well as the venerable Hippocratic Oath, emphasizes the physician’s responsibility to the individual patient. If the patient is insured, that would seem to call for the medical optimum of care – any care that might do the patient some good, regardless of cost.

But the same Physician Charter, now affirmed by physician organizations representing more than half of all U.S. physicians, also says the following, “While meeting the needs of individual patients, physicians are required to provide health care that is based on the wise and cost-effective management of limited clinical resources.” (italics added) In other words, aim for the social optimum of care. In so doing, a physician would be acting in a way that, if all physicians acted in that same way, the welfare of society would be enhanced.
IF ACCOUNTABLE CARE ORGANIZATIONS ARE THE ANSWER, WHO SHOULD CREATE THEM?

By Victor R. Fuchs, Ph.D. and Leonard D. Schaeffer

Public policy experts across the political spectrum agree that health care cost growth must be reduced to avoid major economic problems, including a federal deficit crisis. The need is urgent because the United States cannot continue to decrease funds to support education, postpone crucial maintenance and repair of infrastructure, potentially limit national security, and borrow from international lenders to finance health care. The intimate connection between health care spending and the nation’s debt crisis is well understood by public finance experts. Alice Rivlin, first director of the Congressional Budget Office, former head of the Office of Management and Budget, and former vice chair of the Federal Reserve Board, concluded, “Long-run fiscal policy is health policy.”

Minor tweaks to the current fragmented system will not lower costs or significantly improve quality of care. In short, US health care requires a major change in organization and delivery, but time is running out.

There is little opportunity to experiment with dozens of models of uncertain value. Fortunately, a consensus is building that cost control requires changing how health care is delivered and financed. The fastest route to establishing a new approach is to start with lessons already learned. First, the recent Congressional Budget Office assessment of 10 major Medicare demonstration projects indicated that programs using fee-for-service payment did not reduce federal spending. The one program that produced respectable savings used bundled payment. Also, a recent large-scale study concluded that pay-for-performance does not improve health outcomes. The lesson is do not build new strategies on a fee-for-service platform. Second, managed care has evolved since the 1990s to address the use of specialty care. There should be patient choice of physicians and hospitals but financial incentives to encourage patients to use those affiliated with the network. Third, the mixed history of physician groups and physician-hospital organizations to successfully manage financial risk suggests that this is not a skill that many have mastered. The lesson is that turning these physicians and hospitals into insurers will be difficult for many.

A promising proposal to reform delivery and payment is to shift to accountable care organizations (ACOs). To succeed, however, ACOs must be capable of providing quality care to large, self-defined populations (through prospective annual enrollment) for a risk-adjusted capitation payment. Accountable care organizations must be able to coordinate the full continuum of care across multiple delivery sites using electronic medical records, be capable of creating and adhering to cost-effective guidelines for diagnosis and treatment, and be able to measure, monitor, and report clinical and financial performance.

Capitated ACOs should be able to deliver the same or better quality care at lower cost for several reasons. First, physicians in an ACO would most likely feel significantly less pressure to order tests, visits, or procedures or hospitalize patients to increase revenue for themselves or their institution. One
estimate suggests that additional services that do not improve health may cost as much as $226 billion in 2011.6 Second, prescription drug costs could be reduced through formularies, use of generic products when appropriate, and judicious purchasing. Third, an efficient mix of personnel made possible by the scale of operations would be another advantage. Significant savings could be realized if subspecialists, specialists, primary care physicians, and other practitioners such as nurse practitioners and physician assistants are used in optimal proportions so that each group is focused on what they do best. This readjustment of proportions will increase demand for some professions and specialties and decrease demand for others. Fourth, capitated ACOs would not bill for individual visits, procedures, or hospital stays, thus generating substantial administrative savings.

Who is most able to create and manage this type of organization that can effectively implement both medical and insurance functions? Employers? They would like to bring costs under control while maintaining quality, but most employers have not demonstrated sustained interest or capacity to manage health care for their employees. This task requires knowledge and skills far different from those present in the typical firm.

Physicians? No one knows as much about health care as physicians, and given the right information, infrastructure, and incentives, most physicians would practice cost-effective medicine. With the exception of a few large, multispecialty group practices, however, physicians have undervalued and underinvested in systems, administration, customer service, and financial functions that would be necessary for a successful ACO. In addition, physicians’ traditional emphasis on autonomy runs counter to the standardization, group decision making, measurement of outcomes, and peer review that are important for the success of ACOs.

Hospitals? Although hospitals clearly have more capital and more professional management than physicians, most hospital administrators have been trained and rewarded for their ability to fill beds, not the background and mindset required for an effective ACO. Moreover, more hospital consolidations, unless administrators have been trained and rewarded for their ability to manage health care for their employees. This task requires knowledge and skills far different from those present in the typical firm.

Some health plans are already partnering with physicians and hospitals in different ways to care for patients in a more effective, efficient, and integrated manner. Health plans, including UnitedHealth, Humana, Aetna, WellPoint, and Highmark, have invested billions in information technology, including data-driven tools and services to provide connectivity and information needed for alignment between hospitals, physicians, and payers. Health plans are also purchasing medical groups, independent physician associations, and clinics in an effort to support accountable care. Some traditional health maintenance organizations such as Kaiser Permanente have already formed functioning ACOs.

There are numerous examples of what is possible to enhance value in health care. Intermountain Health System, a hybrid of insurance functions and hospitals, physicians, and multispecialty clinics, has focused on improving value. Its children’s hospital had a cost per case slightly above the national average, but in 15 years has reduced cost per case to 60% of the national average.7 Kaiser Permanente, long a value leader in northern California, has developed successful programs to control sepsis and to avoid pressure ulcers. A joint effort by Aetna and Virginia Mason Medical Center reduced unnecessary and marginal utilization of magnetic resonance imaging and computed tomographic scans, blood tests, and other auxiliary services. The Geisinger Health System, another insurer-hospital-physician hybrid, charges a single-episode price for all services associated with a surgical procedure for 90 days, including rehospitalizations. WellPoint has entered into partnerships with Healthcare Partners in California and with Dartmouth-Hitchcock Medical Center in New Hampshire to create ACO pilots. WellPoint has also teamed with primary care physicians across the country to establish patient-centered medical homes to test new payment models. In Colorado, its patient-centered medical home program experienced an 18% decrease in acute inpatient admissions and a 15% decline in

---


emergency department visits during the first 2 years.8

These gains may just be the beginning. If the present fragmented structure of the US health care system were replaced by well-organized ACOs paid by risk-adjusted capitation, improvement in one aspect of care in one plan would be rapidly copied by others. Health care could begin to resemble industries in other sectors in which productivity gains diffuse rapidly from one firm to another.

Numerous federal and state laws, regulations, and public policies could slow down or bar the formation of ACOs. Prudent modification or removal of these barriers is necessary for health care to adapt to a new world of ACO-based competition. Regulators could accelerate development of ACOs by establishing a regulatory framework that would clearly delineate the responsibilities of the integrated care delivery system vs those of the health plan to ensure that the distinct functions are in place, require ACOs to demonstrate a minimum set of competencies for improving value, and apply solvency/risk-based capital standards to ACOs that assume significant risk.

In nominating health plans as the most feasible candidate to get ACOs rolling within the available time frame, the logic developed by Sherlock Holmes is instructive. When Dr Watson asked Holmes how he arrived at an unusual solution to a difficult case, Holmes replied, “When you have eliminated the impossible, whatever remains, however improbable [italics in original], must be the truth.”9

Conflict of Interest Disclosures: Both authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Mr Schaeffer reported having board memberships and stock/stock options from Amgen, Quintiles, Allergan, and Surgical Care Affiliates. Dr Fuchs reported receiving a grant from the Robert Wood Johnson Foundation.

Online-Only Material: The Author Audio Interview is available at http://www.jama.com.


Experts agree that there is substantial waste in the US health care system. This waste drives up costs, threatens the government’s long-term fiscal stability, suppresses incomes, and reduces resources for public education and other essential services. Similarly, there is substantial waste in the education and training of US physicians. Years of training have been added without evidence that they enhance clinical skills or the quality of care. This waste adds to the financial burden of young physicians and increases health care costs. The average length of medical training could be reduced by about 30% without compromising physician competence or quality of care.

The Obsolete Image of the Ideal Physician

For decades, the ideal academic physician has been the triple threat: an incisive diagnostician and empathetic clinician, a productive researcher, and a scintillating teacher. Similarly, the clinical practitioner was supposed to be omnicompetent, capable of managing all illnesses. The consequence is a broad training regimen that includes mandatory research experience for all physicians, and emphasizes the autonomy of the physician rather than team-based care.

The new model recognizes that with increasing clinical and scientific complexity, no physician can be a competent triple threat; that few clinicians will also be investigators; that no single clinician can know everything even in his or her own specialty; and that effective care requires collaborative, multidisciplinary teams. Medical education in the United States needs to adapt to this changing medical environment and physician ideal. Four elements in the present structure of medical education offer significant opportunities to shorten the training period for most physicians.

Premedical Training

More than 30 medical schools successfully operate 6- or 7-year medical programs in which premedical training is reduced from the typical 4 years of college to 2 or 3 years. Moreover, most medical schools in the United Kingdom and Europe have 6 years of medical school training after graduation from high school. While data are limited, there is no evidence that graduates of 6-year programs perform more poorly on standardized board examinations or as practicing physicians. Students who want the traditional 4 years of college should be free to pursue them, but medical schools should not make it an entrance requirement.

Medical School Training

Why is medical school 4 years in length? The answer probably has to do with the Flexner Report’s recommendation in 1910 for 2 years of preclinical science training followed by 2 years of clinical training. Yet most physicians could be trained in significantly less time. Since 1997, the University of Pennsylvania has only 1-1/2 years of preclinical science training. Duke University medical students focus on the basic sciences in the first year, complete core clerkships during the second year, and devote the third and fourth years to research and electives. While outcomes data on alternative training arrangements are limited, there is no evidence that students...
from either school perform worse on board examinations, placement in residency programs, or other significant metrics of competence.\(^5\)

The important patient care skills can be obtained in less than 2 years of clinical training. The medical school at Harvard University requires students to complete only 15 months of clinical rotations.\(^6\) It is not difficult to eliminate 1 year of medical school training (1/2 year of preclinical and 1/2 year of clinical training) without adversely affecting academic performance. Having 1-1/2 years of clinical training would still give students sufficient exposure to a range of specialties. This change would be consistent with the increasing emphasis on individualized instruction and assessing students on core competencies rather than on time served. Consistent with this proposal, Texas Tech School of Medicine as well as 2 Canadian medical schools now offer 3-year programs.

Residency Training

It is also possible to reduce residency training by 1 year. For internal medicine, pediatrics, and similar 3-year residencies, the third year is not essential to ensure competent physicians. This residency year is mainly engaged in supervising and teaching interns, in taking electives, or in some cases conducting research. While valuable, these activities are hardly essential to becoming a knowledgeable practitioner. Indeed, many trainees are permitted to short track into subspecialty fellowships, reducing their residency from 3 to 2 years. Shortening training in an era of work-week limits will force hospitals to reengineer programs to ensure residents’ clinical competence—a worthwhile exercise.

Most surgical training programs include at least 1 year of research. The most important factor in becoming a competent surgeon is high volume—performing specific procedures many times over.\(^7\) A research year does not add to surgical volume and skills building. A required research year might be relevant if all trainees were destined to become academic physicians. But most trainees will become practitioners; they will not use these gained research skills in their career and their training will be reduced by 1 year. The third year of internal medicine or pediatric residencies or the research year in surgical specialties could be eliminated without compromising the clinical quality of trainees.

Subspecialty Fellowship Training

The typical medical and pediatric subspecialty training is a 3- or 4-year program. The structure involves 1 or 2 years of clinical training that entails caring for patients, performing consultations, and other patient-centered activities, typically followed by 2 years of mentored research with reduced clinical time. This structure indicates that learning the patient care aspects of a medical subspecialty can be accomplished in the 1 or 2 years of intensive clinical training. The time devoted to research is relevant only for trainees destined to become academic researchers. In surgical subspecialties, time could be saved by reducing the amount of training in general surgery. Instead of having reconstructive surgeons become experts in appendectomies, subspecialist surgeons could be trained to achieve clinical competence without spending several years performing general surgery.

A Proposal for the Future of Physician Training

Currently, it takes an average of 14 years of college, medical school, residency, and fellowship to train a subspecialty physician. This period could be reduced to 10 years or by approximately 30%.

Why should a reduction in the training time of physicians be considered? Efficiency has its own value. Waste, especially wasting the time of some of society’s most highly educated and talented people, is unethical. Inevitably in the near future, efforts to reduce the Medicare budget will likely be accompanied by a reduction in the federal government’s support of graduate medical education. Streamlining residencies will save academic health centers money because they would have to spend less on the extra costs associated with training that are now compensated by federal support for medical education. In addition, shortening the length of training would benefit medical students and trainees. With 1 year less of medical school, they would have lower debts from tuition. This reduction could be significant because the average medical student graduates with $160,000 in debt.

Another advantage of shortening the length of training would be to focus attention on the essential content of medical training. Changing the structure of training would force medical leaders to eliminate unnecessary and repetitious material and emphasize training physicians to become part of a care team; enable physicians to recognize their limitations as well as their competencies; enable physicians to use evidence more effectively to improve care; and enable

---


physicians to become comfortable with group decision making, standardization of practices, task shifting to nonphysician providers, and outcomes measurement.

**Objections Considered**

Some physicians may fear that shorter training will not produce high-quality clinicians. However, several first-rate institutions have already shown that each of the reductions discussed can be achieved without decreasing physician competence. Medical schools that provide only 3 years of required classroom and clinical instruction, residencies that certify short-tracking trainees after only 2 years, and subspecialty fellowships that do not require prior specialty training all have shown that added years are not needed.

Shortening medical school training to 3 or 4 years might reduce the maturity, life experience, and socialization of practicing physicians who might start practicing as young as 26 years of age. Certainly clinicians would be younger, but that should not be conflated with immaturity.

Removing time that residents and fellows spend conducting research might affect academic medicine because fewer physicians might choose a research career. Trainees interested in academic careers as researchers will certainly need additional time to develop their basic science, clinical, or policy-related research skills. However, it is wasteful to add years of training for all physicians to ensure the small minority destined to be researchers has the opportunity to engage their interest in research.

A final objection to shortening training will be the coverage that residents and fellows provide in hospitals at night. Residency programs are already grappling with a reduction in work hours. However, the education of residents and fellows should not be held hostage to clinical service responsibilities.

**Conclusions**

Through slow accretion, years have been added to medical training. Yet many medical schools and residency and fellowship programs have already shortened their training in various ways, definitively demonstrating that these added years are not essential to training high-quality, knowledgeable practitioners. In an era when unnecessary medical services are being intensely examined to reduce costs, similar critical attention should be applied to eliminating waste from medical training, with a goal for US medical education to shorten training by 30% by 2020.
Specialization in Medicine. How Much Is Appropriate?

By Allan S. Detsky, M.D., Ph.D., Stephen R. Gauthier, B.Sc., and Victor R. Fuchs, Ph.D.

Professions develop around the delivery of specialized services. Lawyers give legal advice, electricians install wiring, and teachers provide education. At some point in the evolution of a field, licensure or certification defines its area of expertise. Licensure is a legal entity allowing only certain people to perform a task. Certification is a non-legally binding designation that informs consumers of qualifications. Frequently, licensing and certification are performed by professional organizations that oversee education, training/apprenticing, and evaluation through examination.

Specialization in medicine depends on 3 principal factors: advances in medical science and technology, professional preferences, and economic considerations. A new diagnostic tool or procedure may create a need for physicians with special training in its use. Some innovations, like lithotripsy, generate their own demand for specialists, whereas other innovations, like new angioplasty technologies, are endogenous, generated by the experience and needs of specialists. Some physicians are drawn to specialization because it offers defined responsibility, more control over their practice, prestige, and potential remuneration.

Economic considerations play a major role in the development of specialties. Adam Smith begins The Wealth of Nations by stressing the importance of specialization and the division of labor. He notes that its extent is limited by the size of the market. As markets for medical care have expanded through increases in population and income, urbanization, and improvements in transportation and communication, specialization within medicine has increased. Payment differentials between specialists and primary care physicians, which vary across countries, also influence physician choice.

Today, most individual physicians and surgeons are trained and qualified to provide only some kinds of care. Thus, the physician workforce has differentiated into a heterogeneous group of professionals. In 1960, there were only 18 specialty boards and a handful of subspecialties in the United States, but by 2011 there were 158 specialties and subspecialties. Canada has 67 specialties and subspecialties; France, 52; and England, 97. In addition to variation in the number of specialties, the distribution of physicians who are specialists (as opposed to generalists) varies between countries. Recent data from the Organisation for Economic Cooperation and Development show that 12% of physicians in the United States are generalists (including family physicians), compared with 47% in Canada. However, most of this difference is explained by nomenclature, because general internists and general pediatricians, many of whom deliver primary care in the United States, are considered specialists in this report. When these 2 groups of physicians are included, the US proportion of generalists is 36%.

These differences raise the question of how much medical specialization is good for society. The answer depends on the effect of specialization on health outcomes, such as length of life and quality of life, and costs. Circumstantially, it appears that the United States has developed more specialization and higher costs without offsetting gains in health outcomes. The criteria for certifying a new subspecialty appear to have been largely technology driven. There has been no requirement for empirical evidence that creating a specialty will do more good than harm.

This Viewpoint addresses the issues involved in answering the question: How much specialization in medicine is appropriate?

appropriate? Because the United States is about to implement the Affordable Care Act mandating expansion of physician services at a time when the financial credibility of federal and state governments is at stake, the economic effects of specialization are especially important now.

**What Are the Disadvantages of Specialization?**

Specialty certification tends to create professional monopolies—barriers to other physicians who could perhaps perform the same service. Specialization leads to fragmentation of care and discontinuity, even for patients with a single disease. Any improvements in care and reductions in cost resulting from having more highly trained specialists deliver specific services can be offset by the quality-eroding and cost-increasing effects of the multiple communications required when numerous independent specialists treat the same patient. This coordination problem is particularly relevant for older patients, who often have several chronic diseases. A specialist treating a single chronic disease is likely to focus on that disease, perhaps resulting in gaps in other important health care services and potentially leading to adverse outcomes.4

Another disadvantage is the length and cost of training for medical specialists resulting from the requirement that they begin with generalist training. In medicine, it is expected that the specialist know more than the generalist; in most fields, the reverse is true. For example, a specialist in 17th-century French art need not know as much as an art critic with a broader range.

**Why Do Countries Differ?**

The Royal College of Physicians and Surgeons of Canada provides certification for 67 separate specialties or subspecialties. Ten apply to both adult and pediatric medicine, resulting in 77 distinct certifications. The College of Family Physicians grants 2 certifications, for a total of 79 Canadian specialties and subspecialties. In comparison, the American Board of Medical Specialties oversees 24 separate specialty boards; certification is provided in 158 specialties or subspecialties. In some cases, the same subspecialty certification can be obtained through multiple boards (e.g., a subspecialty certification in critical care can be obtained through internal medicine, anesthesia, emergency medicine, obstetrics and gynecology, pediatrics, and surgery). Even if the subspecialties with multiple routes to certification are included only once, there are still 119 distinct specialties or subspecialties in the United States. That is 40 more than in Canada. This difference is relatively simple to calculate; however, comparing the difference in specialized practice between the 2 countries is much more difficult.

Why does the United States have so many more specialties and subspecialties than other countries? One reason may be related to patient volumes. With a population of 35 million people in Canada compared with 313 million in the United States, Canada does not have enough patients to warrant the creation of a subspecialty to treat a disease with limited incidence. The decentralized aspect of separate specialty boards in the United States, as well as local traditions, likely plays a role in differences between these countries. Also, specialization without certification in Canada may reduce the actual differences in specialized care.

**What Does the Evidence Say About Outcomes?**

Several well-known specific examples compare care delivered by specialists with that delivered by nonspecialists. The results are mixed. For example, Pronovost et al5 showed that clinical outcomes in critically ill patients were improved by having intensive care units staffed with mandatory intensivist consultations or care exclusively delivered by an intensivist compared with units in which there was either no intensivist or elective intensivist consultation. The relative reduction in mortality was between 30% and 40%.6 Conversely, Levy et al6 showed in a large sample of US intensive care unit patients that the odds of hospital mortality were higher for patients managed by critical care physicians than for those who were not.

Canadian-US comparisons for patients after myocardial infarction have been well studied, with inconsistent results. GUSTO substudies have shown that better quality-of-life, survival, and functional outcomes for US patients correlated with more specialist care and invasive procedures.7 However, Tu et al8 found no between-country differences in mortality.

**When Should Government Be Involved?**

For the most part, the issue of specialization and certification

---

in health care has been left to self-governing professional bodies. What are the issues that require government involvement?

Patient safety is one. If unsafe practices are occurring because of lack of adequate training and evaluation, government intervention is needed. At the same time, governments in many countries try to ensure adequate access to care. If professional societies create monopolies to increase prices and restrict access, governments could intervene with antitrust actions or other measures, such as providing alternative sources of care. For example, the certification of midwives in some jurisdictions expands the availability of services for delivery of infants in areas where obstetricians are in short supply.

Where government is the payer, the fee differentials it establishes for different kinds of physicians have a significant influence on the extent of specialization. Additionally, in Canada, the provincial governments determine the number of residency positions in each specialty, an instrument of influence that is available but unused in the United States.

**Conclusion**

There has been a long and spirited debate in the United States about the distribution of physicians across specialties. Many experts believe that the mix of physician services in the United States is weighted too heavily toward specialty care, less so in Canada. Escalating costs without clear benefits—at a time when government deficits are reaching crisis levels—suggests the need for careful, objective answers to the questions raised herein. These answers will help determine the appropriate amount of specialization in medicine in the future. Physicians will play a major role in this determination, but they are not the only group in society who have a stake in the outcome.

---

**Author Affiliations:** Institute of Health Policy Management and Evaluation (Dr Detsky) and Department of Medicine (Dr Detsky and Mr Gauthier), University of Toronto, Toronto, Ontario, Canada; Department of Medicine, Mount Sinai Hospital and University Health Network, Toronto (Dr Detsky); and Department of Economics, Stanford University, Stanford, California (Dr Fuchs).

**Corresponding Author:** Allan S. Detsky, MD, PhD, 600 University Ave, Room 429, Toronto, ON M5G 1X5, Canada (adetsky@mtsinai.on.ca).

**Conflict of Interest Disclosure:** The authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflict of Interest. Dr Fuchs reported receiving a grant from the Robert Wood Johnson Foundation. No other authors reported disclosures.

**Additional Contributions:** We thank Jonathan S. Skinner, PhD (Dartmouth College), and Mark R. Cullen, MD (Stanford University), for their comments on an earlier draft of the manuscript. Neither received compensation for his contribution.
The New Demographic Transition: Most Gains in Life Expectancy Now Realized Late in Life†

By Karen N. Eggleston, Ph.D. and Victor R. Fuchs, Ph.D.

The original “demographic transition” describes a process that began in Europe by the early 1800s with decreases in mortality followed, usually after a lag, by decreases in fertility (Davis 1945; for an overview in this journal, see Lee 2003). According to Lee and Reher (2011, p. 1), “this historical process ranks as one of the most important changes affecting human society in the past half millennium.” The increase in life expectancy associated with this demographic transition has been accompanied by rising levels of per capita output, which have in turn spurred further improvements in population health through better nutrition and living standards (Fogel 1994; Barker 1990) and, especially since World War II, through advances in medical care (in this journal, Cutler, Deaton, and Lleras-Muney 2006). At the same time, increases in life expectancy have resulted in a higher proportion of each cohort living long enough to participate in the production of goods and services. Reductions in fertility are also closely linked to higher labor force participation rates among women (Galor and Weil 1996; Costa 2000; Guinnane 2011).

During the original demographic transition, mortality decline prior to fertility decline often led to larger cohorts concentrated in working ages; this transitional change in the age structure of the population provided a boost to income that has been called a “demographic dividend” (Bloom, Canning, and Sevilla 2003). Swift (2011) documents a significant two-way positive relationship between life expectancy and GDP per capita between 1820 and 2001 for 13 high-income countries.

Now, the United States and many other countries are experiencing a new kind of demographic transition. Instead of additional years of life being realized early in the lifecycle, they are now being realized late in life. At the beginning of the twentieth century, in the United States and other countries at comparable stages of development, most of the additional years of life were realized in youth and working ages; and less than 20 percent was realized after age 65. Now, more than 75 percent of the gains in life expectancy are realized after 65—and that share is approaching 100 percent asymptotically. The choice of age 65 to illustrate this new demographic transition is somewhat arbitrary, but if we used 60 or 70 instead, the results would be qualitatively similar.

The new demographic transition is a longevity transition: How will individuals and societies respond to mortality decline when almost all of the decline will occur late in life? This issue is broader and more far-reaching than the issue of cohort size in each age group, with its usual focus on the prospective retirement of the unusually large “baby boomer” cohort, and has important socioeconomic implications independent of patterns of fertility.

When the gains in life expectancy occur mainly towards the end of life, they contribute more to the age bracket that is traditionally mostly retired rather than to the age bracket in prime working years. Retirees are highly dependent on transfers from the working population for living expenses, including large consumption of medical care. Thus, gains in life expectancy concentrated at the end of life can unsettle an economy’s balance between production and consumption in ways that pose a long-run challenge for public policy. The obvious changes needed (at least “obvious” to many economists) would be to raise productivity, the savings rate, and the age of retirement, but how to accomplish such goals is controversial and uncertain.

This paper covers the years 1900–2007 for the United States and 16 other “developed countries,” chosen for the continuity of their mortality data: Australia, Belgium, Canada, Denmark, England and Wales, Finland, France, Iceland, Italy, Spain, Sweden, Switzerland, and the United Kingdom.
Netherlands, Northern Ireland, Norway, Scotland, Spain, Sweden, and Switzerland. We focus on demographic statistics including life expectancy at birth and at age 65, the percent of each birth cohort expected to survive to age 65, and the share of the increase in life expectancy at birth realized after age 65. For the U.S. economy, we also calculate expected labor force participation for each birth cohort, which allows us to investigate how changes in mortality affect labor force participation and worklife as a share of life expectancy. Results on the longevity transition and expected labor force participation for the United States and other high-income countries are followed by consideration of economic and social changes in China and other countries that are experiencing an earlier stage of the original demographic transition. The paper concludes with a brief discussion of the long-run implications of the new demographic transition.

The Longevity Transition

To examine long-term trends in life expectancy at birth, we draw upon the life tables in the Human Mortality Database, which offers high-quality demographic data for selected countries and regions compiled by a respected group of demographers at [www.mortality.org](http://www.mortality.org). We first extract data on life expectancy at birth; in particular, we calculate “period” life expectancy, which is the projected average age of death for a cohort if it experienced the age-specific death rates prevailing at the year of birth. We also look at rates of survival from birth to age 65 and life expectancy at age 65. We use the five-year period life tables since 1900 (or earliest available year) for each of the 17 countries or regions in the Human Mortality Database that have data extending back at least 70 years. The five-year intervals help to smooth annual fluctuations in demographic trends.

We calculate changes for nine overlapping 20-year intervals: 1907–1927, 1917–1937, and so on up to 1987–2007. To calculate the change in years lived past 65, we first multiply survival to 65 by life expectancy at age 65 for each five-year period and then take differences across 20-year intervals. Finally, we calculate the change in years lived past 65 as a percentage of change in life expectancy at birth for each country for each of the nine 20-year intervals.

Figure 1A shows that life expectancy at birth has increased almost continuously for well over a century in high-income countries. Much of this rise in life expectancy was due to a particularly large fall in death rates for infants, children, and young adults, resulting in a sharp rise in the percentage of a cohort surviving to age 65, as indicated in Figure 1B. Survival rates from birth to age 65 more than doubled over the twentieth century from 40.9 percent in 1900–04 to 83.3 percent in 2005–09 in the United States. Similarly, survival rates from birth to age 65 in 16 high-income comparators increased from 42.0 to 87.8 percent over the same period.

Figure 1.
Life Expectancy at Birth and Survival to Age 65, since 1990
(in the United States and 16 other high-income countries)

A: Life Expectancy at Birth

<table>
<thead>
<tr>
<th>Year</th>
<th>United States</th>
<th>16 high-income country average</th>
<th>One standard deviation above and below the 16-country average</th>
</tr>
</thead>
<tbody>
<tr>
<td>1900</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1910</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1920</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1930</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1940</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1950</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1960</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1970</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1980</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Authors’ calculations using data from the Human Mortality Database and other sources as detailed in the online Appendix.
The other major demographic change that contributes to the longevity transition is an increase in life expectancy at age 65, an increase which has become larger in recent decades as shown in Figure 2A. The interaction between the increase in life expectancy at age 65 and the increase in the percentage of the cohort that survives to age 65 has resulted in an exceptionally large increase in the share of the gain in life expectancy that is realized after age 65. As can be seen in Figure 2B, that share was only about 20 percent during each 20-year period at the beginning of the twentieth century, but it was 76 percent in the United States and 78 percent for the 16-country mean by the end of the century, and is approaching 100 percent asymptotically. Our results here are quite similar to, and extend over time, those of Lee and Tuljapurkar (1997) based on the 1995 survival profile of the United States.

Figure 2.
Life Expectancy at Age 65 and Gains in Life Expectancy Realized after Age 65 Since 1990 (in the United States and 16 other high-income countries)

A: Life Expectancy at Age 65

<table>
<thead>
<tr>
<th>Year</th>
<th>United States</th>
<th>16 high-income country average</th>
</tr>
</thead>
<tbody>
<tr>
<td>1900</td>
<td>10.0</td>
<td>12.0</td>
</tr>
<tr>
<td>1920</td>
<td>12.0</td>
<td>14.0</td>
</tr>
<tr>
<td>1940</td>
<td>14.0</td>
<td>16.0</td>
</tr>
<tr>
<td>1960</td>
<td>16.0</td>
<td>18.0</td>
</tr>
<tr>
<td>1980</td>
<td>18.0</td>
<td>20.0</td>
</tr>
<tr>
<td>2000</td>
<td>20.0</td>
<td>22.0</td>
</tr>
</tbody>
</table>

Source: Authors’ calculations using data from the Human Mortality Database and other sources as detailed in the online Appendix.

We can illustrate the shift in survival improvement toward older ages by comparing the age distribution of mortality decline between the first half and second half of the twentieth century for a region with particularly reliable long-run data: England and Wales. Figure 3 shows that between 1900–1904 and 1950–1954, declines in death rates were largest for infants and children, whereas between 1950–54 and 2000–2004, declines were most salient for those over age 70. This pattern of age-specific mortality decline across the twentieth century was similar for Sweden, another country where reliable long-run data is available.

Figure 3.

The actual survival of a given birth cohort will differ from the estimates of life expectancy at birth when survival is changing over time. Remember, estimates of life expectancy at birth (what we earlier called “period” life expectancy) are based on the age-specific death rates prevailing at that year of birth. For example, in 1900–04, life expectancy at birth in England and Wales was 48.6 years. In contrast, the cohort born in 1900–1904 had a cohort life expectancy (actual mean age of death) of 53.8 years, since they experienced part of the increase in survival shown in Figures 1–3. The cohort born only 17 years later experienced a cohort life expectancy of 62.4 years, whereas “period” life expectancy at birth did not reach that level until 1935–1939.

2 For details on Sweden, see the online Appendix. Figure 3 shows a slight increase in death rates for the oldest [90+] age groups between 1900–1904 and 1950–1954, perhaps because of small numbers, less-reliable data, and/or survival of a less-healthy cohort to those ages.

3 Survival gains have been so dramatic that period and cohort survival significantly differs. For example, age-specific death rates for England and Wales in 1900–1904 would have led to only 43.7 percent of women and 36.4 percent of men surviving to 65. But of the cohort born in 1900–1904, 61.3 percent of women and 49.6 percent of men actually survived to age 65.
Nevertheless, we find that estimates based on cohort life tables prepared by the Social Security Administration (Bell and Miller 2005) exhibit a similar trend towards survival gains realized late in life: for men, the share of life expectancy increases realized after age 65 was 28 percent between the 1900 and 1920 birth cohorts, rising to a projected 62 percent between the 1980 and 2000 birth cohorts. For women, the share of life expectancy gains realized after age 65 increased from 30 percent (between the 1900 and 1920 birth cohorts) to an estimated 69 percent (between the 1980 and 2000 birth cohorts).

The century-long demographic trends shown in Figures 1 and 2 have been similar in all 17 countries with available data. From a U.S. perspective, the main difference is lagging survival to 65 compared to the other 16 countries (the U.S. line is below the 16-country average in Figure 1B); also, the United States experienced a larger rise in female life expectancy at age 65 between the 1940s and 1970s than the other countries. The relative differences among countries have decreased over time, especially for life expectancy at birth and survival to age 65.

The Longevity Transition and Expected Labor Force Participation

One of the most significant economic effects of the longevity transition is on expected lifetime labor force participation, partly in terms of total years in the workforce and especially in terms of years in the workforce as a fraction of expected years of life. Two factors affecting the connection from life expectancy to years of work are 1) whether the growing numbers of elderly are healthy enough to work and 2) the economic, social, and political pressures for a period of retirement at the end of life.

Greater longevity can have opposing effects on age-specific health status. If improved survival is correlated with reductions in morbidity for the elderly, then illness may be compressed into the end of life, as posited by the “compression of morbidity” hypothesis (Fries 1980). On the other side, medical interventions do tend to keep alive those who are in worse health (Zeckhauser, Sato, and Rizzo 1985), which suggests the possibility that the longer-lived elderly could be sicker for a longer period. The net effect of rising longevity on age-specific morbidity is an empirical question. According to the National Long-Term Care Survey, the share of elderly Americans with severe disabilities decreased from 26.2 to 19.7 percent between 1982 and 1999 (Manton and Gu 2001). Milligan and Wise (2011) find a strong within-country correlation between declining mortality and improved self-assessed health for several European countries. Thus, the empirical record suggests that better health in terms of both improved survival and reduced morbidity could tend to raise age-specific rates of labor force participation. Changes in occupational structure which lower the physical demands of work also can increase participation.

Higher incomes tend to increase the demand for leisure, in the form of fewer hours of work per week and, especially recently, as a block at the end of life (Costa 1998; Murphy and Topel 2006). Furthermore, several factors might give rise to a negative interaction between improved survival and employment, at least for some subgroups. For example, the reduced selection effect of mortality might also increase the proportion of the cohort that is less valued in employment (because of less stamina, ambition, education, and the like), reducing age-specific labor force participation. Alternatively, if firms have pyramid-like organizational structures with many jobs at entry and fewer at higher levels in the hierarchy—such as the military’s “up or out” policy regarding age and promotion of officers—then increases in survival will lead to crowding at higher levels of the pyramid and lower rates of participation. Moreover, a sharp rise in employment rates for women, at wages that were often below those paid to men, might have led to some decrease in the demand for men’s labor.

On net, which of these forces have predominated over the past century, and which are likely to predominate in the future? Estimates of what we call “expected labor force participation” can help answer this question.

Calculating Expected Labor Force Participation

We define “expected labor force participation” (XLFP) as the total years an individual is expected to participate in the labor force, based on period estimates of survival, and labor force participation by gender and age. That is

\[ XLFP_{ijt} = \sum_{i=1}^{100} \pi_{ijt} L_{ijt}, \]

where \( L_{ijt} \) is the labor force participation rate for age \( i \) and gender \( j \) in year \( t \), weighted by probability of survival to age \( i \) (\( \pi_{ijt} \)). It is necessary to examine men and women separately because of the large upsurge in female labor force participation between the 1950s and 2000 (Goldin 1986, 1990; Costa 2000). Our calculations rely on labor force participation rates from decennial censuses (1900–1930) and the Current Population Survey (1942–2007). As in the earlier estimates of life expectancy, we can calculate both “period” expected labor force participation, which is based on the age-specific labor force participation rates prevailing at a certain point of time, or the actual realized labor force participation rates for a birth
COHORT; THESE ESTIMATES WILL DIFFER WHEN AGE-SPECIFIC LABOR FORCE PARTICIPATION RATES ARE CHANGING OVER TIME.

CHANGES IN LIFETIME EXPECTED LABOR FORCE PARTICIPATION CAN BE DECOMPOSED INTO TWO FACTORS: CHANGES IN SURVIVAL TO GIVEN AGES AND CHANGES IN AGE/SEX-SPECIFIC RATES OF LABOR FORCE PARTICIPATION. FOR EXAMPLE, WE CALCULATE THE EFFECT OF IMPROVING SURVIVAL, HOLDING AGE-SPECIFIC LABOR FORCE PARTICIPATION RATES CONSTANT AT THEIR 2007 VALUES. WE ALSO CALCULATE THE EFFECT OF CHANGING RATES OF LABOR FORCE PARTICIPATION, HOLDING SURVIVAL RATES CONSTANT. 4

OUR WORK IS RELATED TO THE LITERATURE ON EXPECTED LIFETIME WORK HOURS (HAZAN 2009) AND WORK-LIFE EXPECTANCY (SMITH 1982), INCLUDING THE WORK-LIFE ESTIMATES FOR THE U.S. POPULATION FROM THE 1950s THROUGH THE EARLY 1980s FROM THE BUREAU OF LABOR STATISTICS. 5 AS FAR AS WE ARE AWARE, THIS PAPER IS THE FIRST TO PRODUCE WORK-LIFE ESTIMATES FOR THE UNITED STATES COVERING THE PERIOD 1900 TO 2007, DECOMPOSE THOSE CHANGES INTO SURVIVAL AND AGE/SEX-SPECIFIC LABOR FORCE PARTICIPATION EFFECTS, AND TO ESTIMATE WORK-LIFE EXPECTANCY RELATIVE TO LIFE EXPECTANCY AT BIRTH FOR A BROAD RANGE OF COUNTRIES IN RECENT DECADES.

US. EXPECTED LABOR FORCE PARTICIPATION SINCE 1900

IN THE EARLY TWENTIETH CENTURY, MOST OF THE INCREASE IN LIFE EXPECTANCY AROSE FROM THE DRAMATIC DECREASE IN MORTALITY AT YOUNG AGES. THIS CHANGE FIRST INCREASED THE YEARS OF YOUTH DEPENDENCY FOR THESE COHORTS, AND THEN INCREASED EXPECTED LABOR FORCE PARTICIPATION—THE EXPECTED NUMBER OF YEARS THAT AN INDIVIDUAL WILL BE IN THE LABOR FORCE IF HE OR SHE PARTICIPATES AT THE AVERAGE LABOR FORCE PARTICIPATION RATE FOR EACH SEX AND AGE IN A GIVEN YEAR.

FIGURE 4A SHOWS THAT YEARS OF EXPECTED LABOR FORCE PARTICIPATION AT BIRTH FOR U.S. MALES INCREASED BY A THIRD—FROM ABOUT 30 TO 40 YEARS—BETWEEN 1900 AND 1950. FOR THE MOST RECENT HALF CENTURY, HOWEVER, INCREASES IN SURVIVAL HAVE BEEN OFFSET BY DECREASING AGE-SPECIFIC LABOR FORCE PARTICIPATION RATES FOR MEN, CAUSING EXPECTED LIFETIME LABOR FORCE PARTICIPATION TO BE RELATIVELY CONSTANT AT ABOUT 40 YEARS. BECAUSE LIFE EXPECTANCY AT BIRTH HAS CONTINUED TO INCREASE, MALE EXPECTED LABOR FORCE PARTICIPATION AS A FRACTION OF EXPECTED YEARS OF LIFE HAS DECREASED, AS SHOWN IN FIGURE 4B. TABLE 1 SHOWS THAT

IN THE UNITED STATES BETWEEN 1900 AND 2000, MALE LABOR FORCE PARTICIPATION INCREASED FROM 30 TO 40.5 YEARS, FEMALE PARTICIPATION FROM 6.4 YEARS TO 34.4 YEARS, AND FOR THE TOTAL POPULATION FROM 18.5 TO 37.4 YEARS. THIS INCREASE IN YEARS OF EXPECTED LABOR PARTICIPATION IS TWO-THIRDS OF THE TOTAL GAIN IN LIFE EXPECTANCY AT BIRTH OF 28.2 YEARS OVER THE TWENTIETH CENTURY.

FIGURE 4. U.S. EXPECTED LABOR FORCE PARTICIPATION SINCE 1990 AND AS A SHARE OF LIFE EXPECTANCY AT BIRTH

A: EXPECTED LABOR FORCE PARTICIPATION (XLFP)

B: EXPECTED LABOR FORCE PARTICIPATION (XLFP) AS A SHARE OF LIFE EXPECTANCY (LE)

SOURCE: AUTHORS’ CALCULATIONS USING DATA FROM THE HUMAN MORTALITY DATABASE AND OTHER SOURCES AS DETAILED IN THE ONLINE APPENDIX.

HOW MUCH OF THIS CHANGE IS ATTRIBUTABLE JUST TO LONGER LIFE EXPECTANCIES? IF WE HOLD AGE-SPECIFIC RATES OF LABOR FORCE PARTICIPATION CONSTANT BUT ALLOW SURVIVAL RATES TO GROW AT THE ACTUALLY OBSERVED PACE, THE RISE IN LIFE EXPECTANCY ALONE WOULD HAVE INCREASED EXPECTED LABOR FORCE PARTICIPATION BY 13.3 YEARS FOR MALES AND BY 10.8 YEARS FOR FEMALES SINCE 1900 (AS SHOWN IN TABLE 1). THE EFFECT OF MORTALITY DECLINE WAS CONCENTRATED IN THE FIRST HALF OF THE TWENTIETH CENTURY. INDEED, FOR MEN, IF WE

4 THESE ARE DECOMPOSITIONS 1B AND 2B, RESPECTIVELY, IN TABLE 7 OF THE ONLINE APPENDIX. ALTERNATIVE CALCULATIONS, USING 1900 AS THE BASE YEAR (DECOMPOSITIONS 1A AND 2A), SHOW SIMILAR RESULTS.

Table 1. Expected Labor Force Participation in the United States, by Sex, 1900–2007

<table>
<thead>
<tr>
<th>Year</th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male XLFP</td>
<td>Male XLFP adjusted for hours worked</td>
</tr>
<tr>
<td></td>
<td>holding LFP constant</td>
<td></td>
</tr>
<tr>
<td>1900</td>
<td>30.0</td>
<td>25.7</td>
</tr>
<tr>
<td>1910</td>
<td>31.3</td>
<td>27.1</td>
</tr>
<tr>
<td>1920</td>
<td>35.1</td>
<td>30.4</td>
</tr>
<tr>
<td>1933</td>
<td>36.7</td>
<td>32.3</td>
</tr>
<tr>
<td>1942</td>
<td>39.5</td>
<td>34.1</td>
</tr>
<tr>
<td>1950</td>
<td>41.3</td>
<td>35.6</td>
</tr>
<tr>
<td>1960</td>
<td>41.0</td>
<td>36.3</td>
</tr>
<tr>
<td>1970</td>
<td>39.9</td>
<td>36.4</td>
</tr>
<tr>
<td>1980</td>
<td>39.6</td>
<td>37.4</td>
</tr>
<tr>
<td>1990</td>
<td>39.1</td>
<td>37.9</td>
</tr>
<tr>
<td>2000</td>
<td>40.5</td>
<td>38.7</td>
</tr>
<tr>
<td>2007</td>
<td>39.0</td>
<td>39.0</td>
</tr>
<tr>
<td>Change, 1900 to most recent</td>
<td>9.0</td>
<td>13.3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>Total (men and women)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total XLFP</td>
</tr>
<tr>
<td>1900</td>
<td>18.5</td>
</tr>
<tr>
<td>1910</td>
<td>19.8</td>
</tr>
<tr>
<td>1920</td>
<td>22.1</td>
</tr>
<tr>
<td>1933</td>
<td>23.7</td>
</tr>
<tr>
<td>1942</td>
<td>27.4</td>
</tr>
<tr>
<td>1950</td>
<td>29.1</td>
</tr>
<tr>
<td>1960</td>
<td>30.2</td>
</tr>
<tr>
<td>1970</td>
<td>31.3</td>
</tr>
<tr>
<td>1980</td>
<td>33.8</td>
</tr>
<tr>
<td>1990</td>
<td>35.4</td>
</tr>
<tr>
<td>2000</td>
<td>37.4</td>
</tr>
<tr>
<td>2007</td>
<td>36.3</td>
</tr>
<tr>
<td>Change, 1900 to most recent</td>
<td>17.7</td>
</tr>
</tbody>
</table>

Sources: Author calculations based on survival data from the Human Mortality Database (1933–2007), supplemented by data for death registration states for 1900–1920; and labor force participation rates from decennial censuses (1900–1930) and the Current Population Survey (1942–2007). Adjustments for hours worked draw from Hazan (2009). See the online Appendix for details.

Notes: Expected Labor Force Participation (XLFP) is calculated as the total years an individual is expected to participate in the labor force based on period estimates of labor force participation and survival by gender and age. XLFP for a given year represents the expected number of years that an individual would be in the labor force if he or she participates at the average LFP rate for each age in that given year. LE<sub>0</sub> is life expectancy at birth. “XLFP holding LFP constant” uses 2007 age- and sex-specific labor force participation rates, but allows survival to each age to vary as it actually did between 1900 and 2007.
hold age-specific labor force participation rates constant but allow survival rates to vary in calculating expected labor force participation ("male XLFP holding LFP constant"), the ratio of years of expected labor force participation to life expectancy at birth was relatively constant at 54 percent from early in the twentieth century until about 1970 (not shown in the table). At that point, it began a slow but seemingly inexorable decline, now falling to about 50 percent.

Actual years of expected labor force participation, reflecting both survival effects and changes in age-specific labor force participation rates, have also begun to decline. As shown in both Table 1 and Figure 4B, the ratio of years of expected labor force participation to life expectancy at birth (XLFP/LE) has declined for U.S. men from 62.6 percent in 1900 to 51.6 percent in 2007. That same ratio for women increased from 12.7 percent in 1900 to 43.2 percent in 2000, before declining slightly to 41.5 percent by 2007. For the overall U.S. population, years of expected labor force participation divided by life expectancy at birth peaked at 48.6 percent in 2000 and declined slightly to 46.3 percent by 2007.

Since 1950, increases in survival and declines in age-specific participation rates of men tended to offset one another. For example, between 1950 and 2007, labor force participation rates of men ages 45–54 declined from 95.8 percent to 88.2 percent, but survival to age 50 increased from 84.1 to 92.2 percent, so the total expected years in the labor force between ages 45 and 55 remained eight years.6 For women, increases in years of expected labor force participation mostly reflect increases in age-specific rates of labor force participation, especially after 1950. Accordingly, for women, if we hold age-specific labor force participation rates constant but allow survival rates to vary in calculating expected labor force participation ("female XLFP holding LFP constant"), the ratio of years of expected labor force participation to life expectancy at birth has declined slowly but steadily from about 45 percent in the first few decades of the twentieth century to about 40 percent (not shown in the table).

The increase in female labor force participation since the late 1950s could be considered primarily a one-time substitution from unpaid home production to paid work outside the home (Goldin 1990; Costa 2000). If so, then the decrease in years of expected labor force participation for women in the United States since 2000 would reflect completion of the one-off change and the beginning of a similar trend as seen for men—that is, a decline of years in the labor force as a share of life expectancy at birth.

Taking into account the decrease in the intensive margin—annual hours worked per full-time worker—trends to reinforce the conclusion that expected work life has declined as a fraction of life expectancy at birth. Hazan (2009) estimated lifetime work hours over the past century conditional on survival to age five. We adapt Hazan’s data to life expectancy at birth to calculate years of expected labor force participation adjusted for hours worked and show the results in Table 1 (the online Appendix available with this paper at (http://e-jep.org) has details of our calculations).

Calculation of a century-long trend in expected years of labor force participation in other high-income countries is not possible because there is no reliable source for internationally comparable labor force participation rates before 1980. Given the similarities in trends of both survival and labor force participation across these countries for the available years, we suspect the trend of declining expected labor force participation as a share of life expectancy at birth that we found for the United States reflects a broad and robust trend that countries experience as they reach high life expectancy levels. Indeed, with the sole exception of the Netherlands, the ratio of years of expected labor force participation to life expectancy at birth has declined since 1980 for males in all other high-income countries in our analyses.7 Adjusting for a decline in work hours would reinforce this trend.

Demographic Transition across Stages of Economic Development

The demographic transition traces out a pathway, with many societies arrayed along earlier phases of the transition roughly and imperfectly in accordance with their per capita incomes. Many developing countries are currently experiencing the original demographic transition. For example, Table 2 shows that between 1990 and 2010, the share of years lived past 65 as a percentage of increase in life expectancy at birth was only a little over one-third in Vietnam and Brazil, and less than one-quarter in Bangladesh—comparable to levels a century earlier in today’s high-income countries.

6 For the detailed data behind these calculations across the range of ages, for both men and women, see online Appendix Table 7, which offers alternative decompositions of changes in both male and female labor force participation. Online Appendix Table 7 also shows that holding age-specific labor force participation rates constant (at either their 1900 or 2007 values) would have led to a larger increase in male expected labor force participation than actually observed.

7 The online Appendix tables provide calculations of expected labor force participation across 15 countries since 1980; see Appendix Table 8 in the online Appendix available with this paper at (http://e-jep.org). Milligen and Wise (2011, p. 17) examine the age at which male mortality was 1.5 percent in 1977 and 2007, finding that at that age almost 90 percent of UK men were employed in 1977, but by 2007, only 30 percent were.
Improving health and increasing life expectancy at birth clearly can contribute to better living standards for the world’s poor (World Health Organization 2002). Data on labor force participation for developing countries is not always reliably comparable across countries and over time. Nevertheless, the importance of improved survival for gains in expected labor force participation at early stages of the longevity transition can be illustrated with extant data. For example, in 1980 only 70 percent of Indonesian men survived to age 45; by 2007, 90 percent did. This improved survival added 10 years to expected labor force participation rates for Indonesian males between 1980 and 2007. As a result, expected labor force participation rates for Indonesian males rose to 43.7 years, which was 64.5 percent of life expectancy at birth in 2007.

Table 2. The Longevity Transition in Asia and Select Developing Countries

<table>
<thead>
<tr>
<th>Country</th>
<th>Males</th>
<th>Females</th>
</tr>
</thead>
<tbody>
<tr>
<td>Japan</td>
<td>72.7%</td>
<td>82.0%</td>
</tr>
<tr>
<td>South Korea</td>
<td>45.4%</td>
<td>57.1%</td>
</tr>
<tr>
<td>China</td>
<td>51.9%</td>
<td>40.6%</td>
</tr>
<tr>
<td>Philippines</td>
<td>26.2%</td>
<td>36.0%</td>
</tr>
<tr>
<td>Indonesia</td>
<td>26.1%</td>
<td>35.7%</td>
</tr>
<tr>
<td>Brazil</td>
<td>34.2%</td>
<td>35.0%</td>
</tr>
<tr>
<td>Vietnam</td>
<td>32.5%</td>
<td>34.7%</td>
</tr>
<tr>
<td>India</td>
<td>23.6%</td>
<td>25.8%</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>20.7%</td>
<td>25.4%</td>
</tr>
</tbody>
</table>

Source: Authors’ calculations based on the life tables for each country prepared by the International Programs Center of the U.S. Bureau of the Census in its International Data Base.

China and India are especially important cases to consider, given their large populations and relatively rapid economic development. In India, the share of years lived past 65 as a percentage of change in life expectancy at birth was only one-quarter (as shown in Table 2) in the most recent 20-year period. For China, that share was 52 percent for men and 41 percent for women in the 1990–2010 period.

China’s position reflects the rapidity of its demographic transition since the early 1970s and its achievement of relatively high levels of health despite low per capita income by the end of the Mao era (Banister 1987; Wang 2011). Indeed, despite the higher death rates associated with the Great Leap Famine of 1959–61, China’s growth in life expectancy from approximately 35–40 years in 1949 to 65.5 years in 1980 ranks as the most rapid sustained increase in documented global history. These earlier health improvements and growth of the working-age population contributed to China’s unprecedented economic growth for the past quarter-century. Wang and Mason (2008) estimate that between 1982 and 2000, about 15 percent of China’s rapid growth in output per capita stemmed from the demographic dividend. (Bloom and Williamson (1998) estimate that one-quarter to one-third of the growth rates in the “East Asian miracle” stemmed from the demographic dividend.) Although the pace of mortality decline in China has slowed, it continues: Chinese life expectancy increased between 1990 and 2010 from 69.9 to 76.8 years for women and from 66.9 to 72.5 years for men.

With a rapid demographic transition to relatively low mortality and low fertility, China’s population is now aging (Peng 2011). Many policy challenges loom as China establishes social and economic institutions commensurate with its transition to a middle-income, market-based economy with a large elderly population (Eggleston and Tuljapurkar 2010; Chen, Eggleston, and Li forthcoming). One additional challenge for China in reducing the growth-slowing potential of the new demographic transition is China’s increasing burden of chronic disease. Fueled by rapid urbanization, increases in high-fat and calorie-rich diets, reductions in physical activity, unabated male smoking and other factors, prevalence of chronic disease in China has quickly caught up with that of high-income countries. For example, the age-standardized prevalence of diabetes among adults in China was 9.7 percent in 2007–2008, more than three times reported prevalence in 1994 (Yang et al. 2010), comparable to the U.S. rate of 8.3 percent overall in 2010 and 11.3 percent among adults (CDC 2011), and higher than the OECD average (OECD 2011).

The timing and the rapidity of the longevity transition has varied across countries and regions. For example, in Japan between 1950 and 1970, only 13.1 percent of increase in male life expectancy at birth was realized after age 65; for women, that figure was 17.3 percent. During the 1990 to 2009 period, Japan led the world in the new demographic transition, with the share of gains in life expectancy at birth realized after age 65 reaching 72.7 percent for men and 87 percent for women (again, as shown in Table 2).

The original and the new demographic transitions are inextricably intertwined with the evolution of social and

---

8 Miller, Eggleston, and Zhang (2012) assess the relative importance of various explanations proposed for these gains, including better nutrition, widespread public health interventions, improved access to medical care, and increases in educational levels. They find that gains in education and public health campaigns jointly explain 25–32 percent of the crude death rate decline under Mao, and similar proportions of the dramatic reductions in infant and under-five mortality in that period.
economic institutions (Aoki 2011). Evidence is mounting that no society at an advanced stage of economic development can presume that further gains in longevity will contribute to growth of per capita income under currently prevailing institutions. For example, Lee and Mason (2011) compare the “average age of consumption” to the “average age of labor income” across a large group of countries for which they and their international collaborators have collected detailed generational accounts, including the value of assets and transfer wealth from social support programs (but not including bequests or value of nonmarket labor). They find that for developing countries, net transfers flow strongly downward from older to younger ages. However, in a “sea change” analogous to what we call the new demographic transition, “the direction of intergenerational transfers in the population has shifted from downward to upward, at least in a few leading rich nations” including Germany, Austria, and Japan (Lee and Mason 2011, p. 116). Although the Lee–Mason estimates are cross-sectional, the link to the longevity transition is clear: for the 13 countries that overlap between their dataset and ours, there is a strong negative correlation (–0.89) between the share of gains in life expectancy over the past 20 years that were realized after age 65 and the current number of years by which the average age of income exceeds the average age of consumption. In other words, the more the gains in life expectancy are concentrated in traditional retirement years, the closer the intergenerational transfers are to being upward rather than downward.

For a broader group of 107 countries, Bloom, Canning, and Fink (2010) calculate counterfactual annual growth rates of per capita income between 1960 and 2005, using 2005–2050 projections of demographics. The results vary depending on the level of economic development. They find that in most non-OECD countries, declining youth dependency would more than offset increasing old-age dependency. However, about half of countries would have grown more slowly using 2005–2050 projections of demographics. Among 26 OECD countries analyzed, 25 of them (Turkey is the exception) would have had lower economic growth—averaging 2.1 percent per year—under the counterfactual of 2005–2050 demographic change.

Policy Implications of the New Demographic Transition

Historically, adults produced more than they consumed and supported children. With such a pattern in place, the increase in proportion of the population in older years implied by the demographic transition might have been thought to shift out the social budget constraint as people expanded their number of years worked. However, “a funny thing happened along the way: societies invented retirement … and the economic consequences of population aging are now viewed with alarm” (Lee and Mason 2011, p. 115).

Retirement, a relatively new phenomenon in human history, can be viewed as a response to many economic and social changes. Contributing factors include the shift from self-employment on farms or small businesses to wage and salary status; more rapid technological change, resulting in more rapid obsolescence of human capital (alongside compensation packages that often underpay at the beginning and overpay at the end of a career, as discussed in Lazear 1981); the introduction of a variety of health and welfare programs which assist the elderly but also discourage work; an income-driven increase in the demand for leisure, with the diminishing marginal value of an even shorter work week overtaken by the efficiency gains of a block of leisure at the end of life; and, in times of high unemployment, public concern about job opportunities for younger workers.

Will the new demographic transition inevitably lead to slower economic growth? As people foresee longer lives, they might choose to work longer, save more, and/or invest in human capital in sufficient amounts and innovative enough ways that longer lives continue to contribute to increased prosperity. In this spirit, Bloom, Canning, and Fink (2010) assert that “the problem of population ageing is more a function of rigid and outmoded policies and institutions than a problem of demographic change per se” (p. 607).

It is not clear, however, that the United States or other high-income countries even further along in the new demographic transition are reshaping their policies and institutions sufficiently in response to the longevity transition. Although both the United States and France have increased the age of retirement or age to qualify for early retirement, social welfare systems across the high-income countries of the world continue to give strong incentives for earlier, rather than later, retirement (Gruber and Wise 1998). Between 1965 and 2005, the correlation between change in male life expectancy at birth and change in retirement age is actually negative: –0.21 (Bloom, Canning, and Fink 2010, p. 591). This trend cannot continue indefinitely: longer and longer retirement lives are

---

9 They construct the average ages of consumption and labor income as follows: “The average age of consumption is calculated by multiplying each age by the aggregate consumption at that age, summing these products over all ages, then dividing by the total amount of consumption at all ages. An equivalent calculation gives the average age of labor income” (Lee and Mason 2011, p. 123).
not consistent with continued increases in per capita income unless there are significant increases in savings, investment, and productivity. It is ironic that the same phenomenon that led to higher GDP per capita—namely higher life expectancy—could now lead to lower GDP per capita.

Successful navigation of the new demographic transition calls for a combination of policies to give incentives for more savings and investment (including in human capital) earlier in the lifecycle and for additional work later in the lifecycle. Two forces in particular might move the society in that direction: improvement in health, and reductions in the transfers that the elderly can expect to receive from the young.

Public policy should encourage higher labor force participation for the elderly, both by reducing the disadvantages that employers face when employing older workers and by providing enhanced incentives to individuals to continue to work. “People cannot expect to finance 20–25-year retirements with 35-year careers,” Shoven noted (as quoted in Haven 2011). “It just won’t work. Not in Greece [or] the United States. . . Eventually, we are going to have to increase retirement ages.” However, increasing labor force participation for the 65-plus age group alone probably won’t make a big difference: even a doubling of those rates from their 2007 levels of 12.6 for women and 20.5 for men would not bring the U.S. ratio of expected labor force participation to life expectancy at birth back to its 2000 level. Increased labor force participation by men in the 50–64 age bracket is also needed.

Public policy might also seek to improve productivity, with an emphasis on education and building human capital early in the lifecycle, and on investment to reduce morbidity and improve the ability to work later in life. Whether compression of morbidity later in life will continue depends on whether improvements in medical technology and in the socioeconomic determinants of health are offset by adverse trends such as increasing obesity. A potentially promising focus here would be to consider investments in public health and medical technologies that reduce morbidity and improve quality of life, as well as more focus on medical innovations that reduce costs of care. (One example of a policy consistent with both objectives would be expansion of palliative care as a substitute for what can otherwise be extremely expensive end-of-life care in a hospital—especially in countries where the concept of hospice services is relatively new, such as China.)

Finally, increased savings, investment, and capital formation could help in fueling endogenous growth (Lucas 1988; Romer 1990). U.S. personal savings rates have been low for many decades. Increasing the savings rate of individuals before they retire would ameliorate the potential adverse impact of longevity on economic growth. Countries will need to make fiscally realistic structural changes to entitlement programs—such as Medicare and Social Security in the United States—to support acceptable living standards and improvements in health.

High-income societies are now facing a new demographic transition: the longevity transition. They must decide how to respond to mortality decline when almost all of the decline will occur late in life. Additional increases in life expectancy will result in further declines in expected labor force participation as a percentage of life expectancy at birth unless there is a significant rise in labor force participation rates across both middle and older ages. Of course, increased life expectancy has great value independent of its relationship to per capita income (Murphy and Topel 2006). The original demographic transition gave society a “demographic gift” of higher per capita incomes (Bloom and Williamson 1998) without much need for a policy response, but the new demographic transition requires politically difficult policies if societies wish to preserve a positive relationship running from increased longevity to greater prosperity.

References


Eggleston, Karen N., and Shripad Tuljapurkar, eds. 2010. Aging Asia: Economic and Social Implications of Rapid Demographic Change in China, Japan, and South Korea. Shorenstein APRC; distributed by Brookings Institution Press.


Human Mortality Database. University of California, Berkeley; and Max Planck Institute for Demographic Research. Available at: www.mortality.org.


Milligan, Kevin, and David A. Wise. 2011 “Social Security and Retirement around the World: Historical Trends in Mortality and Health, Employment, and Disability Insurance


Karen N. Eggleston is Director of the Stanford Asia Health Policy Program and Center Fellow at the Shorenstein Asia-Pacific Research Center, and Victor R. Fuchs is Henry J. Kaiser, Jr., Professor Emeritus in the Departments of Economics and of Health Research and Policy, and Senior Fellow, Stanford Institute for Economic Policy Research, both at Stanford University, Stanford, California. Their email addresses are (karene@stanford.edu) and (vfuchs@stanford.edu).

The authors gratefully acknowledge constructive comments from the editors, as well as from Judith Banister, Richard Zeckhauser, and participants at a Stanford University seminar. We extend our appreciation and thanks to Loraine West, Daniel Goodkind, and Andrea Miles at the U.S. Bureau of the Census for graciously providing country life tables from the U.S. Bureau of the Census International Data Base. We also thank Shannon Davidson and Shaowen Ang for excellent research assistance.
Abstract

Life expectancy at birth, estimated from United States period life tables, has been shown to vary systematically and widely by region and race. We use the same tables to estimate the probability of survival from birth to age 70 (S₇₀), a measure of mortality more sensitive to disparities and more reliably calculated for small populations, to describe the variation and identify its sources in greater detail to assess the patterns of this variation. Examination of the unadjusted probability of S₇₀ for each US county with a sufficient population of whites and blacks reveals large geographic differences for each race-sex group. For example, white males born in the ten percent healthiest counties have a 77 percent probability of survival to age 70, but only a 61 percent chance if born in the ten percent least healthy counties. Similar geographical disparities face white women and blacks of each sex. Moreover, within each county, large differences in S₇₀ prevail between blacks and whites, on average 17 percentage points for men and 12 percentage points for women. In linear regressions for each race-sex group, nearly all of the geographic variation is accounted for by a common set of 22 socio-economic and environmental variables, selected for previously suspected impact on mortality; R² ranges from 0.86 for white males to 0.72 for black females. Analysis of black-white survival chances within each county reveals that the same variables account for most of the race gap in S₇₀ as well. When actual white male values for each explanatory variable are substituted for black in the black male prediction equation to assess the role explanatory variables play in the black-white survival difference, residual black-white differences at the county level shrink markedly to a mean of −2.4% (+/−2.4); for women the mean difference is −3.7% (+/−2.3).

Introduction

Large differences in life expectancy (LE) between different regions of the country have been long recognized [1–2 3 4 5 6 7]. Higher mortality in large urban areas and in the South may appear at first glance attributable to regional differences in racial composition [8–9 10], but as illustrated by the three maps in Figures 1, 2, and 3 depicting county-level probability of survival to age 70 (S₇₀) separately for white (Figure 1) and black men (Figure 2) and their difference (Figure 3), there are both salient within-race geographic differences and racial differences in mortality; similar gradients are seen for women (see below). Parsing evidence of this type in various ways has led some observers to conclude that there are distinct racial and geographic subpopulations living within the US, possibly with divergent and unique reasons for excess mortality [3, 11–12 13].

Figure 1. Probability of survival to age 70 for white males.
Probability of S₇₀ for white males by county, based on mortality rates 1999–2001. Small counties have been aggregated into Public Use Microdata Areas of >100,000 persons (N = 957). doi:10.1371/journal.pone.0032930.g001
The sources of geographic and racial variation have been the subject of considerable research in social epidemiology, economics, demography, environmental epidemiology, behavioral sciences and health services. Employing approaches and hypotheses along largely disciplinary lines, numerous important sources of the variation have been identified and in many cases confirmed in multiple settings. Factors related to social position, including education, income and job, have been repeatedly shown to correlate strongly with mortality rates, though their causal importance and relative contributions have been subject to extensive debate [4, 14–15, 16, 17]. Region-of-origin (e.g., race-ethnicity), cultural differences (e.g., family structure), urbanization and migration-related factors have been highlighted in other studies [11, 18–19, 20]. The relationships between mortality and so-called life-style choices, such as smoking, diet, and obesity have been examined from many perspectives and implicated as causes of premature mortality in cohort studies, with some evidence they may be on the pathway leading from social to regional differences [17, 21–22, 23, 24]. Differences in the experience of work, both as a psycho-social and possibly physical stressor, has been the focus of several studies [16, 25, 26]. Levels of ambient air pollution, most notably the small particulates generated by motor vehicles and power plants (PM$_{2.5}$), have been implicated in differential mortality [27–28, 29, 30] as have the temperature effects based on data emerging from the climate debate [31–32, 33]. Recent very intense investigation and reporting of regional differences in health care delivery, cost and quality [34–35, 36, 37, 38], as well as evidence of historic and ongoing racial disparities in care between whites and blacks [39, 40], have highlighted the role of these factors, although estimates of their contribution to mortality rates remain uncertain.

In this report we present an ecologic model of premature mortality – death before age 70 – that includes each of the factors that could be adequately measured for both whites and blacks at the county level in order to advance understanding of the disparities in several new ways. Following Deaton, Ezzati, Murray and others [3, 8, 12] we use the whole US population as our study frame, but break the country down to the more granular county level by using as our metric of observation S$_{70}$ rather than LE, avoiding the difficulties of estimating rates in sparse older groups and the widely observed “flattening” of race and geographic disparities observed in the study of mortality among the elderly [17, 41]. Moreover, we incorporate a broader set of predictors to bring socioeconomic, medical, environmental and demographic factors into a single model. To achieve this we employ a simplified regression analysis (weighted OLS) of county-level (ecologic) predictors of sex-specific survival to age 70 from birth separately for the white and black populations of each sex, although our aim is not so much to estimate the role of each specific factor as to describe their overall distribution and the extent to which they may collectively explain regional and race variation. This expansion of potential variables of interest is premised on the notion that racial and geographic variation most likely arises from diverse if inter-related sources. Thirdly, by demonstrating its utility to address these disparities at a granular level, we seek to establish S$_{70}$ as an outcome measure for research beyond the better entrenched metric, life expectancy from birth.
Methods

Outcome measurement

We calculated the probability of survival to age 70 ($S_{70}$) for white males, white females, black males, and black females from the CDC/NCHS Compressed Mortality Files (CMF) for the years 1999–2001 using an average of rates in the three years to reduce the effect of random or transitory circumstances that might have prevailed in 2000. Because of the change in the Census data collection strategy, comparable more recent data are not yet available for many of the predictor variables we use (see below and Table 1). Values were obtained by applying mortality rates for each five- and ten-year interval from birth to age 70 to a child born in that county in 2000. Thus $S_{70}$, derived like LE from period life tables, is a hypothetical statistic. It tells us what percentage of a cohort born in 2000 would survive until age 70 if the cohort experienced the age specific mortality rates that prevailed in that year. Unlike LE, which heavily “weights”

<table>
<thead>
<tr>
<th>Construct</th>
<th>Variable</th>
<th>Variable Name</th>
<th>Data Source</th>
<th>Metric</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low educational attainment</td>
<td>Education &lt;12 years</td>
<td>ED&lt;12</td>
<td>Census</td>
<td>% of subgroup *10^-2</td>
</tr>
<tr>
<td>High educational attainment</td>
<td>Education &gt;12 years</td>
<td>ED&gt;12</td>
<td>Census</td>
<td>% of subgroup *10^-2</td>
</tr>
<tr>
<td>High occupational attainment</td>
<td>Managerial or professional job</td>
<td>PROF&amp;MGR</td>
<td>Census</td>
<td>% of subgroup *10^-2</td>
</tr>
<tr>
<td>Income</td>
<td>Household income per adult equivalent</td>
<td>INCOME</td>
<td>Census</td>
<td>Mean (Household income in$/adult equivalents) *10^-9</td>
</tr>
<tr>
<td>Poverty</td>
<td>Under the poverty line</td>
<td>INPOV</td>
<td>Census</td>
<td>% of subgroup *10^-2</td>
</tr>
<tr>
<td>Wealth (property)</td>
<td>Log of property value</td>
<td>PROPVALUE</td>
<td>Census</td>
<td>Mean log (property value/5x10^4) among homeowners</td>
</tr>
<tr>
<td>Homeownership</td>
<td>Homeowner</td>
<td>HOMEOWNER</td>
<td>Census</td>
<td>% of subgroup *10^-2</td>
</tr>
<tr>
<td>Wealth (property) distribution</td>
<td>Gini coefficient on property values</td>
<td>GINI PROP</td>
<td>Census</td>
<td>Coefficient between 0 and 1</td>
</tr>
<tr>
<td>Between race disparity in (property) wealth</td>
<td>Mean Black/Mean White property value</td>
<td>B/W INCOME</td>
<td>Census</td>
<td>Sex-specific quotient</td>
</tr>
<tr>
<td>Living without a partner</td>
<td>Divorced, separated or never married</td>
<td>SINGLE</td>
<td>Census</td>
<td>% of subgroup *10^-2</td>
</tr>
<tr>
<td>Immigrant status</td>
<td>Not a US citizen</td>
<td>NONCITIZEN</td>
<td>Census</td>
<td>% of subgroup *10^-2</td>
</tr>
<tr>
<td>Urban county</td>
<td>Metro by census definition</td>
<td>METRO</td>
<td>Census</td>
<td>Dummy (yes/no)</td>
</tr>
<tr>
<td>Part urban</td>
<td>Part metro by census definition</td>
<td>PARTMETRO</td>
<td>Census</td>
<td>Dummy</td>
</tr>
<tr>
<td>In the south</td>
<td>Southern by census definition</td>
<td>SOUTH</td>
<td>Census</td>
<td>Dummy</td>
</tr>
<tr>
<td>Population growth rate</td>
<td>Population growth rate (shrinkage) between 1990–2000</td>
<td>GROWTH</td>
<td>Census</td>
<td>%change x10^-2</td>
</tr>
<tr>
<td>Availability of fast food</td>
<td>Proportion of restaurant sales classified</td>
<td>FASTFOOD</td>
<td>Economic census</td>
<td>% sales *10^-2</td>
</tr>
<tr>
<td>Quality of acute hospital care</td>
<td>Proportion of acute MI patients getting beta-blockers</td>
<td>BETABLOCKER</td>
<td>Ref</td>
<td>% hospitals* 10^-2</td>
</tr>
<tr>
<td>Cold climate</td>
<td>Mean January temperature</td>
<td>JANTEMP</td>
<td>Ref</td>
<td>Degrees F*10^-2</td>
</tr>
<tr>
<td>Warm climate</td>
<td>Mean July temperature</td>
<td>JULYTEMP</td>
<td>&quot;</td>
<td>Degrees F*10^-2</td>
</tr>
<tr>
<td>Air pollution</td>
<td>County mean concentration of fine particulate</td>
<td>PM2.5</td>
<td>EPA website</td>
<td>PM$_{2.5}$, in mg/M$^3$</td>
</tr>
<tr>
<td>Proportion of county population that is black</td>
<td>Proportion of adults self-reported as black</td>
<td>%BLACK</td>
<td>Census</td>
<td>% *10^-2</td>
</tr>
<tr>
<td>Black population in surrounding area</td>
<td>Proportion of adults in the State, excluding county, that is black</td>
<td>%STATEBLACK</td>
<td>Census</td>
<td>% *10^-2</td>
</tr>
</tbody>
</table>
events very early or late in life, $S_{70}$ unweighted by age of death, is primarily a summary measure of mortality rates in the 40’s, 50’s and 60’s, as illustrated in Figure 4.

**Figure 4. Age and distribution of deaths before age 70.** The distribution of age at death for all deaths before age 70 for each subpopulation for all US in the year 2000. doi:10.1371/journal.pone.0032930.g004

Our study design would ideally have estimated $S_{70}$ for each sex-race group in every county, but to assure stable mortality estimates requires a minimum of 2000 total sub population in each area in the CMF. This resulted in exclusion of many hundreds of counties that had small black populations. Furthermore, the primary source for variables we used to predict $S_{70}$ is the 5% sample of the 2000 US census, but these data are not geographically matched to the CMF. For privacy reasons the Census defines Public Use Microdata Areas (PUMAs) intended to capture 100,000+ total population areas: for low-density areas, contiguous counties were lumped together; high-density counties were sub-divided. To optimize coverage we created our own area units that match the CMF and Census geographic definitions precisely by using single counties where possible or matching groups of contiguous counties that were already grouped into PUMAs. The result is 510 areas covering 73 percent of the white and 96 percent of the black populations. They include 268 single counties and 242 groups of contiguous counties. For reader convenience, we refer to these 510 areas simply as "counties."

**Predictor Variables**

To analyze geographic differences in $S_{70}$ we examined the relation between $S_{70}$ in each race/sex group in each county as defined above and variation in 22 socio-economic and environmental variables that met two criteria: 1) have been broadly identified in the health literature as likely affecting mortality, hence possibly premature mortality, and 2) could be practically measured at the county level for both white and blacks (Table 1). Variables obtained from the 2000 Census of Population and Housing describe adults in each sex-race group in each county between the ages 30 to 59 with age-adjustment within that range by the direct method. Ten additional predictor variables, obtained from the Census and a wide variety of other sources, describe area characteristics; they are the same for each sex-race group except when variation in population distribution for a sex-race group affects the population weighted mean.

We would have preferred to include in our regression measures of other personal characteristics of the population which are suspect causes of premature mortality and possibly disparities, such as adverse health behaviors, diet, obesity and availability of health insurance. Although such data are sampled in periodic Behavioral Risk Factor Surveillance Surveys (BRFSS) the sample sizes for blacks are too low for all but 50 or so counties. Instead we conducted an additional sensitivity analysis (see below) to assess the importance of these covariates in explaining geographic differences among whites for whom data were adequate.

**Regressions**

Multivariate (population weighted ordinary least squares) regressions of $S_{70}$ on the 22 predictor variables were run for each sex-race group to estimate the contribution of these ecologic-level measures to geographic variation. To assess the degree to which the same 22 predictor variables explain race differences at the county level, we recalculated predicted $S_{70}$ for black men and women after inserting the (counterfactual) corresponding white values for each of the predictor variables in each county, then compared the resulting hypothetical predicted value for blacks to the prediction for whites, county by county.

**Sensitivity Analysis**

Because this study is limited by sample size considerations, availability of desired variables and in other ways, we carried out four complementary analyses to test the sensitivity of our results to these limitations. They are:

**Exclusion of deaths prior to age 30.** Much attention in both popular and professional publications focuses on race or sex differences in infant mortality, homicide, motor accidents, and other causes of death that are particularly important at younger ages. To determine the possible impact of omission of early life characteristics on our results, we repeated the analyses by examining survival to age 70 conditional on reaching age 30 ($S_{70/30}$). Shown in Table 2.
Table 2. Sensitivity Analysis.
Elimination of deaths under age 30 ($S_{70/30}$), N = 510, 22 predictor variables.

<table>
<thead>
<tr>
<th></th>
<th>White males</th>
<th>White females</th>
<th>Black males</th>
<th>Black females</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.73</td>
<td>0.83</td>
<td>0.57</td>
<td>0.72</td>
</tr>
<tr>
<td>(s.d.)</td>
<td>(0.04)</td>
<td>(0.02)</td>
<td>(0.06)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.85</td>
<td>0.76</td>
<td>0.76</td>
<td>0.70</td>
</tr>
</tbody>
</table>

Correlation between predicted $S_{70/30}$ and $S_{70}$>0.99 for all four race-sex groups in 510 counties. doi:10.1371/journal.pone.0032930.t002

**Inclusion of white counties omitted from the basic analysis.**
Because small black population in many counties required exclusion of many white counties, we repeated the calculation for 100 percent of the white population, which we were able to group in 957 areas of which 382 were individual counties and 575 were groups of contiguous counties. For this analysis we omitted the variable B/W INCOME for obvious reasons. Shown in Table 3.

Table 3. Sensitivity Analysis.
Inclusion of 100 percent of whites, $S_{70}$, N = 957, 21 predictor variables.

<table>
<thead>
<tr>
<th></th>
<th>White males</th>
<th>White females</th>
<th>Black males</th>
<th>Black females</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.71</td>
<td>0.82</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>(s.d.)</td>
<td>(0.04)</td>
<td>(0.03)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.83</td>
<td>0.76</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

Correlation between predicted $S_{70}$ for 957 counties from regressions across 957 counties and the original 510 county regression >0.99 for white males and white females. doi:10.1371/journal.pone.0032930.t003

**Inclusion of 8 other health related variables from the Behavioral Risk Factor Surveillance Survey (BRFSS) as predictor variables.** The additional variables are: current smoker, former smoker, obesity, uninsured, consumption of fruits and vegetables, physical activity, cholesterol checked, and the interaction of cholesterol check and obesity. To increase sample size, we average three years of data centered on 2000. Nevertheless, we could only make a direct comparison between the results for this augmented set of predictor variables and the results for the original 22 predictor variables for whites in 188 counties collectively comprising 51% of the US white population. Shown in Table 4.

Table 4. Sensitivity Analysis.
Inclusion of 8 BRFSS health behavior variables, $S_{70}$, N = 188, 29 predictor variables.

<table>
<thead>
<tr>
<th></th>
<th>White males</th>
<th>White females</th>
<th>Black males</th>
<th>Black females</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.72</td>
<td>0.82</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>(s.d.)</td>
<td>(0.04)</td>
<td>(0.02)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.90</td>
<td>0.86</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

Correlation between predicted $S_{70}$ for 188 counties based on 29 predictor variables and predicted $S_{70}$ for same counties based on 21 predictor variables >0.99 for white males and white females. See Table 9 with 30 predictors. doi:10.1371/journal.pone.0032930.t004

Reweighting $S_{70}$ based on the distribution of blacks. To assess the degree to which observed race differences might reflect differences in geographic distribution of the two races, the $S_{70}$ for whites and blacks were weighted for each of the 510 counties by the absolute number of blacks in that county. Shown in Table 5.

Table 5. Sensitivity Analysis.
Weighting $S_{70}$ by black population N = 510.

<table>
<thead>
<tr>
<th></th>
<th>White males</th>
<th>White females</th>
<th>Black males</th>
<th>Black females</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unweighted</td>
<td>0.71</td>
<td>0.82</td>
<td>0.57</td>
<td>0.71</td>
</tr>
<tr>
<td>Weighted</td>
<td>0.69</td>
<td>0.81</td>
<td>0.54</td>
<td>0.70</td>
</tr>
<tr>
<td>Difference</td>
<td>0.02</td>
<td>0.01</td>
<td>0.03</td>
<td>0.01</td>
</tr>
</tbody>
</table>

doi:10.1371/journal.pone.0032930.t005

Results

The outcome variable
Within in each sex-race group, there are striking geographical differences in the probability of survival to age 70 ($S_{70}$) as already suggested by Figures 1 and 2. Table 6 summarizes the extent of these by comparing mean effects as well as the lowest and highest ten percent of counties within each sex-race group. These differences are larger for males than females within each race and larger for blacks than whites within each sex.
Table 6. Population weighted means and standard deviations of $S_{70}$ for all 510 counties and the lowest and highest 10% of counties in each sex-race group.

<table>
<thead>
<tr>
<th></th>
<th>White males Mean (s.d.)</th>
<th>White females Mean (s.d.)</th>
<th>Black males Mean (s.d.)</th>
<th>Black females Mean (s.d.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All 510 counties</td>
<td>0.71 (0.04)</td>
<td>0.82 (0.02)</td>
<td>0.54 (0.07)</td>
<td>0.70 (0.04)</td>
</tr>
<tr>
<td>Lowest 10% of Counties</td>
<td>0.61 (0.02)</td>
<td>0.76 (0.01)</td>
<td>0.45 (0.03)</td>
<td>0.63 (0.01)</td>
</tr>
<tr>
<td>Highest 10% of Counties</td>
<td>0.77 (0.02)</td>
<td>0.85 (0.01)</td>
<td>0.68 (0.03)</td>
<td>0.77 (0.02)</td>
</tr>
</tbody>
</table>

doi:10.1371/journal.pone.0032930.t006

For a more complete picture of inter-county differences, we show in Figure 5 the frequency distribution of $S_{70}$ for the 510 counties for each sex-race group. The means in Table 6 have prepared us to see large differences between groups in the location of the distributions with respect to the $S_{70}$ axis, but the fact that there is so little overlap between the distributions of blacks and whites for either sex is even more striking, as is the absence of a significant overlap of male and female distributions for either race. On average, 82 percent of a cohort of white females born today could expect to live until 70 under the assumption of unchanging mortality rates, whereas only 54 percent of black males may have that expectation. There is a significant interaction between race and sex with respect to $S_{70}$; black-white differences are greater for males than females, and accordingly, male-female differences are greater for blacks than whites.

Figure 5. Frequency distribution (kernel plot) for $S_{70}$.
Frequency distribution (kernel plot) of survival to age 70 county for each subpopulation, 1999–2001. doi:10.1371/journal.pone.0032930.g005

Predictor Variables and Regression Results

The population weighted means and standard deviations for each of the 22 predictor variables for the four sex-race groups in each of the 510 counties are shown in Table 7. Noteworthy are the general similarities between men and women of each race, but striking between-race differences. Also noteworthy is the fact that these predictor variables are neither identically nor independently distributed. Figure 6 illustrates the highly significant inter-correlations among them for each race-sex group.

Results of the bivariate and OLS regression of $S_{70}$ using the 22 predictors are shown in Table 8 for each subgroup, noting the degree to which the estimated coefficients differ from the null. Notably, the percentage of variation in $S_{70}$ accounted for by the predicted values, i.e., the regression $R^2$s, are very high: 0.86 for white males, 0.79 for black males, 0.79 for white females, and 0.72 for black females; i.e., the equations account for most of the inter-county variation in $S_{70}$ within each sex-race group. As can be seen in Figure 7, comparing the predicted and actuals for each county, the predictors are equally relevant for all levels of the distribution, and for all size counties. Figure 8 depicts the t-statistic for each individual variable for each race-sex group that falls outside the window of chance association ($p<.05$) in the full OLS model.

Figure 6. Correlation globes for the predictor and outcome variables for each of the four subpopulations, white males (A), white females (B), black males (C) and black females (D).
All correlations with (absolute value) $r>.36$ are shown. Black lines denote a positive correlation; red negative. The thickness of the line is proportional to the absolute magnitude of the correlation. doi:10.1371/journal.pone.0032930.g006
Figure 7. Actual $S_{70}$ (y-axis) vs. predicted (x-axis) for each subpopulation. Note that circle size is proportional to county population (weight).

Figure 8. T-statistics (by sign and magnitude) for each significant predictor variable. Test statistics from the four weighted OLS regressions. Note that five variables are omitted altogether from the figure because they produced significant associations for none of the four subgroups: PROPVALUE, PARTMETRO, GROWTH, SOUTH and PM$_{2.5}$.

We used the regression results further to examine the extent to which the race differences in distributions may be related to differences in the predictor variables. Figure 9 A and B illustrate one way to assess this. The red and blue bars on the left represent the actual (red) and predicted (blue) distributions of $S_{70}$ for black men minus $S_{70}$ for white men in each of the 510 counties. The green bars to the right of each panel show the results of (counterfactually) replacing the measured black values with the measured white endowments, recalculating the predicted $S_{70}$ for black males under this counterfactual and hence the predicted black-white survival difference if whites...
Table 7. Weighted means and standard deviations for each predictor variable subgroup by county, N = 510.

<table>
<thead>
<tr>
<th></th>
<th>White</th>
<th></th>
<th>Black</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Males</td>
<td>Females</td>
<td>Males</td>
<td>Females</td>
</tr>
<tr>
<td>ED&lt;12</td>
<td>0.13</td>
<td>0.05</td>
<td>0.25</td>
<td>0.08</td>
</tr>
<tr>
<td>ED&gt;12</td>
<td>0.62</td>
<td>0.11</td>
<td>0.42</td>
<td>0.12</td>
</tr>
<tr>
<td>PROF&amp;MGR</td>
<td>0.36</td>
<td>0.09</td>
<td>0.17</td>
<td>0.07</td>
</tr>
<tr>
<td>INCOME</td>
<td>0.04</td>
<td>0.01</td>
<td>0.03</td>
<td>0.01</td>
</tr>
<tr>
<td>INPOV</td>
<td>0.06</td>
<td>0.03</td>
<td>0.15</td>
<td>0.05</td>
</tr>
<tr>
<td>PROPVALUE</td>
<td>1.03</td>
<td>0.48</td>
<td>0.52</td>
<td>0.51</td>
</tr>
<tr>
<td>HOMOWNER</td>
<td>0.77</td>
<td>0.10</td>
<td>0.58</td>
<td>0.12</td>
</tr>
<tr>
<td>GINI PROP</td>
<td>0.35</td>
<td>0.05</td>
<td>0.32</td>
<td>0.07</td>
</tr>
<tr>
<td>B/W INCOME</td>
<td>0.69</td>
<td>0.10</td>
<td>0.66</td>
<td>0.11</td>
</tr>
<tr>
<td>SINGLE</td>
<td>0.30</td>
<td>0.06</td>
<td>0.47</td>
<td>0.08</td>
</tr>
<tr>
<td>NONCITIZEN</td>
<td>0.06</td>
<td>0.06</td>
<td>0.05</td>
<td>0.07</td>
</tr>
<tr>
<td>METRO</td>
<td>0.88</td>
<td>0.30</td>
<td>0.85</td>
<td>0.35</td>
</tr>
<tr>
<td>PARTMETRO</td>
<td>0.05</td>
<td>0.18</td>
<td>0.04</td>
<td>0.18</td>
</tr>
<tr>
<td>SOUTH</td>
<td>0.39</td>
<td>0.49</td>
<td>0.55</td>
<td>0.50</td>
</tr>
<tr>
<td>GROWTH</td>
<td>0.25</td>
<td>0.17</td>
<td>0.18</td>
<td>0.17</td>
</tr>
<tr>
<td>FASTFOOD</td>
<td>0.48</td>
<td>0.07</td>
<td>0.49</td>
<td>0.07</td>
</tr>
<tr>
<td>BETABLOCKER</td>
<td>0.68</td>
<td>0.06</td>
<td>0.68</td>
<td>0.06</td>
</tr>
<tr>
<td>JANTEMP</td>
<td>0.39</td>
<td>0.13</td>
<td>0.39</td>
<td>0.13</td>
</tr>
<tr>
<td>JULYTEMP</td>
<td>0.76</td>
<td>0.07</td>
<td>0.77</td>
<td>0.05</td>
</tr>
<tr>
<td>PM2.5</td>
<td>0.14</td>
<td>0.03</td>
<td>0.14</td>
<td>0.03</td>
</tr>
<tr>
<td>%BLACK</td>
<td>0.15</td>
<td>0.12</td>
<td>0.27</td>
<td>0.16</td>
</tr>
<tr>
<td>%STATEBLACK</td>
<td>0.13</td>
<td>0.08</td>
<td>0.17</td>
<td>0.10</td>
</tr>
</tbody>
</table>

doi:10.1371/journal.pone.0032930.t007

Figure 9. Percent of counties with actual and predicted race differences (black-white) in S70 for men (A) and women (B). Red and black bars represent percent of counties (N = 510) with actual and predicted race differences (black-white) in S70 for men. The grey bars on each panel represent the hypothetical black-white difference in predicted S70 if blacks in each county were assigned the comparable white value for each predictor variable.

doi:10.1371/journal.pone.0032930.g009
and blacks were identical on the attributes. As can be seen in Figure 9A, the race differences in $S_{70}$ at the county level narrow almost to nil: –2.4% (+/–2.4) for men, –3.7% (+/–2.3) for women (Figure 9B). When the procedure is reversed, the conclusion is the same; i.e., when black values for the predictor variables are substituted for white values in the white regressions, the curves for predicted white males (or females) resemble their black counterparts (not shown). Notably, the gender “gap” is not so explained: when female values of the 22 variables are substituted for male values in each county, there is no change in the (large) male-female differences in distribution of predicted $S_{70}$ for both whites and blacks.

### Sensitivity Analyses

The robustness of these results was tested by four alternative approaches that varied the dependent variable, the size of the population covered, the scope of predictor variables (to include health behaviors) and the impact on our results of the different geographic distribution of whites and blacks in the US. In the first test, deaths under age 30 were excluded from the study; the dependent variable was probability of survival to 70 conditional on reaching 30 ($S_{70/30}$). The means, standard deviations, the sex and race differences, the $R^2$s, and the predicted $S_{70}$ ($r>0.99$) for each county all closely match those found for $S_{70}$.
Table 9. $S_{70}$ weighted regressions for white males and females, 8 BRFSS variables added, N = 188 counties.

<table>
<thead>
<tr>
<th></th>
<th>Males</th>
<th></th>
<th>Females</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ED&lt;12</td>
<td>−0.31</td>
<td></td>
<td>**</td>
<td>−0.24</td>
</tr>
<tr>
<td>ED&gt;12</td>
<td>−0.02</td>
<td></td>
<td>0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>PROF&amp;MGR</td>
<td>0.24</td>
<td></td>
<td>**</td>
<td>0.08</td>
</tr>
<tr>
<td>INCOME</td>
<td>−0.78</td>
<td></td>
<td>**</td>
<td>−0.19</td>
</tr>
<tr>
<td>INPOV</td>
<td>−0.37</td>
<td></td>
<td>**</td>
<td>−0.16</td>
</tr>
<tr>
<td>PROPVALUE</td>
<td>0.01</td>
<td></td>
<td>−0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>HOMEOWNER</td>
<td>0.03</td>
<td></td>
<td>0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>GINI PROP</td>
<td>−0.03</td>
<td></td>
<td>0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>B/W INCOME</td>
<td>−0.02</td>
<td></td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>SINGLE</td>
<td>−0.16</td>
<td></td>
<td>**</td>
<td>−0.09</td>
</tr>
<tr>
<td>NONCITIZEN</td>
<td>0.36</td>
<td></td>
<td>**</td>
<td>0.34</td>
</tr>
<tr>
<td>Current Smoker</td>
<td>−0.01</td>
<td></td>
<td>−0.07</td>
<td>**</td>
</tr>
<tr>
<td>Former Smoker</td>
<td>−0.03</td>
<td></td>
<td>0.02</td>
<td>0.00</td>
</tr>
<tr>
<td>Obesity</td>
<td>0.02</td>
<td></td>
<td>0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>Uninsured</td>
<td>0.08</td>
<td></td>
<td>**</td>
<td>0.03</td>
</tr>
<tr>
<td>Fruits&amp;Veg</td>
<td>0.00</td>
<td></td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Physical Activi</td>
<td>0.06</td>
<td></td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Chol Checked</td>
<td>−0.05</td>
<td></td>
<td>**</td>
<td>0.00</td>
</tr>
<tr>
<td>Chol Chk &amp; Obes</td>
<td>−0.01</td>
<td></td>
<td>−0.05</td>
<td>0.00</td>
</tr>
<tr>
<td>METRO</td>
<td>−0.02</td>
<td></td>
<td>−0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>PARTMETRO</td>
<td>−0.01</td>
<td></td>
<td>−0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>SOUTH</td>
<td>0.00</td>
<td></td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>GROWTH</td>
<td>0.00</td>
<td></td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>FASTFOOD</td>
<td>0.00</td>
<td></td>
<td>−0.06</td>
<td>*</td>
</tr>
<tr>
<td>BETABLOCKER</td>
<td>0.08</td>
<td></td>
<td>**</td>
<td>0.04</td>
</tr>
<tr>
<td>JANTEMP</td>
<td>0.02</td>
<td></td>
<td>0.02</td>
<td>0.00</td>
</tr>
<tr>
<td>JULYTEMP</td>
<td>−0.05</td>
<td></td>
<td>**</td>
<td>−0.01</td>
</tr>
<tr>
<td>PM2.5</td>
<td>0.07</td>
<td></td>
<td>0.03</td>
<td>0.00</td>
</tr>
<tr>
<td>%BLACK</td>
<td>−0.08</td>
<td></td>
<td>**</td>
<td>−0.04</td>
</tr>
<tr>
<td>%STATEBLACK</td>
<td>0.02</td>
<td></td>
<td>0.03</td>
<td>**</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.76</td>
<td></td>
<td>**</td>
<td>0.86</td>
</tr>
<tr>
<td>R^2</td>
<td>0.90</td>
<td></td>
<td>0.86</td>
<td>0.00</td>
</tr>
<tr>
<td>N</td>
<td>188</td>
<td></td>
<td>188</td>
<td>0.00</td>
</tr>
</tbody>
</table>

$\approx p<.05$.  * $p<.01$.  ** $p<.001$.

doi:10.1371/journal.pone.0032930.t009

In the second test, all the white counties that had been excluded from the basic analysis because there were insufficient blacks were included, creating a data set of 957 “counties” covering 100 percent of the white population. Again all the relevant results including predicted county $S_{70}$ ($r>0.99$) closely match those obtained when 510 counties covered 73 percent of the white population.

By drawing on the BRFSS data for whites–black sample sizes were too small to allow inclusion in the main analysis–we were able to add 8 predictor variables including smoking, BMI, diet and physical activity for whites. The results, based on 188 counties covering 51% of the US white population, are very similar to those for the same counties with the 22 original predictor variables. The correlation between predicted $S_{70}$ is very high, $r>0.99$. The detailed results of these sensitivity analyses are shown in Tables 2 and 3. Notably only physical activity achieves even marginal significance in a full model, and that only for men.

Finally, by weighting the $S_{70}$ values by the number of blacks in each county, we show that geographic distribution of the races does not explain more than 1–3 percentage points of the race differential.

Discussion

Examining the probability of survival to age 70 for each sex-race group by county we illustrate in a novel way the geographic and race disparities in premature mortality. Figure 5, with its frequency distributions of 510 counties for each of the sex-race groups, illustrates the chasmic difference between blacks and whites, true for both sexes, albeit greater for males than females. Not only are the means of these distributions significantly different as might have been expected, but there is almost no overlap: the counties with the best survival for blacks are little better than the worst counties for whites. Moreover, we have shown that differences in the 22 predictor variables, as a group, account for most of the geographic and black-white disparities in survival to age 70. Figure 7 illustrates the strength of the associations of each with $S_{70}$ within each of the four subpopulations.

Some results, such as the impacts of education, high occupation, and marital status, are highly consistent with expectation from prior work [8, 14, 17, 41] while others—such as the failure of PM$_{2.5}$ or prevalent behavioral factors such smoking and diet (in whites) to achieve significance—may appear surprising. We refrain, however, from drawing strong inferences about the quantitative importance of such individual observations nor do we infer from our results a causal relation between any factor and premature mortality because of limitations in our approach. First among these limitations is measurement error, which could, for example, obscure or diminish the effect of health care quality (because only a single metric was used, and that assigned fairly crudely) or PM$_{2.5}$ that likely varies greatly within topographically diverse
counties such as LA. Further misclassification of exposures are inevitable because of our treatment of time—we have used current exposure in 2000 to “predict” mortality during the same window—which may distort the role of factors with impacts over years, such as smoking and BMI. Likewise, the assumption of a linear relationship implicit in our choice of the OLS model, may be inappropriate for some variables such as income (previously shown to have a diminishing association with health [42]), while omission of other, possibly important variables, such as robust measures of health behaviors for all but the larger counties, is also a significant shortcoming in our approach. We acknowledge that the impact of change in the county composition itself, with in- and out-migration, could bias our results. We have attempted to capture such change with the single variable GROWTH but undoubtedly this is imperfect. Another factor limiting causal inference is the likelihood of reverse causality for some associations, such as health status on subsequent marital or employment status. However, with the exception of this one, most of the other limitations should tend to bias the explanatory power of our model towards the null, hence leading us to underestimate the extent to which the predictor variables as a group account for the observed geographic or black-white disparities.

Perhaps most limiting of all for causal inference is the ecologic, rather than individual level measurement of our key variables in our model because we lack knowledge of the individual characteristics of those who died. The interpretation of such models is inevitably ambiguous. For example, while we find a strong negative relation between percent in poverty and the probability of surviving to age 70, our model cannot distinguish between a) excess pre-70 deaths of individuals who are in poverty vs. b) excess pre-70 deaths of non-poor individuals who live in high poverty areas. Clinearity of some variables, as illustrated in Figure 6, may also lead to partial misattribution. Other potential limitations of this study were addressed in the sensitivity analyses with reassuring results.

While these issues collectively diminish our enthusiasm for drawing strong inferences from estimates for the individual predictors, certain observations merit comment. The very strong effects of education, poverty and occupational status across the race-sex distributions adds premature mortality to the long list of health impacts previously reported. Although the most commonly used measure of distributional disparity within groups (GINI PROP) showed no effect, contrary to some earlier work by [43–44, 45, 46], our result is almost identical to Deaton and others [8, 47], suggesting social disparity between races may be important, as suggested by the negative impact of %BLACK on white survival and the effect of black-white income differentials (B/W INCOME) on black S70. The impact of marital status in not a new observation [20] but the consistency of the effect across race and sex groups is noteworthy. Likewise is the very striking positive effect of NONCITIZEN—proportion immigrants is associated with higher survival in all four groups. This effect is so strong that failure to consider this variable in our model almost completely washes out the effects of education and occupation, as many of the immigrants, both black and white, have very low education attainment despite apparently better survival than their race-matched US born counterparts. This is also not a new observation, but calls further attention to likely strong health-associated selection effects first among those who come to the US who are likely to be healthier than average and later those who return to their country of origin because of poor health. This raises the possibility of statistical measurement errors for assessing mortality among such immigrant groups or possible differential impacts of the other determinants on these subpopulations [48, 49–50].

Among the area variables, it is perhaps surprising that the classic demographic features, e.g., METRO and SOUTH, do not impart much to the aggregate association. Neither do PM2.5 and average temperatures, although measurement may play an important role in the failure to see such effects. FASTFOOD appears to have a measureable association for whites but not blacks, an observation that merits further evaluation. Notably, our single measure of health care quality—BETABLOCKER—shows a consistent and significant effect in all groups despite the fact it was measured at the state, not county level, likely biasing the observed effect towards the null. The impact of health care quality on survival to age 70 in the US has not, to our knowledge, been previously tested.

Taking even these observations cautiously because of the limitations, three conclusions seem inescapable. First, we have shown that geographic disparities are not primarily inherent in location, but are best understood as related to disparities in education, occupations, and the like which are strongly associated with outcome in every country we studied—large, small, urban, rural, southern or not. The absence of even a single strong outlier county (see Figure 7) lends strong support to this notion and suggests that the construct of “8 Americas” based on racial/ethnic and geographic “pockets” of poor health by Ezzati et al [3] and highlighted by others [11–13] is perhaps misguided. Similarly it would appear that most of the black-white gap in health is also related to differences in these well-known socio-economic and environmental variables, with poverty, low education and single marital status appearing particularly disparate between the races (cf. Table 7 and Figure
8). That this observation is not an inevitable consequence of our method is strengthened by the absence of any effect when the independent variables were “switched” between the sexes in an effort to explain the gender gap: women, perhaps due to genetic, biologic or sociologic factors omitted from our analysis, are far less susceptible than men to premature mortality attributable to their social and physical environment, at least as we have measured them.

Finally, we believe that the descriptive clarity and analytic benefits of $S_{70}$ show it to be a useful measure of population health. While life expectancy may be useful for many purposes, such as the study of the impact of care in the elderly or changes in infant mortality to which LE is very sensitive, survival to age 70 provides an alternative measure for elucidating race and sex disparities in health. For example, while white-black difference in male life expectancy in 2006 is 7 percent, the difference in survival to age 70 is 17 percent of the average level. Female life expectancy exceeds male by 7 percent, while survival to age 70 differs by 13 percent of the average level. Not only are the differences magnified, but unlike life expectancy, $S_{70}$ focuses unambiguously on the fact that these disparities occur for the most part in the prime, economically productive years of life (Figure 4). Alternatively we might have looked at survival to early or later ages (e.g. $S_{65}$ or $S_{85}$). However these choices would create other problems, at least for the US population: For $S_{65}$ or smaller compression of the distribution at the right tail becomes a problem, as increasing numbers of counties would have values greater than 90 at least for white women. For $S_{85}$ or greater we would likely run into many of the issues that may limit LE from birth, including flattening of the disparities at older ages and increasing relevance of late life survival factors. We suggest that overall, $S_{70}$ may serve as the most valuable complement to that more familiar statistic used to summarize population mortality rates.

**Author Contributions**

Conceived and designed the experiments: MRC CC VRF. Performed the experiments: MRC CC VRF. Analyzed the data: MRC CC VRF. Contributed reagents/materials/analysis tools: MRC CC VRF. Wrote the paper: MRC CC VRF.

**References**


Acknowledgments

Many colleagues and other professionals have played important roles in the creation of these papers and this booklet. I am grateful to my co-authors, each of whom is acknowledged with the appropriate paper. I thank my long-time assistant, Rossannah Reeves, for help with all of the papers, Michelle Mosman, director of communications at SIEPR for production of this booklet, and John B. Shoven, director of SIEPR for encouragement and advice. I also appreciate the permission to reprint granted by the various journals and the financial support of the Robert Wood Johnson Foundation.