NBER WORKING PAPER SERIES

A SATELLITE ACCOUNT FOR HEALTH IN THE UNITED STATES

David M. Cutler
Kaushik Ghosh
Kassandra Messer
Trivellore Raghunathan
Allison B. Rosen
Susan T. Stewart

Working Paper 27848
http://www.nber.org/papers/w27848

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2020, Revised October 2020

We are grateful to Patricia Berglund, Irina Bondarenko, Paul Imbriano, Marcelo Coca Perraillon, James Schafer, and Rebecca Woodward for research assistance, and to participants in our National Health Accounts Advisory Group and the Harvard University Interfaculty Program for Health Systems Improvement for initial advice and support. We thank Corby Garner, Joe Newhouse, and Louise Sheiner for very helpful comments, and the National Institute On Aging of the National Institutes of Health for research support under award numbers R37AG047312 and P01AG31098. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research or the National Institute On Aging.

A Satellite Account for Health in the United States
David M. Cutler, Kaushik Ghosh, Kassandra Messer, Trivellore Raghunathan, Allison B. Rosen, and Susan T. Stewart
NBER Working Paper No. 27848
September 2020, Revised October 2020
JEL No. E2,I10,I11,I12,I38

ABSTRACT

Estimating medical care productivity is a central economic challenge. This paper develops a satellite account for the US health sector that appropriately measures health care productivity and applies that to the elderly population between 1999 and 2012. The central output of the satellite account is health. The primary input is medical care; we also examine the impact of behavioral risk factors. Our empirical work measures the change in medical spending and health outcomes for a comprehensive set of 80 conditions. We estimate that medical care has positive productivity as a whole, with aggregate productivity growth of 9% over the time period. However, there is significant heterogeneity in productivity by condition. At the upper end, care for cardiovascular disease has been extremely productive. In contrast, care for people with mental illness and musculoskeletal conditions has been costly but not productive.

David M. Cutler
Department of Economics
Harvard University
1875 Cambridge Street
Cambridge, MA 02138
and NBER
dcutler@harvard.edu

Kaushik Ghosh
NBER
1050 Massachusetts Ave.
Cambridge, MA 02138
ghoshk@nber.org

Kassandra Messer
University of Michigan
SRC - Health and Retirement Study
3467 ISR-Thompson
Ann Arbor, MI 48106
kasey@umich.edu

Trivellore Raghunathan
University of Michigan
SPH II
1415 Washington Heights
Ann Arbor, MI 48109
teraghu@umich.edu

Allison B. Rosen
Department of Quantitative Health Sciences
University of Massachusetts Medical School
368 Plantation Street, AS9-1083
Worcester, MA 01605
Arosen006@gmail.com

Susan T. Stewart
NBER
1050 Massachusetts Ave
Cambridge, MA 02138
sstewart@nber.org

A online appendix is available at http://www.nber.org/data-appendix/w27848

"The welfare of a nation can scarcely be inferred from a measure of national income."

- Simon Kuznets (1934)

Estimating medical care productivity is a central economic challenge. The National Income and Product Accounts (NIPAs) report that multi-factor productivity in the health care and social assistance sector declined by 0.4% annually between 1987 and 2018,[1] in contrast to the 0.9% increase in the economy as a whole. Given health care's size in the economy, about 18% of GDP, low productivity in medical care influences aggregate productivity growth. Further, the low rate of measured productivity growth contributes to ongoing concern about wasted resources in medicine (Chandra and Skinner, 2012).

However, official medical care productivity estimates suffer from a variety of problems (Boskin Commission, 1996; Berndt et al., 2001; Dunn et al., 2015). Quality adjustment is an often-noted issue. The primary output of medical care is improved health, yet health is not easily measured or attributed to medical interventions. Thus, health improvements are undercounted in official productivity estimates. This limitation was evident as far back as the 1930s, as the above quotation from Simon Kuznets attests. In addition, the national accounts have difficulties when care from one type of provider substitutes for care from another. Each provider type is treated as a separate industry in the national accounts, for example pharmaceuticals, hospitals, and physicians. Thus, if a medication is developed which avoids the need for hospitalizations, the savings in hospital costs are not credited to the pharmaceutical industry. For each of these reasons, medical care productivity by sector may be significantly misstated.

To address these issues, proposals have periodically been made to establish a 'satellite account' for health. Such an account would measure health outcomes and medical inputs, and thus more accurately account for new goods and substitution of treatments (Schultze and Mackie, 2002; Abraham and Mackie, 2005; Rosen and Cutler, 2007; National Research Council, 2010).[2] The goal of this paper is to develop a satellite account for the US health sector and measure the productivity

[1] This includes health care and social assistance, although social assistance represents less than 10% of the total. Health care services includes inpatient and outpatient facilities but excludes pharmaceuticals and medical devices. Over the 1987-2016 time period, the reported MFP of the pharmaceutical industry was -1.7% annually and the reported productivity of the medical equipment and supplies industry was 0.7% annually.

[2] Health is not the only area with these concerns. Other areas for which satellite accounts are proposed include home production (Landefeld and McCulla, 2000), the environment (Muller et al., 2011), education (Fraumeni et al., 2017), and free goods (Brynjolfsson et al., 2019).

of medical care.

The key design innovation in our accounts is to use medical conditions as the industries, not providers of care. For example, there is an industry for heart disease and a second one for lung cancer. Our accounts make no distinction based on the type of care provided; all that matters to people is how much they spend on care and their resulting health.

The data needs for a satellite account are substantial, including measures of population health and medical spending matched to clinical characteristics and other sociodemographic factors. Our primary data are from the Medicare Current Beneficiary Survey (MCBS), an ongoing survey of the elderly population with data from 1999–2012. The elderly account for one-third of total medical spending.

We start by identifying a set of 80 conditions that are prevalent and severe enough to affect the elderly population. Of these conditions, 78 are clinical, for example heart disease and lung cancer. The remaining two are moderate and severe frailty. Frailty is not a disease that clinicians code, but we show that it is vital for both spending and health. We measure the prevalence of each condition using information from medical claims combined with imputations based on self-reports in the National Health and Nutrition Examination Survey.

Medical spending is straightforward to measure, but health is not. The concept we wish to measure is quality-adjusted life expectancy (QALE), defined as the expected number of years a person will live adjusted for the health-related quality of those years. To match our focus on the elderly, we measure QALE as of age 65. Mortality by age is readily available; quality of life is not. We develop an empirical measure of quality of life based on survey data asking about necessary physical activities, sensory impairments, and how much health interferes with social interactions. We estimate that over the 1999–2012 time period, quality-adjusted life expectancy increased about one year age at age 65. The increase in QALE is the net effect of two offsetting trends: the prevalence of chronic disease is rising as obesity increases, while people with chronic diseases are on average living longer.

The bulk of our empirical analysis compares the change in spending for each condition over time with the change in health outcomes for people with that condition. Measuring the growth in medical expenditures and health implications of each condition is complicated because of comorbidities. If a person has a heart attack, should the spending for the hospital stay be attributed to the coronary event itself or to the mental illness, which made taking preventive medication

difficult? Since both conditions are at fault, we need a method of decomposing spending and health decrements that partials out the spending and health decrement to both conditions appropriately.

Our assumption, consistent with clinical literature, is that medical care for a condition affects the health of people with that condition but not the prevalence of other conditions. That is, care for a heart attack influences the length and quality of life for people with a heart attack but not whether the person develops cancer. The only exceptions we make to this assumption are for conditions with well-identified clinical links, such as high blood pressure leading to heart disease. The implication of our assumption about there being few clinical risk factors is that most prevalence changes are assumed to fall outside of the medical sector. Thus, the productivity of medical care is mostly determined by health trends for people once they have acquired diseases.

Even focusing on the outcomes per case of disease, there are still empirical challenges in attributing outcomes to conditions. Most people have more than one condition, and thus a method must be used to decompose spending among conditions. We use a propensity score framework to do this. We compare spending, mortality, and quality of life for a person with each condition to similar people who do not have the condition. This yields an estimate of spending and health outcomes that result from that condition. The estimates are then further adjusted so that total spending and mortality match national totals. The result is condition-specific spending, mortality, and quality of life for each year 1999–2012.

To test the quality of our estimates, we compare the health changes we estimate to disease models in the literature. For both heart disease and some cancers, there are well-validated disease models indicating how specific treatments have affected survival. We show that the impact of medical treatment changes using our method is very close to that in the literature.

Using our estimates of spending and health changes by condition, we estimate productivity for medical care alongside productivity for each of the 80 conditions. On net, we estimate positive productivity growth in medical care. Aggregate productivity growth is 9% over the time period or 0.7% annually. In comparison, official data indicate a productivity change of 0.1% annually over this time period.

However, there is considerable heterogeneity in productivity by condition. At the upper end, care for cardiovascular disease has been extremely productive; health has significantly improved with minimal cost changes or in some cases cost reductions. Similarly, there is improved health in areas as diverse as kidney failure, lung and colorectal cancer, and frailty. At the other

end, productivity growth is low for patients with mental illness, arthritis, and other musculoskeletal conditions. There are some areas with negative productivity growth, but the magnitude of these productivity reductions is not large. Instead, the data show some conditions with very high productivity and a number indicating little or no productivity growth. This heterogeneity in productivity is indicative of a medical system where input allocation is not always optimized.

The paper is structured as follows. The first section discusses the conceptual issues in health accounting. The second section describes our data. The third section details our methodology for estimating condition-specific spending growth and health improvement. The fourth section estimates spending changes by condition, and the fifth section does the same for health improvement. The sixth section presents productivity estimates. The last section concludes. There are two appendices: one describing our data set and the adjustments we make to it and a second giving details on our spending and health attribution models.

I. Accounting for Health

A. Current Treatment of Medical Care in NIPAs and Problems

The United States has two sets of medical spending accounts, one a part of the NIPAs and a second maintained by the Centers for Medicare and Medicaid Services (CMS). The two have many features in common (Sensenig and Wilcox, 2001). Both measure medical spending, not health outcomes. Both divide spending into industries based on the type of provider, not the disease. In the NIPAs, there are 13 industries relevant to health care, reflecting a mix of goods (prescription drugs, nonprescription drugs, and medical equipment) and services (e.g., hospital care and physician care).[3]

The Bureau of Labor Statistics (BLS) measures prices in each industry via surveys. Providers in the price surveys are asked to price the same bill over time, holding constant patient characteristics, payment source, and method of payment (for example, a hospitalization for a heart attack without significant comorbidities, covered by the Blue Cross-Blue Shield HMO plan). As in most industries, quality adjustment is designed to occur at the price index stage. Real output is nominal spending deflated by quality-adjusted prices (BLS, 2007).

[3] The full list of services includes professional services (physicians' services, dental services, eyeglasses and eye care, and other medical professionals) and hospital and related services (hospital services, and nursing homes and adult day care). Health insurance is its own category. CMS further decomposes spending by payer (public insurance, private insurance, out-of-pocket, etc.), though that is not our concern in this paper.

In practice, there are several difficulties with the national account treatment of health. We demonstrate these issues with an example consisting of a single condition, heart disease. The first problem is adjusting for changing quality. Imagine that a new surgical procedure is developed that allows people with heart disease to live longer. The impact of the new surgery on the medical price index will depend on how it is paid for. If the new service is placed in its own payment bucket, it will be linked to the existing price index and its introduction alone will not affect price trends. If the new surgery is placed in the same billing category as an older surgery or is put there by BLS,[4] the price index will rise or fall depending on the relative reimbursement for the old and new surgery. However, neither of these methods has an explicit quality adjustment.

BLS does not have a good quality adjustment in health care. Respondents to the BLS pricing surveys are asked to indicate if the good has changed materially, but few say so. And many changes in quality result from a series of small steps more than one large change. This is true in many industries, so other quality-adjustment techniques are standard. In many industries, quality adjustment is done through hedonic analysis (Groshen et al., 2017). For example, prices for each procedure or medication over time could be regressed on characteristics of the therapy, including outcomes and side effects (Cockburn and Anis, 2001; Comaner et al., 2005), with the non-quality component of a price change being the price index.

For several reasons, however, hedonic analysis is unlikely to work well in medical care (Berndt et al., 2000): consumers rarely pay the full cost of care they use, so marginal valuations are not equal to price; hedonics do not work well in life-or-death decisions; information asymmetries mean that consumers are unlikely to know the real value of services when they receive them; and health outcomes are difficult to parcel out to particular interventions, making it difficult for doctors and patients to know the consequences of what they are doing. As a result, hedonic analysis is not used in health care.

The impact of omitting quality change could be substantial. Life expectancy at birth increased 3.2 years between 1990 and 2017, at a time when estimated medical care productivity was falling. If even a portion of the health improvement were due to medical care, and this effect was captured, the overall evaluation of medical care productivity could change considerably.

A second problem with current methods is that they do not account for substitution across industries. Returning to the heart disease example, suppose that a new class of pharmaceuticals is

[4] For example, BLS made brand and generic drugs equivalent in the medical spending account.

developed that substitutes for surgery in people with heart disease. No lives are saved, but the new medication is sufficiently inexpensive that less is spent in total. Regardless of whether the new drug is linked into the existing pharmaceutical index or included in a basket with other medications, the national accounts will not show any productivity gain from reduced surgery. That is because pharmaceuticals and hospital care are different industries. In essence, the shift from surgery to medication will be attributed to consumer tastes, not to medical care productivity.

To make progress on both of these issues, we need to partition medical care into more appropriate industries. The assumption we make is that the patient cares about their health and the total amount they spend on treatment, but not what particular treatment they receive. That is, pills and surgery differ only in their cost and whether they lead to different health outcomes. On all other metrics, they are the same. As a result, the relevant industries are not the providers of therapy but rather the conditions being treated.

The Bureau of Economic Analysis has reached a similar conclusion and has recently started experimentally measuring medical spending on a condition basis (Dunn et al., 2015). However, the methods to do this are complex, as we explain below. And no attempt has been made to measure health outcomes systematically.

B. A Satellite Health Account

To understand the concept behind a health account, we start with individual behavior and then introduce technological changes. Our framework parallels that of Murphy and Topel (2006), though it is richer in the treatment of medical advances.

Individual utility

Consider a population of *ex-ante* identical individuals who receive utility from two goods: health and non-medical consumption.[5] Health is denoted $H_t(a)$ for a person aged a in year t. We denominate health so that it ranges from 0 (death) to 1 (perfect health). In any period, the utility is multiplicative in health and a concave function of non-medical consumption, x_t: $U_t(a) = H_t(a)\, U(x_t)$, where $U' > 0$ and $U'' < 0$. Thus, greater non-fatal health increases the value of all other consumption, a reflection of quality-of-life improvement. Lifetime utility is the present value of flow utility for each year the person is alive. Denote $S_t(a_0)$ as an indicator for whether the person

[5] For convenience, we omit labor supply. Adding a labor-leisure choice would not change the results materially.

is alive in year t conditional on being alive at age a_0 in t=0. ρ is the discount rate. Lifetime welfare is then given by:

$$W(a_0) = \sum_{t=0}^{\infty} (1+\rho)^{-t} S_t(a_0) H_t(a) U(x_t) \quad . \tag{1}$$

If people are fully insured or nearly so, medical spending in any period will be the average spending in the population, denoted $m_{ave,t}$.[6] $m_{ave,t}$ may vary over time because the quantity of treatments change or the price paid per treatment changes; given our external measure of quality, we do not need to differentiate these. The budget constraint is then that the present value of medical and non-medical consumption must be less than or equal to the present value of income plus initial assets (A_0). With perfect annuity markets, this will be an equality:

$$\sum_{t=0}^{\infty} (1+\rho)^{-t} S_t(a_0) (Y_t - x_t - m_{ave,t}) + A_0 = 0 \quad . \tag{2}$$

For convenience, we assume the interest rate is equal to the discount rate, but this is not necessary if individuals can borrow and lend. Individuals maximize utility (1) subject to the budget constraint (2). The formulation of this problem is given in equation (3)

$$W(a_0) = \sum_{t=0}^{\infty} (1+\rho)^{-t} \left\{ S_t(a_0) H_t U(x_t) + \mu S_t(a_0) (Y_t - x_t - m_{ave,t}) \right\} + \mu A_0 \, . \tag{3}$$

Equation (3) is denominated in utils. To translate this into dollars, we divide both sides by the marginal utility of wealth, μ.

There are two objects of choice in equation (3): medical spending and non-medical consumption.[7] While medical care decisions are the outcome of some optimization problem, it is not clear that it is a rational decision. In any event, our empirical work does not rely on the optimality of medical care decision-making. To derive the relevant accounting, we assume only optimization of non-medical consumption. This implies that the marginal utility of non-medical consumption is the same in all periods; $H_t U_t^{'} = \mu$.

The Impact of Medical Treatment

[6] In the absence of this assumption, the value of medical spending varies because marginal utility of income changes with spending on medical care.

[7] One might imagine that some consumption goods also affect health, for example smoking. Our empirical analysis does not assume that such goods are chosen optimally, and there is much evidence that they are not.

We now provide more detail on medical conditions and treatments for them. People have one or more medical conditions, each one denoted by c. The prevalence of condition c at time t is $prev_{ct,}$ and spending is m_{ct}. Prevalence may be exogenous or endogenous; we specify which in the empirical section. Similarly, we omit for now the problem of decomposing spending among comorbid conditions. Bold terms denote vectors, where each element is the value for each condition. Thus, condition prevalence and medical spending per condition at time t are given by $\mathbf{prev_t}$ and $\mathbf{m_t}$. Population average medical spending is the product of these two terms: $m_{ave,t} = \mathbf{prev}_t \cdot \mathbf{m}_t$.

The vector of mortality rates for each condition is denoted $\boldsymbol{\gamma}_{at}$. This varies by age and may change over time as medical treatments or other factors change. Aggregate mortality is the prevalence-weighted average of the mortality rates for individual conditions: $\Gamma_{at} = \mathbf{prev}_{at} \cdot \boldsymbol{\gamma}_{at}$.[8] The survival rate in equation (3) is the product of the survival rate at each age up to a: $S_t(a_0) = \prod_{\tau=0}^{t-1}(1-\Gamma_{\tilde{a}\tau})$, for $\tilde{a}=a_0+\tau$.

Medical care also affects quality of life. We assume quality of life is linear in the decrement associated with each condition: $H_t(a) = \bar{h}_t(a) - \mathbf{prev}_{at} \cdot \mathbf{h}_{at}$, where $\bar{h}_t(a)$ is the level of health for people with no conditions – which may be below one due to non-captured conditions – and $\mathbf{h}_{at}$ is the vector of health decrements associated with each condition, which depend on medical treatments.

The question we ask is how welfare is affected by changes in medical treatments. Consider treatment changes for a single condition, c. For notational simplicity, consider a situation where the amount paid for each service does not change, so that treatment changes are synonymous with changes in condition spending.[9] Changes in m_c affect welfare in up to three ways. The first is the direct monetary effect of changes in spending. The lifetime spending impact of treatment changes is $-\sum_{t=0}^{\infty}(1+\rho)^{-t}\ S_t(a_0)\ prev_{ct}\Delta m_c$, the discounted sum of spending increases among people who are alive at each age.

Second, treatment changes may affect survival or quality of life. $\left(\frac{dS_t(a)}{d\gamma_c}\right)\left(\frac{d\gamma_c}{dm_c}\right) prev_{ct}\Delta m_c$

[8] By adding across conditions, we are assuming that mortality rates are independent; i.e., the mortality rate from cancer does not depend on whether the person also has heart disease. Our data are not rich enough to allow us to estimate interaction effects in mortality rates.

[9] If prices were changing as well, we would need to distinguish the effect of quantity changes from the effect of price changes.

is the impact of medical care on survival to age a. The welfare value of this is found by multiplying the survival gain by $\left(\frac{U_t}{U'} + (Y_t - x_t - m_{ave,t})\right)$. The first term is flow utility in year t and the second is the increment to future consumption. The welfare impact of treatment changes through their impact on quality of life is given by $- S_t \frac{U_t}{H_t U'} \left(\frac{dh_c}{dm_c}\right) prev_{ct} \Delta m_c$ in year t. The change in health due to medical care $\left(\frac{dh_c}{dm_c}\right)$ is multiplied by the utility value of consumption in that year.

Finally, changes in medical care for some conditions might affect the prevalence of other conditions. We term such conditions risk factors, as opposed to 'direct conditions', which have an immediate effect on health. In our empirical analysis, a condition can be both a risk factor and a direct condition, as with diabetes. Consider a risk factor c that influences the prevalence of direct condition k. The welfare impact in year t of treatment changes for c is given by $\left\{\left(\frac{dS_t}{dprev_k}\right)\left(\frac{U_t}{U'} + (Y_t - x_t - m_{ave,t})\right) - S_t(a_0) \frac{U_t}{H_t U'} h_{kat}\right\}\left(\frac{dprev_k}{dm_c}\right)\Delta m_c$.

Combining these three effects and adding over the lifetime, the dollar value of the total change in welfare is given by:

$$\begin{aligned}\frac{\Delta W(a_0)}{\mu} = &\sum_{t=0}^{\infty}(1+\rho)^{-t}\left\{\left(\frac{dS_t}{d\gamma_c}\right)\left(\frac{d\gamma_c}{dm_c}\right)\left(\frac{U_t}{U_t'} + (Y_t - x_t - m_{ave,t})\right) - S_t(a)\frac{U_t}{H_t U_t'}\left(\frac{dh_c}{dm_c}\right)\right\} prev_{ct}\Delta m_c \\ &+ \sum_{t=0}^{\infty}(1+\rho)^{-t}\left\{\left(\frac{dS_t}{dprev_k}\right)\left(\frac{U_t}{U_t'} + (Y_t - x_t - m_{ave,t})\right) - S_t(a)\frac{U_t}{H_t U_t'} h_{kat}\right\}\left(\frac{dprev_k}{dm_c}\right)\Delta m_c \qquad (4) \\ &- \sum_{t=0}^{\infty}(1+\rho)^{-t} S_t(a)\, prev_{ct}\Delta m_c\end{aligned}$$

The first row is the direct effect on health, the second row is the impact on health through risk factors, and the third row is the loss from increased spending.

Equation (4) is the multi-factor productivity (MFP) of medical treatment changes, or what Abraham and Mackie (2006) term 'social profit'. MFP can be positive or negative. We express MFP in dollars and as a growth rate, dividing by the value of life in the base period, $\left(\frac{W(a_0)}{\mu}\right)$.

An alternative metric of treatment impacts is the quality-adjusted price of medical care. The price of providing medical care in the base period is the present value of lifetime medical spending: $P_0 = \sum_{t=0}^{\infty}(1+\rho)^{-t} S_t(a_0)\, m_{ave,t}$. Over time, this rises to P_1 as treatments change. The

quality-adjusted price change is the unadjusted price change less the dollar value of the improvement in health, $(P_1 - \Delta\tilde{V})/P_0$, where $\Delta\tilde{V}$ is the value of health in period 1 prices.[10]

Equation (4) is written from the perspective of the average person in society (a 'per capita' measure). An alternative is to estimate productivity for a person with condition c (a 'per case' measure). The per capita and per case measures are approximately related by the prevalence rate.[11] In evaluating each condition individually, we present the per case metric. That is, we ask the question: how is the welfare of people with condition c affected by treatment changes for that condition, assuming the person had the condition in the first and all subsequent years? We do this so that we can more easily compare across conditions. In contrast, when we evaluate medical care as a whole, we consider treatment changes per capita: what has been the impact on social welfare of all the medical care changes that have occurred?

To estimate productivity, we measure each term of equation (4) empirically. The primary difficulty in doing so is isolating the impact of treatment changes from other factors influencing health and medical spending. We detail our approach to this issue after first describing the data we employ.

II. Data and Population Trends

The data needs in constructing a satellite health account are substantial, including information on overall medical spending matched to disease prevalence, risk factors, and other socioeconomic information. Our primary dataset is the Medicare Current Beneficiary Survey (MCBS), an ongoing survey of the Medicare population. We describe the data briefly here and in more detail in Appendix A.

Medicare is nearly universal in the elderly population and covers disabled people below age 65. To ensure a representative sample of the population, we utilize only the population age 65 and over, roughly 13,000 people per year. MCBS has data on all medical services – both Medicare and non-Medicare. Spending on the elderly is about one-third of national medical spending.

We modify MCBS data in two ways. First, claims data, and hence costs of care incurred, are only available for beneficiaries in traditional Medicare (i.e., not in Medicare Advantage HMO

[10] See Sheiner and Malinoskaya (2018). Note that nothing constrains prices in period 1 to be positive. A negative price is equivalent to an innovation sufficiently valuable that people would have to be paid in period 1 to take the treatment in period 0, given that period 1 treatment is available.

[11] They are not exactly the same because survival is non-linear in prevalence.

plans). To account for this, we reweight the population in traditional Medicare to reflect total Medicare enrollment, adjusting for the different demographics and health status of the two groups. Second, the MCBS spending totals differ slightly from national estimates of medical spending in the elderly (by roughly 10% on average; CMS, 2018). We adjust MCBS spending by type of care to reconcile the MCBS to national totals.

Figure 1 shows per capita spending for elderly Medicare beneficiaries, in real $2010 dollars (all spending data in this paper are in real dollars). Spending rose from $13,096 per capita in 1999 to $17,736 per capita in 2012, an increase of $4,700 or 2% annually. There is a slowdown in spending growth around 2005, which we return to below (see also Cutler et al., 2019).

MCBS has a variety of sociodemographic and behavioral controls, shown in Table 1. This includes demographics (age, gender, and race/ethnicity), socioeconomic information (poverty status, the highest grade of education completed, marital status, and service in the armed forces), behavioral risk factors (smoking and Body Mass Index [BMI]), self-reported health (one year prior and relative to others of the same age), and difficulty with everyday activities such as eating, shopping, walking, preparing meals, and doing light housework.

A. Population Health

Our primary measure of health is quality-adjusted life expectancy or QALE. Using the notation from section IA above, QALE at age a_0 is defined as:[12]

$$QALE(a_0) = \sum_{t=0}^{T} S_t(a_0) \cdot H_t(a) \, . \qquad (5)$$

Consistent with our focus on the elderly population, we take a_0 to be age 65.

Actual QALE will depend on a cohort's realized survival and quality of life. Following conventional methods, we measure health using period methods. We assume that the survival rate and quality of life for a person as they age will equal average survival rates and quality of life for people of that age in the base year.

Survival rates are straightforward to measure, as mortality is tracked in national data. Figure 2(a) shows the age-adjusted mortality rate in the elderly population from 1999 to 2012. Mortality fell by 23 percent between 1999 and 2012.

[12] QALE is not discounted in this equation, while welfare discounts health. That is largely a semantic distinction. QALE is a physical concept – years of quality-adjusted life. Such a unit is better expressed without discounting. The value of those years, in contrast, is discounted.

There is no universally accepted metric of quality of life. We thus develop a metric using available data. We measure quality of life based on a person's ability to engage in essential aspects of living, such as caring for oneself or engaging in social activities. We term these health indicators 'symptoms and impairments' and denote the vector of prevalence rates as $\mathbf{si_t}$. These may change over time as medical care improves or a health condition worsens. $\mathbf{h_{si}}$ is the corresponding vector of health decrements. In principle, $\mathbf{h_{si}}$ could also vary over time; in practice, we do not have the data to allow for this. Quality of life is thus defined as $H_t(a) = \bar{h}_t(a) - \mathbf{si_t}(a) \cdot \mathbf{h_{si}}$.

Our choice of symptoms and impairments, along with the health decrement weights, are determined by data availability. The symptoms and impairments from MCBS we analyze are those in the upper part of Table 1. A first metric is functional limitations, including difficulty walking, stooping/crouching/kneeling, lifting or carrying heavy objects, reaching or extending arms, and writing or grasping small objects. People are also asked whether they have impairments in any of 6 Activities of Daily Living (ADLs, including dressing, bathing, or toileting) or Instrumental Activities of Daily Living (IADLs, such as managing money, shopping, or doing light housework). In addition, people are asked whether they have difficulty seeing or hearing. Finally, people are asked whether health limits their social interactions. Appendix Figure A3 shows the trend in each of these measures of health over time. While this is a rich set of metrics, it is worth noting that several domains of health are not asked about explicitly, including pain, cognition, and mental health.

To estimate the $\mathbf{h_{si}}$ vector of health decrements for each symptom or impairment, we use data from the 2002 Medical Expenditure Panel Study (MEPS). This survey asks people questions about their health similar to the MCBS and also included an overall health assessment on a scale from 0 ('worse imaginable health state') to 100 ('best imaginable health state').[13] To form the disutility weight associated with each impairment, we regress the scale response on the symptoms and impairments.[14] The results of the regression are shown in Appendix Table A5. The constant term, an estimate of $\bar{h}$, is .85, and is statistically different from 1. The deviation from 1 indicates

[13] The specific question is from the EQ-5D: "To help people say how good or bad a health state is, we have drawn a scale (rather like a thermometer) on which the best state you can imagine is marked by 100 and the worst state you can imagine is marked by 0. We would like you to indicate on this scale how good or bad your own health is today, in your opinion. Please do this by drawing a line from the circle below to whichever point on the scale indicates how good or bad your current health state is." https://euroqol.org/eq-5d-instruments/

[14] We do not include demographic variables such as age because we do not want them to pick up the effect of symptoms and impairments. That said, the estimates associated with the symptoms and impairments do not differ greatly if demographics are included.

unmeasured symptoms and impairments and possibly health decrements not due to articulated problems, such as unmeasured frailty. All of the coefficients for the impairments are negative and statistically significant. The largest decrements are for health interfering with social interactions (-.16 for moderate impairment and -.20 for severe impairment) and for severe walking difficulty (-.12).

We use these coefficients to calculate quality of life for each person in the MCBS. The lowest imputed quality of life is 0.26; the average is 0.69. Figure 2(b) shows the trend in quality of life between 1999 and 2012. Quality of life was generally unchanged with a very modest increase. Put another way; survival has increased without creating a larger impaired population.

The combination of falling mortality and relatively constant quality of life implies that quality-adjusted life expectancy increased. This is shown in Figure 2(c), which plots QALE at age 65 over time. In 1999, the typical person aged 65 could expect to live 12.2 quality-adjusted years. In 2012, the value was 13.2 years, a 1-year increase in just over a decade.

B. Identifying Conditions

There is an enormous number of medical conditions; over 68,000 are noted in the latest International Classification of Diseases (ICD, version 10). This is too many for our analysis. The Agency for Health Care Research and Quality has consolidated this list into 285 Clinical Conditions (termed CCS; Elixhauser et al., 2014). However, even this condensed list is too granular for our purposes. For example, over 20 codes are for pregnancy, and many others are for childhood conditions. In consultation with clinical experts, we reclassified the AHRQ summaries into 80 conditions (Raghunathan et al., 2020), shown in Appendix Table A6.[15]

One way to characterize these 80 conditions is by chapters of the ICD manual: circulatory disease, respiratory disease, and so on. There are 19 chapters in the 9th edition of the ICD. We delineate highly prevalent conditions within each chapter and generally create an 'other' category for the remainder. For example, our cardiovascular disease conditions include ischemic heart disease, congestive heart failure, stroke and cerebrovascular disease, and other heart and vascular disease – the latter of which is the residual. In some cases, there are no apparent subsets to pull out, and the entire chapter is left as one condition (as with gastrointestinal and liver disease). Some

[15] These conditions largely combine CCS categories, though some CCS categories are disaggregated given relatively high prevalence and differing disease courses in the elderly.

catchall categories are a bit broader. Two of our conditions are 'aftercare' (general sequelae to a hospitalization, for example physician visits for medication monitoring after surgery), and 'signs and symptoms' (adverse medical events of unknown origin, such as fainting or nausea).

In addition to clinical conditions, we create conditions for moderate and severe frailty. We think of these as akin to 'old age'. The frailty measures are based on self-reports of difficulty or inability to do 5 activities: stooping/kneeling, lifting 10 lbs, reaching over one's head, writing, and walking 2-3 blocks. People with a lot of difficulty or an inability to do activities in 1-3 areas are classified as moderately frail, and people with difficulty in 4-5 areas are severely frail.

As noted in section I, an important classification of conditions is by whether they directly influence health (direct conditions) or mostly affect the prevalence of other conditions (risk factors). Seventy-one of the conditions are posited to have a direct effect on mortality or quality of life. Five of the conditions are exclusively risk factors. These include hypertension, high cholesterol, eye and ear disorders, and infectious disease immunization (flu and pneumococcal vaccine). Two conditions, diabetes and mood disorders, are both direct conditions and risk factors for other diseases. Finally, there are four cancer screening variables: breast, colorectal, prostate, and cervical cancer. Conceptually, these are closely related to risk factors, in that they influence health by affecting direct conditions. The difference is in the nature of the health impact: cancer screening does not reduce the incidence of cancer and may increase it (Welch and Browley, 2018), while the treatment of hypertension reduces the risk of heart disease.[16] Thus, the empirical treatment of screening and risk factors is different. Figure 3 shows the direct conditions that are matched to each risk factor condition.

Most of the conditions that we measure have long-term impacts on health and medical spending, for example heart disease or kidney failure.[17] Thus, our measure of prevalence is whether the person ever had the condition. In principle, one might want to distinguish cases that are entirely chronic from those where an event occurred in the indicated year. However, the annual incidence of many conditions is too small to permit this.

To measure the prevalence of each condition, we start with diagnoses in the claims data in MCBS for all conditions except frailty, which is based on self-reports. Claims data are known to

[16] Colon cancer screening through colonoscopy is a hybrid; it both detects cancer earlier and reduces the development of colon cancer, for example by removing polyps.

[17] Exceptions include conditions with an explicitly temporary impact, such as acute infectious disease and cancer screening.

underreport prevalence, especially for less severe conditions. For example, 23% of people in the MCBS have a claim for arthritis in a year, but nearly half of the aged population in the National Health and Nutrition Examination Survey (NHANES) reports having been diagnosed with arthritis. To estimate true prevalence, we impute 'calibrated conditions' using self-reports of ever having had a condition in NHANES.

The procedure is described in detail in Appendix A.4. When a disease that has a self-report question in NHANES matches one of our categories (e.g., ischemic heart disease), we impute prevalence in the MCBS to match the prevalence in NHANES, using demographic and health measures to match individuals across surveys. When there is not an exact self-report, we impute prevalence using conditions that have a similar overall prevalence rate to the relevant MCBS condition. In each case, we form five multiply-imputed datasets, based on imputing different people to have each condition.

One important addition that comes from using the NHANES is that we can impute undiagnosed hypertension, high cholesterol, and diabetes, which are included in the NHANES from measurement and blood draws. As we show below, the treatment of cardiovascular risk factors is a significant part of health and spending changes for the elderly, so this imputation is quite valuable. Since the NHANES data are from 1999 on, our analyses are limited to 1999–2012.

For presentation purposes, we generally aggregate the 80 conditions into 30 groups of relatively similar conditions. For example, acute myocardial infarction (heart attack) and coronary atherosclerosis (build-up of plaque in the coronary arteries) are grouped into ischemic heart disease. Within the 30 categories, trends tend to be relatively similar for the individual conditions. Appendix B shows all of our results for the 80 conditions.

The list of aggregated conditions, along with prevalence in 2012, is shown in Figure 4. Cardiovascular risk factors are the most common condition; 86% of the elderly population has one or more. Other common conditions include musculoskeletal disorders (e.g., arthritis and back pain), eye and ear problems, respiratory problems, and uncategorized signs and symptoms. The presence of multiple chronic conditions is very common. The average elderly person has 13 conditions, and 95 percent of people have three or more conditions.

The prevalence of most conditions is rising over time (Appendix Table A6). The largest increases are for cardiovascular risk factors, chronic renal failure, and musculoskeletal issues, all of which are partly a consequence of obesity. We return to the role of obesity in section IV. The

major exception to the trend of rising prevalence is frailty. Moderate frailty was constant over time, and severe frailty fell.

III. Methodology for Estimating Spending, Survival, and Health Decrements by Condition

Our primary empirical challenge is to determine the effect of changes in treatment for each condition on medical spending for that condition and health outcomes. As with any production function, we make several assumptions to allow us to identify the impact of a particular input. A first assumption was noted above: disease prevalence is determined by factors outside of the medical system. The only exceptions to this are clinically identified risk factors, shown in Figure 3.

Even with that assumption, there are still empirical issues to overcome. The primary concern is comorbidities. When people have multiple conditions, it is difficult to know which condition to attribute spending and health outcomes to and thus which conditions are more and less productive over time. There are two general approaches in the literature to addressing comorbidities (Ghosh et al., 2020a). The first approach relies on physician identification. Most medical claims have a diagnosis, coded by the treating physician. These diagnoses can be used to distribute the dollars associated with an event to conditions. For example, some analyses attribute all spending to the primary diagnosis, while others divide among primary and secondary diagnoses based on some type of rule. Similarly, the physician or coroner certifying death indicates an underlying cause, which can be used to classify the cause of death. There is a long history of researchers using physician attribution to measure condition-specific spending and health outcomes.[18]

However, the physician attribution methodology has several drawbacks. As noted above, not all diseases are coded on claims records. Thus, spending cannot be parceled out to all conditions. Death certificates are subject to similar error. Because heart stoppage is the definition of death, heart disease is overcoded on death records relative to other causes (Messite and Stellman, 1996). Further, there is no obvious methodology for allocating quality of life to conditions, since physicians do not rate quality of life decrements for patients. Finally, the

[18] See, for example, see Rice (1967) and Thorpe et al. (2004) for models of spending. The BEA uses this methodology in their satellite spending accounts (Dunn et al., 2015; Dauda et al., 2019).

physician attribution approach cannot account for non-medical contributors to outcomes. If people are more obese and thus cost more when they have a disease, the physician attribution approach has no way to distinguish this from the provision of less valuable medical care.

We thus pursue an alternative approach to condition attribution. We relate medical spending, survival, and quality of life in a year to all the conditions the person had that year along with demographic and behavioral characteristics. Comparing people with a condition to otherwise similar people without the condition gives an estimate of the spending and health outcomes that result from that condition. We estimate these models each year and calculate productivity over time as the dollar value of the change in health less the increased spending.

The specific methodology we employ is described briefly here, and in more detail in Appendix B. Before presenting details, it is worth noting that we do not have random variation in treatments with which to measure the impact of medical treatments. As a result, our evaluation of health and spending changes is a form of 'residual analysis' – the difference for people with and without each condition after controlling for other factors that influence health. To validate our most important results, we show below that our attribution of the sources of health improvement matches simulation models in the clinical literature where such models exist.

Step 1: The first step is to estimate the effect of each condition on spending and the impact of the direct conditions (i.e. conditions that are not exclusively risk factors) on mortality and quality of life. To do this, we use a propensity score model.[19] A separate logistic regression for each condition and year is used to predict whether a person has each condition. The covariates include the demographic and behavioral risk factor variables shown in Table 1 along with dummies for the other 79 conditions.[20] The results of the propensity score models are used to stratify the population into quintiles based on the probability of having the indicated condition, or quartiles/terciles for conditions with relatively low prevalence. The difference in the within-strata outcomes of people with and without the relevant condition, averaged across strata, is the first stage estimate. We apply this methodology to each of the five multiply-imputed data sets and

[19] Trogdon et al. (2008) do this in a regression approach. The methodology we propose does somewhat better in out-of-sample prediction than the regression method; see Ghosh et al. (2020a).

[20] We exclude those conditions that have a deterministic or extremely tight correlation with the condition of the interest. For example, undiagnosed hypertension is not included in the model for diagnosed hypertension, and gender is not included in the model for prostate cancer. Appendix Table B1 provides more detail.

average the resulting condition-specific outcomes across replicates.

Table 2 shows the cohort balance tables for coronary atherosclerosis and acute myocardial infarction in 1999 as an example. The overall F-statistic fails to reject equality of the means between cases and controls, and only 2 of the 40 comparisons are statistically different. This is representative of all the conditions in all of the years.

Step 2. Adjusting for national totals and outliers. Because each condition model is estimated separately, the prevalence-weighted average of the spending and mortality estimates in step 1 will not necessarily add to population average spending and recorded deaths.[21] There is an additional concern for spending: summing condition-specific costs at the personal level does not predict anywhere near the number of high spenders that are in the actual data.[22] High spenders generally do not have different conditions than low spenders. Instead, they have relatively common conditions that become expensive.

We address each of these issues in a second stage. For mortality, we adjust the mortality rate of each condition proportionately so that the aggregate mortality rate in each of three age groups (65-74, 75-84, 85+) equals the actual mortality rate in that year. In general, the adjustment is about 20% (Table B4). For medical spending, the procedure is somewhat more complex, since we also want to reflect the spending of high spenders. We do so with a regression analysis. We relate the actual cost for each person to the predicted cost based on summing the first stage estimates to the individual level, interacted with a series of variables capturing high use of services ($\mathbf{X_{it}}$): $\text{ActualCost}_{it} = \text{PredictedCost}_{it} \cdot (\mathbf{X_{it}}\boldsymbol{\alpha_t}) + \varepsilon_{it}$.

$\mathbf{X_{it}}$ includes the number of conditions, that number squared, an indicator for any hospitalization, the number of nights in the hospital, the number of hospital admissions, the number of days institutionalized, survival in the indicated year, the number of months survived for those who are deceased, and the number of outpatient claims. We average the $\mathbf{X_{it}}\hat{\boldsymbol{\alpha}}_{\mathbf{t}}$ estimates across people with each condition to form a condition-specific correction factor. This is then multiplied by the first stage cost estimates to estimate spending for each condition.

This adjustment is estimated for each year separately, to allow the impact of high utilization

[21] The quality of life adjustment estimates do not use such an adjustment because there is no national quality of life to adjust to.

[22] In the data, the ratio of the 90th/10th percentile of spending is 30. Using predicted condition-specific spending, in contrast, the ratio is only 5. See Appendix Figure B2.

to vary over time. Appendix Table B3 shows the estimates of the equation for the year 2009. The **X** variables lead to increased spending in the expected direction; for example, people who spend more time in a hospital or an institution cost more than people who do not.

Step 3. Smoothing over time. The estimates from step 2 are subject to noise as the sample sizes of people with each condition in each year vary and are not always large. To account for this, we take the annual estimates of spending, mortality, and quality of life decrement for each condition and fit a second-order polynomial. The fitted values for the beginning and ending years are then used to estimate productivity changes.[23]

Figure 5 shows an example of the spending and health outcomes for coronary atherosclerosis and heart attack. Together, these form the category of ischemic heart disease, which is among the categories where we find the highest productivity growth (as discussed below). The first panel shows the trends in cost per case, along with the 95% confidence interval (discussed below) and the fitted values. While there are some outlier years, the trend is generally smooth. The second and third panels show the same for quality of life and mortality.

Step 4. Outcomes for Risk Factors and Screening. Our final step is to estimate outputs for risk factor conditions and screening. To determine the impact of treating risk factor conditions on downstream conditions, we estimate regression models for each of the final conditions for which there is an identified risk factor condition. The prevalence of the direct condition (i.e., ischemic heart disease) is related to the prevalence of the risk factor conditions (i.e., hypertension, high cholesterol, and diabetes), demographics, and health behaviors such as obesity and smoking. The resulting coefficients indicate the impact of each risk factor condition on the direct condition over time. We do this in an early period (1999-2001) and a later period (2010-12). Multiplying these impacts by the risk factor condition prevalence gives the share of the total population that has the direct condition as a result of having the risk factor condition. For example, 57% of people in 1999-2001 had hypertension, and people with hypertension were 7.5% more likely to have ischemic heart disease in those years. Thus, 4.3% (= 57% x 7.5%) of the population had ischemic heart disease as a result of hypertension. Where possible, we compare our estimates of the impact

[23] Analyses using the average of data in the first three years of the sample and the last three years of the sample provide similar results, as shown in Appendix Figure B4.

of each risk factor condition on the direct condition to estimates in the clinical literature; Appendix Table B7 shows that they match up reasonably well.

To reflect this impact in the health accounts, we increase the total spending, mortality, and quality of life decrement associated with the risk factor condition by the indicated amount. To continue the example above, the effect of hypertension is incremented by 4.3% of the health and spending impacts estimated to be due to ischemic heart disease.[24] This is offset in the ischemic heart disease condition accounting by reducing the prevalence of heart disease; some of those cases have been moved to other conditions.

Appendix Table B6 shows the dollars and deaths that are reallocated from direct conditions to risk factors. In our heart disease example, one-third of ischemic heart disease cases in 1999-2001 are estimated to be due to risk factors.

To account for cancer screening, the ideal data would disaggregate cancers by the stage of detection; greater screening should reduce late-stage cancer and increase early-stage cancer. However, our data do not contain cancer stage.[25] In the absence of better data, we take the mortality impacts of screening from the clinical literature. As shown in Appendix Table B7, the clinical literature shows clear benefits of breast and colorectal screening for reducing mortality, but unclear benefits of cervical and prostate cancer screening, especially in the elderly. We thus model only mortality benefits from breast and colorectal cancer screening. We assume no impact of cancer screening on quality of life, since quality of life likely varies with any treatment, and treated cases rise with screening. For these reasons, we note that the productivity of cancer screening is subject to more uncertainty than are other conditions.

Step 5: The Impact of Smoking and Obesity. To estimate the contribution of obesity and smoking to spending growth and health changes, we use a methodology similar to that for clinical risk factors. We first estimate the impact of BMI and smoking on the prevalence of conditions that are listed in clinical reviews as resulting from these risk factors.[26] Using data on changes in obesity and smoking in our sample, we then simulate the impact on the prevalence of direct conditions.

[24] This assumes that the cases that result from hypertension have similar clinical outcomes to those that do not.
[25] An alternative would be to estimate the impact of cancer screening on mortality by interacting the presence of cancer in the mortality equation with whether the person was screened. A screened person would be expected to have lower mortality given a cancer diagnosis. Because many of the cancers we examine are reasonably rare, our data do not permit such an interaction.
[26] Appendix Table B9 compares the values we estimate to those in the clinical literature.

The overall smoking change in the elderly was small in this period (see Appendix Table B10). While current smoking declined slightly, former smoking increased, especially at the oldest ages. Thus, the net impact of changes in smoking was minor. Obesity increased rapidly among the elderly in all age sub-groups.

Productivity estimates. In estimating productivity, two additional parameters are needed: the value of a year of life $V = \left(\frac{U_t}{U_t'} + (Y_t - x_t - m_{ave,t})\right)$, and the discount rate (ρ). Our baseline assumption is that the value for a year of life in good health equal to $100,000 (Viscusi and Aldy, 2003). We conduct sensitivity analyses assuming the value is alternatively half and twice that amount. We use a baseline discount rate of 0% and perform sensitivity to a rate of 3% (Gold et al., 1996).

Standard Errors. Our estimates of the spending and health impacts of each condition in each year are a combination of several different estimates. To calculate standard errors for them and the resulting productivity estimates, we use a bootstrap technique. We draw 1,000 simulated MCBS samples, taken from the actual MCBS population with replacement. We do this separately for each multiply-imputed dataset. For each draw, we estimate all four steps above. We use this to calculate the empirical 95% confidence interval for the spending, health, and productivity impacts of each condition.

IV. Change in Medical Spending by Condition

In this section, we present our estimates of spending by condition. The first two columns of Table 3 show spending per case for the 30 aggregated conditions in 1999 and 2012 (all in 2010 dollars). Appendix Table B5 shows direct spending for each of the 80 conditions; Appendix Table B6 shows the transfers from direct conditions to risk factor conditions.

Almost all conditions are associated with higher spending. The major outlier to this is cancer screening, where the estimates suggest slightly lower spending for people who are screened relative to those who are not. It may be true that cancer screening saves money, though it is also possible that elderly people who are screened for cancer are healthier in ways that our model does not capture. In some studies (Trogdon, 2008), condition-specific spending is constrained to be non-

negative. That may not be necessary for our application, however, as we are interested in the change in spending at the condition level. If the bias from unmeasured health status is constant over time, the *change* in spending will be accurate even if the *level* of spending is not. Thus, we do not constrain spending to be positive. That said, caution is warranted about the estimated cost changes for cancer screening.

The conditions for which per case spending is highest are generally the ones we would expect: heart disease, certain cancers (especially lung cancer), dementia, kidney failure, and frailty. The case-specific cost estimates may seem low, but it is important to recall that these are the cost per prevalent case of disease, not per incident case. For example, spending on heart attacks is the average among everyone who ever had a heart attack, not new heart attack cases. Because past cases generally add a smaller increment to spending than new events, our estimates are smaller than are typically found in studies looking at acute disease onset.

Spending for most conditions rose over time. The largest increases were for lung cancer (a $2,500 per case increase) and arthritis and musculoskeletal problems (an $800 per case increase). Notably, the conditions that cost the most at a point in time are not always the ones for which spending rose the most. For example, spending for ischemic heart disease rose only modestly over the time period, despite starting at a high level.

This fact has important implications for the understanding of medical spending increases in this time period. Figure 6 shows the change in per capita spending—per case spending times the prevalence rate—accounted for by each of our 30 aggregated conditions. Acute cardiovascular diseases together account for 17% of spending in 1999, but only 4% of the spending increase. The largest contributors to increased spending are cardiovascular risk factor conditions, musculoskeletal diseases (which includes arthritis and back pain), and aftercare. Together, these three conditions account for 40% of the increase in spending.

Spending per capita can increase because the prevalence of disease increases or because each case costs more. The importance of each of these is shown by disease in Appendix Figure B5 and overall in the first column of Table 4. Sixty percent of the increase in spending is due to the greater cost per case, and one-third is due to greater prevalence (the residual is the covariance term). As noted above, prevalence changes are particularly significant for cardiovascular risk factors, renal failure, and musculoskeletal problems. These three conditions are also among the group most likely to be affected by obesity.

Table 4 shows the simulated impact of smoking and obesity changes on medical spending. Increased obesity led to an increase in spending of $342 per person while smoking declines led to a reduction of $64 per capita. There is also a sizeable residual increase in the prevalence of many conditions not associated with smoking or obesity changes, $1,382, shown in the last row of the table.[27]

V. Change in Health by Condition

In this section, we decompose changes in health into the same conditions for which we estimate the change in spending. Estimates of mortality and quality of life impact for our 30 aggregated conditions are shown in the third–sixth columns of Table 3; results for all conditions are in Appendix Tables B8a and B8b. The direct conditions are almost all associated with higher mortality and lower quality of life than not having the condition. Further, the relative magnitude makes sense; people with lung cancer have an annual mortality rate 15% greater than comparable controls (this includes past as well as current cases of lung cancer); acute kidney failure adds a 3% increment to mortality; a person with both would have an 18% higher mortality rate.

The mortality estimates are worthy of some note, as our method of assessing cause of death differs from Vital Statistics. The Vital Statistics estimates are based on a single cause judged most important by a physician or coroner. Our estimates, in contrast, allow each of the conditions a patient has to contribute to death.

Our methodology attributes far fewer deaths to heart disease than do Vital Statistics (see Ghosh et al., 2020b, for details). For example, Vital Statistics data indicate a death rate from ischemic heart disease of nearly 1,400 per 100,000; our estimates are half as large. The deaths that we do not code as heart disease are spread across several conditions: kidney failure and blood disorders, for example. Frailty is also a major cause of death, which is not recorded in the official data at all. Effectively, our data say that systemic failure is a cause of death more than heart disease by itself.

The mortality and quality of life data allow us to estimate QALE for people with each condition in 1999 and 2012, and thus the change in QALE by condition. These results are shown in Figure 7. The largest increase in QALE was for cardiovascular diseases—particularly ischemic

[27] It is possible that the impact of obesity on spending is understated, as we only have data on current weight and conditions are likely to be affected weight in the past in addition to current weight (Preston, et al., 2018).

heart disease but including other cardiovascular diseases as well. People with cardiovascular disease in 2012 experienced an increase in QALE of 1 year over similar people in 1999. Many other conditions had increases in QALE that were somewhat smaller than this but still substantial. For example, people with lung cancer, colorectal cancer, and acute kidney failure all experienced QALE increases of around 0.4 years. Frail people lived an additional 0.3 quality-adjusted years over the time period.

Table 4 shows the combined impact of medical treatments and other factors on QALE at age 65. The overall increase in QALE noted above (1.0 years) is the net effect of two offsetting trends. Holding prevalence constant, QALE would have increased by 1.7 years. However, the rising prevalence of conditions reduced QALE by -.9 years. Some of this reduction in QALE was a result of increasing obesity (-.13 years). This effect, while only a portion of the prevalence increase, is quantitatively very large. For example, the impact of obesity on QALE is roughly equivalent to the entire QALE increase from improved care of musculoskeletal conditions and mental health conditions combined.[28]

VI. The Productivity of Medical Care

Having estimated the cost and health consequences of changes in medical treatments, this section brings the two together to measure productivity. Figure 8 and Table 5 show the calculation of productivity by condition and for the medical system as a whole. The first column in Table 5 shows the increase in the present value of lifetime costs, the capitalized version of the changes in Table 3. As with our estimates of quality of life changes, we assume that future spending on each condition as a person ages is equal to current average spending on that condition at older ages.

The second column is the increase in QALE. The remaining columns show productivity estimates for different values of V (the value of a year of life) and ρ (the discount rate). Standard errors are presented in Table 5 for the productivity estimate for medical care as a whole and in Appendix Table B12 for each of the conditions.

Several results are apparent. The overall estimate of medical care productivity, shown in the first row, is positive. On net, medical care increased in value by $113,000. This estimate is statistically significantly different from zero. The value of health improvement is roughly

[28] To make these comparisons, we convert the QALE increase per case for these conditions into the QALE increase per capita. This is done by multiplying by the prevalence rate.

$170,000, offset by an increase in spending of $84,000. The value of remaining life at age 65 was about $1.22 million in 1999 (12.2 years x $100,000 / year). Thus, medical care increased in productivity by 9% over the time period ($113,000 / $1.22 million), or 0.7% per year.

The high overall productivity estimate is a mix of heterogeneous estimates for different conditions, shown visually in Figure 8. The most productive part of medical care is treatment for cardiovascular disease, both acute conditions and risk factors. Productivity estimates for acute cardiovascular diseases are $89,000 in aggregate—79% of the total increase. Other conditions have more modest but still large increases in value-added. For example, the net benefit from treatment changes is $38,000 for people with colorectal cancer, $30,000 for people with acute kidney failure, and $22,000 for frail people. Very few conditions have negative productivity growth, but many are near 0. For example, there was little productivity change for people with mental illness ($4,000 on net) or arthritis and musculoskeletal conditions (-$3,000 on net), even though treatment changed rapidly for both conditions. The declining productivity of treating musculoskeletal conditions likely reflects the increased use of opioids over this period. These drugs are expensive and do not appear to be very effective in reducing pain (Phillips, Ford, and Bonnie, 2017). The poor productivity of mental illness is more difficult to understand. Despite a vast increase in the number of people treated with drugs for mental illness, the population's mental health showed essentially no change over time.

Columns 4 and 5 of Table 5 show the impact of varying the value of a life-year, to $50,000 in column 4 and $200,000 in column 5. These changes affect the net value of medical care as one would expect. Medical care as a whole has a value of about $25,000 at the lower value, with a 95% confidence interval that includes 0. The higher value of life raises the net benefit to just above $280,000.

The final column of Table 5 shows the impact of using a 3% discount rate for both cost and utility discounting. The results are in the same pattern, though the overall magnitude is somewhat smaller. The net benefit of medical care is $50,000 per person.

We can alternatively express the impact of treatment changes as the change in quality-adjusted prices. Real spending per person net of prevalence increases grew by 1.6% annually. However, taking account of the value of quality-adjusted life, the price index fell by 2.4% annually.

A. Cardiovascular disease

Because cardiovascular disease is so prominent as a cause of death, researchers have developed clinical models of cardiovascular disease mortality reductions to which we can compare our results. The most commonly referenced heart disease model is the IMPACT model (Ford et al., 2007; Ogata et al., 2019). The IMPACT model sorts people into cells based on the presence or absence of heart disease and uses clinical trial evidence on the impact of different treatments to simulate how mortality changed over time with treatment changes. For example, one cell of the model estimates the differential mortality rate for people with prior heart disease who are and are not taking antihypertensive medications. Using this coefficient and the share of people taking anti-hypertensive medications over time, the model estimates the change in mortality that is predicted to have occurred because of the expanded treatment of hypertension. We use the parameters of the IMPACT model and apply them to treatment changes in the elderly population.[29]

In addition to looking at heart disease, we replicated the structure of the IMPACT model for congestive heart failure and cerebrovascular disease. In some cases, the studies that provide the impact of treatments on mortality are the same as those for ischemic heart disease. In other cases, we find comparable meta-analyses (see Cutler et al., 2019).

Table 6 shows the comparison of our results to the IMPACT model. The first column shows national totals for cardiovascular disease mortality. Relative to a constant age and sex-specific death rate over time, cardiovascular disease mortality in the elderly declined by 300,000 deaths between 1999 and 2012. The results of the IMPACT model are shown in the second column. The model suggests there should be 219,000 fewer deaths resulting from medical treatments over the time period we examine.

The estimates from our models, presented in the third column, are that medical care accounts for 233,000 fewer deaths over the time period. This is within 17% of the IMPACT model estimates. One would expect our results to be somewhat greater than the IMPACT model since we allow for all effects of cardiovascular risk factors, which include non-cardiovascular complications such as kidney disease. Thus, the estimates are very close.

In addition to matching model totals, our results add richness to them. In the national data, very few deaths are attributable to cardiovascular risk factors, and in the IMPACT model, deaths due to poor risk factor management are not separated. Our data shows that much of the

[29] The impact of treatments on outcomes is generally not estimated differently for the elderly and non-elderly populations.

improvement in cardiovascular disease mortality comes from better control of risk factor conditions. That is, the prevention industry is critical alongside the treatment industry.

B. Cancer

We can also compare our results for certain cancers to those in the clinical literature. Table 5 shows large productivity gains in lung, colorectal, and 'other' cancers (most common skin cancer, benign neoplasms, and cervical cancer). The quality-adjusted life expectancy for each of these conditions increased about 0.4 years. We compare our results to the clinical literature on lung and colorectal cancer.

Survival data for people with cancer are collected by the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) program.[30] Case survival is generally equivalent to the mortality rate in our results. The trend in mortality that we attribute to lung cancer matches well with national data. SEER data indicate that mortality per diagnosed case fell 10% between 1999 and 2012; our estimate is a reduction of 8%. For colorectal cancer, there is a larger discrepancy: SEER data indicate an 8% reduction in per case mortality, compared to 35% in our data. This may reflect some additional cases of colorectal cancer detected in our data that are not in SEER data, perhaps because of polyp removal being coded as cancer. Thus, our colorectal cancer rates may indicate productivity growth that is too high.

We attribute the survival gains to medical treatment advances; the clinical literature helps us evaluate that as well. The major alternative to treatment advances is that improved survival is a result of detecting smaller cancers that would never become fatal or detecting cancers that would be fatal but doing so earlier in the process, so it appears that people survive with them longer.[31] Bias from increased diagnosis is unlikely to be an issue in lung cancer, where screening was minimal over the time period, and most tumors grow rapidly (Noone et al., 2017). Further, treatments have improved over time. Lung surgery is generally less invasive, and chemotherapy is more common, as are biomarker-targeted therapies. van der Drift et al. (2012) and Gunn et al.

[30] The SEER data give survival in each year after diagnosis for people diagnosed in each calendar year. To calculate the mortality rate in a calendar year, for example 1999, we find mortality rates in the first year among people diagnosed in 1999, in the second year among people diagnosed in 1998, and so on. We weight these mortality rates by the share of people alive in the relevant year who would have been diagnosed in each past year.

[31] This goes by several names, including lead time bias (the treatment has no impact but the cancer is known about for longer), length bias (screening detects more slowly growing cancers so that the average mortality rate among clinically treated cancers falls) and overdiagnosis (the detection of preclinical cancers that would never have become clinically relevant).

(2018) argue that these treatment changes have improved lung cancer survival in the Netherlands and Finland.

Screening for colorectal cancer grew over the time period, therefore diagnosis changes could be an explanation for improved survival. Other evidence suggests that the attribution to treatment is likely accurate, however. Treatment changes for colorectal cancer include greater use of pre-surgical radiation, increased employment of adjuvant chemotherapy, new medications for metastatic disease, and development of new targeted biologic agents. Colorectal cancer models suggest that greater use of therapies known about in 2000 could lead to reduced mortality on the order of what occurred (Vogelaar et al., 2006). Lemmens et al. (2010) and Brouwer et al. (2018) argue that these treatment changes can explain the improvement in case-specific colorectal cancer survival that was observed in the Netherlands. However, there is no definitive evidence as yet for the US.

C. Comparison to Existing Economic Estimates

There are several estimates of health care productivity in the economic literature (Sheiner and Malinskaya, 2018). However, making comparisons between our analysis and other papers is difficult. Conceptually, some studies look at changes in the average cost of a year of life over time (e.g., Romley et al., 2015, 2020), in contrast to the marginal value of additional spending that we analyze. An increased average cost per unit of health is judged to reflect declining productivity. Our data show such an increase.[32]

However, this conclusion is only valid if one could have bought additional years of quality-adjusted life at the same price in the base period. In that case, spending more on health gains in a later time period is a productivity reduction. If that is not the case, spending more to get more health may reflect a quality improvement if the health benefit is greater than its cost.

Empirically, comparisons across studies are difficult because different studies use different measures of quality. Quality metrics in past studies include clinical estimates of the value of innovation (Lakdawalla et al., 2015), predicted survival based on biomarkers (Eggleston et al., 2011), and expert opinion about the value of treatment changes (Berndt et al., 2002). It is unclear

[32] In 1999, lifetime medical spending of $245,000 bought 12.2 years of quality-adjusted life expectancy in 1999. In 2012, $303,000 of lifetime spending bought 13.9 years (ignoring reductions due to prevalence increases). Thus, the cost per year of life rose from roughly $20,000 to roughly $22,000.

how these relate to actual changes in outcomes. Even studies that look at health outcomes (Romley et al., 2015, 2020) do so over different time horizons.

The most comparable papers to ours are Cutler et al. (1998) and Dauda et al. (2019). Cutler et al. estimate price reductions of over 14% annually for elderly heart attack sufferers. Dauda et al. estimate price reductions of over 7% annually for elderly people hospitalized with heart attack, heart failure, or pneumonia conditions. The price reductions in both of these studies are larger than we estimate for the average elderly person, reflecting the very high benefits of treatment change for acute cardiovascular events.

VII. Conclusion

Measuring the productivity of medical care is a central policy concern. In this paper, we model the productivity of the medical sector as a whole and also of specific conditions. We do so in the context of a satellite health account, an account designed to show the costs and benefits of medical treatment changes over time.

Our analysis of medical care productivity yields three primary conclusions. First, medical treatment changes as a whole have brought benefits greater than their cost. We estimate that the medical treatment sector expanded in value by $113,000 per elderly person between 1999 and 2012, corresponding to a productivity growth rate of 9% percent. This is below other estimates for specific conditions, and likely reflects the heterogeneity in the costs and benefits of medical treatments across conditions.

The considerable heterogeneity in productivity across conditions is our second major conclusion. Conditions with significant increases in value include cardiovascular disease, colorectal cancer, lung cancer, acute kidney injury, and frailty. The major exceptions to the rising value of care are for people with mental illness and arthritis and other musculoskeletal disorders, where cost increases have been large, but health outcomes only modest.

Third, the net benefits of medical treatment changes are partly offset by adverse changes in the prevalence of diseases. Taken as a whole, rising obesity and other factors leading to increased disease prevalence reduced quality-adjusted life expectancy by 0.9 years between 1999 and 2012.

One limitation of our analysis is that we have estimated the value of medical care for the typical elderly person. However, health trajectories have not been the same for all people. For

example, Case and Deaton (2017) show disparate trends in mortality by race and ethnicity in the adult population. How such findings affect the experience of different socioeconomic groups in the elderly is unknown and is a worthy topic for future research.

REFERENCES

Abraham, Katherine G., and Christopher Mackie, eds., *Beyond the Market: Designing Nonmarket Accounts for the United States,* Washington, D.C.: National Academy Press, 2005.

Abraham, Katherine G., and Christopher Mackie, "A Framework for Non-Market Accounts," in Dale Jorgenson, J. Steven Landefeld, and William D. Nordhaus, eds., *A New Architecture for the U.S. National Accounts*, Chicago: University of Chicago Press, 2006, 161–192.

Aizcorbe, Ana M., "Recent Research on Disease-Based Price Indexes: Where Do We Stand?" *Survey of Current Business*, 93 (July) 2013, 9–13.

Aizcorbe, Ana M., Eli B. Liebman, Sarah Pack, David M. Cutler, Michael Chernew, and Allison B. Rosen, "Measuring Health Care Costs of Individuals with Employer-Sponsored Health Insurance in the United States: A Comparison of Survey and Claims Data." *Statistical Journal of the International Association for Official Statistics*, 28, 2012, 43–51.

Berndt, Ernst R., David M. Cutler, Richard Frank, Zvi Griliches, Joseph P. Newhouse, and Jack E. Triplet, "Medical Care Prices and Output," in Anthony Culyer and Joseph Newhouse, eds., *Handbook of Health Economics, Volume I*, Amsterdam: Elsevier, 2000, 119–180.

Berndt, Ernst R., Anupa Bir, Susan H. Busch, Richard G. Frank, and Sharon-Lise T. Normand, "The Medical Treatment of Depression, 1991-1996: Productive Inefficiency, Expected Outcome Variations, and Price Indexes," *Journal of Health Economics,* 2002, 21(3): 373–396.

Boskin, Michael J. 1996. "Toward a More Accurate Measure of the Cost of Living." In: *Final Report from the Advisory Commission to Study the Consumer Price Index*. Washington, DC: US Senate Finance Committee.

Brouwer, Nelleke P.M., Amand C.R.K. Bos, Valery E.P.P. Lemmens, et al., "An Overview of 25 Years of Incidence, Treatment, and Outcome of Colorectal Cancer Patients," *Int J Cancer*, 2018, 143(11): 2758–2766.

Brynjolfsson, Erik, Avinash Collis, W. Erwin Diewert, Felix Eggers, and Kevin J. Fox, "GDP–B: Accounting for the Value of New and Free Goods in the Digital Economy," NBER Working Paper No. 25695, March 2019.

Bureau of Labor Statistics, "Technical Information about the US Multifactor Productivity Measures", mimeo, September 26, 2007.

Case, Anne, and Angus Deaton. 2017. "Mortality and Morbidity in the 21st Century." *Brookings Papers on Economic Activity*, Spring 2017.

Chandra, Amitabh, and Jonathan Skinner, "Technology Growth and Expenditure Growth in Health Care," *Journal of Economic Literature*, September 2012, 50(3): 645–80.

Cockburn, Iain, and Aslam H. Anis, "Hedonic Analysis of Arthritis Drugs," in David M. Cutler and Ernst Berndt, eds., *Medical Care Output and Productivity,* Chicago: University of Chicago Press, 2001.

Comaner, William S., Stuart O. Schweitzer, and Tanja Carter, "A Hedonic Model of Pricing of Innovative Pharmaceuticals," in Marco R. Di Tommaso and Stuart O. Schweitzer, eds., Health Policy and High–Tech Industrial Development, Edward Elgar Publishing, 2005.

Cutler, David M., Mark McClellan, Joseph P. Newhouse, and Dahlia Remler, "Are Medical Prices Declining? Evidence from Heart Attack Treatments." *The Quarterly Journal of Economics,* 113(4), November 1998, 991–1024.

Cutler, David M., Allison B. Rosen, and Sandeep Vijan, "Value of Medical Spending in the United States: 1960–2000," *New England Journal of Medicine, 2006,* 355(9): 920–927.

Cutler, David M., Kaushik Ghosh, Kassandra Messer, Trivellore Raghunathan, Susan T. Stewart, and Allison B. Rosen. "Explaining the Slowdown in Medical Spending Growth among the Elderly, 1999–2012." *Health Affairs*, 38(2), February 2019, 222–229.

Dauda, Seidu, Abe Dunn, and Anne Hall, "Are Medical Care Prices Still Declining? A Systematic Examination of Quality–Adjusted Price Index Alternatives for Medical Care," mimeo, 2019.

Dunn, Abe, Lindsey Rittmueller, and Bryn Whitmire. 2015, "Introducing the New BEA Health Care Satellite Account," *Survey of Current Business*, 95(1), January 2015, 1–21.

Eggleston, Karen N., Nilay D. Shah, Steven A. Smith, Amy E. Wagie, Arthur R. Williams, Jerome H. Grossman, Ernst R. Berndt, Kirsten H. Long, Ritesh Banerjee, and Joseph P. Newhouse. "The Net Value of Health Care for Patients With Type 2 Diabetes, 1997 to 2005." Annals of Internal *Medicine*, 2009, 151(6): 386–393.

Ettehad, Dena, Connor A Emdin, Amit Kiran, Simon G Anderson, Thomas Callender, Jonathan Emberson, et al. "Blood Pressure Lowering for Prevention of Cardiovascular Disease and Death: A Systematic Review and Meta–Analysis". *Lancet*, March 2016, 5;387(10022): 957–967.

Elixhauser Anne, Claudia Steiner, and L. Palmer. Clinical Classifications Software (CCS). U.S. Agency for Healthcare Research and Quality, https://www.hcup-us.ahrq.gov/toolssoftware/ccs/CCSUsersGuide.pdf Internet accessed, 2018.

Fenton, Joshua J., Meghan S. Weyrich, Shauna Durbin, Yu Liu, Heejung Bang, and Joy Melnikow, "Prostate-Specific Antigen-Based Screening for Prostate Cancer: Evidence Report and Systematic Review for the US Preventive Services Task Force" JAMA, 2018, 319(18): 1914–1931.

Finkelstein, Amy, Erzo F.P. Luttmer, and Matthew J. Notowidigdo. "What Good Is Wealth Without Health? The Effect of Health on the Marginal Utility of Consumption," *Journal of the European Economic Association*, January 2013, 11(1): 221–258.

Ford, Earl S., Umed A. Ajani, Janet B. Croft, Julia A. Critchley, Darwin R. Labarthe, Thomas E. Kottke, et al., "Explaining the Decrease in U.S. deaths from Coronary Disease, 1980–2000". *N Engl J Med*, 2007, 356(23): 2388–2398.

Frank, Richard G., Ernst R. Berndt, and Susan M. Busch. 1999. "Price Indexes for the Treatment of Depression." In *Measuring the Prices of Medical Treatments*, edited by Jack E. Triplett, 72–102. Washington, DC: The Brookings Institution.

Fraumeni, Barbara M., Michael S. Christian, and Jon D. Samuels, "The Accumulation of Human and Nonhuman Capital, Revisited," *Review of Income and Wealth*, 2017, 63: S381–S410.

Ghosh, Kaushik, Irina Bondarenko, Kassandra L. Messer, Susan T. Stewart, Trivellore Raghunathan, Allison B. Rosen, and David M. Cutler. "Attributing Medical Spending to Conditions: A Comparison of Methods." PLoS One, 2020a, 15(8) :e0237082.

Ghosh, Kaushik, David Cutler, Trivellore Raghunathan, Kassandra Messer, Susan T. Stewart, and Allison B. Rosen. "Attributing Mortality to Medical Conditions: Levels and Trends in the United States, 1999–2012." mimeo, 2020b.

Gold, Marshe R., Joanna E. Siegel, Louise B. Russell, and Milton C. Weinstein, eds. *Cost-Effectiveness in Health and Medicine*. New York, NY: Oxford University Press; 1996.

Groshen, Erica L., Brian C. Moyer, Ana M. Aizcorbe, Ralph Bradley, and David M. Friedman, "How Government Statistics Adjust for Potential Biases from Quality Change and New Goods in an Age of Digital Technologies: A View from the Trenches," *Journal of Economic Perspectives,* 2017*,* 31(2): 187–210.

Gunn, Jarmo, Johanna Valo, Jussi Sipilä, Päivi Rautava, Eero Sihvo, and Ville Kytö, "Trends and Results of Lung Cancer Surgery in Finland between 2004 and 2014," *European Journal of Cardio-Thoracic Surgery*, July 2018, 54 (1): 127–133.

Hall, Anne E., and Tina Highfill, "Calculating Disease-Based Medical Care Expenditure Indexes for Medicare Beneficiaries: A Comparison of Method and Data Choices." Bureau of Economic Analysis Working Paper. Washington, DC: BEA, June 2014.

Kuznets, Simon, *National Income 1929–1932*. Senate Document No. 124, 73rd Congress, 2nd Session. Washington, D.C.

Lakdawalla, Darius, Jason Shafrin, Claudio Lucarelli, Sean Nicholson, Zeba M. Khan, and Tomas J. Philipson, "Quality-Adjusted Cost of Care: A Meaningful Way to Measure Growth in Innovation Cost Versus the Value of Health Gains." Health Affairs, 2015, 34(4): 555–61.

Landefeld, J. Steven, and Stephanie H. McCulla. "Accounting for Nonmarket Household Production within a National Accounts Framework." *Review of Income and Wealth,* September 2000, 46(3): 289–307.

Lemmens Valery, Liza van Steenbergen, Maryska Janssen-Heijnen, Hendrik Martijn, Harm Rutten, and Jan Willem Coebergh. "Trends in colorectal cancer in the south of The Netherlands 1975-2007: rectal cancer survival levels with colon cancer survival," *Acta Oncol,* 2010, 49:784–96.

Lin, Jennifer S., Margaret A. Piper, Leslie A. Perdue, Carolyn M. Rutter, Elizabeth M. Webber, Elizabeth O'Connor, Ning Smith, and Evelyn P. Whitlock. "Screening for Colorectal Cancer: Updated Evidence Report and Systematic Review for the US Preventive Services Task Force". *JAMA*, June 2016, 315(23):2576–94.

Melnikow, J., Jillian T. Henderson, Brittany U. Burda, Caitlyn A. Senger, Shauna Durbin, and Meghan S. Weyrich. "Screening for Cervical Cancer with High-Risk Human Papillomavirus Testing: Updated Evidence Report and Systematic Review for the US Preventive Services Task Force". *JAMA*, August 2018, 320(7):687–705.

Messite, Jacqueline, and Steven D. Stellman, "Accuracy of Death Certificate Completion: The Need for Formalized Physician Training." *JAMA*, 1996, 275(10):794–6.

Moulton, Brent R., Brian C. Moyer, and Ana Aizcorbe. 2009. "Adapting BEA's National and Industry Accounts for a Health Care Satellite Account." Strategies for a BEA Satellite Health Care Account: Summary of a Workshop. Washington, DC: The National Academies Press.

Muller, Nicholas Z., Robert Mendelsohn, and William Nordhaus, "Environmental Accounting for Pollution in the United States Economy." *American Economic Review*, 2011, 101(5): 1649–75.

National Research Council, Panel to Advance a Research Program on the Design of National Health Accounts. Washington (DC): National Academies Press (US); 2010.

Nelson, Heidi D., Rochelle Fu, Amy Cantor, Miranda Pappas, Monica Daeges, and Linda Humphrey, "Effectiveness of Breast Cancer Screening: Systematic Review and Meta-analysis to Update the 2009 U.S. Preventive Services Task Force Recommendation". *Ann Intern Med*, February 2016, 164(4): 244–55.

Noone, Anne-Michelle, Kathleen A. Cronin, Sean F. Altekruse, Nadia Howlader, Denise R. Lewis et al., "Cancer Incidence and Survival Trends by Subtype Using Data from the Surveillance Epidemiology and End Results Program, 1992–2013." *Cancer Epidemiol Biomarkers Prev*. 2017, 26(4):632–641.

Ogata Soshiro, Kunihiro Nishimura, Maria Guzman-Castillo, Yoko Sumita, Michikazu Nakai, Yoko M. Nakao et al., "Explaining the Decline in Coronary Heart Disease Mortality Rates

in Japan: Contributions of Changes in Risk Factors and Evidence-Based Treatments between 1980 and 2012." *International Journal of Cardiology*, September 2019, 291:183–188.

Phillips Jonathan K., Morgan A. Ford, and Richard J. Bonnie, eds., *Committee on Pain Management and Regulatory Strategies to Address Prescription Opioid Abuse*, National Academies of Sciences, Engineering, and Medicine; Health and Medicine Division; Board on Health Sciences Policy; Washington (DC): National Academies Press (US); 2017 Jul 13.

Preston, Samuel H., Yana C. Vierboom, and Andrew Stokes, "The Role of Obesity in Exceptionally Slow US Mortality Improvement," *Proceedings of the National Academy of Sciences*, 2018, 115, 201716802.

Raghunathan, Trivellore E., Kaushik Ghosh, Allison B. Rosen, Paul Imbriano, Susan T. Stewart, Irina Bondarenko, Kasey Messer, Patricia Berglund, James Shaffer, and David M. Cutler, "Combining Information from Multiple Data Sources to Assess Population Health," *Journal of Survey Statistics and Methodology*, March 2020, smz047.

Rice, Dorothy P., "Estimating the Cost of Illness," *American Journal of Public Health,* 1967, 57(3): 424–440.

Romley, John A., Dana P. Goldman, and Neeraj Sood, "US Hospitals Experienced Substantial Productivity Growth During 2002–11." *Health Affairs*, 2015, 34(3): 10–1377.

Romley, John A., Dana P. Goldman, Neeraj Sood, and Abe Dunn, "Productivity Growth in Treating a Major Chronic Health Condition," mimeo, Brookings Institution, 2020.

Rosen, Allison B., and David M. Cutler. "Measuring Medical Care Productivity: A Proposal for U.S. National Health Accounts." *Survey of Current Business*, June 2007, 87: 54–58.

Rosen, Allison B., Eli Liebman, Ana M. Aizcorbe, and David M. Cutler. "Comparing Commercial Systems for Characterizing Episodes of Care." Bureau of Economic Analysis (BEA) Working Paper WP2012–7. 2012, Washington, DC: BEA.

Schultze, Charles and Christopher Mackie, eds. *At What Price?: Conceptualizing and Measuring Cost-of-Living and Price Indexes.* Washington, D.C.: National Academies Press, 2002.

Sensenig, Arthur, and Ernest Wilcox, "National Health Accounts/National Income and Product Accounts Reconciliation—Hospital Care and Physician Services," in David Cutler and Ernst Berndt, eds., *Medical Care Output and Productivity*, Chicago: University of Chicago Press, 2001, 271–302.

Shapiro, Irving, Matthew D. Shapiro, and David W. Wilcox, "Measuring the Value of Cataract Surgery," in David Cutler and Ernst Berndt, eds., *Medical Care Output and Productivity*, Chicago: University of Chicago Press, 2001, 411–437.

Sheiner, Louise, and Anna Malinovskaya, "Measuring Productivity in Health Care: An Analysis of the Literature," Brookings Institution, May 2016.

Thorpe, Kenneth E., Curtis S. Florence, and Peter Joski, "Which Medical Conditions Account for the Rise in Health Care Spending?" *Health Affairs,* 23(S1), August 2004, W4-437–W4-445.

Trogdon, Justin, Eric A. Finkelstein, and Thomas J. Hoerger, "Use of Econometric Models to Estimate Expenditure Shares." *Health Services Research,* August 2008, 43(4): 1442–1452.

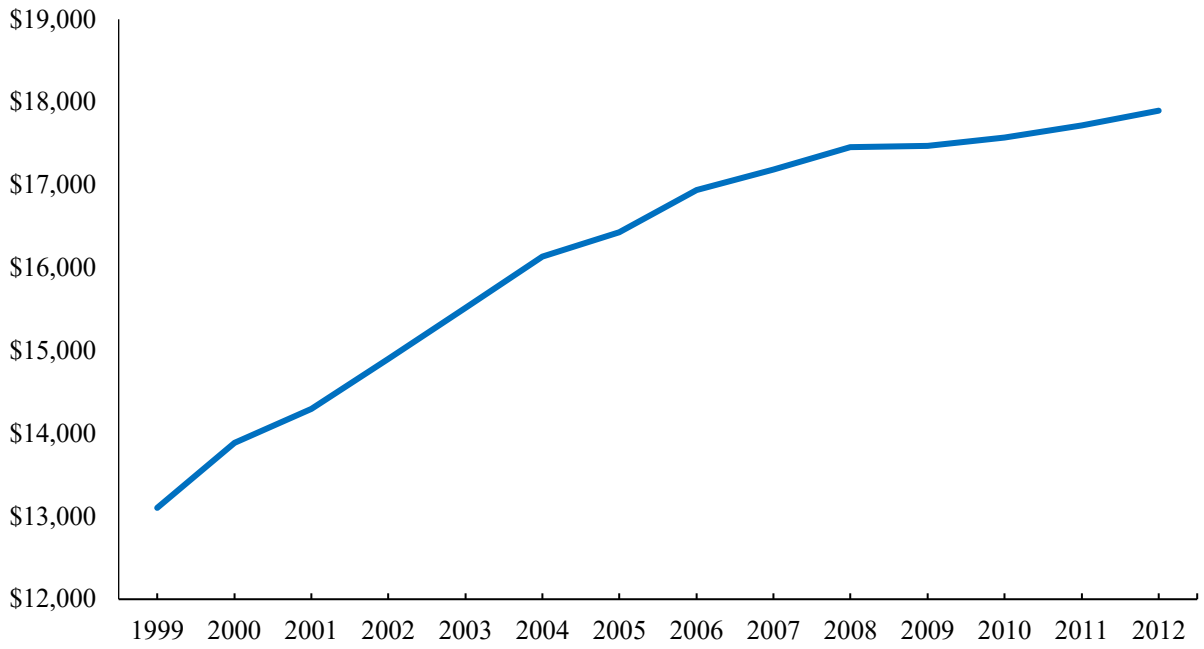
Van der Drift, Miep A., Henrike E. Karim-Kos, Sabine Siesling, et al., "Progress in Standard of Care Therapy and Modest Survival Benefits in the Treatment of Non-Small Cell Lung Cancer Patients in the Netherlands in the Last 20 Years," *Journal of Thoracic Oncology,* February 2012, 7(2): 291–298.

Vogelaar, Iris, Marjolein van Ballegooijen, Deborah Schrag, et al., "How Much Can Current Interventions Reduce Colorectal Cancer Mortality in the U.S.?" *Cancer*, 2006, 107: 1624–1633.

Viscusi, W. Kip, and Joseph E. Aldy, "Value of a Statistical Life: A Critical Review of Market Estimates Throughout the World," *Journal of Risk and Uncertainty*, 2003, 27(1), 5–76.

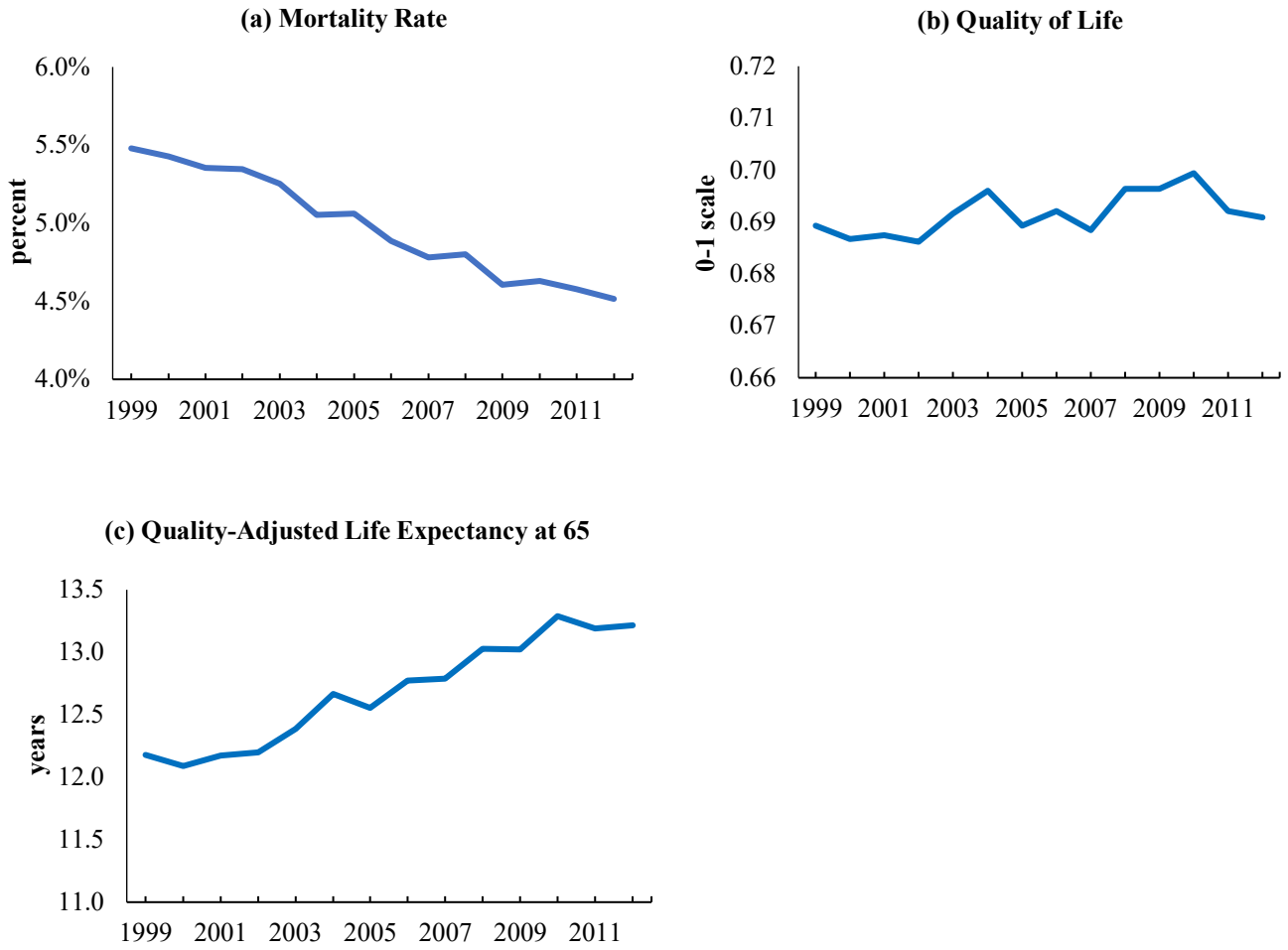
Welch, H Gilbert and Otis W. Brawley, "Scrutiny-Dependent Cancer and Self-fulfilling Risk Factors," *Ann Intern Med*, January 2018, 168:143–144.

Figure 1: Real, Per Capita Medical Spending in the Elderly Population, 1999–2012



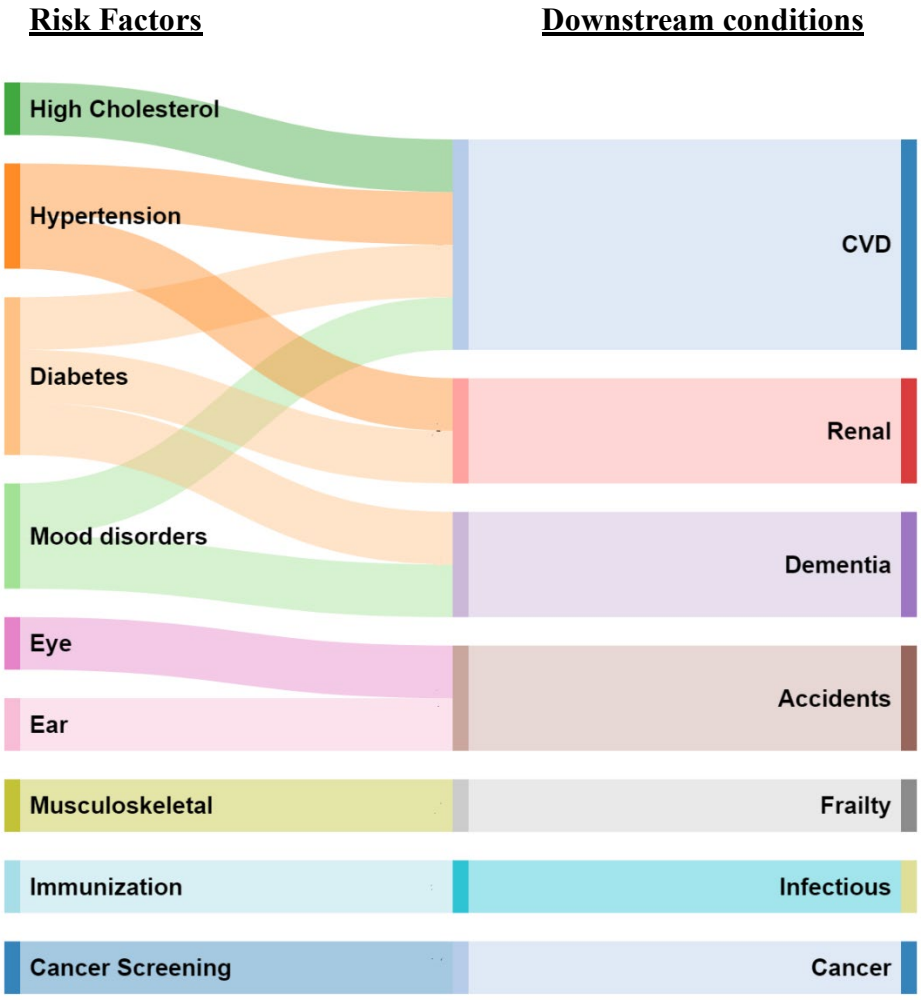
Note: Data are from the Medicare Current Beneficiary Survey with totals matching estimated national spending on the elderly. Spending is in real ($2010) dollars.

Figure 2: Measures of Health in the Elderly Population, 1999–2012



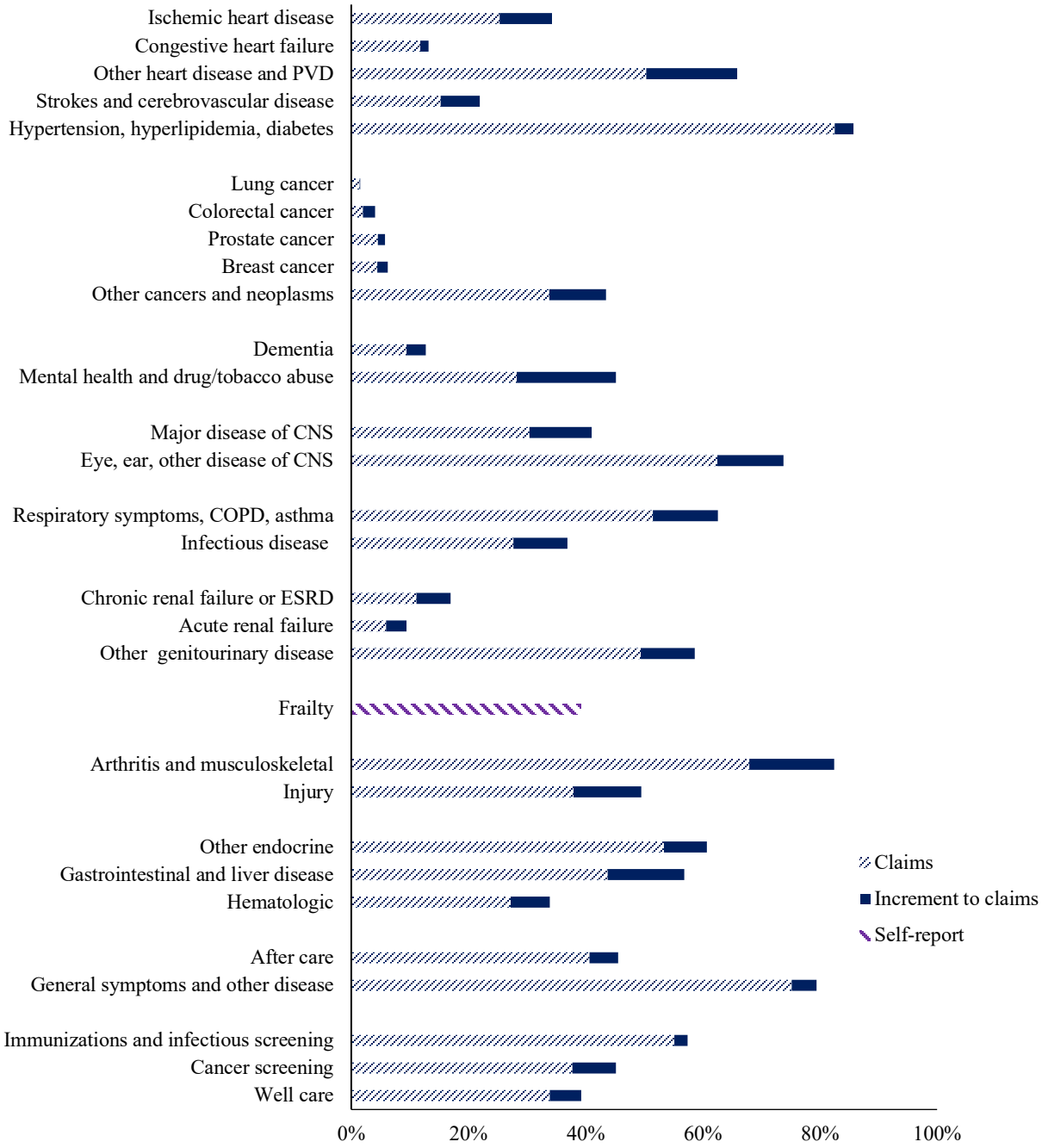
Note: Mortality rates and quality of life in charts (a) and (b) are age-adjusted to the 2010 population, using data on population shares aged 65-74, 75-84, and 85+. Mortality rates are from Vital Statistics. Quality of life disutilities are based on a regression model relating self-rated health on a 0-100 scale in MEPS 2000-2002 to measures of ADL impairments, IADL impairments, functional limitations, sensory problems, and whether health limits social activities. These impairments are available in the MCBS each year from 1999–2012. Weights are assumed to be constant over time. Quality-adjusted life expectancy is the expectation of quality-adjusted years of life remaining, using age- and year-specific mortality and quality of life weights.

Figure 3: Relationship Between Risk Factors on Downstream Conditions



Note: Each of the risk factors on the left is matched to downstream conditions on the right, based on studies in the clinical literature. Mood disorders include depression, bipolar disorder, and anxiety/PTSD. CVD includes ischemic heart disease, congestive heart failure, other heart disease, and peripheral vascular disease, strokes, and cerebrovascular disease. Renal includes acute and chronic renal failure or end-stage renal disease. Eye and ear include disorders of the eye and ear. Accidents includes all injury: accidents, falls, poisoning. The full list of conditions in each category is shown in Appendix Table A6.

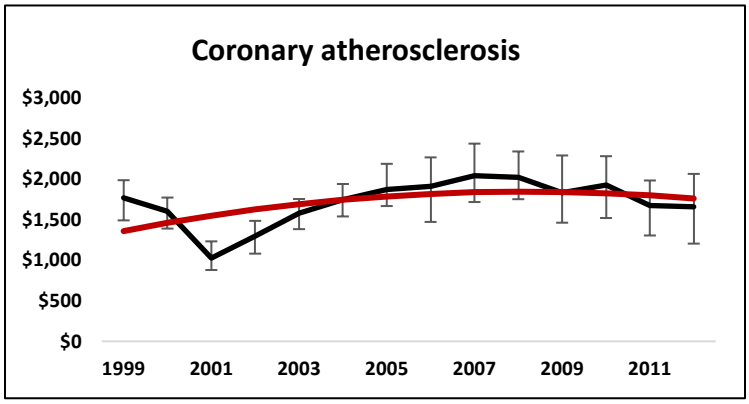
Figure 4: Conditions and Prevalence Rates, 2012

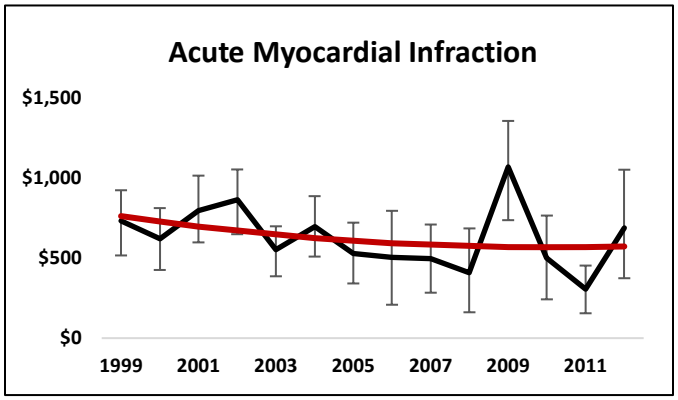


Note: The figure shows the prevalence of different conditions in the elderly population in 2012. Except for frailty, each condition has two parts: the prevalence as measured by claims, and the additional prevalence added via calibration to prevalence in the NHANES. Frailty is based on self-reports of frailty in MCBS. Data are weighted using population weights. CNS = Central nervous system; COPD = Chronic obstructive pulmonary disease; ESRD = End-stage renal disease; ID = Infectious disease.

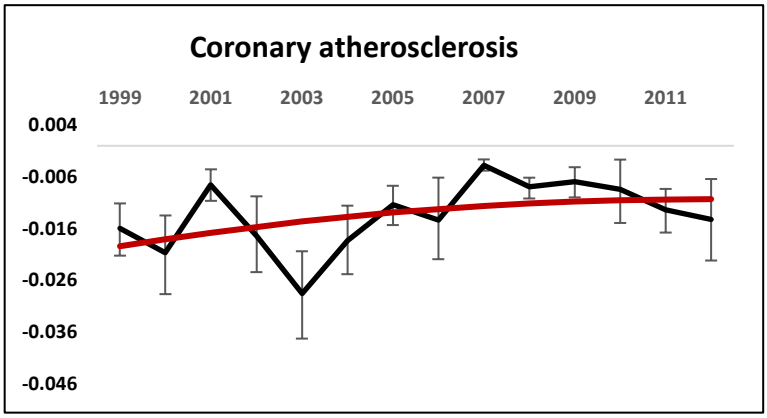
Figure 5: Example of Cost and Health Estimates for Ischemic Heart Disease

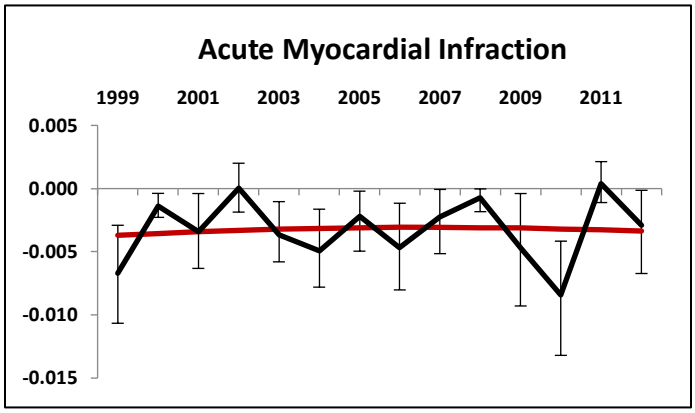
A. Cost per case



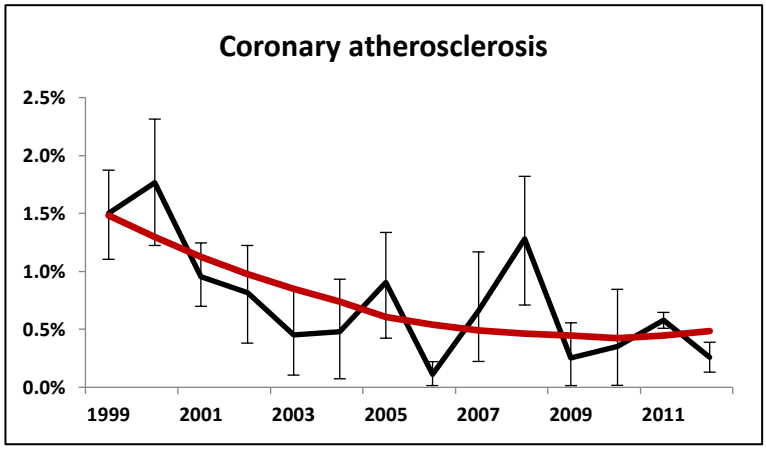


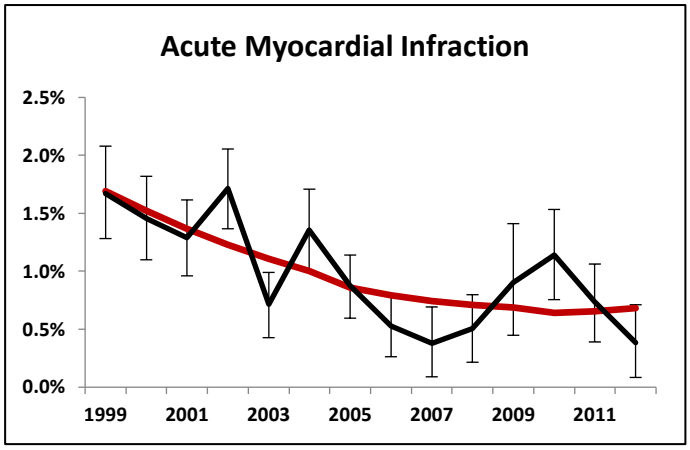
B. Quality of Life





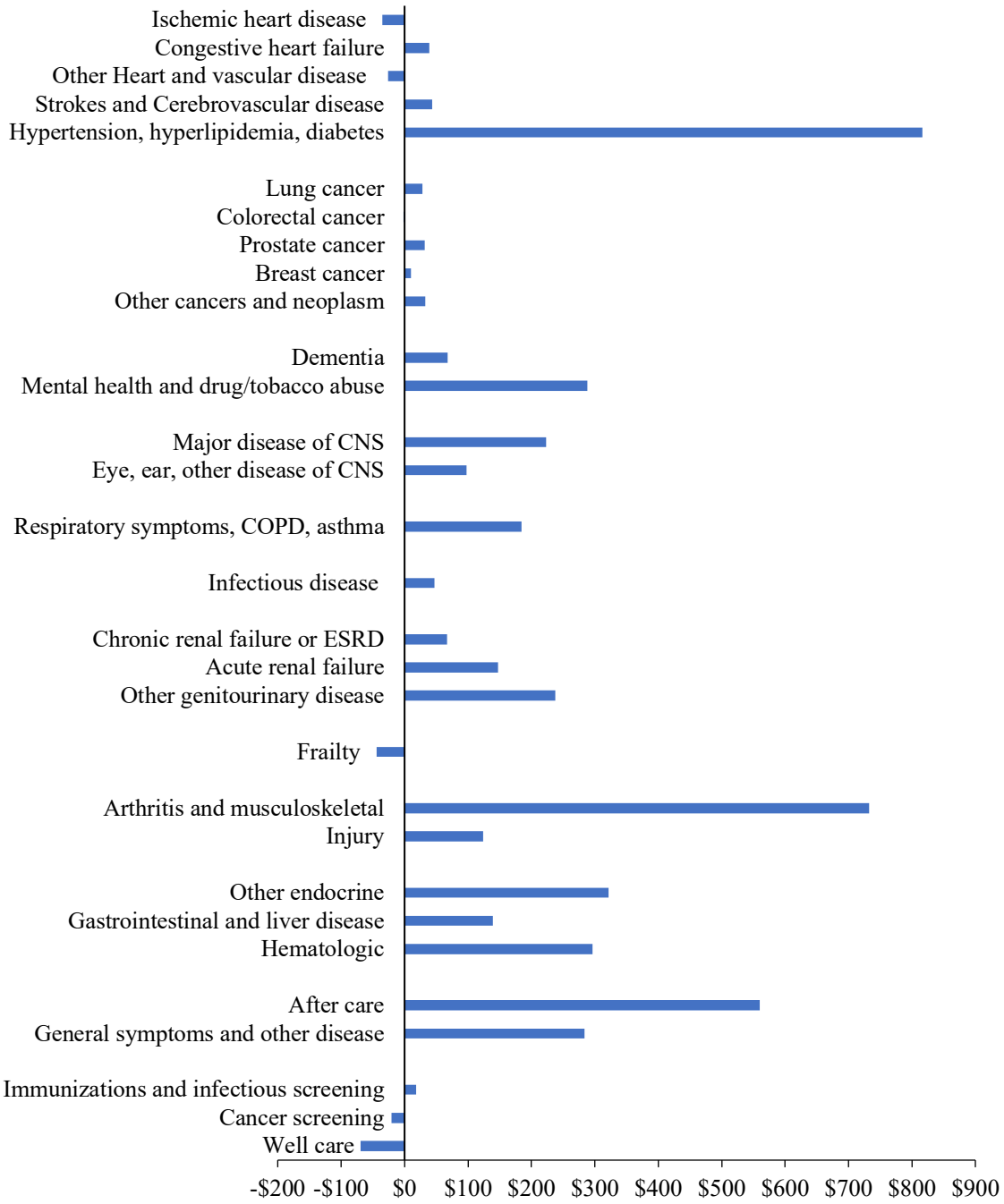
C. Mortality rate





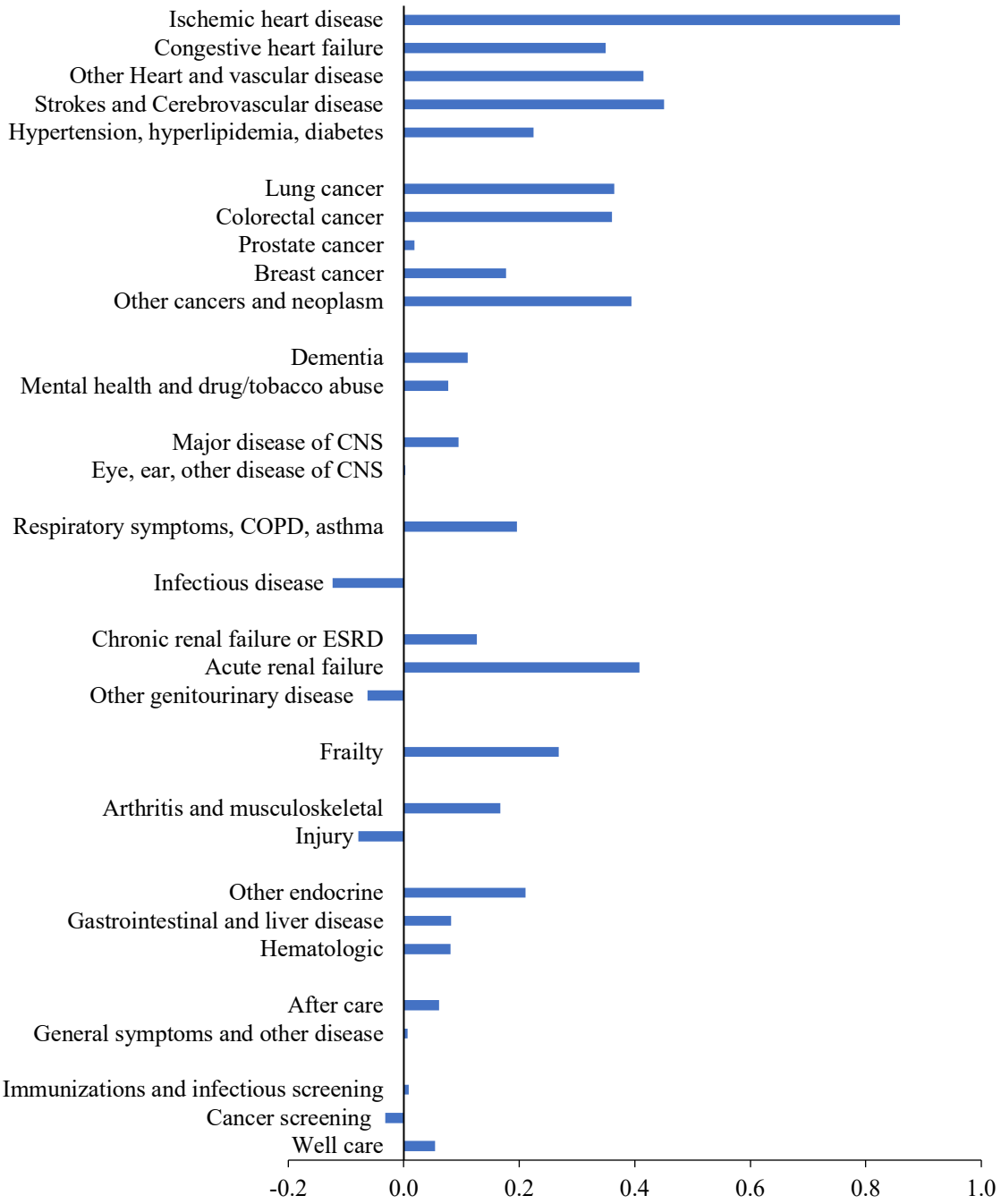
Note: ischemic heart disease includes coronary atherosclerosis and acute myocardial infarction. The outcomes are based on model estimates as described in the text. The black line shows the estimates for each year, along with the 95% confidence intervals for the estimates. The red line shows the smoothed trend, fitted from a second order polynomial.

Figure 6: Change in Per Capita Medical Spending by Condition, 1999–2012



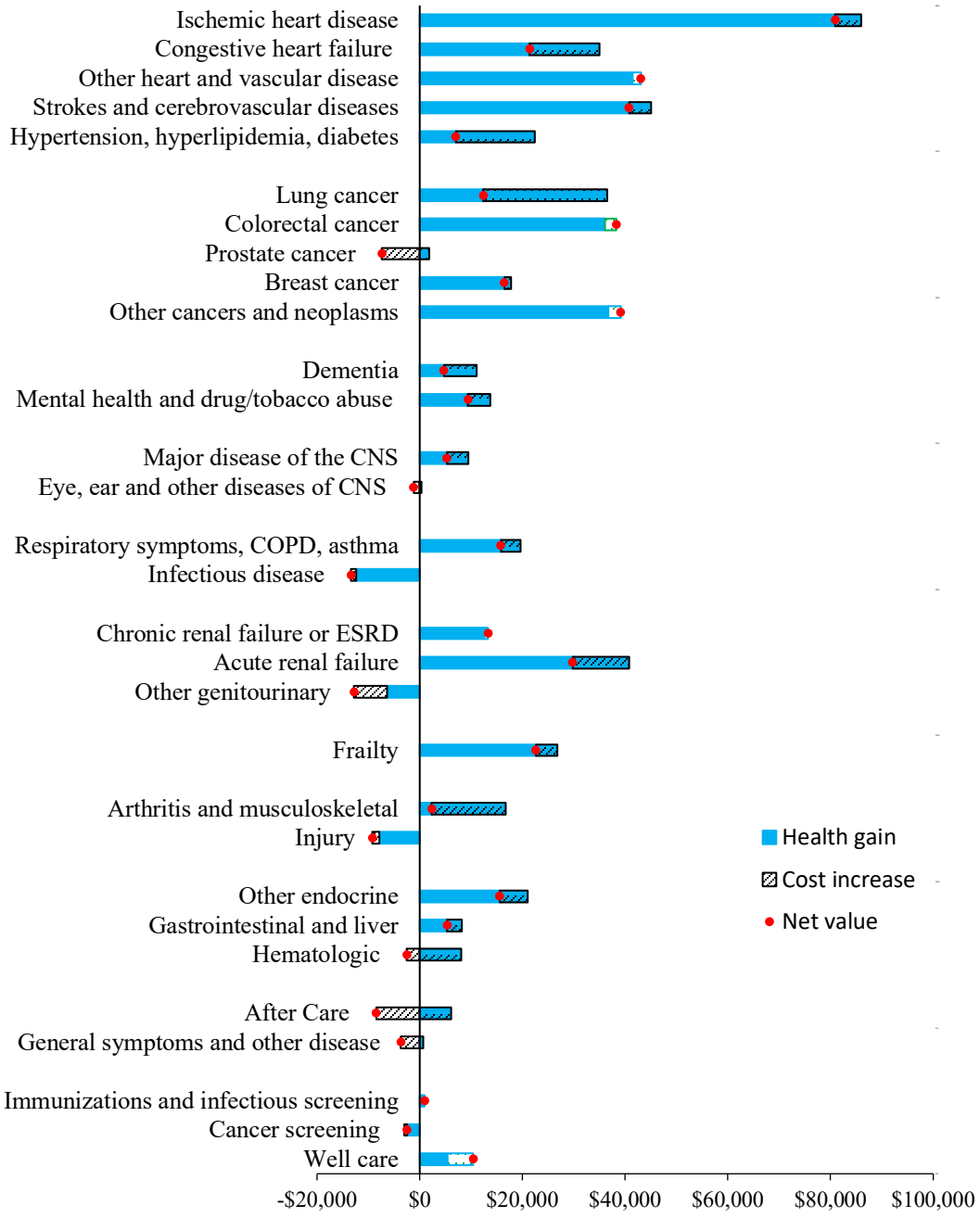
Note: Data are from the Medicare Current Beneficiary Survey, adjusted so that total spending matches national estimates of spending on the elderly. Per capita spending is spending per person with the disease multiplied by the disease prevalence rate. Spending is in real ($2010) dollars.

Figure 7: Change in QALE by Condition, 1999–2012



Note: Data are from the Medicare Current Beneficiary Survey. QALE is as of age 65. Mortality data are matched to vital statistics. Quality of life is based on survey responses about symptoms and impairments. Disutility weights are based on analysis of the Medical Expenditure Panel Study.

Figure 8: Net Value of Medical Spending Change by Condition, 1999–2012



Note: Data are from the Medicare Current Beneficiary Survey with totals matching estimated national spending on the elderly. Spending is in real ($2010) dollars. The blue bar depicts improvement in health outcomes over the period, expressed in dollars. Health change is the change in QALE attributed to medical care and not changes in the prevalence of the condition. The hatched bar shows the change in medical spending. The red dot shows the net productivity estimate, defined as the dollar value of health improvement minus the increase in spending.

Table 1: Characteristics of the Sample, 2010-2012

Metric	Mean
Mortality rate (annual)	4.6%
Components of quality of life	
Difficulty seeing	33%
Difficulty hearing	44%
Any ADL impairment	16%
Any IADL impairment	25%
Difficulty walking (moderate, severe)	9%, 28%
Difficulty stooping/crouching/kneeling	48%
Difficulty lifting or carrying objects as heavy as 20 pounds	26%
Difficulty reaching or extending arms above shoulder level	17%
Difficulty writing or handling and grasping small objects	17%
Health limits social interactions (moderate, severe)	7%, 5%
Sociodemographics	
Male	44%
Race (white / black / other)	79%, 8%, 13%
Married	54%
Poverty category (1=worse, 5=best)	3.7
Education (<9 years, 9-11 years, HS/GED, some college, CG)	10%, 13%, 28%, 27%, 23%
Veteran	24.1%
Behavioral risk factors	
Smoking (never, current, former)	37%, 15%, 48%
BMI (underweight, normal, overweight, obese)	3%, 34%, 38%, 26%
Health Status	
Self-reported health compared to one year previous (much/somewhat better, about the same, somewhat/much worse)	15%, 63%, 22%
Self-reported health compared to others of the same age (excellent, very good, good, fair, poor)	18%, 32%, 30%, 15%, 5%

Note: Data are the elderly sample in the Medicare Current Beneficiary Survey. Means are weighted using sample weights.

Table 2: Cohort balance table for 1999: Ischemic Heart disease

	Acute Myocardial Infraction			Coronary Atherosclerosis		
Metric	**Yes**	**No**	**p-value**	**Yes**	**No**	**p-value**
Age	77.5	77.4	0.37	77.0	77.1	0.71
Male	40.4%	40.3%	0.94	41.7%	41.6%	1.00
Education (9^{th}-12th)	16.8%	17.1%	0.76	15.9%	17.0%	0.16
Education (High school)	28.4%	28.3%	0.90	29.6%	28.7%	0.35
Education (Asso./some College)	19.9%	19.8%	0.92	22.0%	21.2%	0.40
Education (College and Above)	14.3%	14.3%	0.92	14.9%	15.6%	0.43
Military service	25.4%	25.8%	0.65	29.4%	28.1%	0.19
Smoke now	9.4%	9.9%	0.42	9.5%	10.1%	0.33
Ever smoke	54.2%	54.3%	0.93	58.4%	57.9%	0.68
Race (Black)	7.3%	7.9%	0.32	6.4%	7.0%	0.25
Race (Hispanic)	5.0%	5.3%	0.50	5.8%	6.3%	0.43
Race (Other)	2.8%	2.7%	0.77	4.6%	4.8%	0.64
Marital status (Widowed)	41.2%	40.9%	0.76	37.9%	37.0%	0.40
Marital status (Divorced/Separated)	6.4%	6.4%	0.94	7.1%	7.3%	0.73
Marital status (Never married)	4.5%	4.4%	0.77	4.8%	3.9%	0.06
Pneumonia shot	65.6%	64.7%	0.37	66.1%	65.3%	0.45
Flu shot	71.9%	71.4%	0.62	72.6%	71.3%	0.19
Hysterectomy	25.2%	25.1%	0.88	25.2%	26.0%	0.42
Income / FPL<1.5	9.9%	9.7%	0.86	9.5%	9.4%	0.86
Income / FPL<2.0	20.0%	20.2%	0.84	20.9%	21.2%	0.76
Income / FPL<4.0	27.7%	28.1%	0.73	29.9%	29.2%	0.44
Income / FPL<5.0	15.1%	15.1%	0.92	15.8%	16.1%	0.66
Other private insurance	74.2%	73.3%	0.31	76.0%	75.4%	0.51
Difficulty lifting (scale 1-4)	1.78	1.75	0.15	1.63	1.61	0.40
Difficulty stooping (scale 1-4)	2.11	2.09	0.39	1.99	1.96	0.19
Difficulty walking	44.1%	43.1%	0.34	39.5%	38.5%	0.35
Difficulty dressing	17.4%	16.1%	0.13	12.6%	12.5%	0.80
Difficulty eating	9.4%	8.6%	0.24	7.0%	6.4%	0.35
Health since last year – same	55.2%	56.1%	0.40	58.3%	59.2%	0.39
Health since last year – worse	22.4%	21.3%	0.22	21.6%	20.7%	0.30
General health status – very good	25.0%	25.4%	0.72	25.8%	27.2%	0.13
General health status – good	32.8%	33.0%	0.83	32.5%	32.9%	0.69
General health status – fair	20.4%	20.4%	0.98	19.8%	18.7%	0.22
General health status – poor	9.3%	8.4%	0.17	7.7%	7.1%	0.30
Hearing aid	12.4%	11.8%	0.34	10.7%	10.8%	0.91
Institutional days	32.7	31.2	0.48	23.1	20.7	0.18
Inpatient nights	3.0	2.6	0.01	2.6	2.1	0.00
Inpatient stays	0.5	0.4	0.00	0.5	0.4	0.00
Height	166.8	166.7	0.75	166.9	167	0.78
Weight(kg)	71.6	71.7	0.76	72.3	72.5	0.64

Note: The table shows the cohort balance table for analysis of the impact of acute myocardial infarction and coronary atherosclerosis. The p-value for the joint F-test of equality between treatment and controls is 0.628 for acute myocardial infraction and 0.929 for coronary atherosclerosis.

Table 3: Estimates of Spending and Health Decrements by Condition

	Spending per case		Mortality Rate		Quality of Life	
Condition	**1999**	**2012**	**1999**	**2012**	**1999**	**2012**
Cardiovascular diseases						
Ischemic heart disease	$1,300	$1,502	1.4%	0.5%	-0.011	-0.007
Congestive heart failure	$1,248	$2,049	1.5%	1.1%	-0.009	-0.010
Other heart and vascular disease	$2,264	$2,092	0.8%	0.4%	-0.007	-0.005
Strokes and cerebrovascular diseases	$1,432	$1,632	0.8%	0.4%	-0.011	-0.008
Cardiovascular risk factors	$1,212	$2,102	0.8%	0.5%	-0.014	-0.014
Cancers						
Lung cancer	$2,948	$5,436	9.9%	9.9%	-0.034	-0.015
Colorectal cancer	$1,222	$1,042	1.3%	1.0%	-0.008	-0.008
Prostate cancer	$340	$873	0.4%	0.4%	0.005	0.000
Breast cancer	$285	$352	0.5%	0.2%	0.000	-0.004
Other cancers and neoplasm	$1,184	$1,005	0.6%	0.3%	-0.003	-0.002
Mental Health						
Dementia	$2,132	$2,491	1.3%	1.1%	-0.023	-0.027
Mental health and tobacco/drug abuse	$1,740	$1,990	0.7%	0.7%	-0.026	-0.021
Central Nervous System						
Major disease of the CNS	$1,224	$1,472	1.0%	0.9%	-0.013	-0.013
Eye, ear, other diseases of the CNS	$743	$827	0.0%	0.0%	0.000	0.000
Respiratory System						
Respiratory symptoms, COPD, asthma	$2,177	$2,359	0.5%	0.4%	-0.007	-0.007
Infectious disease	$1,526	$1,605	0.6%	0.7%	-0.006	-0.008
Kidney Disease						
Chronic renal failure or ESRD	$1,543	$1,484	1.1%	1.0%	-0.013	-0.012
Acute renal failure	$2,592	$3,260	3.2%	2.7%	-0.010	-0.008
Other genitourinary diseases	$563	$929	-0.1%	-0.1%	-0.004	-0.007
Frailty	$2,044	$2,249	1.3%	1.0%	-0.056	-0.055
Musculoskeletal						
Arthritis and musculoskeletal	$1,190	$2,017	0.3%	0.2%	-0.023	-0.016
Injury	$1,743	$1,830	0.2%	0.2%	-0.007	-0.008
Endocrine, GI, Liver, Hematologic						
Other endocrine	$1,071	$1,371	0.4%	0.2%	-0.006	-0.005
Gastrointestinal and liver disease	$934	$1,091	0.5%	0.4%	-0.004	-0.006
Hematologic	$1,645	$2,245	0.3%	0.3%	-0.008	-0.008
Miscellaneous						
After care	$961	$1,807	0.4%	0.4%	-0.008	-0.006
General symptoms and other disease	$1,528	$1,770	0.0%	0.0%	-0.008	-0.009
Prevention and screening						
Immunizations and infectious screening	$251	$243	0.0%	0.0%	0.000	0.000
Cancer screening	-$135	-$165	-0.1%	0.0%	0.000	0.000
Well Care	$680	$409	-0.4%	-0.4%	-0.001	0.001

Note: The table shows the impact of each condition on medical spending, mortality, and quality of life. The year 1999 pools data from 1999–2001, and the year 2012 pools data from 2010–2012. PVD = Peripheral vascular disease. COPD = Chronic obstructive pulmonary disease. ID = Infectious disease. ESRD = End stage renal disease

Table 4: Factors Explaining Change in Medical Spending and Health, 1999–2012

	Change in	
Measure	**Annual Medical spending**	**QALE**
Overall	$4,640	1.04
Treatment changes	$2,918	1.69
Prevalence change	$1,661	-0.87
Behavioral risk factors		
Smoking	-$64	-0.01
Obesity	$342	-0.13
Other changes in prevalence	$1,382	-0.73

Note: The change in medical spending and QALE are based on calculations described in the text. The components do not add to totals because of the covariance.

Table 5: Productivity of Medical Care Per Capita

Condition	Spending Change	QALE Change	Value of a Year of Life			Disc rate
			$100,000	$50,000	$200,000	(3%)
Overall	$57,893	1.70	$112,575	$27,341	$283,042	$66,220
(95% confidence interval, 000s)	($46-$69)	(0.7-2.5)	($14-$188)	($-18-$64)	($87-$440)	($0-$116)
Cardiovascular diseases	$5,157	**0.94**	**$89,327**	**$42,085**	**$183,811**	**$56,731**
Ischemic heart disease	$5,054	0.86	$80,888	$37,917	$166,830	$51,361
Congestive heart failure	$13,569	0.35	$21,395	$3,913	$56,358	$12,042
Other heart and vascular disease	-$1,573	0.41	$43,058	$22,316	$84,543	$27,848
Strokes and cerebrovascular disease	$4,354	0.45	$40,750	$18,198	$85,855	$25,588
Cardiovascular risk factors	$15,439	**0.22**	**$6,995**	**-$4,222**	**$29,429**	**$2,289**
Cancers	-$80	**0.38**	**$37,987**	**$19,033**	**$75,895**	**$23,821**
Lung cancer	$24,099	0.36	$12,384	-$5,857	$48,868	$8,994
Colorectal cancer	-$2,196	0.36	$38,275	$20,235	$74,353	$24,610
Prostate cancer	$9,211	0.02	-$7,374	-$8,292	-$5,537	-$6,918
Breast cancer	$1,272	0.18	$16,470	$7,599	$34,213	$9,041
Other cancers and neoplasm	-$2,423	0.37	$39,116	$20,770	$75,809	$24,909
Mental Health	$5,568	**0.10**	**$4,397**	**-$493**	**$14,175**	**$3,126**
Dementia	$6,355	0.11	$4,725	-$815	$15,805	$1,672
Mental health and tobacco/drug abuse	$4,343	0.08	$3,519	-$307	$11,170	$3,020
Central Nervous System (CNS)	$3,446	**0.05**	**$1,591**	**-$927**	**$6,629**	**$729**
Major disease of the CNS	$4,213	0.10	$5,305	$546	$14,822	$3,155
Eye, ear, other disease of the CNS	$1,484	0.00	-$1,197	-$1,341	-$911	-$955
Respiratory System	$3,815	**0.10**	**$6,553**	**$1,369**	**$16,921**	**$3,535**
Respiratory symptoms, COPD, asthma	$3,828	0.20	$15,789	$5,981	$35,406	$9,397
Infectious disease	$1,082	-0.12	-$13,404	-$7,243	-$25,726	-$8,687
Kidney Disease	$7,286	**0.01**	**-$5,796**	**-$6,541**	**-$4,307**	**-$4,722**
Chronic renal failure or ESRD	-$645	0.13	$13,281	$6,963	$25,917	$8,844
Acute renal failure	$11,009	0.41	$29,783	$9,387	$70,575	$19,183
Other genitourinary disease	$6,520	-0.06	-$12,826	-$9,673	-$19,132	-$9,386
Frailty	$4,168	**0.27**	**$22,615**	**$9,224**	**$49,399**	**$14,458**
Musculoskeletal	$14,172	**0.11**	**-$2,792**	**-$8,482**	**$8,589**	**-$2,061**
Arthritis and musculoskeletal	$14,416	0.17	$2,304	-$6,056	$19,024	$1,355
Injury	$1,375	-0.08	-$9,204	-$5,290	-$17,033	-$6,194
Endocrine, GI, Liver, Hematologic	$10,012	**0.23**	**$13,146**	**$1,567**	**$36,304**	**$6,977**
Other endocrine	$5,525	0.21	$15,553	$5,014	$36,631	$9,287
Gastrointestinal and liver disease	$2,836	0.08	$5,325	$1,245	$13,485	$2,767
Hematologic	$10,616	0.08	-$2,543	-$6,580	$5,531	-$2,963
Miscellaneous	$10,660	**0.03**	**-$7,280**	**-$8,970**	**-$3,900**	**-$5,934**
After care	$14,564	0.06	-$8,486	-$11,525	-$2,407	-$6,685
General symptoms and other disease	$4,365	0.01	-$3,680	-$4,023	-$2,994	-$3,110
Prevention and screening	-$2,557	**0.02**	**$4,261**	**$3,409**	**$5,964**	**$3,024**
Immunizations and infectious screening	-$134	0.01	$949	$541	$1,763	$580
Cancer screening	-$519	-0.03	-$2,437	-$959	-$5,393	-$1,369
Well care	-$4,975	0.05	$10,425	$7,700	$15,876	$7,188

Note: Spending change uses the present value of expected lifetime costs for each disease per person in the population. QALE change uses expected QALE at age 65 for each disease per person in the population. The discount rate is 0% in the columns varying the value of a year of life. The value of a year of life is $100,000 in the column varying the discount rate.

Table 6: Comparison of Mortality Changes for Cardiovascular Diseases to Simulation Models, 1999–2012

Condition	Simulated total decline	IMPACT model Estimate	NHA - Treatment
Ischemic heart disease	-219,721	-129,568	-114,150
Congestive heart failure	-10,560	-41,901	-40,634
Cerebrovascular disease	-68,850	-47,488	-37,545
Cardiovascular risk factors	-2,496	—	-95,813
Total	-300,749	-218,957	-233,004

Note: The simulated total decline in mortality is relative to a counterfactual where age-specific rates were assumed to be constant over time. Estimates from the IMPACT model are based on the most recent adaptation of the model (Ford et al., 2007; Ogata et al., 2019). The NHA estimate is from the analytic work underlying the national health accounts. It includes change in mortality attributed to medical care and not changes in the prevalence of the condition.