Hutchins Center Working Paper #61

May 2020

Productivity Growth in Treating a Major Chronic Health Condition

John A. Romley

University of Southern California

Neeraj Sood

University of Southern California

Dana Goldman

University of Southern California

Abe Dunn

Bureau of Economic Analysis

ABSTRACT

Recent assessments of acute episodes of care point to productivity growth stemming from improvements in patient outcomes. This study assesses productivity growth in the treatment of a major chronic condition, specifically, type 2 diabetes. Analyzing traditional Medicare beneficiaries with new diagnoses over 2004-2012, we find that the productivity of health care improved at an annualized rate of 2.2%. In this context productivity growth translated into only modest improvement in patient outcomes; most of the growth was realized in the form of lower treatment costs. These findings are robust to a range of sensitivity analyses.

THIS PAPER IS ONLINE AT

https://www.brookings.edu/research/productivity-growth-in-treating-a-major-chronic-health-condition/

Contact John Romley at romley@usc.edu. This research was supported by the Productivity Measurement Initiative (PMI) at the Brookings Institution's Hutchins Center on Fiscal & Monetary Policy. We thank participants in the June 2019 PMI conference for helpful comments. Dana Goldman received support from the Schaeffer Center for Health Policy & Economics. The views expressed in this paper are those of the authors and do not necessarily represent the U.S. Bureau of Economic Analysis or the U.S. Department of Commerce. The authors did not receive financial support from any firm or person for this article or from any firm or person with a financial or political interest in this article. None is currently an officer, director, or board member of any organization with an interest in this article.

Hutchins Center on Fiscal & Monetary Policy at BROOKINGS

Economic Studies at BROOKINGS

USC Schaeffer Leonard D. Schaeffer Center for Health Policy & Economics

1. Introduction

Productivity in health care delivery is the subject of renewed attention. Following the literature on geographic variation in health care costs and quality (The Center for the Evaluative Clinical Sciences 1996, Fisher, Wennberg et al. 2003a, Fisher, Wennberg et al. 2003b, Institute of Medicine 2013, Newhouse and Garber 2013a, Newhouse and Garber 2013b), (Romley, Trish et al. 2019) shows that U.S. regions differ markedly in the productivity of acute care for heart attacks. Based on an alternative measure of productivity, (Chandra, Finkelstein et al. 2015) demonstrate that the variation in productivity across U.S. hospitals (also in the treatment of heart attack) is substantial, but not exceptional in comparison to plant-level variability in other industries.[1]

Changes over time in the productivity of health care delivery are also of considerable interest. Health care accounted for 17.7 percent of U.S. GDP in 2018, a share that is projected to rise to 19.4 percent by 2027. (Hartman, Martin et al. 2019, Sisko, Keehan et al. 2019) The trustees of the Medicare program recently assessed that the trust fund supporting hospital insurance (Medicare Part A) will become insolvent in 2026. (OASDI Board of Trustees 2019) In the face of these pressures, the National Academy of Medicine (NAM) has taken the position that "the only sensible way to constrain costs is to increase the value of the health care system, thus extracting more benefit from the dollars spent." (Institute of Medicine 2013, Newhouse and Garber 2013a, Newhouse and Garber 2013b)

The NAM's position is motivated by a common concern that cost containment may come at the expense of quality of care. One broad approach to improving value in the health care system focuses, put somewhat crudely, on demand. For example, value-based insurance design favors high-value services with lower out-of-pocket costs, while the Choosing Wisely Initiative discourages physicians and patients from the use of costly services with limited clinical benefit.

An alternative approach to enhanced value emphasizes supply. Established by the Affordable Care Act (ACA), the Center for Medicare and Medicaid Innovation (CMMI) has been experimenting with a wide array of payment and delivery reforms. Initiatives for bundled payment and accountable care organizations are well-known examples. In addition, the ACA modified the formula by which the Centers for Medicare and Medicaid Services (CMS) update reimbursement rates to health care facilities each fiscal year. Previously, payment rates for hospitals and other facility providers grew according to estimates of inflation in the cost of producing care. Under the revised formula, the existing inflation adjustment is offset by the rate of productivity growth in the American economy as a whole. This policy change recognizes that the previous formula allocated more resources than would have been necessary to maintain quality of care, if providers had been able to improve their productivity.

Yet this policy change effectively asks health care providers to improve their productivity at the same rate as the rest of the U.S. economy. The Medicare Trustees have forecast that productivity in the health care sector will grow over the long term by 0.4 percent per year, substantially more slowly than their forecast for the economy as a whole, specifically, 1.1 percent per year. These forecasts are consistent with the notion that health care (like certain other service industries) suffers from a "cost disease" stemming from a limited flow of cost-reducing technological progress, coupled with a substantial reliance on labor.

. . .

1. (Chandra, Finkelstein et al. 2016) consider a wider range of conditions and procedures and show that market share is positively associated with hospital quality.

(Baumol and Bowen 1965, Newhouse 1992, Baumol, de Ferranti et al. 2012) Indeed, the Bureau of Labor Statistics (BLS) has estimated that productivity among hospitals and nursing and residential care facilities actually declined, by 0.9 percent per year, from 1987 through 2006.

BLS applies a theoretically grounded measurement framework in a consistent manner across industries, measuring output based on deflated revenues. While one would expect the revenues of an automobile manufacturer to embody the quality of its vehicles, it is plausible that the revenues earned by providers in the U.S. health care system, aggregated across payers, have been less reflective of product quality, both for differences across providers and over time. Until recently, a limited number of empirical studies in health economics and policy has estimated productivity growth across providers, while neglecting the issue of quality of care.

(Romley, Goldman et al. 2015) analyze productivity growth at U.S. hospitals in treating heart attack, heart failure and pneumonia over 2002-2011. Ignoring the health outcomes experienced by patients, productivity growth is estimated to be negative for all three conditions, and worst for heart failure at -0.9 percent per year. After accounting for quality of care (measured by patient survival and avoidance of unplanned readmissions), productivity growth is found to be positive and substantial, exceeding a rate of 0.5 percent per year, for all three conditions.

(Gu, Sood et al. 2019) also find that quality adjustment transforms a negative assessment into a positive one, in this case in the context of skilled nursing facilities. Recognizing that health care is delivered by multiple providers and that coordination of care is a major concern in health care practice and policy, (Romley, Dunn et al. 2019) analyze episodes of care for conditions and procedures involving a hospitalization, but including all services covered by A and B spending, with the select conditions studied accounting for more than 10 percent of Part A and B spending. They find that productivity improved for most of the eight episode types studied and that productivity growth tended to manifest itself in the form of better health outcomes.

This evidence is limited to acute care, and broader perspective is needed. Much of the spending in the U.S. health care system is for chronic conditions such as cardiovascular disease and diabetes. Progression of and complications from chronic disease sometimes eventuate in acute care, but high-acuity services are only part of the overall picture of caring for these populations. Moreover, relatively high-tech and invasive interventions are sometimes viewed as the success story of American health care (in terms of effectiveness if not costs) compared to primary care and population health. Indeed, CMMI models that focus on primary care (such as Comprehensive Primary Care Plus) are arguably motivated by, and intended to moderate, this point of view.

It is therefore important to know whether apparent improvements in the productivity of higher-acuity care translate to the health care system more broadly. To the best of our knowledge, this study is a first attempt to address this issue.[2] To do so, we apply the analytic framework from prior studies to elderly fee-for-service Medicare beneficiaries newly diagnosed with type 2 diabetes over the period 2004-2012. (Romley, Goldman et al. 2015, Gu, Sood et al. 2019, Romley, Dunn et al. 2019, Romley, Trish et al. 2019)

. . .

2. To be clear, there is compelling evidence on the performance of the health care system in treating chronic disease, and specifically diabetes (Eggleston, Shah et al. 2009, Eggleston, Shah et al. 2011, Eggleston, Chen et al. 2019). As discussed below, this evidence addresses the question of social value. Our focus here is on the related but distinct matter of the systemic productivity of health care providers.

2. Approach

The starting point for the analysis is a set of CMS medical claims files for a random 20% sample of Medicare beneficiaries for the period 2002-2014. (Research Data Assistance Center) We identify beneficiaries with diabetes by applying a validated and widely used algorithm to diagnosis codes reported in the claims. (Hebert, Geiss et al. 1999, Rector, Wickstrom et al. 2004, Centers for Medicare and Medicaid Services 2020)[3] We limit the resulting sample to cases of type 2 diabetes.[4] We also limit our sample to beneficiaries who were enrolled in fee-for-service (FFS) Medicare at the time of first diagnosis. As Table 1 reports, we identify 2,855,063 cases among the 20% sample of beneficiaries over 2002-2014.

Following the literature, we define a case as newly diagnosed if there were no diabetes claims in the prior six months. (Kim, Schneeweiss et al. 2015, Sohn, Talbert et al. 2015, Kreft, McGuinness et al. 2018) We exclude beneficiaries who were not continuously enrolled in FFS for the six months prior to first diagnosis, as well as beneficiaries who were not continuously enrolled in the twenty-four months after diagnosis (or until time of death, if sooner). Based on these criteria, our sample of new cases was 1,558,394, as shown in Table 1. In sensitivity analysis, we consider alternative "look back" and "follow up" windows (twelve and thirty-six months, respectively).

Because our claims data started in 2002, some of the apparently new cases identified in the early years of the period 2002-2014 may have actually been previously diagnosed existing cases (rather than new diagnoses) based on unobserved diabetes claims prior to 2002. The pattern of diagnoses over time is consistent with this kind of misclassification, as can be seen with the monthly number of new cases in Appendix Figure 1. The number of apparently new cases stabilized as of 2004, and so we limit our sample to the 1,142,747 cases first diagnosed in 2004 or later (see Table 1). In a sensitivity analysis, we include cases diagnosed in 2002 and 2003. Our base analysis considers cases diagnosed between 2004 and 2012, with the latter allowing for sufficient follow up through the end of the claims data (2014).

Following prior research (The Center for the Evaluative Clinical Sciences 1996, Fisher, Wennberg et al. 2003a, Fisher, Wennberg et al. 2003b, Institute of Medicine 2013, Newhouse and Garber 2013a, Newhouse and Garber 2013b, Romley, Trish et al. 2019), we measure quality, cost and productivity at the regional level. Specifically, we use the 306 hospital referral regions first defined by the Dartmouth Atlas of Health Care.(The Center for the Evaluative Clinical Sciences 1996) In a sensitivity analysis, we consider hospital service areas, also from the Atlas. Then, to quantify productivity in treating newly diagnosed diabetes, we estimate the following relationship:

$$(1)\quad \ln(Y_{rt}/C_{rt}) = \alpha + \boldsymbol{S_{rt}\beta_S} + g(t) + \epsilon_{rt},$$

in which Y_{rt} is the total output of cases diagnosed in hospital referral region r during year t, C_{rt} is the total cost delivering care for these cases, and $\boldsymbol{S_{rt}}$ is severity factors for the newly diagnosed patients. The left-hand side of this equation is the ratio of output to inputs, or more colloquially "bang for the buck." This metric is commonly used in economic assessments of health-system performance. (Berndt, Bir et al. 2002, Howard, Bach et al. 2015, Lakdawalla, Shafrin et al. 2015, Sheiner and Malinovskaya 2016, Hult, Jaffe et al. 2018)

. . .

3. The algorithm requires at least one diabetes diagnosis in the Inpatient, Skilled Nursing or Home Health Files, or 2 claims in the Hospital Outpatient or Carrier Files, within a two-year window.

4. Specifically, we exclude cases with a claim whose ICD-9 code ended with a 1 or a 3.

On the right-hand side of the equation, our object of interest is the function $g(t)$, a common-across-regions but year-specific residual between measured determinants of production and measured output. As is conventional, we will interpret this residual as productivity and changes in the residual over time as productivity improvement (or decline). As is well recognized, the validity of this interpretation depends on the validity of the measurement of production determinants and output, and also the functional form of the production function in equation (1).

We measure output in each region-year based on newly diagnosed cases with "high-quality" outcomes during the follow up window. Under this approach, the health care system does not receive full credit (in terms of output) for a low-quality case, yet is still responsible for the cost of scare resources used to treat the case. One metric is the *number* of high-quality cases. That is, in equation (1), $Y_{rt} = N_{rt}Q_{rt}$, in which N_{rt} is the number of cases and Q_{rt} is the proportion that are high-quality. This metric has a natural interpretation, and has been applied in prior studies. (Romley, Goldman et al. 2015, Gu, Sood et al. 2019, Romley, Dunn et al. 2019, Romley, Trish et al. 2019) However, this metric implies that the elasticity of substitution in production between quantity and quality is -1.[5] This elasticity implies that for providers to produce a 10-percent (relative) increase in quality they would need to reduce the number of treated cases (i.e., N_{rt}) by 10 percent. However, evidence on the tradeoff between quantity and quality in health care is remarkably scarce. (Grieco and McDevitt 2016) recently investigated the provision of kidney dialysis services, and their findings imply an elasticity of quantity with respect to quality of -1.4, which implies that quality is more costly to produce.[6] This estimate has been applied in prior research on productivity in treating acute episodes (Romley, Dunn et al. 2019), and we use it here.[7] In sensitivity analysis we also consider a value of -1.

In prior studies we have aggregated multiple dimensions of quality into a single discrete measure of a treatment being successful or not. For example, in analyzing acute episodes of care we considered survival, avoidance of readmission and return to community, to be counted as a successful treatment, and episodes with a death, a readmission or not returning to the community as unsuccessful. A challenge to aggregating these outcomes was that not enough was known about the marginal rate of technical substitution (MRTS) in the production of these distinct health outcomes. Put more simply, we didn't have direct evidence on how health care providers could deliver better survival at the expense (or, potentially, in combination with) avoidance of readmission or return to community. The solution to this problem was to use information on consumer valuation of the outcomes to quantify the trade-off. In well-functioning markets, trade-offs in the production of goods are driven by economic incentives to match how consumers value goods in relative terms (i.e., the MRTS equals the marginal rate of substitution.) While health care suffers from market failures, consumer valuation was used as an approximation.

Specifically, outcomes were valued using the quality-adjustment framework applied in the extensive literature on quality-adjusted life years (QALYs). The QALY score ranges from zero to one, with the former corresponding to death and the latter to perfect health. While evidence on the "quality decrement" from institutionalization was limited, in the case of diabetes there is a well-developed base of evidence on

. . .

5. Fixing Y_{rt} and differentiating with respect to cases and quality, we find that $\frac{\partial \ln N_{rt}}{\partial \ln Q_{rt}} = -1$.

6. (Grieco and McDevitt 2016) report a semi-elasticity of quality with respect to quantity of -0.016 percent, where quality is measured based on the rate of infection. To obtain the elasticity of quality with respect to quantity, we multiply this value with the mean success rate of no infections of 87.5 percent calculated from the study to obtain an elasticity estimate of -1.4.

7. In this case, $Y_{rt} = N_{rt}Q_{rt}^{1.4}$.

quality of life, much of it linked to the United Kingdom Prospective Diabetes Study. (UK Prospective Diabetes Study Group 1998) A person with diabetes but no complications from the disease experiences a QALY score of 0.711.(Currie, Morgan et al. 2006) We follow the diabetes QALY literature in specifying important complications and measuring their impacts on quality of life. For example, amputation involves a decrement of 0.280. For each case, we identify complications using diagnoses reported in claims and calculate a QALY score, then average across new cases in each region-year.[8] Additional details are provided in Appendix Table 1; this QALY-based approach was employed in one of the two studies most closely related to this one. (Eggleston, Shah et al. 2009) This framework for incorporating quality is a version of what has been called the "redefine the good" approach[9], in contrast with the "cost of living" approach. (Hall 2016, Sheiner and Malinovskaya 2016)

The latter was used to develop the heart attack inflation measure referenced previously. (Cutler, McClellan et al. 1998) These two approaches are related but distinct. The cost of living approach determines the compensating variation associated with improved outcomes from treatment. However, (Sheiner and Malinovskaya 2016) note that the rate of productivity change is the relevant metric for assessing whether providers could treat the same number of cases with same quality when their reimbursement rates are reduced, as the Affordable Care Act mandates according to the rate of productivity growth outside the heath-care sector. As with the BLS conceptualization of productivity (Harper, Khandrika et al. 2010), our focus here is on producers / firms.[10] We therefore consider the inputs needed by health care providers to deliver care, and the cost to producers of using scarce inputs to treat patients with diabetes. In addition to consistency with BLS, the focus on producer's cost is more closely aligned to societal cost of treatment, as payment amounts from commercial insurers and public insurers, such as Medicare and Medicaid, may deviate substantially from the actual cost of providing treatment.

The value of additional quality from the firm's perspective may be different from that of consumers. For example, (Grieco and McDevitt 2016) measure the production function from dialysis treatments and found that from a producer's perspective, the opportunity cost of reducing one infection was approximately $75,000. However, the societal benefit of reducing one infection in terms of the value of life years saved was estimated to be $92,000. In this example the value of quality as perceived by consumers, greatly exceeds that of producers. In this case, if a technology existed so that an additional infection could be prevented for a cost of $76,000, consumer would perceive this as a large benefit, but producers would not as $76,000 exceeds their current opportunity cost. Similar to this example, it is possible for our productivity measures to fall, but for consumers to be better off if, for example, they value quality improvements more than producers.

. . .

8. If, hypothetically, half of cases in a region-year involved amputation but no other complications, then $Q_{rt} = 0.711 - 50\% \times 0.28 = 0.571$. In measuring quality, complications are treated additively.

9. When the elasticity of quantity with respect to quality is specified to -1.4, our version places extra weight on quality, based on the evidence described above, in comparison to the standard version of the redefine-the-good approach. In addition, while the approach typically defines success dichotomously, we allow success to be polychotomous according to the quality of life associated with distinct patient outcomes (see above.)

10. (Dauda, Dunn et al. 2019) shows that a cost of living index indicates greater improvement than our approach here when the value of the health improvement exceeds its incremental cost, consistent with other evidence from health care.(Cutler and McClellan 2001) For more on the cost-of-living approach, see the relevant appendix.

The returns to capital, labor and other factors are not of interest, and so we combine the resources used in providing care (see, e.g., (Chandra and Staiger 2007, Doyle 2011, Skinner and Staiger 2015, Chandra, Finkelstein et al. 2016)), aggregating all cases in each region-year. Thus we seek to measure cost, that is, the opportunity cost to producers of allocating scarce resources to the treatment of diabetes cases instead of alternative goods and services. To do so, we identify claims that occurred within the follow up window for each new case, including inpatient (short-term acute-care hospitals but also long-term care hospitals and inpatient rehabilitation facilities), outpatient facilities, professional (e.g., a claim submitted by a doctor for an inpatient surgery or an office visit), skilled nursing facilities (SNFs), home health, durable medical equipment, and hospice.[11] Where a claim in any file did not fall entirely within the follow up window, we allocate costs based on the proportion of days with overlap.

CMS claims do not directly report costs, but instead provide a measure of resource use. For example, total charges are reported for hospital stays. To estimate costs, we use the financial reports that institutional providers participating in Medicare are required to submit to CMS. (Centers for Medicare and Medicaid Services 2019) Hospitals, for example, report not only their actual costs, but the ratio of their charges to their costs (CCRs). So a hospital's cost for a claim is measured by linking reported charges on the claim to the hospital's reported CCR based on Medicare provider number and then multiplying the former by the latter, as is commonly done in the literature. (Cutler and Huckman 2003) SNF cost reports include revenue-to-cost ratios, and so these ratios are multiplied by claim-reported revenues to measure the cost of the claim.[12]

CCRs are sometimes unavailable, and the base analysis excludes cases for which any CCR is missing during the follow up window. As Table 1 shows, this criterion excludes about 20% of cases. In a sensitivity analysis, we also include cases with 1 or more institutional claims that could not be matched to cost data, and whose payments for claims with missing cost data as a share of total payments was less than or equal to the median.[13] We then inflate total measured costs of these cases according payments for claims with missing costs as a share of total payments for all cases that initiated within the same calendar year.

Professional claims report Relative Value Units (RVUs), a measure of the resources needed by a provider to efficiently deliver a particular service.[14] (Medicare Payment and Advisory Committee 2018) The reimbursement received by a professional is equal to the number of RVUs multiplied by a dollar-denominated "conversion factor" (CF) specified annually in CMS's Medicare Physician Fee Schedule Final Rule, adjusted for geographic differences in the cost of care. (Medicare Payment and Advisory Committee 2018) One objective in setting the CF is to ensure that professional providers offer accessible care to beneficiaries, yet federal policy makers have intervened in the CF-setting process to postpone reductions in professional payments mandated by statute for the purpose of controlling cost growth.(Guterman 2014) We assume that the CF in 2004 equated total Medicare payments to professional providers (that is,

. . .

11. These types of claims correspond to the Inpatient, Outpatient, Carrier, Skilled Nursing Facility, Home Health Agency, Durable Medical Equipment, and Hospice Files, respectively.

12. Charges are not in general equal to payments in health care, due, for example, to contractual discounts off list price for commercial insurers as well as administrative pricing for Medicare and other public payers. (Reinhardt 2006)

13. We include payments from all sources.

14. RVUs are updated over time based in part on advice from an American Medical Association committee. (Chan and Dickstein 2018)

revenues received by them) to the total costs to the professional providers of caring for this population in that year. (Romley, Dunn et al. 2019)[15]

The measurement of treatment costs presents challenges. First, some of the care received by patients is unrelated to diabetes. The base analysis includes all costs, whether directly related or not. In a sensitivity analysis, we exclude claims for which diabetes is not the principal diagnosis.

A second challenge is that prescription drugs are commonly used in diabetes treatment. A challenge for this analysis is that Medicare Part D was introduced in 2006, and so prescription drug utilization is unavailable before then. In the base analysis, we exclude prescription drugs from treatment cost. In additional analysis, we limit the sample to individuals continuously enrolled in a standalone Part D plan throughout the follow up window, and include drug costs as measured by total reimbursement (beneficiary as well as plan). Necessarily, the beginning year of this analysis was 2006.

We wish to measure changes over time in the real cost to providers of treating cases, which provides a consistent measure of the quantity of inputs into treatment over time. Accordingly, we adjust our nominal cost estimates for the inflation of inputs using input cost indexes from CMS. Specifically, as an input into its reimbursement policy making, CMS constructs and reports "market basket indices" and the Medicare Economic Index (MEI). The Inpatient Hospital market basket index, for example, measures changes in the cost of providing inpatient hospital care. We use this index and those for other institutional settings to deflate nominal costs into real 2014 dollars. The MEI is used for professional payment, and measures inflation in the cost of providing professional services, less an adjustment for productivity growth in the economy at large. (2012 Medicare Economic Index Technical Advisory Panel 2012) We inflation-adjust professional costs by reversing the productivity adjustments to the MEI; durable medical equipment costs are similarly deflated.

Turning to patient severity ($\boldsymbol{S_{at}}$), we exploit diagnostic information in the data by measuring the proportion of cases in each region-year with 26 comorbidities at the time of diabetes diagnosis, for example, ischemic heart disease as well as stroke / transient ischemic attack (see Table 2 for a complete list). These comorbidities are identified according to validated claims algorithms. (Centers for Medicare and Medicaid Services 2020) In addition, we use average age at diagnosis, and the proportion of patients who were female and of various races, as reported in the Beneficiary Summary Files. These files also report the zip code in which each beneficiary resides, which we link to zip code-level data from the 2000 Census on a variety of community sociodemographic characteristics used as proxies for patient severity in prior literature (Fisher, Wennberg et al. 2003a, Fisher, Wennberg et al. 2003b, Romley, Jena et al. 2011, Romley, Goldman et al. 2015); examples include the poverty rate and the proportion of elderly individuals residing in an institution. As Table 1 shows, about 8,000 of 909,400 cases were diagnosed in region-years in which *none* of the patient zip codes matched to the Census data. To deal with possible seasonality in severity, we include the proportion of cases diagnosed in each calendar month as covariates in equation (1). Finally, it is worth noting that our focus on newly diagnosed cases helps to minimize unmeasured severity.

Our regressions cluster standard errors at the level of the index hospital. For representativeness, regressions weight region-year observations by their number of cases. In further sensitivity analysis, we include fixed effects for the regions in which cases were diagnosed, based on the beneficiary's zip code. This specification aims to deal with the possibility that unmeasured heterogeneity between providers

. . .

15. Ideally, we would like to observe actual practice expenses of physicians, but these are not available. RVUs and the associated CF are the best approximation available.

(including productivity differences) was systematically related to patient severity, leading to bias in estimates of the trajectory of productivity over time.

3. Findings

Before reviewing the regression results, we first describe the cases studied. We focus on diabetes cases diagnosed in January, 2004 or later, in an effort to exclude cases which were identified in our 2002 or 2003 claims but which were first diagnosed prior to then (see Appendix Figure 1). Table 2 reports sample statistics for the 901,472 cases newly diagnosed in our 20% sample of Medicare fee-for-service beneficiaries throughout all 306 hospital referral regions in the U.S. (2,742 region-years). The average date of diagnosis was the second half of 2007. The average cost of treatment during the two years after diagnosis was $21,200 in 2014 dollars, with a standard deviation of $4,500.[16] 85.2% of the newly diagnosed beneficiaries lived beyond the two-year window. Among survivors, the most common complication of diabetes during the follow-up window was ischemic heart disease (2.7% of cases had a relevant diagnosis), while the rarest was blindness (0.001%). Overall, quality of life averaged 0.597 on the 0-1 scale described in the preceding section (a score of 1 indicates perfect health, while diabetes patients without complications have a score of 0.71).

We characterize case severity in terms of beneficiary demographics, existing comorbidities, neighborhood characteristics, and month of diagnosis. The average age at diagnosis was 75.3 years, and slightly more than half of cases (54.3%) involved females. More than four out of five cases (84.3%) involved white beneficiaries. Hypertension was the most common comorbidity (72.0% of beneficiaries with newly diagnosed diabetes had a history of hypertension), followed by hyperlipidemia (55.6%). Endometrial cancer was the rarest comorbidity (0.2%), and lung cancer and hip fracture were the next rarest (1.1%). The average beneficiary resided in a zip code with a poverty rate of 11.2% in the 2000 Census; the standard deviation of the neighborhood poverty rate was 4.2%. On average, 5.3% of elderly residents within the diagnosed beneficiary zip codes were institutionalized, while 28.4% of non-institutionalized residents had a physical disability. The share of diagnosed cases across months ranged from 7.6% (December) to 9.1% (March).

A simple measure of productivity is the cost of a newly diagnosed case of diabetes, without regard to outcomes or severity. Figure 1 shows this measure for cases diagnosed between 2004 and 2012. The cost of treating a beneficiary in the two years after diagnosis was $21,600 in 2004, measured in 2014 dollars. This cost increased from 2004 to 2005 and again from 2007 to 2008, but the overall pattern was one of modest but steady decline, reaching a cost of $19,500 in 2012, that is, 9.9% less than 2004.

This simple measure of productivity ignores the quality of the health outcomes experienced by the newly diagnosed patients. Figure 2 shows that quality of life averaged 0.587 (again on a 0-1 scale) among cases diagnosed in 2004, and that quality of life improved modestly by 2012, reaching an average of 0.600. This improvement reflected better rates of survival, as well as lower rates of complications. The middle panel of Figure 2 shows that survival beyond the two years after diagnosis was 85.3% among cases diagnosed in 2012, up from 84.0% in 2004. The average number of diabetes-related complications among survivors during the follow-up window was 0.054 per case in 2012, down from 0.091 in 2004.

. . .

16. Below we consider alternative windows for the follow up as well as a new diagnosis.

In this context, a productivity measure that reflects patient outcomes is the cost per high-quality case, in which the patient survived and did not experience a diabetes-related complication through the end of the follow-up window.[17] Figure 1 shows that the cost of a high-quality case declined from $26,200 in 2004 to $23,100 in 2012, an 11.9% reduction. The overall pattern is similar for the cost of a case (without regard to quality) due to the modest change in quality of life compared to the change in treatment costs.

These changes in the cost of cases may have reflected trends in the severity of new cases. Figure 3 shows that the age of a newly diagnosed beneficiary averaged 75.8 years in 2004, and that diagnosis age tended to decline over the analysis period, reaching an average of 74.5 in 2012. The middle panel of Figure 3 further shows that the average number of existing comorbidities at diagnosis tended to rise, from 3.64 among new cases in 2004 to 4.03 in 2012. Appendix Table 2 compares the 2012 and 2004 averages for all sample variables, including our severity measures based on demographics, comorbidities, neighborhood characteristics and month of diagnosis. Using all of our severity measures and the results of our baseline regression, we can construct a patient severity index.[18] Figure 3 shows that severity rose from its baseline of 100 in 2004, was flat through 2006 before declining to 96.2 in 2007, then alternately rose and declined through 2012, reaching a value of 104.1 then. One interpretation of this pattern is that treatment productivity would have had to grow by 4.1% in order to provide patients diagnosed in 2012 with the same outcomes at the same cost experienced by patients diagnosed in 2004, due to the greater severity of the former than the latter.

Turning to our regression analyses, the trajectory of estimated treatment productivity is shown in Figure 4. Compared to the initial year of diagnosed cases (2004), productivity decreased by 5.5% in 2005, then rose through 2007, reaching a level 12.9% above baseline. Productivity decreased again in 2008, then rose at an accelerating rate through 2012.

For cases diagnosed in that final year, treatment productivity was 19.1% higher than in 2004. An initial decline followed by ultimately positive growth emerged in prior studies of hospital treatment of cardiac conditions as well as episodes of care for heart attack (Romley, Goldman et al. 2015, Romley, Dunn et al. 2019).[19] Figure 5 shows how the 19.1% in productivity growth as of 2012 was allocated to better quality, lower costs, and greater severity. 10.9% of the 19.1% in productivity improvement manifested itself in the form of lower treatment costs. That is, lower costs accounted for three fifths of the productivity improvement (59.9% = 10.9% / 19.1%). Greater severity consumed the next largest share of the productivity growth, slightly more than one fifth (22.7% = 4.1% / 18.2%). Quality improvement accounted for the smallest share of improvement, slightly less than one fifth of the total (17.4% = 3.2% / 18.2%).

The 19.1% improvement in productivity from 2004 through 2012 corresponds to an annualized growth rate of 2.2%. This rate is substantial, and considerably exceeds the forecast rate for the U.S. economy as a whole from the Medicare Trustees. (OASDI Board of Trustees 2012) This rate also exceeds estimated productivity growth in health care for acute conditions. If the current analysis had ignored improvement in the quality of patient outcomes, estimated growth would have been 1.9% per year instead

. . .

17. That is, a high-quality case corresponds to the quality of life experienced by a diabetes patient living without complications from the disease.

18. The construction of the index is described in the note to Figure 3.

19. Prior research on hospital care for pneumonia did not show a downturn in the mid-2010s.

of 2.2%, as shown in Figure 6. If case severity had been ignored, estimated growth would have been slower still, at 1.6% per year.

For the elasticity of quantity of cases with respect to quality, the baseline value is -1.4, based on the evidence discussed in the preceding section. Recall that this elasticity implies that for providers to increase quality by 10% (relative) with a fixed level of inputs and patient severity, it would require a 14% reduction in the number of cases treated.[20] We also consider an elasticity value of -1.0, consistent with prior studies. (Romley, Goldman et al. 2015, Gu, Sood et al. 2019, Romley, Dunn et al. 2019, Romley, Trish et al. 2019) With this alternative elasticity, a 10% (relative) improvement in case quality would require a 10% reduction in the number of cases treated. In the context of quality improvement, as seen in Figure 2, quality is less costly to deliver, and so the productivity growth needed to sustain an improvement, all else equal, is lesser, resulting in a slower rate of measured growth. With the elasticity assumed to be -1.0, Figure 7 shows that estimated growth is indeed slower, albeit modestly, at 1.9% year.

As noted previously, some studies of incident diabetes have used a look back window of twelve months, and our two-year follow up limits the generalizability of our baseline findings for this chronic long-term disease. Figure 8 compares a twelve-month look back and a three-year follow up to the baseline approach, over the period 2004-2011 for consistency across specifications. In the base analysis, annualized growth was 1.81% over this period. With a twelve-month look back, estimated growth was slightly higher, at 2.03% per year. With a three-year follow up, annualized growth was indistinguishable, at 1.82% instead of 1.81%. Our findings are also insensitive to the inclusion of cases from 2002-2003, including potentially prevalent cases (see Appendix Figure 2).

Figure 9 considers alternative specifications of treatment costs. The top panel limits costs to claims with a diagnosis of diabetes, because many of these individuals have significant co-morbidity burdens. In this specification, productivity growth over 2004-2012 is considerably faster, at 4.1% per year. The middle panel addresses prescription drug costs, because the Part D benefit was introduced in 2006. Over the period 2006-2012, annualized productivity growth under the baseline approach was 2.2% per year. The population of beneficiaries who were continuously enrolled in Part D plans during follow up was systematically different, and productivity growth for these beneficiaries, still excluding drug costs, was larger, at 2.6% per year. When the cost of diabetes drugs for these beneficiaries is included, measured growth was still higher, at 2.9% per year.

As noted previously, cost data is unavailable for some facility claims (20.4% of cases had at least 1 such claim). We assess the sensitivity of estimated productivity growth to the inclusion of cases with some (but relatively limited) missingness, with their measured total costs inflated according to payments on claims with missing costs in comparison total payments for such cases in each year (see notes to Figure 9 for additional detail). The bottom panel of the figure shows that the effect on estimated productivity growth is negligible (2.1% versus 2.2% per year), something also seen in prior analysis of acute episodes of care. (Romley, Dunn et al. 2019)

Figure 10 shows the result of additional sensitivity analyses. In the top panel, we include fixed effects for hospital referral regions. Estimated productivity growth is not sensitive to this alternative (2.3% versus 2.2% per year). Finally, we define areas by subunits of hospital referral regions, namely, hospital service areas. Under this alternative, measured productivity growth was meaningfully faster, at 2.7% versus 2.2% per year.

. . .

20. In this case, $\frac{\partial \ln \quad_{rt}}{\partial \ln Q_{rt}} = -1$, compared to $\frac{\partial \ln N_{rt}}{\partial \ln Q_{rt}} = -1.4$ in the baseline analysis.

4. Conclusion

In this study, we assessed the productivity of health care providers in treating elderly fee-for-service Medicare beneficiaries newly diagnosed with type 2 diabetes between 2004 and 2012. In the base analysis, the rate of productivity growth was substantial, at 2.2% per year. The growth rate remained substantial under a variety of alternative specifications.

Notably, measured growth in diabetes treatment exceeds recent (and relatively favorable) estimates of productivity growth in acute care. For example, (Romley, Dunn et al. 2019) analyzed eight episodes for conditions or procedures that began with a hospitalization, and found growth rates slightly in excess of 1% for two episodes (stroke and chronic obstructive pulmonary disease). In general, this evidence calls into question the notion that health care suffers from a "cost disease," and casts some doubt on forecasts that health care productivity growth will lag that of the broader economy as a whole. (Baumol and Bowen 1965, Newhouse 1992, Baumol, de Ferranti et al. 2012)

To our knowledge, this is the first study that assesses productivity growth among health care providers in treating a major chronic condition. Nevertheless, it is important to recognize that compelling evidence exists with respect to the *social value* of evolving treatments for chronic diseases, and to place the current findings in that broader context. (Cutler and McClellan 2001, Eggleston, Shah et al. 2009, Eggleston, Shah et al. 2011, Grabowski, Lakdawalla et al. 2012) The two most closely related studies to this one are (Eggleston, Shah et al. 2009) and (Eggleston, Shah et al. 2011). Both of these studies analyzed people in a particular health plan offered by a particular employer.[21] A strength of these studies is the availability of medical records together with the use of a high-quality disease model to quantify the benefits of improved care. The studies considered the periods 1997-2005 and 1999-2009, and in both cases found in the aggregate that the economic value of better care dominated the sizable increases in the cost of treatment. Notably, both studies determined that the benefits of improved care were concentrated among individuals with a longer history of diabetes.

The productivity improvement found in this study manifested itself partly in the form of better health outcomes, but mainly as a substantial decline in treatment costs. Specifically, costs decreased in the base analysis by 11.5% over 2004-2012, for an annualized rate of 1.3%. By contrast, recent studies of the delivery of acute care have found flat-to-increasing costs, with productivity improvement manifested in the form of improved outcomes for patients.

The validity of the current findings is an important issue but one that is not straightforward to resolve. Another recent study based on the Medical Expenditure Panel Survey has found decreases (albeit modest) in the cost of diabetes treatment. (Highfill and Bernstein 2019) In addition to (Eggleston, Shah et al. 2009) and (Eggleston, Shah et al. 2011), others have found increasing costs. (Dunn, Whitmire et al. 2018, Squires, Duber et al. 2018, Wamble, Ciarametaro et al. 2019) In general, the estimates are not directly comparable, due to differences in populations, timeframes, and disease stages.

(Squires, Duber et al. 2018) provides some insight into this issue by disaggregating costs, by driver (e.g., prevalence), type of care (e.g., drugs), and patient age. Overall, health care spending on diabetes in the U.S. increased by $64.7 billion from 1996 to 2013, measured in 2015 dollars. Of this total, $24.0 billion resulted from changes in care (service utilization plus spending per encounter). Pharmaceutical

. . .

21. (Eggleston, Chen et al. 2019) recently extended their social value analysis to large populations with diabetes in Asia and Europe. In related studies, (Huang, Zhang et al. 2007) and (Bealieu, Cutler et al. 2006) assess the cost effectiveness of care improvement in federal clinics and a health plan, respectively.

spending accounted for $24.3 billion of this increase; spending on other types of care actually declined slightly. Considering age, spending on the elderly accounted for $24.9 billion of the total increase of $64.7 billion. For the elderly, the spending increase attributable to changes in care was $6.9 billion.

Another reason that our cost trend is not directly comparable is that we sought to measure the producers' cost of treatment, while other studies have focused on health care spending, including on populations with commercial insurance. Outside of Medicare and Medicaid, trends in the prices paid to providers potentially have a weaker relationship with trends in costs. As just one example, (Akosa Antwi, Gaynor Martin et al. 2009) found that the commercial prices paid to California hospitals nearly doubled between 1999 and 2006, but that costly new state regulations (seismic retrofitting and minimum nurse staffing ratios) were not responsible. While this particular study did not link increased prices to growing market concentration, there is a large body of evidence on concentration and prices in health care (Gaynor and Town 2012b, Gaynor and Town 2012a), as well as an ongoing wave of consolidation in the U.S. health care sector. Our focus here is on the cost of resources used in treatment, not the prices paid by purchasers in reimbursing providers, and for this reason our cost trend can diverge from trends in diabetes spending from other studies.

The focus in the current study on a representative sample of individuals with an important chronic condition and coverage from a large national insurance program promises some generalizability of our findings to other contexts. Even so, there is clearly a critical need and opportunity to deepen our understanding of productivity in health care delivery. A worthwhile next step would be to explore how the use of primary and other types of health care have evolved among patients with early as well as advanced diabetes.

REFERENCES

2012 Medicare Economic Index Technical Advisory Panel (2012). Report to the HHS Secretary: Review of the Medicare Economic Index.

Agency for Healthcare Quality and Research (2013). Lower-Extremity Amputation among Patients with Diabetes Rate.

Akosa Antwi, Y., S. Gaynor Martin and B. Vogt William (2009). A Bargain at Twice the Price? California Hospital Prices in the New Millennium. Forum for Health Economics & Policy. **12**.

Baumol, W. J. and W. G. Bowen (1965). "On the Performing Arts: The Anatomy of Their Economic Problems." The American Economic Review **55**(1/2): 495-502.

Baumol, W. J., D. de Ferranti, M. Malach, M. Pablos, xe, A. ndez, H. Tabish and L. G. Wu (2012). The Cost Disease: Why Computers Get Cheaper and Health Care Doesn't, Yale University Press.

Bealieu, N., D. M. Cutler, K. Ho, G. Isham, T. Lindquist, A. Nelson and P. O'Connor (2006). "The Business Case for Diabetes Disease Management for Managed Care Organizations." Forums for Health Economics & Policy **9**(1).

Berndt, E. R., A. Bir, S. H. Busch, R. G. Frank and S.-L. T. Normand (2002). "The medical treatment of depression, 1991–1996: productive inefficiency, expected outcome variations, and price indexes." Journal of Health Economics **21**(3): 373-396.

Centers for Medicare and Medicaid Services. (2019). "Hospital Cost Report Public Use File." from https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Medicare-Provider-Cost-Report/HospitalCostPUF.html.

Centers for Medicare and Medicaid Services. (2020). "Chronic Conditions Data Warehouse." from https://www2.ccwdata.org/web/guest/home/.

Chan, D. and M. Dickstein (2018). Industry Input in Policymaking: Evidence from Medicare. National Bureau of Economic Research.

Chandra, A., A. Finkelstein, A. Sacarny and C. Syverson (2015). Health Care Exceptionalism? Performance and Allocation in the US Health Care Sector, National Bureau of Economic Research,.

Chandra, A., A. Finkelstein, A. Sacarny and C. Syverson (2016). "Health Care Exceptionalism? Performance and Allocation in the US Health Care Sector." American Economic Review **106**(8): 2110-2144.

Chandra, A. and D. O. Staiger (2007). "Productivity Spillovers in Health Care: Evidence from the Treatment of Heart Attacks." Journal of Political Economy **115**(1): 103-140.

Currie, C. J., C. L. Morgan, C. D. Poole, P. Sharplin, M. Lammert and P. McEwan (2006). "Multivariate models of health-related utility and the fear of hypoglycaemia in people with diabetes." Current Medical Research and Opinion **22**(8): 1523-1534.

Cutler, D. M. and R. S. Huckman (2003). "Technological development and medical productivity: the diffusion of angioplasty in New York state." Journal of Health Economics **22**(2): 187-217.

Cutler, D. M. and M. McClellan (2001). "Is technological change in medicine worth it?" Health Aff (Millwood) **20**(5): 11-29.

Cutler, D. M., M. McClellan, J. P. Newhouse and D. Remler (1998). "Are Medical Prices Declining? Evidence from Heart Attack Treatments." The Quarterly Journal of Economics **113**(4): 991-1024.

Dauda, S., A. Dunn and A. Hall (2019). Are Medical Care Prices Still Declining? A Systematic Examination of Quality-Adjusted Price Index Alternatives for Medical Care.

Doyle, J. J. (2011). "Returns to Local-Area Health Care Spending: Evidence from Health Shocks to Patients Far from Home." American Economic Journal: Applied Economics **3**(3): 221-243.

Dunn, A., B. Whitmire, A. Batch, L. Fernando and L. Rittmueller (2018). "High Spending Growth Rates For Key Diseases In 2000–14 Were Driven By Technology And Demographic Factors." Health Affairs **37**(6): 915-924.

Eggleston, K., B. Chen, C.-H. Chen, Y. Chen, T. Feenstra, T. Iizuka, J. Lam, G. Leung, J. Lu, B. Rodriguez-Sanchez, J. Struijs, J. Quan and J. P. Newhouse (2019). Are Quality-Adjusted Medical Prices Declining for Chronic Disease? Evidence from Diabetes Care in Four Health Systems, National Bureau of Economic Research,.

Eggleston, K. N., N. D. Shah, S. A. Smith, E. R. Berndt and J. P. Newhouse (2011). "Quality Adjustment for Health Care Spending on Chronic Disease: Evidence from Diabetes Treatment, 1999-2009." American Economic Review **101**(3): 206-211.

Eggleston, K. N., N. D. Shah, S. A. Smith, A. E. Wagie, A. R. Williams, J. H. Grossman, E. R. Berndt, K. H. Long, R. Banerjee and J. P. Newhouse (2009). "The Net Value of Health Care for Patients With Type 2 Diabetes, 1997 to 2005." Annals of Internal Medicine **151**(6): 386-393.

Fisher, E. S., D. E. Wennberg, T. A. Stukel, D. J. Gottlieb, F. L. Lucas and E. L. Pinder (2003a). "The implications of regional variations in Medicare spending. Part 1: the content, quality, and accessibility of care." Ann Intern Med **138**(4): 273-287.

Fisher, E. S., D. E. Wennberg, T. A. Stukel, D. J. Gottlieb, F. L. Lucas and E. L. Pinder (2003b). "The implications of regional variations in Medicare spending. Part 2: health outcomes and satisfaction with care." Ann Intern Med **138**(4): 288-298.

Gaynor, M. and R. Town (2012a). Competition in Health Care Markets. Handbook of health economics, volume 2. M. V. Pauly, T. McGuire and P. Barros, Elsevier**:** 501-637.

Gaynor, M. and R. Town (2012b). The impact of hospital consolidation—Update. The Synthesis Project.

Grabowski, D., D. N. Lakdawalla, D. Goldman, M. Eber, L. Liu, T. Abdelgawad, A. Kuznik, M. E. Chernew and T. Philipson (2012). "The Large Social Value Resulting From Use Of Statins Warrants Steps To Improve Adherence And Broaden Treatment." Health Affairs **31**(10): 2276-2285.

Grieco, P. L. E. and R. C. McDevitt (2016). "Productivity and Quality in Health Care: Evidence from the Dialysis Industry." The Review of Economic Studies **84**(3): 1071-1105.

Gu, J., N. Sood, A. Dunn and J. Romley (2019). "Productivity growth of skilled nursing facilities in the treatment of post-acute-care-intensive conditions." PLOS ONE **14**(4): e0215876.

Guterman, S. (2014). "The “Doc Fix” — Another Missed Opportunity." New England Journal of Medicine **370**(24): 2261-2263.

Hall, A. E. (2016). "Adjusting the Measurement of the Output of the Medical Sector for Quality: A Review of the Literature." Medical Care Research and Review **74**(6): 639-667.

Harper, M., B. Khandrika, R. Kinoshita and S. Rosenthal (2010). "Nonmanufacturing industry contributions to multifactor productivity, 1987–2006." Monthly Labor Review.

Hartman, M., A. B. Martin, J. Benson and A. Catlin (2019). "National Health Care Spending In 2018: Growth Driven By Accelerations In Medicare And Private Insurance Spending." Health Affairs **39**(1): 8-17.

Hebert, P. L., L. S. Geiss, E. F. Tierney, M. M. Engelgau, B. P. Yawn and A. M. McBean (1999). "Identifying Persons with Diabetes Using Medicare Claims Data." American Journal of Medical Quality **14**(6): 270-277.

Highfill, T. and E. Bernstein (2019). "Using disability adjusted life years to value the treatment of thirty chronic conditions in the U.S. from 1987 to 2010: a proof of concept." International Journal of Health Economics and Management **19**(3): 449-466.

Howard, D. H., P. B. Bach, E. R. Berndt and R. M. Conti (2015). "Pricing in the Market for Anticancer Drugs." Journal of Economic Perspectives **29**(1): 139-162.

Huang, E. S., Q. Zhang, S. E. S. Brown, M. L. Drum, D. O. Meltzer and M. H. Chin (2007). "The Cost-Effectiveness of Improving Diabetes Care in U.S. Federally Qualified Community Health Centers." Health Services Research **42**(6p1): 2174-2193.

Hult, K. J., S. Jaffe and T. J. Philipson (2018). "How Does Technological Change Affect Quality-Adjusted Prices in Health Care? Systematic Evidence from Thousands of Innovations." American Journal of Health Economics **4**(4): 433-453.

Institute of Medicine (2013). Variation in Health Care Spending: Target Decision Making, Not Geography. Washington, DC, The National Academies Press.

Kim, S. C., S. Schneeweiss, R. J. Glynn, M. Doherty, A. B. Goldfine and D. H. Solomon (2015). "Dipeptidyl peptidase-4 inhibitors in type 2 diabetes may reduce the risk of autoimmune diseases: a population-based cohort study." Annals of the Rheumatic Diseases **74**(11): 1968-1975.

Kreft, D., M. B. McGuinness, G. Doblhammer and R. P. Finger (2018). "Diabetic retinopathy screening in incident diabetes mellitus type 2 in Germany between 2004 and 2013 - A prospective cohort study based on health claims data." PLoS One **13**(4): e0195426.

Lakdawalla, D. N., J. Shafrin, C. Lucarelli, S. Nicholson, Z. Khan and T. Philipson (2015). "Quality-Adjusted Cost Of Care: A Meaningful Way To Measure Growth In Innovation Cost Versus The Value Of Health Gains." Health Affairs **34**(4): 555-561.

Leal, J., A. J. Hayes, A. M. Gray, R. R. Holman and P. M. Clarke (2013). "Temporal Validation of the UKPDS Outcomes Model Using 10-Year Posttrial Monitoring Data." Diabetes Care **36**(6): 1541-1546.

McEwan, P., M. Evans and K. Bergenheim (2010). "A population model evaluating the costs and benefits associated with different oral treatment strategies in people with type 2 diabetes." Diabetes, Obesity and Metabolism **12**(7): 623-630.

Medicare Payment and Advisory Committee (2018). Physician and other health professional payment system.

Newhouse, J. P. (1992). "Medical Care Costs: How Much Welfare Loss?" Journal of Economic Perspectives **6**(3): 3-21.

Newhouse, J. P. and A. M. Garber (2013a). "Geographic variation in health care spending in the United States: insights from an Institute of Medicine report." JAMA **310**(12): 1227-1228.

Newhouse, J. P. and A. M. Garber (2013b). "Geographic variation in Medicare services." N Engl J Med **368**(16): 1465-1468.

OASDI Board of Trustees (2012). The 2012 Annual Report of the Boards of Trustees of the Federal Hospital Insurance Trust Fund and the Federal Supplementary Medical Insurance Trust Fund.

OASDI Board of Trustees (2019). The 2019 Annual Report of the Boards of Trustees of the Federal Hospital Insurance Trust Fund and the Federal Supplementary Medical Insurance Trust Fund.

Rector, T. S., S. L. Wickstrom, M. Shah, N. Thomas Greeenlee, P. Rheault, J. Rogowski, V. Freedman, J. Adams and J. J. Escarce (2004). "Specificity and Sensitivity of Claims-Based Algorithms for Identifying Members of Medicare+Choice Health Plans That Have Chronic Medical Conditions." Health Services Research **39**(6p1): 1839-1858.

Reinhardt, U. (2006). "The Pricing Of U.S. Hospital Services: Chaos Behind A Veil Of Secrecy." Health Affairs **25**(1): 57-69.

Research Data Assistance Center. "Inpatient (Fee-for-Service)." from https://www.resdac.org/cms-data/files/ip-ffs.

Romley, J., A. Dunn, D. Goldman and N. Sood (2019). Quantifying Productivity Growth in the Delivery of Important Episodes of Care Within the Medicare Program Using Insurance Claims and Administrative Data, National Bureau of Economic Research,.

Romley, J., E. Trish, D. Goldman, M. Beeuwkes Buntin, Y. He and P. Ginsburg (2019). "Geographic variation in the delivery of high-value inpatient care." PLOS ONE **14**(3): e0213647.

Romley, J. A., D. P. Goldman and N. Sood (2015). "US hospitals experienced substantial productivity growth during 2002-11." Health Aff (Millwood) **34**(3): 511-518.

Romley, J. A., C. Gong, A. B. Jena, D. P. Goldman, B. Williams and A. Peters (2015). "Association between use of warfarin with common sulfonylureas and serious hypoglycemic events: retrospective cohort analysis." BMJ **351**.

Romley, J. A., A. B. Jena and D. P. Goldman (2011). "Hospital spending and inpatient mortality: evidence from California: an observational study." Ann Intern Med **154**(3): 160-167.

Sheiner, L. and A. Malinovskaya (2016). Measuring Productivity Growth in Healthcare: An Analysis of the Literature, Brookings Institution.

Sisko, A. M., S. P. Keehan, J. A. Poisal, G. A. Cuckler, S. D. Smith, A. J. Madison, K. E. Rennie and J. C. Hardesty (2019). "National Health Expenditure Projections, 2018–27: Economic And Demographic Trends Drive Spending And Enrollment Growth." Health Affairs **38**(3): 491-501.

Skinner, J. and D. Staiger (2015). "Technology Diffusion and Productivity Growth in Health Care." Review of Economics and Statistics **97**(5): 951-964.

Sohn, M., J. Talbert, K. Blumenschein and D. C. Moga (2015). "Atypical antipsychotic initiation and the risk of type II diabetes in children and adolescents." Pharmacoepidemiology and Drug Safety **24**(6): 583-591.

Squires, E., H. Duber, M. Campbell, J. Cao, A. Chapin, C. Horst, Z. Li, T. Matyasz, A. Reynolds, I. B. Hirsch and J. L. Dieleman (2018). "Health Care Spending on Diabetes in the U.S., 1996–2013." Diabetes Care **41**(7): 1423-1431.

The Center for the Evaluative Clinical Sciences (1996). The Dartmouth Atlas of Health Care in the United States.

UK Prospective Diabetes Study Group (1998). "Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes." The Lancet **352**(9131): 837-853.

Wamble, D., M. Ciarametaro, K. Houghton, M. Ajmera and R. W. Dubois (2019). "What's Been The Bang For The Buck? Cost-Effectiveness Of Health Care Spending Across Selected Conditions In The US." Health Affairs **38**(1): 68-75.

Whitmer, R. A., A. J. Karter, K. Yaffe, C. P. Quesenberry and J. V. Selby (2009). "Hypoglycemic Episodes and Risk of Dementia in Older Patients With Type 2 Diabetes Mellitus." JAMA **301**(15): 1565-1572.

TABLES AND FIGURES

Table 1: Sample Construction

Cases	*Hospital referral regions*	*Description*
2,855,063	306	All type 2 diabetes cases in Medicare FFS, 2002-2014, based on random 20% sample of beneficiaries
1,558,394	306	Excluding cases lacking continuous enrollment before incident diagnosis, or after diagnosis without death
1,142,747	306	Excluding potentially prevalent cases (2002-2003)
909,442	306	Excluding cases with any missing cost-to-charge ratios
901,472	306	Excluding cases with no Census sociodemographic data available

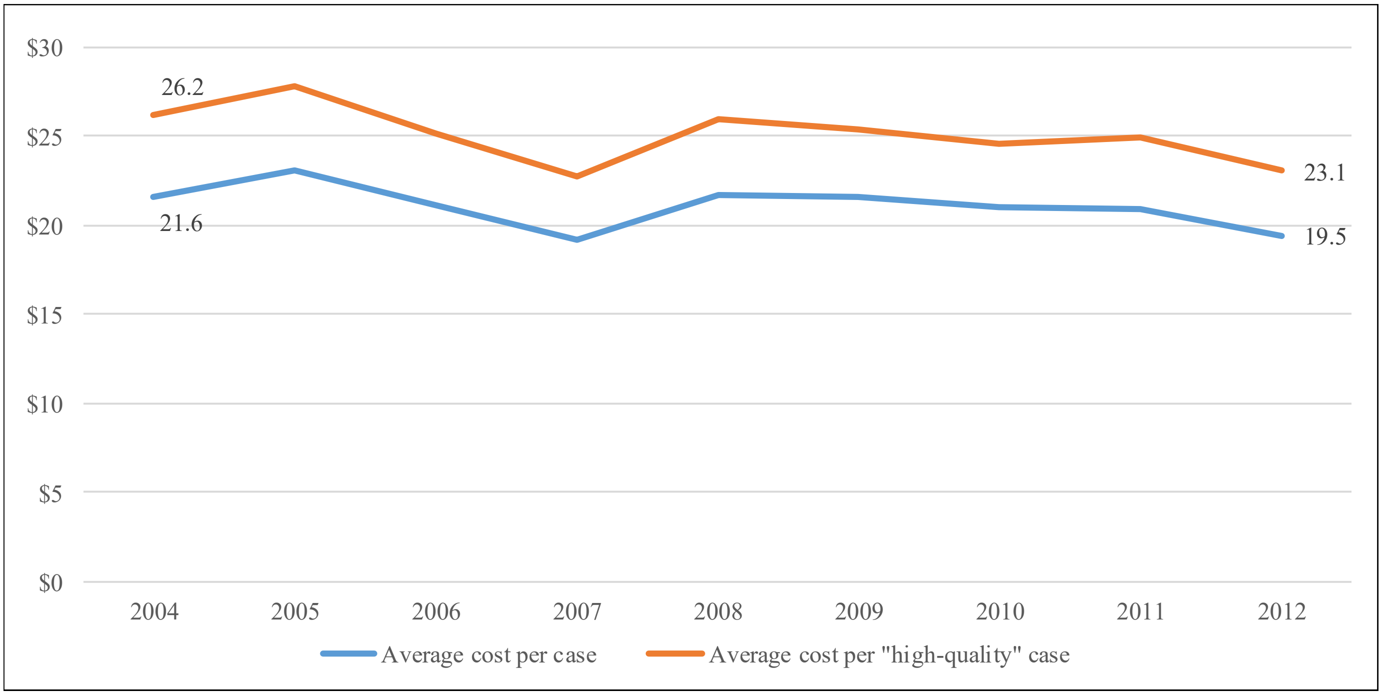
Table 2: Sample Statistics for Newly Diagnosed Diabetes Cases

Variable	Mean (SD)
Sample	
Cases, n	901,472
Areas, n	306
Area-years, n	2,742
Year of diagnosis	2007.7 (2.6)
Outcomes within 2 years of diabetes diagnosis	
Cost (000s of 2014 dollars)	$21.2 ($4.5)
Survival	85.2% (3.2%)
Amputation *among survivors*	0.03% (0.10%)
Blindness	0.001% (0.015%)
End stage renal disease	0.6% (0.5%)
Heart attack	1.3% (0.7%)
Heart failure	2.0% (1.0%)
Hypoglycemia	0.02% (0.08%)
Ischemic heart disease	2.7% (1.5%)
Stroke	1.8% (0.9%)
Quality of life	0.597 (0.023)
Patient demographics	
Age	75.3 (1.1)
Female	54.3% (4.1%)
White	84.3% (11.6%)
African American	9.1% (8.3%)
Hispanic	1.8% (3.5%)
Other race	4.7% (7.4%)
Patient comorbidities prior to diabetes diagnosis	
Acquired hypothyroidism	8.9% (2.4%)
Alzheimer's disease	5.6% (2.1%)
Alzheimer's related disorders or senile dementia	11.7% (3.2%)
Anemia	30.6% (10.7%)
Asthma	4.7% (1.4%)
Atrial fibrillation	10.4% (2.3%)
Benign prostatic hyperplasia	6.9% (2.4%)
Cancer, breast	2.9% (1.0%)
Cancer, colorectal	1.6% (0.7%)
Cancer, endometrial	0.2% (0.3%)
Cancer, lung	1.1% (0.6%)
Cancer, prostate	4.0% (1.2%)
Cataract	21.1% (4.0%)
Chronic kidney disease	12.9% (3.9%)
Chronic obstructive pulmonary disease	14.1% (3.3%)
Depression	10.3% (3.0%)
Glaucoma	10.6% (2.8%)
Heart attack	1.5% (0.7%)
Heart failure	20.9% (4.4%)
Hip fracture	1.1% (0.6%)
Hyperlipidemia	55.6% (8.6%)
Hypertension	72.0% (5.1%)
Ischemic heart disease	39.5% (7.6%)
Osteoporosis	6.7% (2.5%)
Rheumatoid arthritis / osteoarthritis	28.1% (5.3%)
Stroke / transient ischemic attack	5.8% (1.6%)

Table 2, Continued: Sample Statistics for Newly Diagnosed Diabetes Cases

Variable	Mean (SD)
Patient zip code characteristics	
Median household income ($000)	$46.2 ($10.3)
Social Security income ($000)	$11.5 ($0.8)
Poor	11.2% (4.2%)
Employed	94.4% (1.6%)
Less than high school education	19.0% (5.2%)
Urban	78.3% (17.3%)
Hispanic	10.2% (12.4%)
Single	42.0% (3.6%)
Elderly in an institution	5.3% (1.3%)
Non-institutionalized elderly with physical disability	28.4% (3.8%)
Mental disability	10.7% (2.3%)
Sensory disability among elderly	14.0% (1.9%)
Self-care disability	9.5% (1.9%)
Difficulty going-outside-the-home disability	20.3% (2.8%)
Month of diagnosis	
January	8.5% (1.6%)
February	8.1% (1.6%)
March	9.1% (1.7%)
April	8.7% (1.6%)
May	8.6% (1.6%)
June	8.5% (1.6%)
July	8.1% (1.6%)
August	8.3% (1.6%)
September	8.1% (1.6%)
October	8.5% (1.6%)
November	7.9% (1.6%)
December	7.6% (1.5%)

Figure 1: Cost of Diabetes Cases (000s of 2014 dollars), by Year of Diagnosis



Notes: In this figure, a "high-quality" case means that the patient lived and avoided any complication from diabetes during the twenty-four months after diagnosis. That is, a diagnosed patient maintains a quality of life equivalent to that of diabetes without complications. See Figure 3 for more information.

Figure 2: Case Quality, by Year of Diagnosis

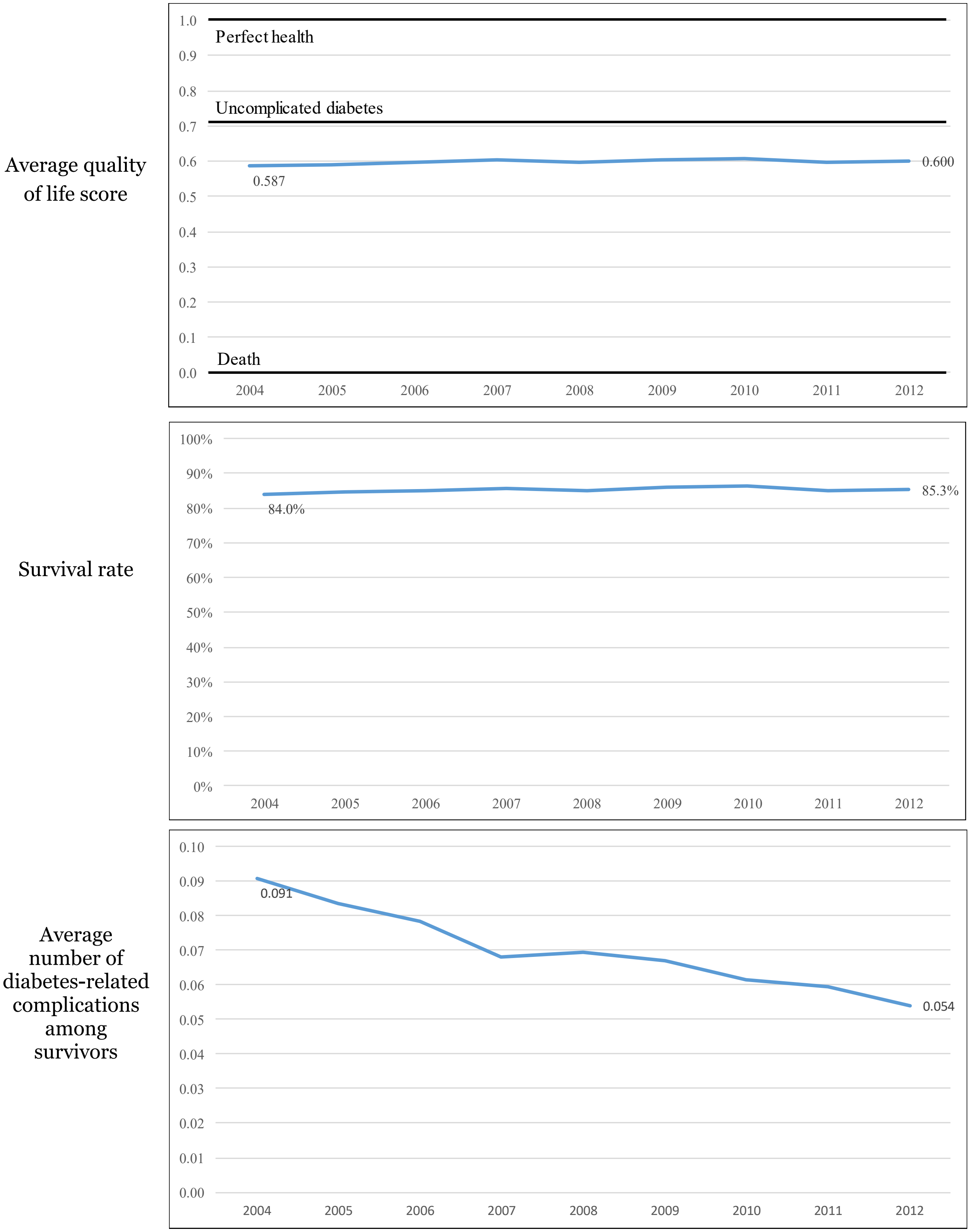
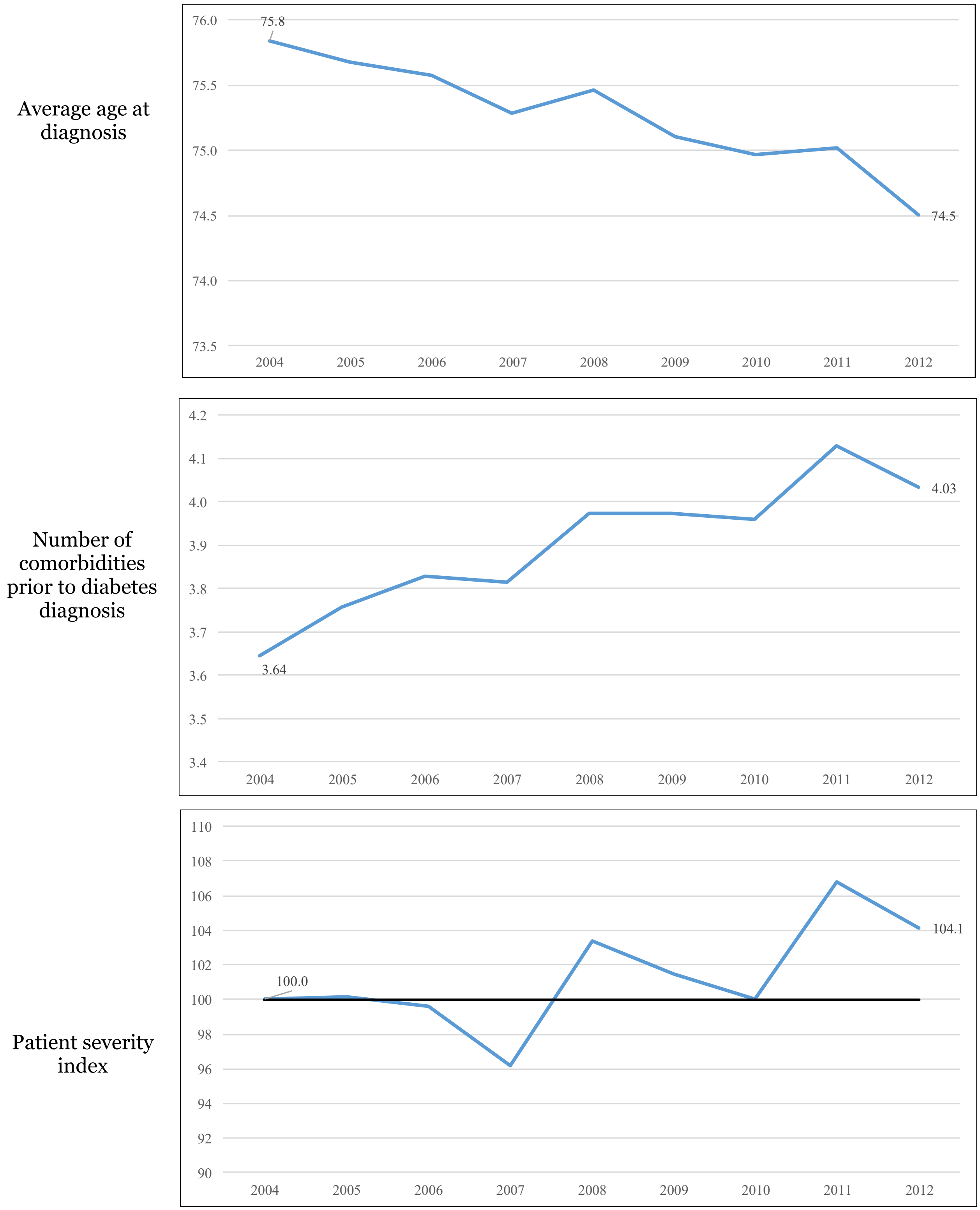


Figure 3: Select Patient Severity Measures, by Year of Diagnosis



Note: I construct the patient severity index by exponentiating $-\bar{S}_{ht}\hat{\beta}_S$, obtaining $\hat{\beta}_S$ from the regression results corresponding to Figure 4 and normalizing the index to a value of 100 in 2004.

Figure 4: Cumulative Change in Productivity Since 2004

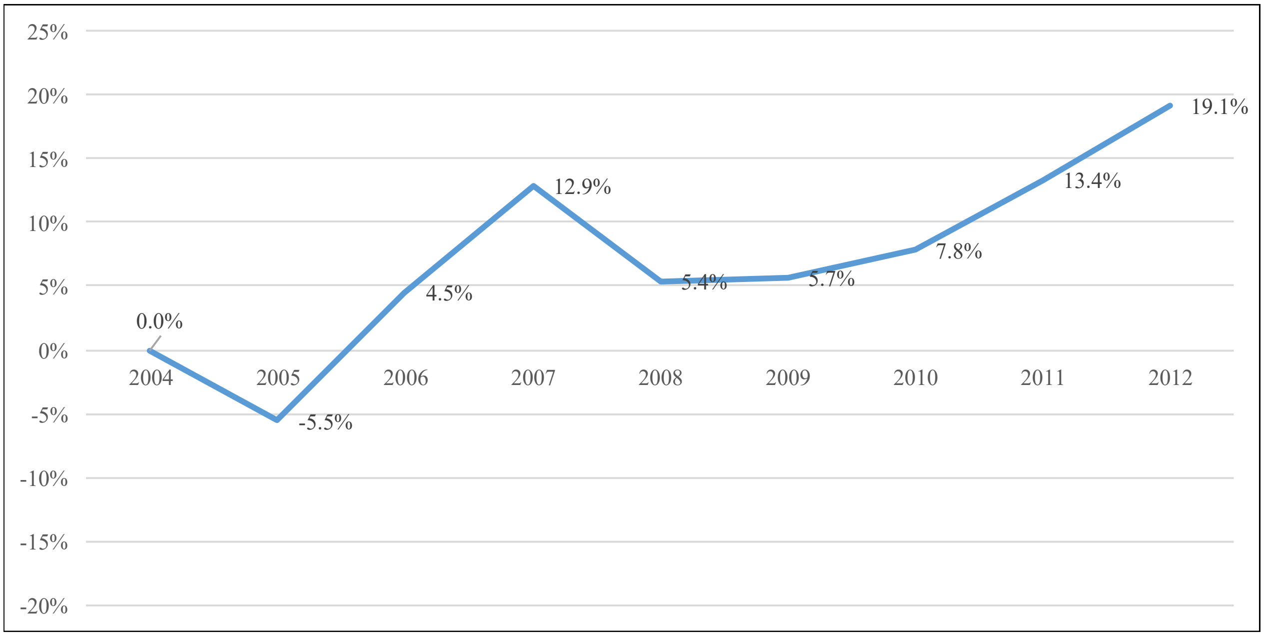
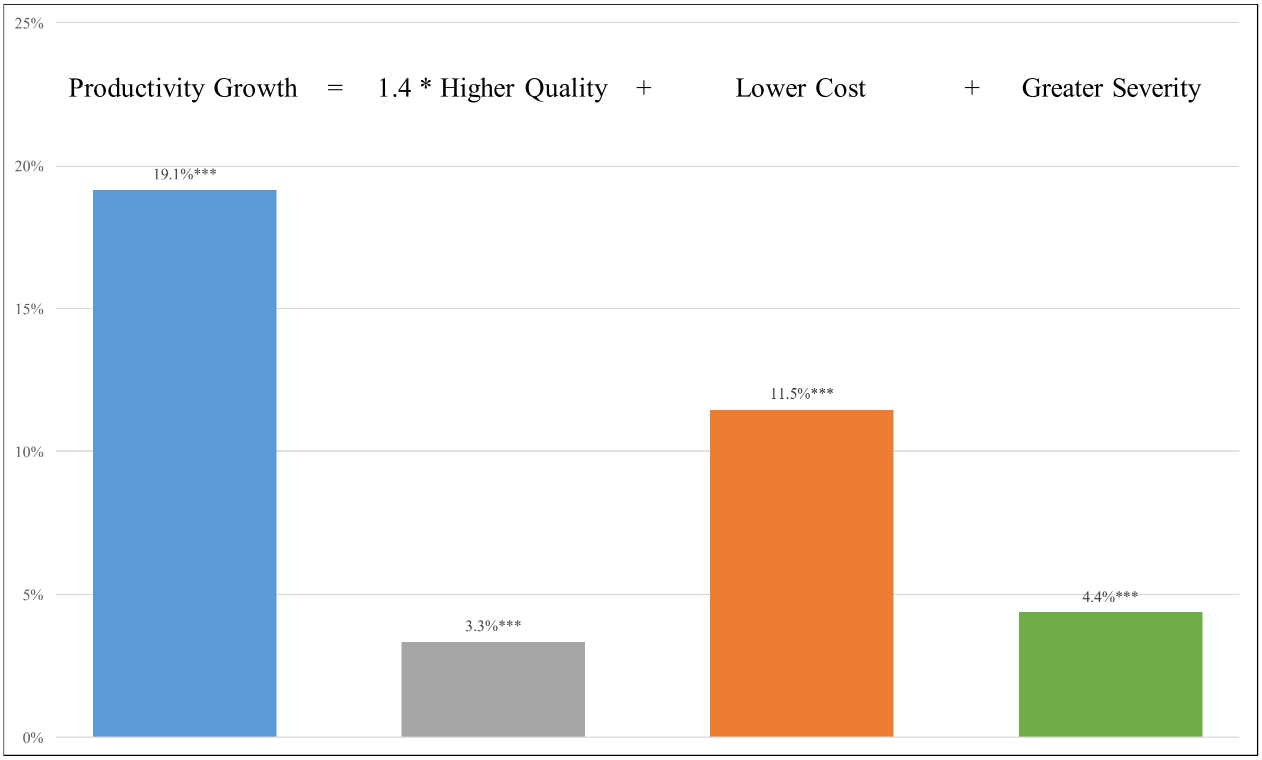
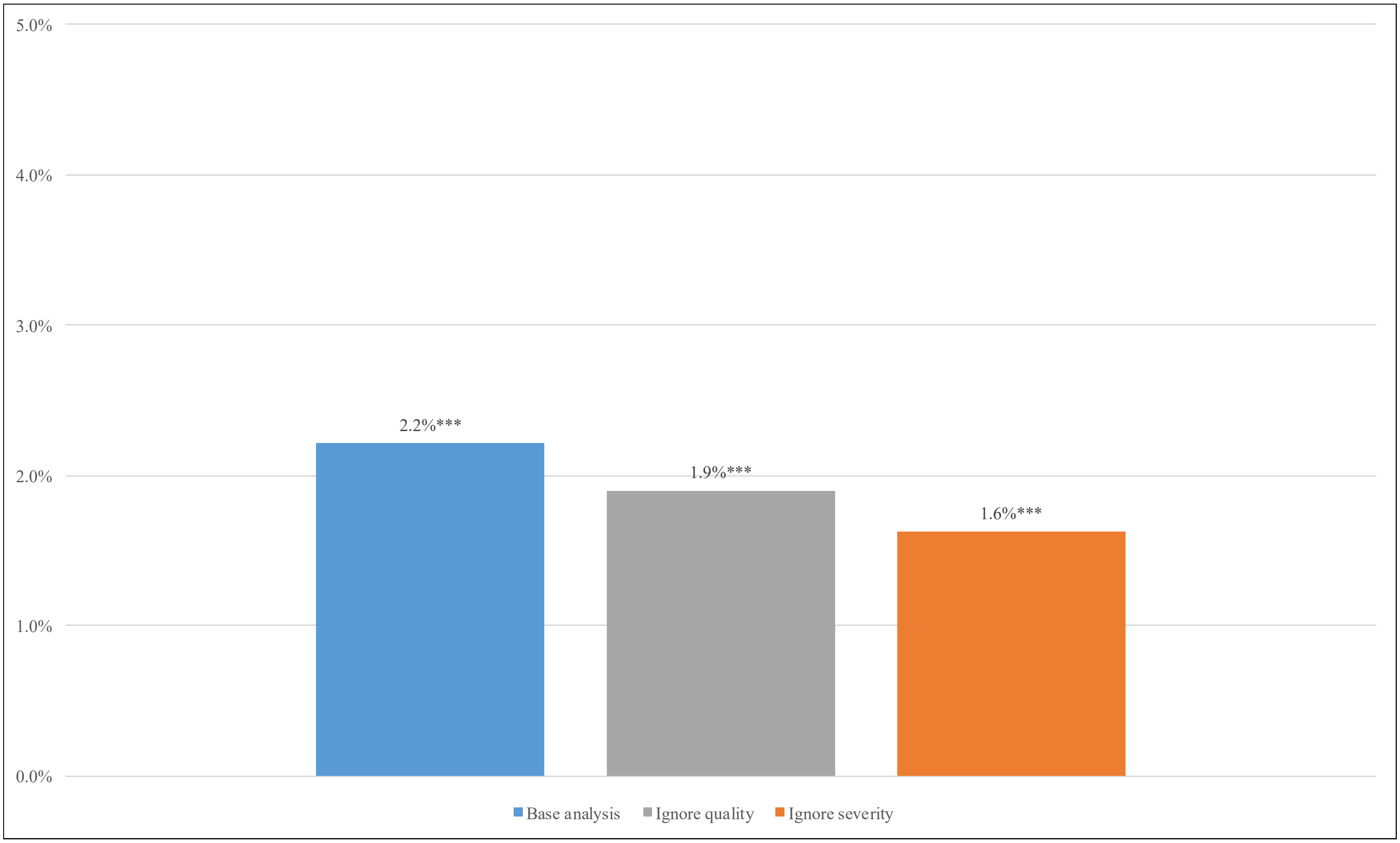


Figure 5: Allocation of Productivity Growth



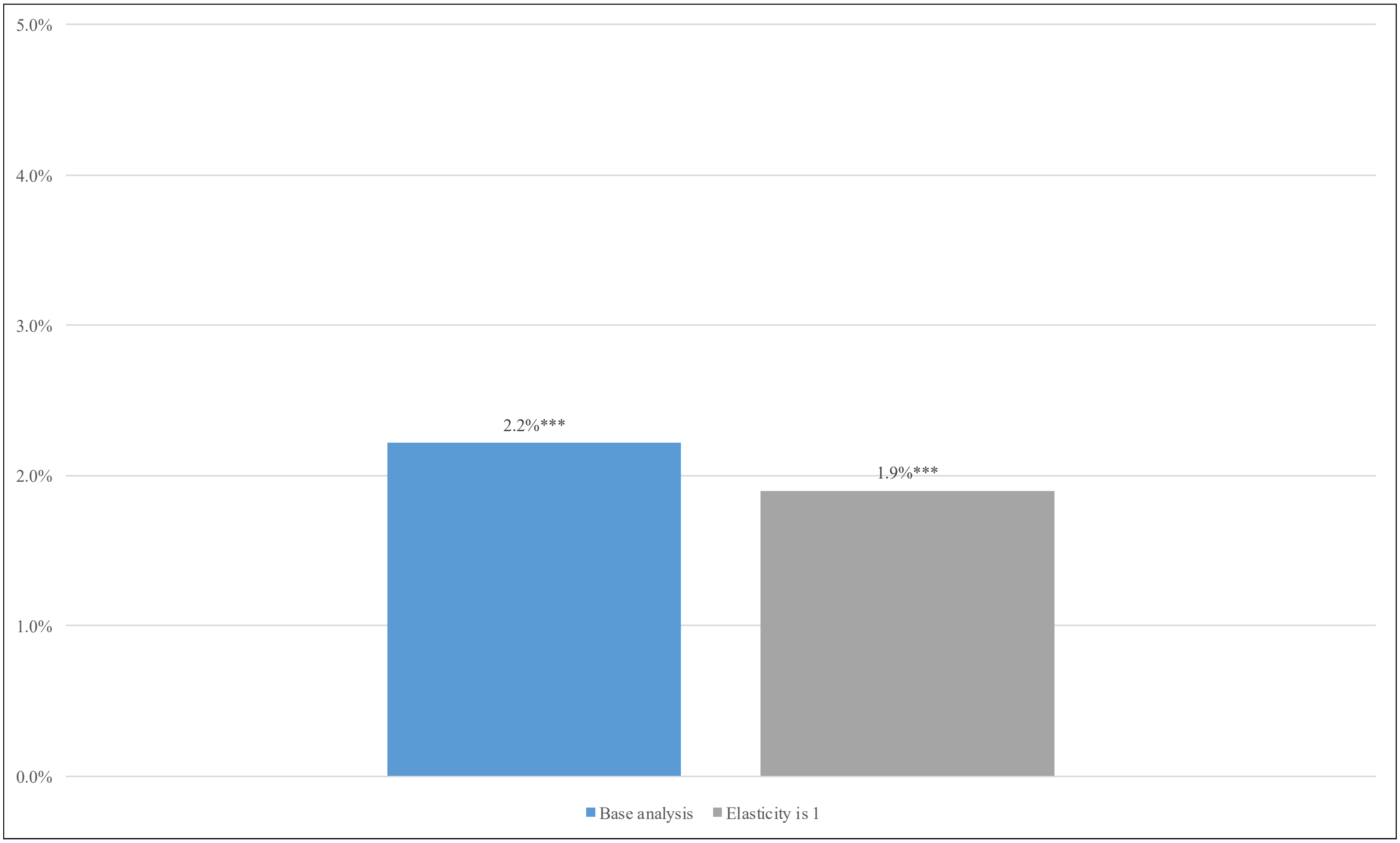
Note: *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Figure 6: Impacts of Adjustment for Outcome Quality and Patient Severity on Annualized Productivity Growth Estimates



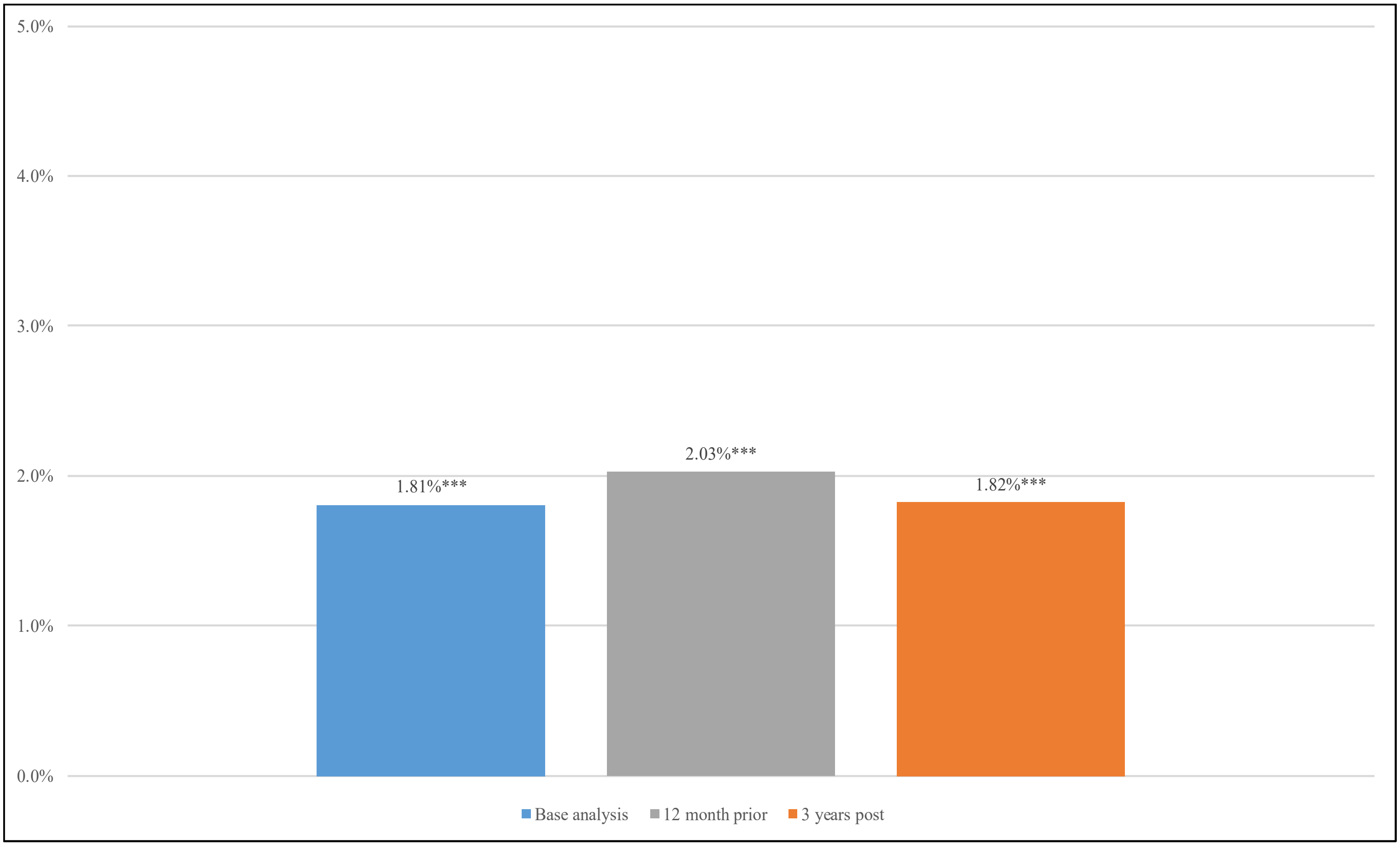
Note: *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively

Figure 7: Sensitivity of Annualized Productivity Growth Estimates to Elasticity of Case Quality with Respect to Quantity



Note: *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively

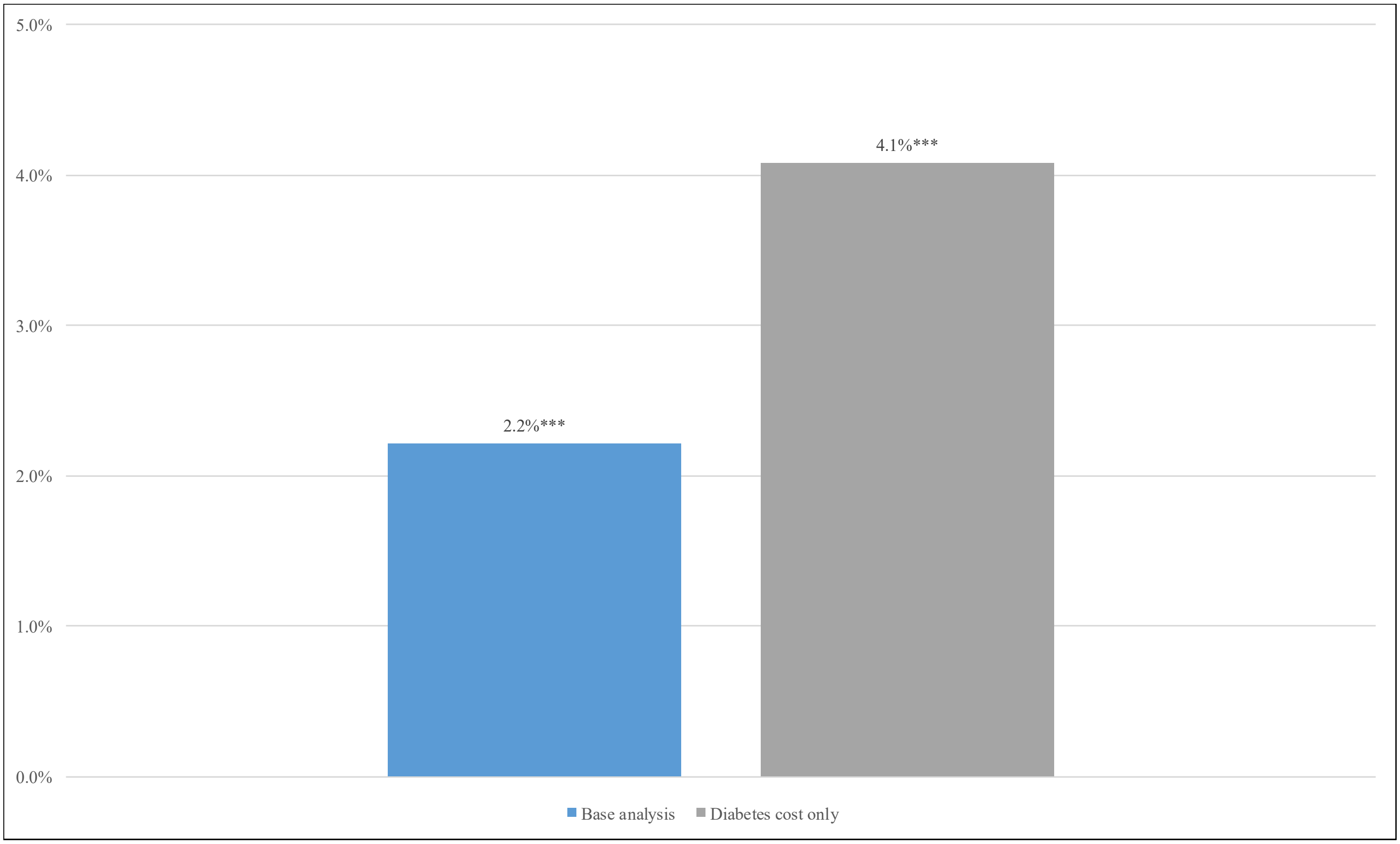
Figure 8: Alternative Case Windows, 2004-2011



Note: *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively

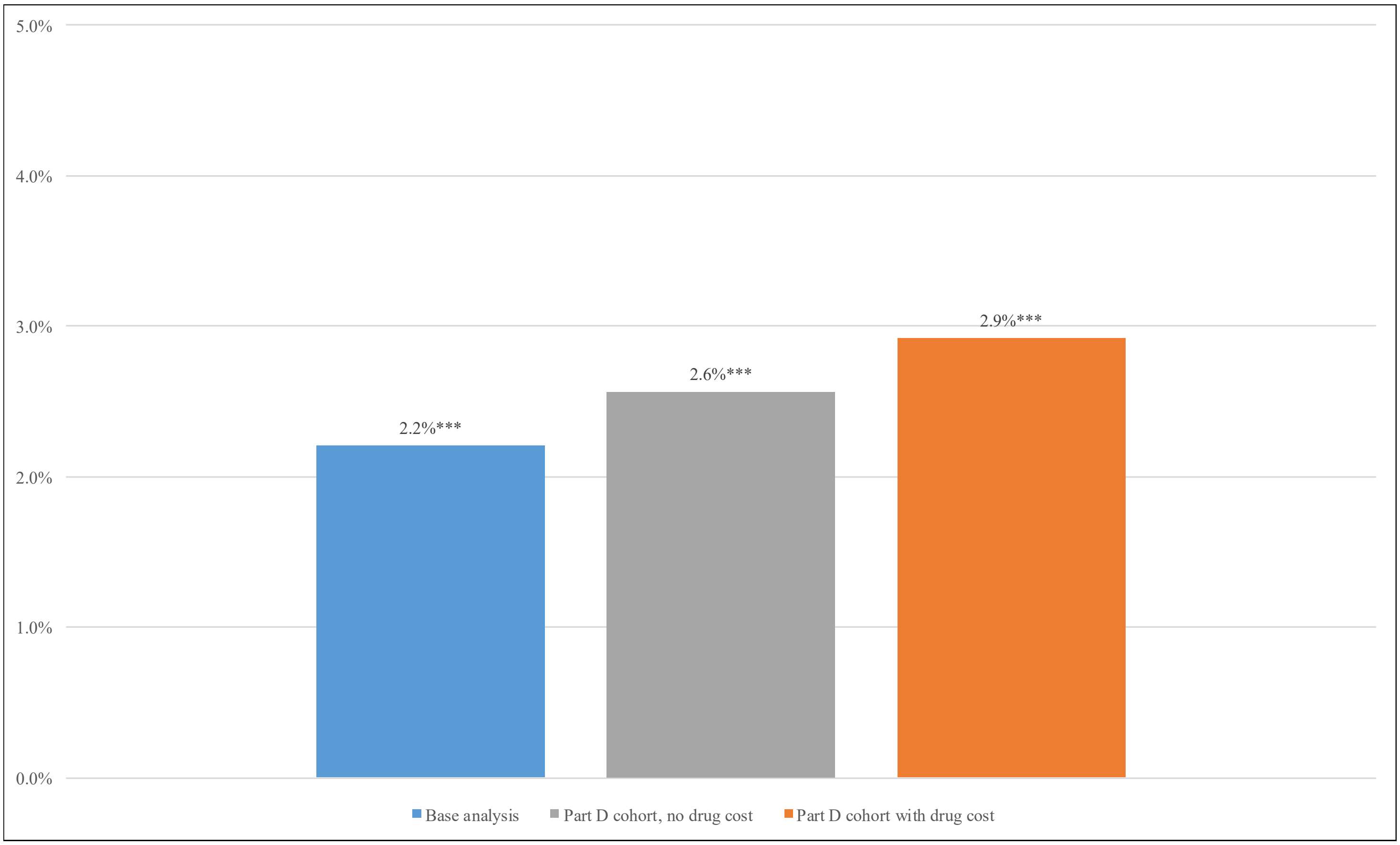
Figure 9: Annualized Productivity Growth under Alternative Specifications of Costs

(1)



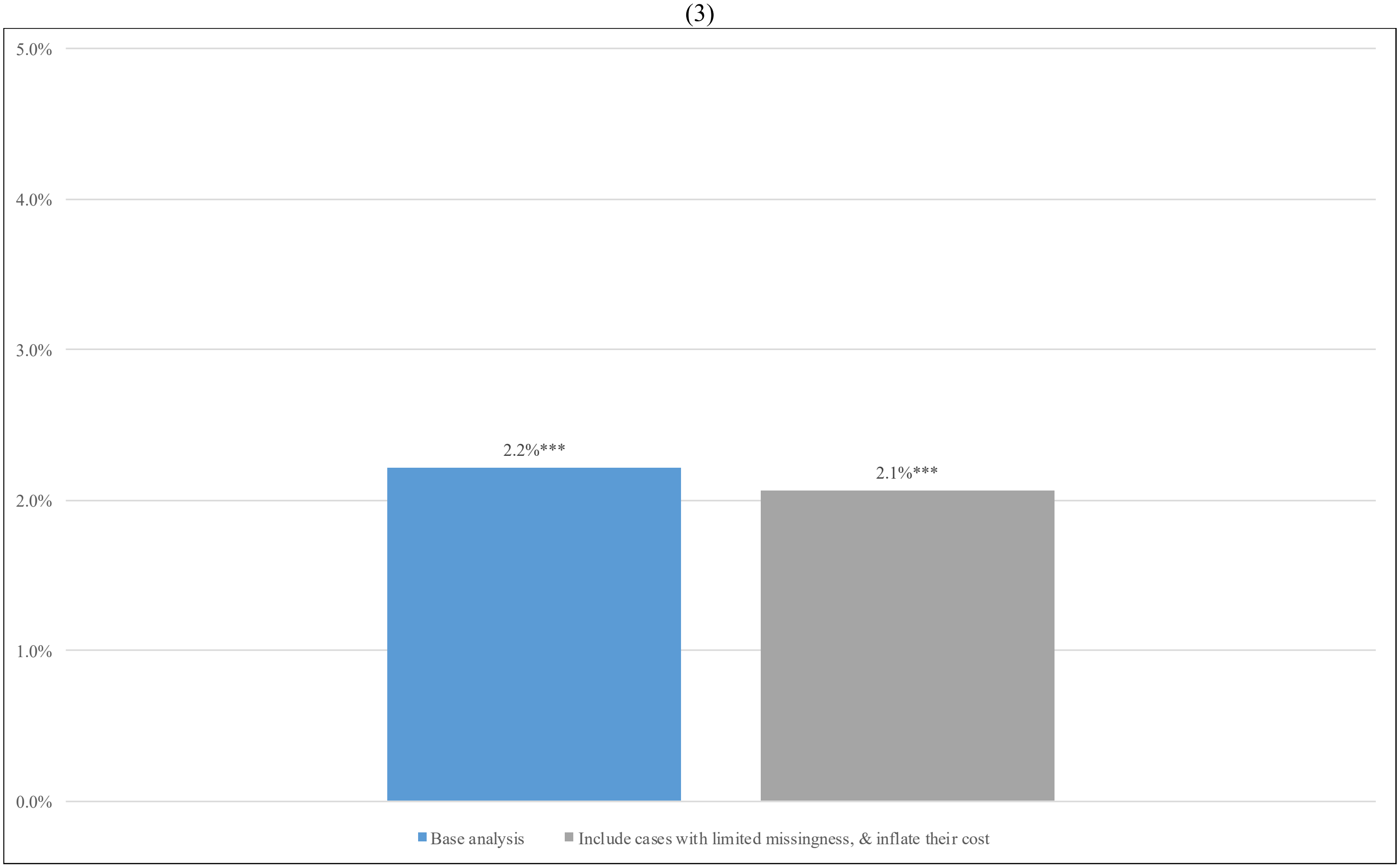
Exclusion of costs whose claim lacks a diabetes diagnosis

(2)



Inclusion of drug (Part D) costs

(3)

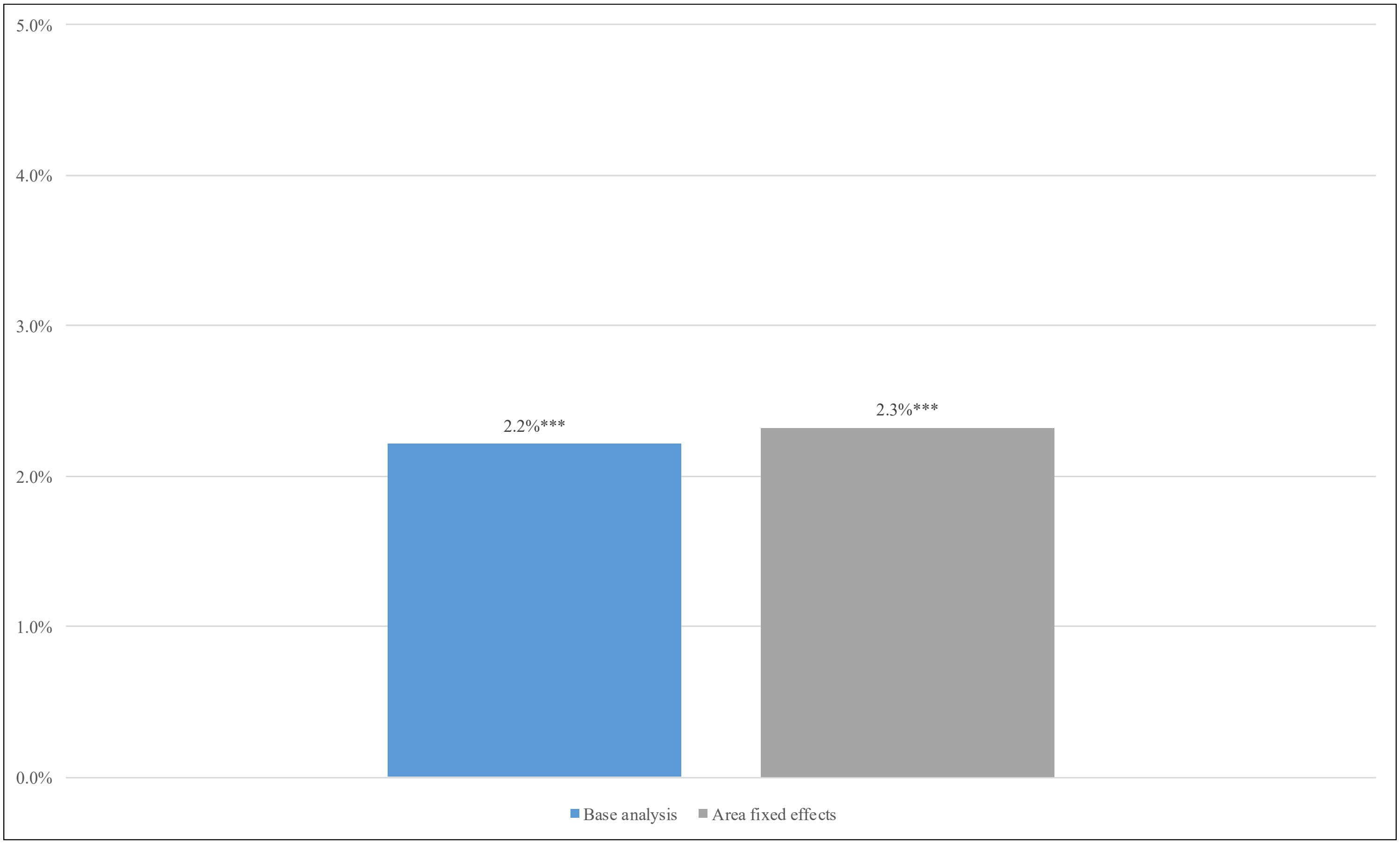


Inclusion of cases with limited cost missingness, with measured costs inflated

Notes: *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively. The middle panel is limited to the period 2006-2012. "Limited" missingness refers to cases with a) 1 or more institutional claims that could not be matched to cost data, and b) whose payments for claims with missing cost data as a share of total payments for the case was at or below the median. Total measured costs for these cases were inflated according to payments for claims with missing costs as a share of total payments for all cases diagnosed in the same calendar year.

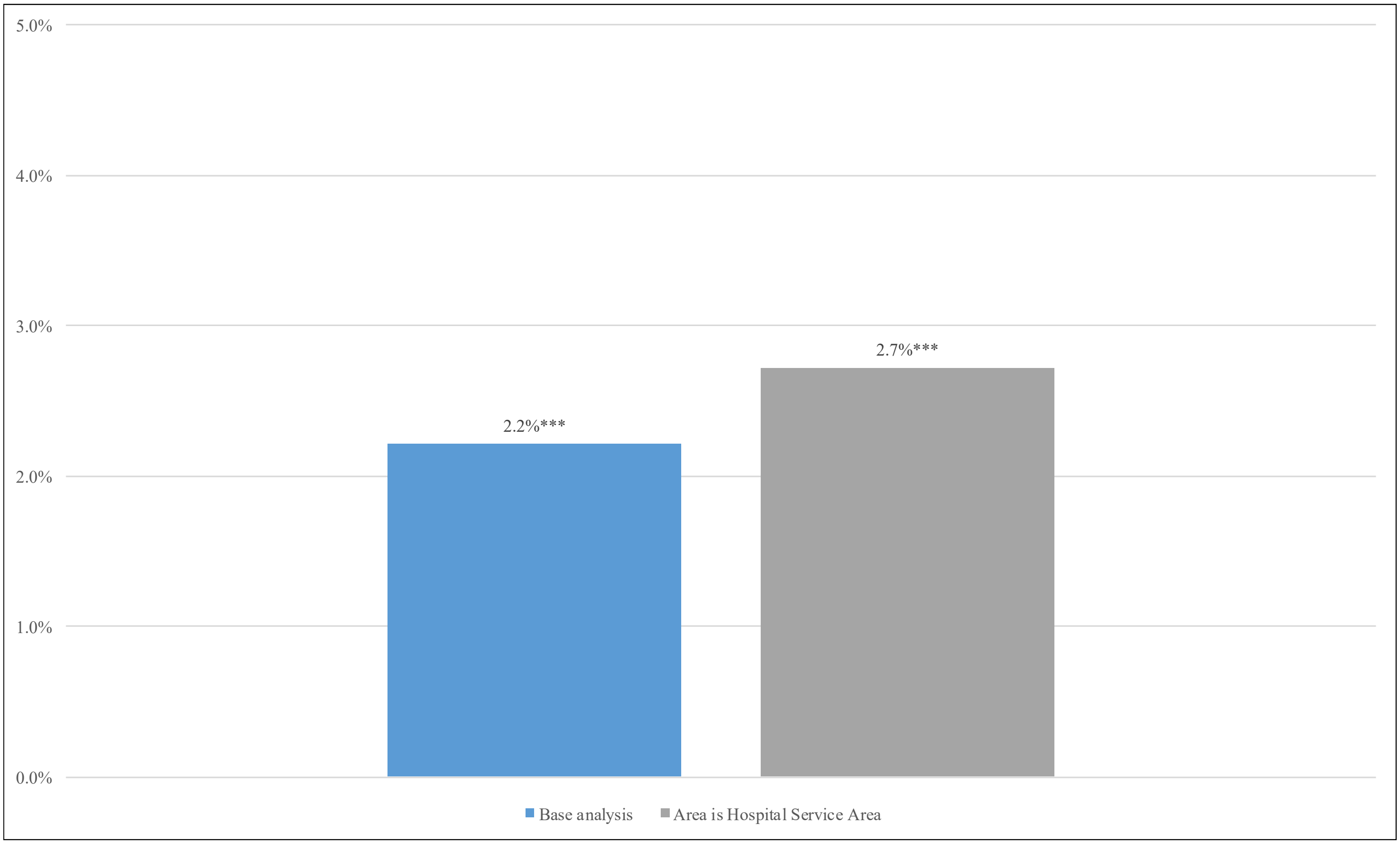
Figure 10: Annualized Productivity Growth under Additional Sensitivity Analyses

(1)



Inclusion of Area Fixed Effects

(2)



Alternative Specification of Geographic Area

Note: *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

APPENDIX ON COST-OF-LIVING APPROACH

The cost-of-living approach focuses on consumer well-being, and measures the change in costs (that is, consumer prices) that would have left consumers equally well off after some other change, typically in the health care context an improvement in a treatment outcome.

Following (Cutler, McClellan et al. 1998), the measure of the compensating variation from changing treatment is:

$$\frac{\partial U/\partial H}{\partial U/\partial X}(H_t - H_{t-1}) - (C_t - C_{t-1}),$$

where t indexes time, $\frac{\partial U}{\partial H}$ is the marginal utility of health H, $\frac{\partial U}{\partial X}$ is the marginal utility of the numeraire X,

and C is the cost of the treatment to the consumer.

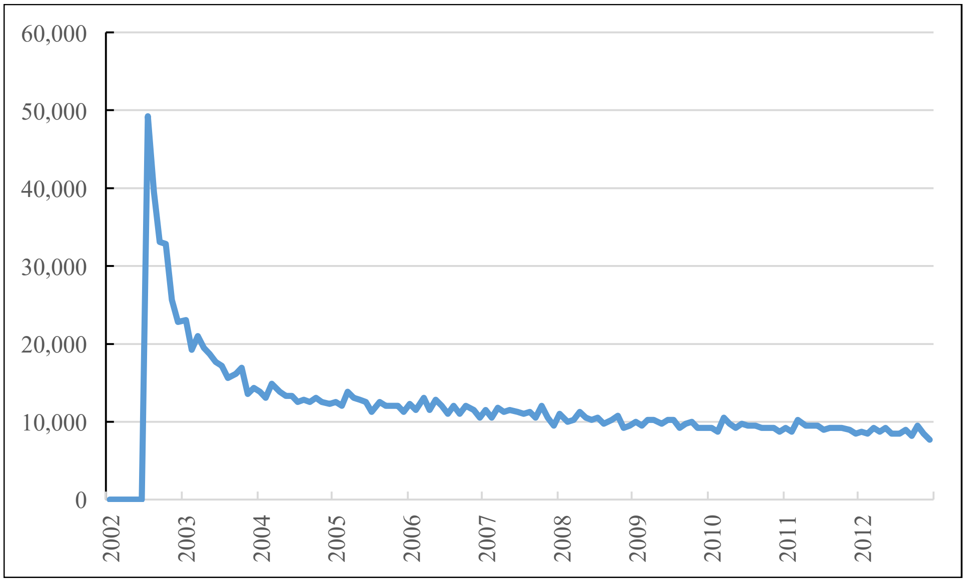
The ratio of the marginal utilities values health improvements in dollar terms, and so the first term is the value of a better health outcome. Where this benefit exceeds its cost (the second term), consumer welfare has increased.

Turning to our production framework, suppose that health is determined according to the production function $H(K, L, A)$, where K is capital, L is labor, and A is multifactor productivity (MFP). Then the above equation becomes:

$$\frac{\partial U/\partial H}{\partial U/\partial X}\big(H(K_t, L_t, A_t) - H(K_{t-1}, L_{t-1}, A_{t-1})\big) - (C_t - C_{t-1})$$

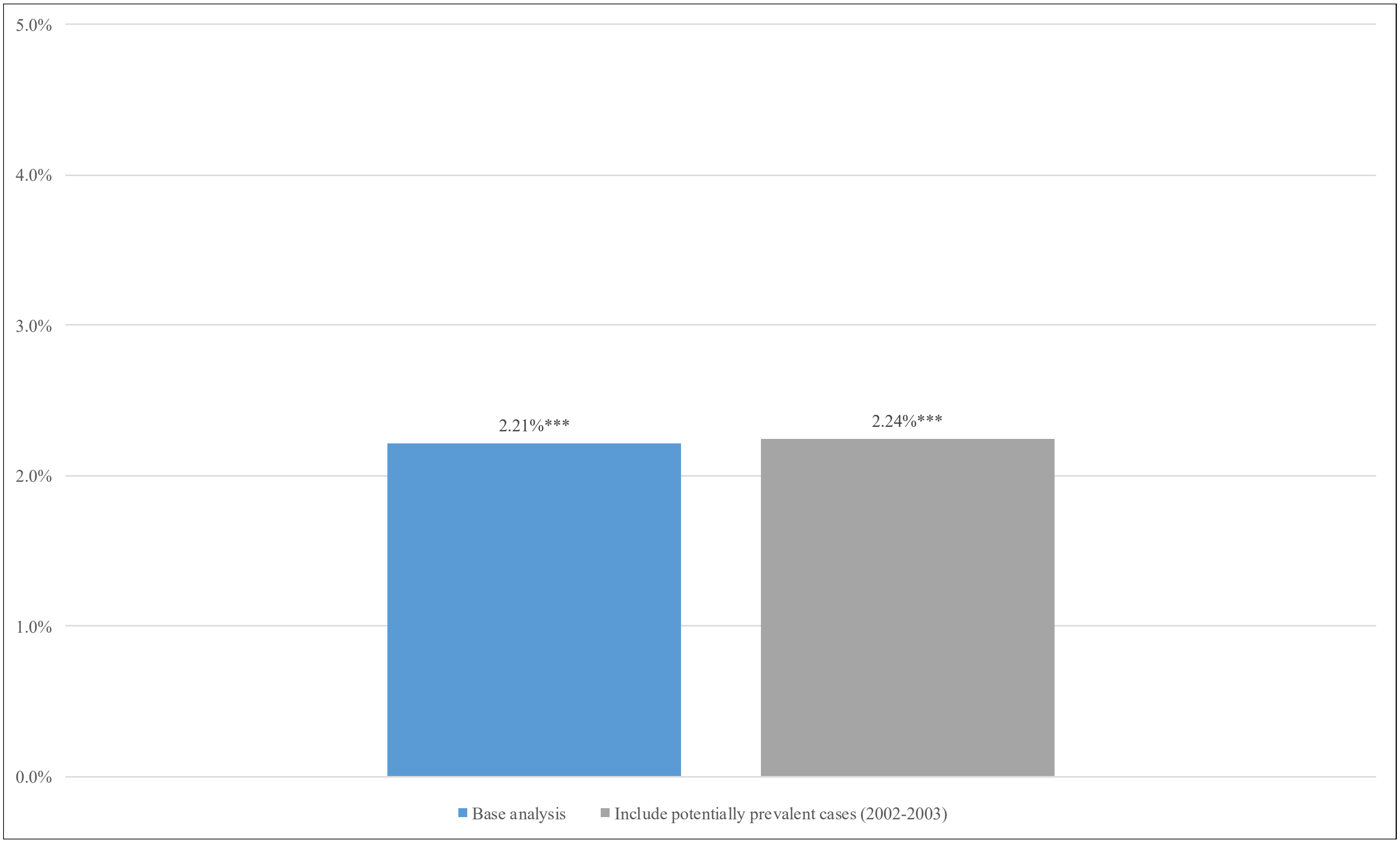
Clearly, an increase in MFP will, all else equal, improve consumer welfare. Nevertheless, the relationship between consumer welfare and the productivity of the health care system is complex. For example, an improvement in productivity could be consistent with a decline in welfare if health outcomes decline, but the capital and labor resources needed to produce those health outcomes fall by a much greater amount.

Appendix Figure 1: Number of Diabetes Diagnoses by Month



Notes: These counts correspond to the second row of Table 1 (n=1,558,394).

Appendix Figure 2: Including Potentially Prevalent Cases



Notes: *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively. Growth rates are annualized over 2004-2012.

Appendix Table 1: Diabetes Complications

Complication	*QALY decrement*	*Source(s)*	*ICD-9 code(s)*
Amputation	-0.280	(McEwan, Evans et al. 2010, Agency for Healthcare Quality and Research 2013)	250xx and (841xx** but not 841.1**)
Blindness	-0.074	(McEwan, Evans et al. 2010, Leal, Hayes et al. 2013)	369xx
End-stage renal disease	-0.263	(McEwan, Evans et al. 2010)	Inpatient or Home Health service provided by clinic or hospital-based renal dialysis facility in Outpatient File
Heart attack*	-0.055	(McEwan, Evans et al. 2010, Leal, Hayes et al. 2013)	410xx
Heart failure	-0.108	(McEwan, Evans et al. 2010, Leal, Hayes et al. 2013)	428xx
Ischemic heart disease	-0.090	(McEwan, Evans et al. 2010, Leal, Hayes et al. 2013)	411-414
Hypoglycemia	-0.244	(Currie, Morgan et al. 2006, Whitmer, Karter et al. 2009, Romley, Gong et al. 2015)	251.0x-251.2x
Stroke	-0.164	(McEwan, Evans et al. 2010, Leal, Hayes et al. 2013)	430-434 or 436xx

*Acute myocardial infarction. **Procedure (rather than diagnosis) codes.

Appendix Table 2: Trends in Covariates

Variable	2004 mean	2012 mean	2012 / 2004 ratio
Outcomes within 2 years of diabetes diagnosis			
Cost (000s of 2014 dollars)	$21.6	$19.5	0.901
Mortality	16.0%	14.7%	0.916
Amputation *among survivors*	0.04%	0.03%	0.640
Blindness	0.001%	0.000%	0.000
End stage renal disease	0.64%	0.58%	0.897
Heart attack	1.5%	1.2%	0.758
Heart failure	2.5%	1.7%	0.654
Hypoglycemia	0.02%	0.01%	0.599
Ischemic heart disease	4.0%	1.3%	0.339
Stroke	2.1%	1.6%	0.754
Quality of life	0.587	0.600	1.023
Patient demographics			
Age	75.8	74.5	0.982
Female	55.4%	53.2%	0.960
White	84.6%	82.8%	0.978
African American	9.7%	9.4%	0.972
Hispanic	1.9%	1.8%	0.962
Other race	3.8%	6.0%	1.573
Patient comorbidities prior to diabetes diagnosis			
Acquired hypothyroidism	7.7%	10.6%	1.376
Alzheimer's disease	5.3%	5.5%	1.046
Alzheimer's related disorders or senile dementia	11.2%	11.8%	1.058
Anemia	28.4%	33.3%	1.176
Asthma	4.2%	5.2%	1.233
Atrial fibrillation	9.7%	10.8%	1.104
Benign prostatic hyperplasia	0.0%	5.2%	n/a
Cancer, breast	2.8%	3.3%	1.199
Cancer, colorectal	1.7%	1.6%	0.945
Cancer, endometrial	0.2%	0.3%	1.311
Cancer, lung	1.0%	1.3%	1.243
Cancer, prostate	3.8%	4.1%	1.058
Cataract	22.6%	18.0%	0.798
Chronic kidney disease	9.0%	17.5%	1.941
Chronic obstructive pulmonary disease	14.0%	13.9%	0.991
Depression	8.8%	12.8%	1.461
Glaucoma	10.5%	10.1%	0.964
Heart attack	1.6%	1.6%	0.966
Heart failure	22.6%	19.1%	0.847
Hip fracture	1.0%	1.1%	1.037
Hyperlipidemia	46.8%	60.3%	1.290
Hypertension	68.5%	73.0%	1.066
Ischemic heart disease	40.3%	36.9%	0.914
Osteoporosis	5.6%	7.3%	1.316
Rheumatoid arthritis / osteoarthritis	25.9%	30.1%	1.162
Stroke / transient ischemic attack	6.1%	5.6%	0.914

Appendix Table 2, Continued: Trends in Covariates

Variable	2004 mean	2012 mean	2012 / 2004 ratio
Patient zip code characteristics			
Median household income ($000)	$44.9	$47.4	1.055
Social Security income ($000)	$11.4	$11.5	1.007
Poor	11.7%	10.9%	0.937
Employed	94.2%	94.5%	1.003
Less than high school education	19.6%	18.6%	0.950
Urban	78.3%	78.6%	1.004
Hispanic	10.3%	10.3%	1.005
Single	42.6%	41.7%	0.979
Elderly in an institution	5.3%	5.2%	0.971
Non-institutionalized elderly with physical disability	28.7%	28.3%	0.986
Mental disability	10.8%	10.7%	0.989
Sensory disability among elderly	14.1%	13.9%	0.986
Self-care disability	9.6%	9.5%	0.985
Difficulty going-outside-the-home disability	20.5%	20.2%	0.986
Month of diagnosis			
January	8.7%	8.4%	0.968
February	8.4%	8.3%	0.986
March	9.5%	8.9%	0.935
April	8.9%	8.4%	0.946
May	8.3%	8.8%	1.052
June	8.4%	8.2%	0.977
July	7.9%	8.2%	1.044
August	8.0%	8.6%	1.066
September	8.0%	7.8%	0.974
October	8.2%	9.0%	1.096
November	7.9%	8.1%	1.026
December	7.8%	7.3%	0.946

Appendix Table 3: Complete Results from Baseline Regression

Coefficient (Standard Error)	
Constant	-0.163 (2.473)
Diagnosis year	
2005	-0.057*** (0.008)
2006	0.044*** (0.012)
2007	0.121*** (0.016)
2008	0.053*** (0.017)
2009	0.055*** (0.019)
2010	0.075*** (0.020)
2011	0.125*** (0.026)
2012	0.175*** (0.028)
Patient demographics	
Age, logged	-0.727 (0.553)
Female	0.005 (0.105)
White	0.038 (0.111)
African American	-0.119 (0.129)
Hispanic	0.269 (0.303)
Patient comorbidities prior to diabetes diagnosis	
Acquired hypothyroidism	-0.014 (0.173)
Alzheimer's disease	-0.350 (0.281)
Alzheimer's related disorders or senile dementia	-0.366* (0.215)
Anemia	0.047 (0.093)
Asthma	-0.267 (0.250)
Atrial fibrillation	-1.045*** (0.168)
Benign prostatic hyperplasia	-0.559** (0.233)
Cancer, breast	-0.263 (0.305)
Cancer, colorectal	-0.787** (0.397)
Cancer, endometrial	-0.484 (0.851)
Cancer, lung	-0.795* (0.465)
Cancer, prostate	-0.147 (0.251)
Cataract	0.164 (0.116)
Chronic kidney disease	-0.907*** (0.150)
Chronic obstructive pulmonary disease	-0.868*** (0.153)
Depression	-0.392** (0.159)
Glaucoma	0.106 (0.151)
Heart attack	-0.422 (0.392)
Heart failure	-0.412*** (0.146)
Hip fracture	-1.666*** (0.455)
Hyperlipidemia	0.503*** (0.080)
Hypertension	-0.036 (0.107)
Ischemic heart disease	-0.491*** (0.099)
Osteoporosis	-0.200 (0.211)
Rheumatoid arthritis / osteoarthritis	-0.308*** (0.108)
Stroke / transient ischemic attack	-1.321*** (0.230)

Appendix Table 3, Continued: Complete Results from Baseline Regression

Coefficient (Standard Error)	
Patient zip code characteristics	
Median household income ($000)	-0.006*** (0.001)
Social Security income ($000)	-0.020 (0.014)
Poor	0.503 (0.384)
Employed	0.878 (0.567)
Less than high school education	-0.599*** (0.215)
Urban	-0.171*** (0.056)
Hispanic	-0.276*** (0.097)
Single	-0.616*** (0.228)
Elderly in an institution	1.162** (0.505)
Non-institutionalized elderly with physical disability	-0.998** (0.489)
Mental disability	-1.518* (0.773)
Sensory disability among elderly	1.637** (0.732)
Self-care disability	-0.276 (1.088)
Difficulty going-outside-the-home disability	1.937*** (0.661)
Month of diagnosis	
February	-0.114 (0.209)
March	-0.195 (0.215)
April	-0.011 (0.210)
May	0.305 (0.226)
June	-0.101 (0.216)
July	0.239 (0.222)
August	0.079 (0.210)
September	0.398* (0.225)
October	0.202 (0.220)
November	0.245 (0.244)
December	0.295 (0.214)
Other Statistics	
Area-years, n	2,742
R squared	0.677

Note: *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

The mission of the Hutchins Center on Fiscal and Monetary Policy is to improve the quality and efficacy of fiscal and monetary policies and public understanding of them.

The USC-Brookings Schaeffer Initiative for Health Policy is a partnership between the Economic Studies Program at Brookings and the USC Schaeffer Center for Health Policy & Economics and aims to inform the national health care debate with rigorous, evidencebased analysis leading to practical recommendations using the collaborative strengths of USC and Brookings.

Questions about the research? Email communications@brookings.edu.
Be sure to include the title of this paper in your inquiry.