NBER WORKING PAPER SERIES

THE EFFICIENCY OF MEDICARE

Jonathan Skinner
Elliott S. Fisher
John E. Wennberg

Working Paper 8395
http://www.nber.org/papers/w8395

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2001

Support by the Robert Wood Johnson Foundation and the National Institute on Aging is gratefully acknowledged. We are grateful for very helpful comments from Duncan Thomas, Frank Vella, Douglas Staiger, the NBER Summer Institute, and seminar participants at UCLA/Rand, Dartmouth College, Georgia State University, and the Universities of Virginia, Maryland, and North Carolina. Dan Gottlieb provided excellent data analysis. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research or the Department of Veteran's Affairs.

The Efficiency of Medicare
Jonathan Skinner, Elliott Fisher, and John E. Wennberg
NBER Working Paper No. 8395
July 2001
JEL No.H5, I1

ABSTRACT

Technological advances in health care have been shown to yield large average health benefits for the U.S. elderly population. However, less is known about the marginal or incremental benefits of health care spending. We use geographical variations in health care spending to measure the marginal value of greater health care intensity among the elderly Medicare population. To correct for the reverse causation problem -- that sicker areas tend to require more health care -- we use regional averages of physician visits in the last six months of life as a natural randomization for health care intensity. Using linear and semiparametric instrumental variables, we find that a large component of Medicare expenditures -- \$26 billion in 1996 dollars, or nearly 20 percent of total Medicare expenditures -- appears to provide no benefit in terms of survival, nor is it likely that this extra spending improves the quality of life. While secular trends in health care technology have delivered large health benefits, variation in health care intensity at a point in time have not.

Jonathan Skinner
Department of Economics
Dartmouth College
Hanover, NH 03755
and NBER
jonathan.skinner@dartmouth.edu

Elliott Fisher
Dartmouth Medical School
Strasenburgh 7251
Hanover, NH 03755
elliott.fisher@dartmouth.edu

John E. Wennberg
Dartmouth Medical School
Strasenburgh 7251
Hanover, NH 03755
john.wennberg@dartmouth.edu

I. Introduction

The U.S. spends more on health care in per capita terms and as a percent of GDP than any other developed country (OECD, 1998). This can be interpreted in two ways. One is that the elevated spending is symptomatic of failure in the health care system. Money is wasted through administrative overhead, the overuse of fully insured health care, or the provision of expensive tertiary but only marginally useful technology. [1] A different view is that U.S. citizens demand, and get, a higher quality level of health care than anywhere else in the world. It may be expensive, but the technological advances provided in the U.S. have led to dramatic improvements in functioning and life expectancy.[2]

Knowing which story holds true is crucial for any kind of health care reform, and particularly for Medicare reform. Unfortunately, the answer is elusive. While the evidence strongly suggests that technological gains in the treatment of heart attacks and low birthweight infants are substantial (e.g., Cutler et al, 1998; Cutler and Meara, 1999), it is not clear how well these specific paradigms generalize to the entire health care system where medical progress has not been nearly so robust. More importantly, the secular improvements in mortality are "average" effects of technology rather than the marginal impact of greater health care intensity on health outcomes (see Cutler, 2000).

Researchers have attempted to exploit "natural randomization" in outcomes data to estimate the marginal effectiveness of specific medical technologies on outcomes such as mortality for people with heart attacks (McClellan et al, 1994) or for infants (Gruber and Currie,

[1] For a good presentation of this view, see Evans and Stoddart (1994).

[2] For a general exposition of the view that current high spending levels for medical technology will yield benefits that could even lower costs in the future, see Pardes et al (1999). For specific measures of improvements in outcomes following the use of more intensive technology, see Cutler et al (1998), and Cutler and Meara (1999).

1996). For example, in McClellan et al, the "treatment" group were people experiencing heart attacks who lived relatively near a hospital equipped with diagnostic laboratory facilities that helped physicians decide whether to proceed with surgery, while the "control" group were those living further away. They found minimal benefits among the population treated most intensively for heart attacks, but it is not know how well this finding specific to heart attacks generalizes to the average Medicare patient, particularly those with chronic illnesses who account for a large fraction of overall expenditures.

In this paper, we use the idea of "natural randomization" to evaluate the efficiency of the Medicare program more generally. The macro-level equivalent of living near a hospital with advanced diagnostic laboratories for heart attack patients is whether the health care *system* provides a higher-than-average intensity of health care; hospitalization instead of outpatient care, surgery instead of watchful waiting, and 3-month physician appointments instead of 6-month appointments.

Simply comparing outcomes between regions with higher-than-average and lower-than-average Medicare expenditures risks the reverse causality problem; the sickest regions tend to experience more spending on health care.[3] To avoid these reverse causality issues, we focus on the practice and intensity of health care among elderly people in their *last six months of life.* By definition, these patients are similar across regions in terms of one primary indicator of health status, their six-month survival rate. Because there are no well-developed medical guidelines on how aggressively one should treat illnesses such as chronic heart failure, chronic obstructive pulmonary disease, and cancer near death, we suspect that community-level norms or standards

[3] The first study measuring the influence of Medicare spending on outcomes in an instrumental variables context was by Hadley (1988). He found a positive impact of Medicare spending on survival using nursing home residence rates (for example) as an instrument. However, nursing home residence rates may well be correlated with unmeasured health status in a region.

of care would be revealed most accurately in how such patients are treated. Obviously, one cannot use treatment patterns in the last six months to make inferences about Medicare efficiency in this group, since they are all dead by the end of the period. Instead, the intensity of care in the last six months is used as an instrument to measure whether regions with higher levels of Medicare expenditures yield benefits in terms of survival or health status.

Measuring intensity is difficult, since there are so many dimensions along which regions can differ; more physician visits, more tests, or simply more "best practice" health care, which may in fact be quite inexpensive (e.g., administering aspirin following heart attacks). We show below that the intensity that seems to account for most of the variation in Medicare expenditures is not the use of intensive or "high tech" surgery, but is instead inframarginal intensity in the sense of more laboratory tests, physician visits, specialist consults, inpatient care, and admission to intensive care units.

Using data from the *Dartmouth Atlas of Health Care* as well as supplemental information from the U.S. Census and the Centers for Disease Control, we estimate an instrumental variables model of Medicare spending, both in a linear model and in a nonlinear framework (Newey, Powell, and Vella, 1999). Briefly, we find little evidence that the greater spending observed in high-intensity regions leads to better health outcomes; in other words, we find "flat of the curve" benefits from the higher expenditures. On the other hand, regional indicators of effective practice – rates of screening for breast cancer, influenza vaccinations, and appropriate treatment of heart attacks – are associated strongly with improved survival. In short, our estimates imply that the Medicare program provides too little in the way of inexpensive and effective care, while at the same time spending $26 billion annually, or nearly 20 percent of its budget for health care of questionable value.

II. The Nature of the Problem: Per Capita Medicare Expenditures in the United States

We first consider the magnitude and extent of regional differences in Medicare expenditures in the United States. Primary data sources are samples of either 20 percent (outpatient) or 100 percent (inpatient) of the Medicare claims data during 1995 and 1996. The basic unit of analysis is the Hospital Referral Region (HRR) of which there are 306 in the United States. The HRR was constructed in the *Dartmouth Atlas of Health Care* as a unit of analysis that reflected the actual hospital migration patterns of Medicare patients for tertiary care. An HRR must include at least one hospital that performs cardiac surgery and neurosurgery. Each zip code in the United States is assigned to an HRR depending on what hospital the majority (or in some cases, the plurality) of Medicare enrollees seek their hospital care, see Wennberg and Cooper (1999) for details. Thus the HRR may cross county or state boundaries, or in some cases follow interstate highways.

The important thing to note about HRRs for this study is that all rates are based upon the zip code of residence and not where the person actually sought care. Thus if an individual lives in Lebanon NH and is admitted to a hospital in Boston, all utilization is assigned to the Lebanon HRR. Most care is delivered locally; 80 percent of the US population lives in HRRs in which over 85 percent of care is delivered by providers within the HRR. In the analysis that follows, utilization rates have been adjusted for differences across HRRs in the age, sex, and racial composition of the population, and (where necessary) differences in the price level. We restrict our attention to the fee-for-service Medicare population that during the study period accounted for more than 85 percent of the total Medicare population.[4]

[4] One concern with using the fee-for-service population is the selection problem; healthier patients tend to enroll in HMOs. We control for this in part by measuring health status (discussed below) for the same fee-for-service population. Thus if a healthy region experiences a high rate of HMO enrollment, leaving an unhealthy fee-for-service population, this will be reflected in both health status measures and in per capita spending. In practice, either including the percentage of the HMO population in the HRR, or excluding regions with more than 10 percent HMO enrollment, has little impact on the results.

Figure 1 uses these data from Wennberg and Cooper (1999) to construct a map showing the distribution of per capita Medicare expenditures across the United States in 1996; these are adjusted for differences across regions in age, sex, and race (black and nonblack). There are clearly wide variations in the extent of spending, with per capita expenditures ranging from \$3,341 in Minneapolis, for example, to \$8,414 in Miami. There are clusters of high-expenditure regions largely concentrated in Florida, the deep South, and urban areas on the East and West coasts. There are exceptions as well; inexpensive regions in Florida, and low-cost cities on the West Coast (e.g., Portland Oregon) and the East Coast (Richmond, VA).

This Figure raises some immediate policy issues. If the higher expenditures in some regions actually lead to better health, then the Medicare program may be inequitable to the extent that taxpayers in the low-expenditure regions are paying for the better health of those in the high-expenditure regions (Feenberg and Skinner, 2000). Conversely, if the higher expenditures yield nothing in health benefits, then there is tremendous waste in the program; reducing spending in the high expenditure areas can save enough money to preserve the solvency of the Medicare trust fund by a decade (Skinner and Wennberg, 2000). It could also be the case that people in the high expenditure areas simply prefer more intensive care. One might then ask why other regions should be subsidizing their preferences.

To consider the efficiency of Medicare, we provide a simple graph in Figure 2 that shows per-capita Medicare expenditures (age, sex, race, and price-adjusted) on the horizontal axis and survival, which we define as the (1- age-sex-race-adjusted mortality) x 100, on the vertical axis. There is a clear negative correlation between expenditures and survival. This in itself is not too surprising; spending should be higher in regions with poorer levels of health, so we might expect to observe Mobile, Alabama spending more than Grand Junction, Colorado. In the next section, we consider a simple model that formalizes how to evaluate the efficiency of Medicare given that Medicare spending is likely to be higher in regions with poor health.

III. Medicare Expenditures and Outcomes: Theoretical and Measurement Issues

In this section, we first develop a theoretical model of Medicare expenditures and outcomes, and show that the observed negative correlation between Medicare expenditures and survival in Figure 2 is at least consistent with the Medicare program meeting stringent efficiency conditions. We then develop criterion for determining what one *should* expect in terms of better survival, as well as discussing alternative measures of health care intensity.

1. A Theoretical Model of Regional Differences in Medicare Expenditures

Suppose that the value function of the Medicare "social planner" is written

$$V = V(\mathbf{S}(\mathbf{M}), \mathbf{Q}(\mathbf{M}), \mathbf{Y}(1-\tau), \mathbf{P}) \qquad (1)$$

where V is the concave value function, the bold-faced **S** denotes the vector of (regional) per capita survival measures for regions i = 1,…k, **M** is the level of per capita Medicare expenditures (and S_i depends only on M_i and not spending in other areas), **Q** the vector measuring quality of life (where Q_i is also dependent on M_i), and $\mathbf{Y}(1-\tau)$ the vector of per capita income after the Medicare tax has been paid; we assume that the tax rate τ is proportional to income.[5] Finally, the population of each region is given by **P**, this is to allow for larger regions to receive a larger weight in the social welfare function. While the Medicare program is a complex intergenerational transfer mechanism in which younger workers pay most of the taxes ultimately consumed by the elderly, we assume for analytic simplicity that the people paying the taxes in region i are the same ones experiencing the benefits in region i.[6]

[5] The Medicare payroll tax that funds Part A, or the hospital component, is proportional to earnings. While the Part B (physician) premium is regressive, the larger proportion funded by general tax revenue is progressive; overall, the tax is not far from proportional. See McClellan and Skinner (1999).

[6] See for example Feenberg and Skinner (2000).

Increasing Medicare spending in just one region i is assumed to result in an increase in the overall Medicare tax rate τ; $\Delta M_i = \Sigma_j P_j/P_i Y_j \Delta\tau$. Thus the balanced budget change in the tax rate necessary to fund an extra (per capita) Medicare dollar spent in region i is

$$\frac{d\tau}{dM_i} = \left[\sum_{j=1}^{k} \frac{P_j}{P_i} Y_j \right]^{-1} \qquad (2)$$

and the first-order conditions for Medicare expenditures across regions is, for each i

$$\frac{dV}{dM_i} = V_{1i} \frac{dS_i}{dM_i} + V_{2i} \frac{dQ_i}{dM_i} - \sum_{j=1}^{k} V_{3j} Y_j \frac{d\tau}{dM_i} \qquad (3)$$

where V_{1i} is the contribution of an incremental increase of survival in region i to the social welfare function, V_{2i} the contribution of quality of life (conditional on survival), and V_{3i} the impact of after-tax non-medical income on social welfare of the entire country. In the simplest case, where $V_{1i} = P_i$ and $V_{2i} = 0$ for all i, the objective of the welfare function is to maximize the national survival rate conditional on the overall Medicare budget. When V" < 0, the redistributive component also matters in the first-order condition; there is a greater social benefit to raising the survival rates of the sickest regions.

Ignore for the moment the impact of Medicare expenditures on health status or health functioning, so that $dQ_i/dM_i = 0$ for all i. Then the first-order condition can be written

$$\frac{dS_i}{dM_i} - \left[\frac{\sum_{j=1}^{k} V_{3j} Y_j \frac{d\tau}{dM_i}}{V_{1i}} \right] = 0 \qquad (4)$$

The first term on the LHS measures the marginal productivity of Medicare expenditures on survival in region i. (We ignore for the moment the fact that the incremental M_i can be used to increase a variety of health inputs; for the moment we assume there is just a single dimension of spending.) Suppose that the value function is linear, so there is a uniform social tradeoff between increasing survival by one unit and reducing after-tax income by β. Thus β is the conventional cost-effectiveness "hurdle," or how much is society willing to spend to increase survival rates by a given amount. In this special case, $\beta = P_j V_{3j}/V_{1i}$, $\forall$ i,j, which allows us to simplify equation (4) to

$$\frac{dS_i}{dM_i} = \beta^{-1}$$

In other words, all regions should devote expenditures up to the point where the marginal gains are equal. This can be shown graphically in Figure 3, with the same dimensions as those shown in Figure 2; expenditures (or intensity) on the horizontal axis and survival on the vertical axis. Combinations of expenditures and survival rates are shown for three regions, A, B, and C, as well as each of their concave health "production function" $S_i(M_i)$. The slopes of each of the straight tangential lines are equal to β so that $dS_i/dM_i = \beta$ across regions. Furthermore, this graph replicates the general pattern of spending and survival shown in the empirical data in Figure 2. Accounting for the concavity of the value function V would imply efficiency conditions that would move region A further along its production function to A', and would move C to C', resulting in a reduction in health disparities across regions.

Suppose that some of the benefits of the expenditures come in the form of better quality of life, and not just survival. Then the first-order condition becomes

$$\frac{dS_i}{dM_i} - \frac{V_{2i}}{V_{1i}}\frac{dQ_i}{dM_i} - \left[\frac{\sum_{j=1}^{k} V_{3j} Y_j \frac{d\tau}{dM_i}}{V_{1i}}\right] = 0 \qquad (5)$$

Thus even when the social welfare function is linear, we may not necessarily find that regions with greater spending would experience an increase in survival of β^{-1} per dollar spent. For example, it is often argued that "discretionary" surgical procedures are designed to yield better health functioning without affecting survival rates. However, regional Medicare expenditures are not closely correlated with "discretionary" surgery such as angioplasty, bypass surgery, back surgery, or radical mastectomies; thus it seems unlikely that the incremental higher Medicare expenditures are necessarily targeted towards procedures associated with improved health functioning (Wennberg, Fisher, and Skinner, 2001).

In our econometric specification, we consider a model in which survival S_i in region i is a general nonlinear function of Medicare expenditures in that region, $S(M_i)$, i.e.,

$$S_i = S(M_i, Z_i) + u_i \qquad (6)$$

where Z_i is a vector of underlying observable health characteristics,. Thus S(M, Z) traces out the family of productivity curve shown in Figure 3. A simplified version of this equation is

$$S_i = S(M_i) + \Gamma Z_i + u_i \qquad (6')$$

where Γ is the corresponding vector of coefficients; thus $Z_i\Gamma$ shifts the productivity curves vertically with respect to observable differences in health status. To the extent that *unobservable* health status is reflected in the error term u_i, it will be correlated with Medicare spending, leading to inconsistent estimates in a single equation model. We therefore model Medicare expenditures as a nonlinear function of instrument(s) X as well as the other variables Z;

$$M_i = M(X_i) + Z_i\Pi + \epsilon_i \qquad (7)$$

where Π is a vector of coefficients and ϵ_i the error term. This block-diagonal structure is well suited to estimation using the methods developed in Newey, Powell, and Vella (1999); we return to estimation issues below in Section IV.

2. How Much Should Survival Rates Differ Across Regions?

The theoretical model suggested that it is important in assessing efficiency to measure the impact of the marginal dollar of Medicare spending on health outcomes. As a first step, we would like to know how much difference in survival rates should we expect to see under the null hypothesis that incremental Medicare expenditures yield first-order health benefits. In the short term, we would expect to see a jump in survival. If we viewed the social β to be \$100,000 per additional year of life, then spending an extra \$1,000 per capita in Medicare spending should, in the short term, yield a drop in mortality rates (or increase in survival rates) of 1.0 percentage points. Over the long term, the mortality rate would climb back up as those patients saved early on ultimately die.

To quantify the change in survival and mortality rates at the level of the population that is consistent with the micro-level cost-effectiveness benchmarks, we develop a simple model using the lifetables for 1991, from Wilmoth (2001). Figure 4 shows the benchmark survival pattern for the US population in 1991 (the leftmost curve); the average mortality rate is 5.2 percent in this population, which is consistent with mortality in the Medicare population. Next, suppose an innovation is introduced that reduces annual mortality rates by 25 percent, for example, leading to the rightward shift in the survival curve (Figure 4). The population-weighted decline in the mortality rates is 1.3 percentage points, down to 3.9 percent. In the steady state (assuming no population growth), the population is larger by 11 percent, the area between the two curves. In other words, the implicit numerical derivative of the percentage change in steady-state life years per one percentage point change in the mortality rate is equal to 11/(5.2-3.9) or 9.0.[7] Conversely, a "hurdle" rate of \$100,000 per life year implies that every increase of \$1,000 in per capita Medicare expenditures should increase survival rates by 1/9, or 0.11 percentage points.

This hurdle rate may be too low. As noted above, short run effects are larger than long-run effects, and to the extent that Medicare spending has risen dramatically in the last decade, we

---

[7] We chose the 25 percent mortality decline to make the differences apparent in Figure 4. Smaller changes in mortality yield the same numerical derivatives, however.

might expect to observe a larger impact on survival. Second, the $100,000 hurdle is sometimes given for a year of life in perfect health, and one is rarely extending life years in perfect health; indeed the increased survival may be at the expense of quite poor health functioning. On the other hand, if functioning is improved as well as survival, one may require less improvement in survival to justify the incremental Medicare expenditures, an issue to which we return below.

IV. An Instrumental Variables Approach to Measuring Health Care Intensity

In this section, we consider a variety of instrumental variables that we argue are correlated with health care intensity, but not associated with underlying unobservable community illness levels. Our primary instrument will be physician visits per decedent in the last six months of life. We focus on this group because these individuals are equally sick by one very powerful marker – their 6-month survival rate – so that even in generally healthy regions, we can measure the intensity of health care provided for these patients, and by extension all patients (those who survived as well as those who died) in the region. Figure 5 shows a turnip graph of the distribution of physician visits in the last six months of life; each dot represents average rates for an HRR. The rates vary widely, from 8.5 in Grand Junction, CO to 43 in Ridgewood NJ and nearly 48 in Miami. We also consider two additional instruments. One is the percentage of Medicare enrollees who in their last six months are admitted to an intensive care unit (ICU). This value again varies widely, from 14 percent in Sun City, AZ to 49 percent in Miami, FL and Munster, IN. The second takes a prospective, rather than retrospective approach and considers 6-month Medicare expenditures for elderly patients who were admitted during 1993-95 for a hip fracture, an injury that nearly always results in admission to the hospital. These expenditures are not correlated with underlying rates of comorbidities at admission, nor are they correlated with subsequent survival (Chau, Fisher, and Skinner, 2001; Fisher, Stukel, and Wennberg, 2001).

1. What Specific Medical Services do Residents in High-Intensity Regions Get?

What do patients get from the incremental dollars spent in the regions with higher levels of Medicare spending? We provide a heuristic impression of differences in health care intensity by considering utilization rates for two groups of regions in the U.S. The first (Decile 10) samples the top decile of regions as ranked by physician visits in the last six months of life (weighted by the size of the Medicare population) during 1995/96. These regions include much of the New York and New Jersey metropolitan areas, Takoma Park, MD, Philadelphia, and McAllen, Texas; average visits per decedent range from 36 to 48.

The second sample is for HRRs that are classified in the lowest decile (Decile 1); these are all regions with fewer than 16 visits per decedent. These include Lynchburg, VA, Minneapolis, Portland OR, and Salt Lake City, UT. In these samples, we compiled selected measures of utilization for patients in their last six months of life that tell us something about how characteristics of care for these (typically) chronically ill patients differ across regions of the U.S.

In Table 1, we consider rates of physician procedures from Part B Medicare claims data, using a sample of about 32,000 decedents who died during 1995-96 between the ages of 70 and 90. The two groupings of HRRs were nearly identical with regard to the age distribution of death and broadly similar with regard to (state-level) causes of death.[8] Rates are expressed as average counts per person in the sample; thus a higher number may reflect more people receiving the treatment or a larger number of treatments per person.

It is perhaps not surprising that physician visits per decedent are higher in Decile 10 than Decile 1, since that is how the categories were chosen. However, the types of visits show quite different patterns. Outpatient office visits were 46 percent higher in Decile 10 regions, relative to our benchmark of the lowest (Decile 1) regions, and 79 percent higher for the initial visit by the

[8] Cause of death data are from the Centers for Disease Control web site (www.cdc.gov). Rates are adjusted for age and sex for the population over age 65 and are drawn from state measures and merged (by zip code) to the relevant hospital referral regions (HRRs). The age-adjusted mortality rates were roughly identical in the two deciles, 4.97 percent in Decile 1, and 4.96 in Decile 10. While the percentage of deaths due to cancer in the two groups was similar (22.7 percent in Decile 10, 22.1 percent in Decile 1), as was diabetes (2.6 and 2.7 percent), cardiovascular diseases were higher in Decile (49.1 percent versus 44.1 percent), and chronic obstructive pulmonary diseases were lower (4.9 versus 5.8 percent).

physician when the patient was admitted to the hospital. The real differences occur for the initial visit by a specialist newly brought on to the case (392 percent more in Decile 10 regions). In other words, there is greater *scope intensity* – more specialists treating separate organs or systems – in Decile 10 regions.

Regions in the top decile are also characterized by their greater use of diagnostic techniques such as endoscopies, X-rays, Doppler echocardiograms, and electrocardiograms; their use is between 106 and 235 percent greater in Decile 10 regions. Finally, the greatest divergence in specific medical procedures comes in those that are used to maintain survival among seriously and chronically ill patients: insertions of emergency airways, dialysis for failing kidneys (hemodyalisis), feeding tubes inserted into the stomach (gastrostomy tube placement), and mechanical breathing assistance (continuous ventilator management). Rates in Decile 10 range from 233 percent to 674 percent above those in Decile 1. Note that these procedures are not designed to improve quality of life, but instead are directed towards maintaining short-term survival.

2. Instrumental Variables and Natural Randomization

We have provided a brief theoretical justification for why physician visits in the last six months might proxy for a different degree of health care intensity, both on the intensive and extensive margin. We now consider whether physician visits in the last six months can be judged to provide a "natural randomization" across regions in the U.S.

A necessary requirement for a valid instrument is that it be uncorrelated with the error term u_i in equation (6), conditional on the covariates Z_i. Suppose that the error term $u_i = Z_i^*\Theta + u^*$ where Z_i^* represents unmeasured components of health, for example childhood health histories of elderly residents of region i, Θ is a vector of coefficients, and u_i^* the "true" error term. Under conventional identification strategies, the researcher argues that the instrument X may be correlated with Z, but is uncorrelated with Z*|Z (and u*). This can be a difficult argument to make in explaining survival, since there are a variety of causal factors in population-based survival rates, some of which happen to be observed and collected by government agencies, and some of which are not.

A more stringent requirement is "natural randomization" in which the instrumental variable X is not even correlated with the measured covariates Z (McClellan, Newhouse, and McNeil, 1994, Hearst, Newman, and Hulley, 1986; Angrist, 1990). In this case, the argument that the instrument X is correlated with unmeasured components of health Z* is harder to make, since one must argue that X is correlated *only* with Z* (and not Z), and furthermore that it is correlated with the components of Z* that are themselves orthogonal to Z, that is, Z*|Z. There are a variety of ways to consider this notion of natural randomization. One is to ask whether adding or subtracting covariates Z affect the estimated function S(M), a strategy pursued in Altonji, Elder, and Taber (2001). Another is to test the extent to which the instrument is correlated with Z. We consider both strategies below.

The basic set of covariates Z include age-sex-race-adjusted population-based measures at the HRR level of hospital admissions (per 1,000 elderly population) for heart attacks (AMI), stroke (CVA), gastrointestinal bleeding, hip fractures, colon cancer, and lung cancer (Wennberg and Cooper, 1999). Hospital admission rates for these diseases are accurate measures of the population incidence since nearly every person with one of these 6 diseases will be admitted to the hospital. In addition, we include rates of poverty in the elderly population and average Social Security Income among households receiving this source of income, measured using the CensusCD data (based on primary Census data) at the zip code level in 1990. Since Social Security benefits are based on lifetime earnings, these provide a good (albeit nonlinear) index of lifetime earnings. We also use the Census data to measure the fraction of elderly people living alone, since such patients may have fewer potential caretakers and thus would be more likely to spend their last six months receiving inpatient care.[9] We consider measures of behavioral health status from the general population such as obesity, binge drinking, cigarette smoking, and seatbelt use; these are derived from Centers for Disease Control data (MMWR, 1999) at the state level, and assigned to the HRR level according to the relative state population weights in each HRR. These variables are summarized in Table 2.

Table 3 demonstrates that these variables combined are strongly associated with survival rates, with an R^2 of 0.56 for the full model. Interestingly, the poverty rate enters with a positive coefficient (that is, a higher poverty rate implies a higher survival rate, holding constant income). In part, this is the consequence of including comorbidities such as heart attacks and hip fractures; poverty works primarily through its impact on disease categories.[10]

To what extent are physician visits in the last six months correlated with these measures of health Z? We again create groups of HRRs based on physician visits in the last six months, only this time we consider in Table 4 the entire sample of 306 HRRs separated into terciles of the population, with mean values of selected health measures by tercile. There does not appear to be a consistent pattern in health status across the terciles; some health measures are higher and some are lower. Rather than using as a summary statistic the simple average of health measures, we create a weighted average, where the weights are the coefficient estimates in Column C of Table 2, or in other words the predicted survival rate of each community; this predicted measure, as

[9] This is defined as the ratio of people over age 65 living alone divided by the total number of people over 65 not in institutions or living with unrelated people.

[10] Holding income constant, a larger poverty rate is consistent with a widening of the income distribution; e.g., more income inequality. However, the hypothesis that income inequality is bad for health (e.g., Wilkinson, 1997) would predict a negative coefficient on poverty rates, not a positive coefficient. See also Deaton and Paxson (1999).

noted above, explains 56 percent of the variation in mortality.[11] Predicted survival is nearly identical across the terciles, and the correlation coefficient is 0.01 with a p-value of 0.88. In other words, physician visits in the last six months does not meet the strictest definition of natural randomization, which would require it to be orthogonal to any element of Z. It does meet the less stringent condition that it is uncorrelated with the linear combination of Z that best predicts survival. The two other instruments, the percent decedents admitted to ICU beds in the last six months, and average 6-month expenditures in the hip fracture cohort, exhibit modest but significant correlation coefficients (of about 0.20) with the predicted survival measure. We consider these additional instruments below in our sensitivity analysis.

3. Alternative Measures of Medical Care Intensity

To this stage, we have considered just a single index of health care M equal to overall Medicare per capita expenditures. In the regressions below, we also explore an alternative measure of health care that focuses not on overall expenditures, but on "process quality" of care. Wennberg and Cooper (1999) adopted a measure of "effective" care that measured the frequency of use where appropriate for 11 treatments or screening methods that are generally agreed upon to be effective in medical care; this include the fraction of women in the Medicare population who were screened for breast cancer (mammography), the percentage of diabetics administered blood tests and eye exams, and the percentage receiving a pneumonia vaccination. Averaged in with these measures are quality measures for the treatment of heart attacks; the percentage of appropriate patients in each HRR who received effective drug treatment such as aspirin, beta blockers, and ACE inhibitors, for example.[12] One would not expect 100 percent compliance for

[11] Note that this predicted measure of mortality is a linear function only of Z, and not of S or u.

[12] These latter indicators are drawn from the Cooperative Cardiovascular Project, or CCP, a detailed study of more than 200,000 heart attacks in the US. See Jencks et al (2000) for a more comprehensive measure of quality by state.

a variety of reasons, but in general the quality indexes were quite low, with a (weighted) mean of 48 percent and a range from 32 to 57 percent (see Table 2), with a higher index indicating better quality care.

V. Estimating the Efficiency of Medicare Expenditures

We first consider factors that can affect Medicare expenditures in the context of a first-stage regression, as in (7). Here we use expenditures that have been adjusted for age, sex, race, and differences in cost-of-living reimbursement rates for Medicare services (Wennberg and Cooper, 1999). To show the extent to which variations in Medicare expenditures across regions are correlated with covariates Z, we first consider a parsimonious model that includes just the six measures of health status (heart attack rates, stroke rates, etc.), with results in Column C, Table 3. The set of underlying health measures explain 27 percent of the total variation in Medicare spending. Column D of Table 4 includes additional socioeconomic and behavioral explanatory variables, which raises the R^2 to 0.44. The counterintuitive signs of some coefficients (such as for smokers and elderly living alone), as well as the robust correlation between income and Medicare spending suggests that the coefficients may reflect factors other than health status *per se*. (see McClellan and Skinner, 1997, 1999).

Average physician visits in the last six months are positively and significantly correlated with Medicare expenditures; the R^2 rises to 0.61 (Table 3, Column E). To give some sense for how expenditures differ by physician visits in the last six months, we also report predicted Medicare spending (controlling for Z) from a regression with dummy variables reflecting the regional decile ranking for physician visits in the last six months. These coefficient estimates are shown visually in Figure 6, where Decile 1 (the lowest decile) is anchored at the average Medicare spending in that decile. We can use these estimates to calculate overall Medicare expenditures that are explained by end-of-life physician visits; noting that each decile contains 2.77 million elderly people in 1996, we find net spending relative to the bottom decile is $26 billion larger. Since overall expenditures during the same year were about $138 billion, the variation in expenditures attributable to end of life care accounts for nearly 20 percent of overall spending.

Note that the quality index is *negatively* associated with Medicare expenditures; in other words, Medicare spending does not appear to buy better quality of care (Wennberg, Fisher, and Skinner, 2001). Finally, Column F shows the first-stage regression when all three instruments are included; the R^2 rises to 0.71.

We next consider the linear two-stage regression, in which Medicare expenditures are just identified by physician visits in the last six months of life, presented in Table 5. The first two columns present regression results for the limited set of covariates (the six health conditions, Social Security income, and poverty rates), with Column A weighted by the Medicare population and Column B unweighted. The estimated linear coefficient estimates in Columns A and B (Table 5) of Medicare expenditures on survival are 0.009 and –0.047, respectively. Neither is significantly different from zero. In other words, these results imply that the roughly $26 billion in Medicare expenditures explained by the instrument generate no clinically important (or economically important) influence on survival. Using the Altonji, et al (2001) approach, we also compare the regression coefficient when all covariates Z are excluded. The results are similar (the coefficient for Medicare expenditures is –0.015, with a t-statistic of 0.25), providing further support for the view that the results are not systematically biased by unobservable health status.

Recall from above that one should expect to find in the steady state an improvement in survival rates of 0.11 from an incremental $1,000 in Medicare expenditures per capita. However, we can reject the null hypothesis that the coefficient is 0.11 at the 0.05 significance level for both

Columns A and B. Similar results obtain for the full set of explanatory variables, shown in Columns C and D; the coefficients on Medicare expenditures are not significantly different from zero, and we can reject the null hypothesis of 0.11 at the 0.05 level, thus rejecting the hypothesis that regional variations in Medicare spending satisfy conditions for efficiency.[13]

Note also that in these four columns of Table 5, the coefficient on the index of effective care is positive and significant; a 10 percent increase in the index is associated with survival rates roughly 0.2 percentage points higher. This is large in both a clinical sense and in an economic sense; it implies that increasing the index by 4 percentage points, or one standard deviation (Table 2), would increase survival rates by 0.8 percentage points and, using the benchmark of \$100,000 per life-year, yield benefits equivalent to about \$7,000 per person. This estimate should be interpreted cautiously, since it may be the case that the effectiveness index is associated with a variety of other characteristics of the health care system that we have not controlled for in the regression analysis.

We include the additional instruments (percent with ICU days in the last six months, and hip fracture Medicare expenditures) in Column E of Table 5. Once again, the coefficient on Medicare expenditures is not significant and near zero in magnitude. Furthermore, the model does not reject an overidentification test for the additional two instruments ($p = 0.48$).

To this point, we have not allowed for nonlinearities in either S(M, Z) or in M(X) as in equations (6) and (7). We adopt the general model developed in Newey, Powell, and Vella (1999), but because of the limitations on the size of the data set, we adopt a simple splined

[13] We also experimented with a variety of other variables, such as the percentage in HMOs and the percentage urban population. These variables tended to have similar effects on the coefficient estimates (since they were quite closely correlated), and tended to increase the coefficient estimate to between 0.04 and 0.06, but not significantly different from zero, or 0.11. (Using the urban variable resulted in the loss of one HRR; Ocala FL revamped its zip codes substantially between the 1990 Census and the 1996 HRR crosswalk.) Excluding HRRs with more than 10 percentage point HMO enrollment resulted in a *negative* estimated effect of expenditures on survival (-0.147, significant at the 10 percent level).

function S(M) and M(X) where the five intervals of the spline are evenly distributed along the ranges of M and X.

We estimated the simplified S(M) function conditional on $Z\Gamma$ as shown in Figure 7a (weighted) and 7b (unweighted), with the incremental gain in survival set at zero for the minimum level of Medicare expenditures. The 95 percent confidence intervals are also shown in Figure 7; these are estimated by bootstrapping the combined first and second stage regressions.[14] Rather than a concave shape, as is suggested by diminishing returns (as in Figure 3), the predicted function is convex in Figure 7a, and nearly flat in Figure 7b. The jump in predicted survival for the final quintile in the weighted regression is explained almost entirely by the presence of Miami, which, conditional on other factors, is a surprisingly healthy region (e.g., Fuchs, McClellan, and Skinner, 2001). However, neither estimated model rejects the hypothesis that Medicare expenditures have zero incremental marginal effectiveness.[15]

Discussion and Conclusion

In this paper, we have attempted to test whether the Medicare program is broadly consistent with the efficiency criterion commonly used in public economics where the marginal social value of the last dollar spent on health care (in each region) is equal to the marginal social benefits of the dollar that could have been spent for other worthy causes. We used data on survival rates, Medicare expenditures, and health status measures across 306 hospital referral regions in the United States to test these hypotheses. Our best estimates of the incremental value

---

[14] This is done by predicting the entire S(M) curve for each of the bootstrapped iterations, and then graphing the 2.5^{th} and 97.5^{th} percentiles of the bootstrapped values for S(M) along the entire range of M. For computational convenience, the true distribution is evaluated for ten evenly spaced values of M, and interpolated linearly.

[15] One concern with the specification estimated in Figure 6 is that because S(M) is additively separable from $Z\Gamma$, the productivity curve is restricted to shift just vertically. One might expect that the curve could also shift horizontally; i.e., that the marginal productivity of Medicare spending would be higher (for a given M) in sicker regions. Because a full interaction between M and Z rapidly uses up degrees of freedom, we tested this hypothesis by stratifying the data according to the predicted survival rate (from Table 3, Column 2). Once again, the null hypothesis of zero effects could not be rejected.

of Medicare spending with regard to survival are essentially zero across all regions. The "marginal" patients in this study are not the ones enjoying much improved levels of functioning and survival as a result of technological improvements in the treatment of heart attacks, for example. Instead, they are the ones receiving greater levels of care for chronic illness, particularly those who are sick enough to result commonly in death.

One explanation for these results is that while survival may not be improved, patients either enjoy better health functioning or they may simply prefer such care. These hypotheses are difficult to address using Medicare claims data alone. Available evidence, however, does not support this view (e.g., Guadagnoli et al, 1995). In the SUPPORT study, seriously ill patients were asked about their preferences regarding dying in the hospital and intensive life-saving care, and efforts were made to ensure that they got what they wanted (Lynn et al, 2001). However, there was no correlation between expressed preferences and what they actually got; indeed Pritchard et al (1998) found the only predictor of whether patients died in the hospital was the supply of hospital beds in the area. Finally, a recent study using the Medicare Current Beneficiary Survey (MCBS) found no correlation between HRR-level health care intensity and either patient satisfaction or access to care (Fisher et al, 2001).

We are not able to provide a compelling story of why some regions are so much more intensive than others. The Medicare program is federal, with uniform prices paid for procedures (apart from cost-of-living adjustments), so that one cannot appeal to differences in relative prices to explain why Miami spends so much more than Minneapolis. Indirect prices of care may be lower in urban areas, for example with regard to specialization by physicians (i.e., higher quality and hence lower "real" costs) or travel costs for patients. Still, these differences are not likely to explain differences in utilization between urban areas such as Minneapolis and Miami, or between rural areas that show similar variations.

This "macro" level study does not provide an easy prescription for how to fix the Medicare program, or how to encourage the greater use of effective care. It does suggest, however, that while average effects of technology may be large, incremental benefits of using those technological advances for a wider range of patients is minimal, leading to nearly 20 percent of Medicare expenditures spent for health care with no measurable survival benefit.

References

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber, “Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools,” mimeo, Northwestern University (March 2001).

Angrist, Joshua D., “Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records,” *American Economic Review* 80(3) (June 1990), 313-336.

Bolen, Julie C., Luann Rhodes, Eve E. Powell-Griner, Shayne D. Bland, and Deborah Holtzman, “State-Specific Prevalence of Selected Health Behaviors, by Race and Ethnicity – Behavioral Risk Factor Surveillance System, 1997” *MMWR* 49 (SS02) (March 24, 2000), pp. 1-60.

Chau, Derek, Fisher, Elliott, and Jonathan Skinner, “The Importance of Regional Practice Style in a Cohort of Hip Fracture Patients,” mimeo, Dartmouth Medical School (2001).

Cutler, David M., “The Incidence of Adverse Medical Outcomes under Prospective Payments,” *Econometrica* 63(1) (January 1995): 29-50.

Cutler, David M. "Walking the Tightrope on Medicare Reform" *Journal of Economic Perspectives* 14(2) (Spring 2000), 45-56.

Cutler, David, Mark McClellan, Joseph Newhouse, and Dahlia Remler, “Pricing Heart Attack Treatments,” *Quarterly Journal of Economics* 113(4) (November 1998), 991-1024.

Cutler, David, and Ellen Meara, “The Technology of Birth: Is it Worth It?” *NBER Working Paper* No. 7390 (October 1999).

Currie, Janet, and Jonathan Gruber, “Saving Babies: The Efficacy and Cost of Recent Changes in the Medicaid Eligibility of Pregnant Women,” *Journal of Political Economy* 104(6) (December 1996): 1263-1296.

Deaton, Angus, and Christina Paxson, “Mortality, Education, Income, and Inequality among American Cohorts,” NBER Working Paper No. 7140 (May 1999).

Evans, Robert G., and G. S. Stoddart, “Producing Health, Consuming Health Care,” in Robert G. Evans, Morris L. Barer, and Theodore R. Marmor (eds.) *Why are Some People Healthy and Others Not?* Hawthorne NY: Aldine de Gruyter (1994).

Feenberg, Daniel, and Jonathan Skinner, “Medicare Transfers Across States: Winners and Losers,” *National Tax Journal* (September 2000).

Fisher ES, Wennberg JE, Stukel TA, et al. Associations among hospital capacity, utilization, and mortality of US Medicare beneficiaries, controlling for sociodemographic factors. *Health Services Research* 34(6) (2000) :1351-62.

Fisher, Elliott S., David Wennberg, and Therese Stukel, " " , mimeo, Dartmouth Medical School (2001).

Fisher, Elliott S., Etoile Pinder, Daniel Gottlieb, and Therese Stukel, "Associations Between Regional Variations in Medicare Expenditures, Access to Care, and Patient Satisfaction," mimeo, Dartmouth Medical School (2001).

Fuchs, Victor, Mark McClellan, and Jonathan Skinner, "Area Differences in Utilization of Medical Care and Mortality Among U.S. Elderly," mimeo, Stanford University (2001).

Guadagnoli E, Hauptman PJ, Ayanian JZ, Pashos CL, McNeil BJ, Cleary PD. "Variation in the use of cardiac procedures after acute myocardial infarction" *New England Journal of Medicine*. 333 (1995) :573-8.

Hadley, Jack, "Medicare Spending and Mortality Rates of the Elderly." *Inquiry* 25 (1988): 485-93.

Hearst, Norman, Thomas B. Newman, and Stephen B. Hulley, "Delayed Effects of the Military Draft on Mortality," *New England Journal of Medicine* 314(10) (March 6, 1986): 620-624.

Lynn, J., Arkes, HR, Stevens, M., et al. "Rethinking Fundamental Assumptions: SUPPORT's Implications for Future Reform." *Journal of the American Geriatric Society* 48 (2000): S214-21.

McClellan Mark, McNeil Barbara J, Newhouse Joseph P., "Does More Intensive Treatment of Acute Myocardial Infarction in the Elderly Reduce Mortality? Analysis using Instrumental Variables," *JAMA* 272 (September 1994), 859-866.

McClellan, Mark, and Jonathan Skinner, "The Incidence of Medicare," NBER Working Paper (1997), revised 2000.

McClellan, Mark, and Jonathan Skinner, "Medicare Reform: Who Pays and Who Benefits?" *Health Affairs* 18(1) (Jan/Feb 1999), 48-62.

Newey, Whitney K., James L. Powell, and Francis Vella, "Nonparametric Estimation of Triangular Simultaneous Equations Models," *Econometrica* 67(3) (May 1999). 565-603.

Pardes, H., Manton, KG, Lander, ES et al, "Effects of Medicare Research on Health Care and Economy," *Science* 283 (January 1, 1999): 36-37.

Pritchard RS, Fisher ES, Teno JM, et al. Influence of patient preferences and local health system characteristics on the place of death. SUPPORT Investigators. Study to Understand Prognoses and Preferences for Risks and Outcomes of Treatment . *Journal of the American Geriatric Society* 46 (1998):1242-50.

Schneider, EL, "Aging in the Third Millennium," *Science* 283 (February 5, 1999): 796-97.

Skinner, Jonathan, and John E. Wennberg, “Regional Inequality in Medicare Spending: The Key to Medicare Reform?” in A. Garber (ed.) *Frontiers of Health Economics 3.* Cambridge, MA: MIT Press (2000).

Wennberg John E and Megan M Cooper, *The Quality of Medical Care in the United States: A Report on the Medicare Program. The Dartmouth Atlas of Health Care in the United States.* American Health Association Press, Chicago. (1999).

Wennberg, John E, Elliott Fisher, and Jonathan Skinner, “Geography and the Debate Over Medicare Reform,” mimeo, Dartmouth Medical School (May 2001).

Wilkinson, Richard, *Unhealthy Societies: The Affliction of Inequality*. Routledge (1997).

Wilmoth, John R., http://www.demog.berkeley.edu/wilmoth/mortality/ (2001).

Table 1: Rates of Specific Procedures per 1000 Decedents, by Regional Frequency of Physician Visits in the Last Six Months

Physician Service Code	Regions in highest decile of MD Visits During Last Six Months	Regions in lowest decile of MD Visits During Last Six Months	Ratio
Physician Visits			
Physician Office Outpatient Visits	4,453	3,051	1.46
Physician Visit for Initial Hospital Care	1,442	804	1.79
Specialist Initial Inpatient Consult	3,087	628	4.92
Diagnostic Testing			
Upper Gastrointestinal Endoscopy	132	64	2.06
Cat Scan of Head/Skull/Brain	492	236	2.08
Chest X-Ray	6,631	2,700	2.46
Doppler/Echocardiagram of Heart	799	268	2.98
Electrocardiogram & Report	3,888	1,161	3.35
Treatments for Serious Chronic Illnesses			
Insertion of Emergency Airway	140	42	3.33
Hemodialysis (related to kidney failure)	384	87	4.41
Gastrostomy Tube Placement/ Change (feeding tube)	136	25	5.44
Continuous Ventilator Management (mechanical breathing apparatus)	387	50	7.74

Notes: Age 75-90, N = 15,097 (Decile 1), N =17,225 (Decile 10). CPT codes: Initial Hospital Care (99221-99223), Office Outpatient Visits (99201-99215), Initial Inpatient Consult (99251-99223), Cat Scan of Head/Skull/Brain (70450 - 70482), Chest X-Ray (71010 - 71035), Doppler/Echo of Heart (93320-93325), Electrocardiogram & Report (93000 - 93010), Laboratory Pathology Consultant (80500-80502), Upper GI Endoscopy (43234-43243,) Hemodialysis (90935, 90937), Gastrectomy (43246, 43830, 43832, 43750, 43760, 43761), Emergency Airway (31500), Continuous Ventilator Management (94657)

Table 2: Summary Statistics

Dependent Var.	Mean	Standard Deviation	Minimum	Maximum
Medicare Expenditures (in $1,000)	5.08	0.86	3.12	8.86
Heart Attack (AMI) Rate (per 1000)	19.45	2.91	11.44	29.44
Stroke (CVA) Rate (per 1000)	22.95	2.84	15.24	32.47
Gastrointestinal Bleeding Rate (per 1000)	15.16	1.64	10.54	20.43
Colon Cancer Rate (per 1000)	4.74	0.54	2.83	6.34
Lung Cancer Rate (per 1000)	1.42	0.28	0.50	2.28
Hip Fracture Rate (per 1000)	15.55	1.53	9.20	19.62
MD visits in the Last Six Months (per decedent)	24.20	7.12	8.5	47.9
Decedents Admitted to ICU in Last 6 Months (Percent)	31.26	5.71	14.2	49.3
Average Medicare Expend., Hip FX patients ($1,000)	11.20	0.84	9.23	15.09
Fraction Living in Poverty	13.10	5.59	4.66	32.73
Average Social Security Income (in $1,000)	7.760	0.61	6.056	9.532
Percent of elderly living alone	35.02	3.62	20.8	41.9
Cigarette Smokers (Percent)	23.47	2.47	13.7	30.8
Obese (Percent)	17.09	1.84	11.9	22.0
Seatbelt users (Percent)	28.39	8.23	12.8	59.6
Binge Drinkers (Percent)	13.41	3.61	6.3	23.2
High School Graduates (Percent)	49.37	4.68	31.91	58.89
College Graduates (Percent)	25.55	6.23	11.49	52.24
Effectiveness Index	47.08	4.06	32.21	56.74
Survival Rate (percent)	94.70	0.34	93.78	95.99

Table 3: Regression Models of Survival and Medicare Expenditures

Dependent Var.	Survival	Survival	Medicare Expend.	Medicare Expend.	Medicare Expend.	Medicare Expend.
	A	B	C	D	E	F
AMI Rate (per 1000)	-0.010 (1.2)	-.007 (0.8)	-0.061 (2.3)	-0.017 (0.6)	0.024 (1.1)	0.028 (1.4)
Stroke Rate (per 1000)	-0.038 (4.5)	-0.043 (4.4)	0.048 (1.8)	-0.008 (0.3)	-0.032 (1.2)	-0.033 (1.4)
Gastrointestinal Bleeding Rate (per 1000)	-0.051 (2.6)	-0.054 (3.4)	0.190 (3.6)	0.177 (4.0)	0.091 (2.4)	0.028 (0.7)
Colon Cancer Rate (per 1000)	-0.011 (0.4)	-0.006 (0.2)	-0.062 (0.6)	0.141 (1.3)	-0.164 (1.6)	-0.076 (0.8)
Lung Cancer Rate (per 1000)	0.133 (1.4)	0.140 (1.9)	0.467 (2.3)	0.499 (2.9)	0.421 (2.8)	0.412 (3.0)
Hip Fracture Rate (per 1000)	-0.047 (3.8)	-0.031 (2.0)	0.056 (1.4)	0.119 (3.0)	0.153 (5.3)	0.151 (5.2)
Fraction Living in Poverty		.026 (3.0)		0.073 (3.1)	0.053 (3.2)	0.052 (3.4)
Average Social Security Income (in $1,000)		0.205 (2.8)		0.432 (2.4)	0.054 (0.4)	0.026 (0.2)
High School Graduate (Percent)		-0.008 (1.6)		-0.043 (2.3)	0.012 (0.9)	0.022 (1.7)
College Graduate (Percent)		-0.003 (0.6)		-0.023 (1.9)	-0.010 (1.1)	0.002 (0.3)
Binge Drinking (Percent)		0.003 (0.6)		0.058 (3.6)	0.065 (4.8)	0.066 (5.7)
Cigarette Smokers (Percent)		-0.027 (2.8)		-0.078 (3.2)	-0.082 (3.8)	-0.059 (3.0)
Obesity (Percent)		-0.002 (0.1)		0.008 (0.3)	0.020 (0.8)	0.017 (0.7)
Seatbelt Use (Percent)		0.011 (4.6)		-0.001 (0.1)	-0.001 (0.2)	0.007 (1.2)
Elderly Living Alone (Percent)		-0.017 (2.2)		-0.047 (2.2)	-0.026 (1.7)	-0.036 (3.2)
Index of Effective Care					-0.022 (1.6)	-0.016 (1.2)
MD Visits in the Last Six Months (Average)					0.071 (7.6)	0.045 (5.4)
Percent Admitted to to ICU in Last Six Months						0.027 (3.2)
Six -Month Medicare Spending: Hip Fx Cohort						0.362 (7.8)
Constant	97.124	96.372	1.052	0.717	1.306	-3.682
R^2	0.46	0.56	0.27	0.44	0.61	0.71

Notes: N = 306. Robust standard errors; absolute value of t-statistics in parentheses. All sample sizes weighted by Medicare population. Low variation conditions (e.g., AMI, stroke) and effectiveness index are from 1995/96, Medicare expenditures data is for 1996, poverty and Social Security data from 1990 Census, and CDC behavioral data from 1997.

Table 4: Variations in Population-Based Health Status, by Tercile of Physician Visits in the Last Six Months of Life

Dependent Var.	First (Lowest) Tercile	Second Tercile	Third (Highest) Tercile	Correlation Coeff. with MD Visits in Last 6 Months (p-value)
Average Physician Visits in the Last 6 Months	17.49	22.90	32.15	1.00
Medicare Expenditures (in $1,000)	4.486	5.187	5.576	0.53 (<0.01)
Heart Attack (AMI) Rate (per 1000)	19.09	19.87	19.38	0.003 (0.96)
Stroke (CVA) Rate (per 1000)	21.52	23.85	23.45	0.23 (< 0.01)
Hip Fracture Rate (per 1000)	15.52	16.10	15.04	-0.13 (0.02)
Fraction Living in Poverty	12.93	14.76	11.60	-0.063 (0.23)
Average Social Security Income (in $1,000)	7.68	7.57	8.03	0.21 (<0.01)
Fraction of elderly living alone	35.95	35.30	33.82	-0.26 (<0.01)
Cigarette Smokers (Percent)	23.13	24.03	23.25	-0.03 (0.66)
Surgery Index	38.62	39.90	37.85	-0.20 (<0.01)
Predicted Survival (from Table 3, Column B)	**94.74**	**94.61**	**94.73**	**0.01 (0.88)**

Table 5: Instrumental Variable Regression Estimates of Factors Affecting Survival (Dependent Variable is the One-Year Survival Rate)

Column	A	B	C	D	E*
Medicare Expenditures (in $1,000)	0.009 (0.2)	-0.047 (0.9)	0.002 (0.0)	-0.049 (1.0)	0.005 (0.1)
AMI Rate (per 1000)	-.014 (1.7)	-.015 (2.2)	-.014 (1.7)	-.010 (1.4)	-.014 (1.7)
Stroke Rate (per 1000)	-0.039 (4.3)	-0.037 (5.0)	-0.035 (3.7)	-0.030 (4.0)	-0.035 (3.7)
Gastrointestinal Bleed Rate (per 1000)	-0.048 (2.5)	-0.044 (2.7)	-0.049 (2.6)	-0.047 (3.0)	-0.050 (2.9)
Colon Cancer Rate (per 1000)	0.002 (0.1)	0.003 (0.1)	-0.012 (0.3)	-0.012 (0.4)	-0.013 (0.4)
Lung Cancer Rate (per 1000)	0.051 (0.7)	0.019 (0.4)	0.089 (1.3)	0.072 (1.4)	0.087 (1.3)
Hip Fracture Rate (per 1000)	-0.052 (3.8)	-0.056 (4.9)	-0.033 (2.3)	-0.035 (2.8)	-0.034 (2.3)
Fraction Living in Poverty	0.030 (3.8)	0.037 (5.2)	0.029 (3.3)	0.035 (4.3)	0.029 (3.4)
Average Social Security Income (in $1,000)	0.205 (2.9)	0.215 (3.3)	0.200 (2.9)	0.215 (3.3)	0.198 (2.9)
Effectiveness Index	0.019 (3.4)	0.013 (2.7)	0.021 (3.9)	0.016 (3.4)	0.021 (4.2)
Constant	94.51	94.83	95.38	95.26	94.53
Extended set of variables included?**	NO	NO	YES	YES	YES
Regression weighted by Medicare population?	YES	NO	YES	NO	YES
R^2	0.52	0.52	0.58	0.58	0.58

Notes: N = 306. * Instrument is physician visits in the last six months, except Column E that includes as instruments percentage of decedents admitted to the ICU in their last six months, and Medicare expenditures in the last six month. ** Set of extended health measures include percent high school graduates, percent college graduates, percent of elderly population living alone, and percentage of population who are binge drinkers, cigarette smokers, obese, and seatbelt users. Robust standard errors; absolute value of t-statistics in parentheses.All sample sizes weighted by Medicare population. Low variation conditions (e.g., AMI, stroke) and effectiveness index are from 1995/96, Medicare expenditures data is for 1996, poverty and Social Security data from 1990 Census, and CDC behavioral data from 1997.

Figure 1: Non-Capitated Medicare Expenditures Per Enrollee, 1996

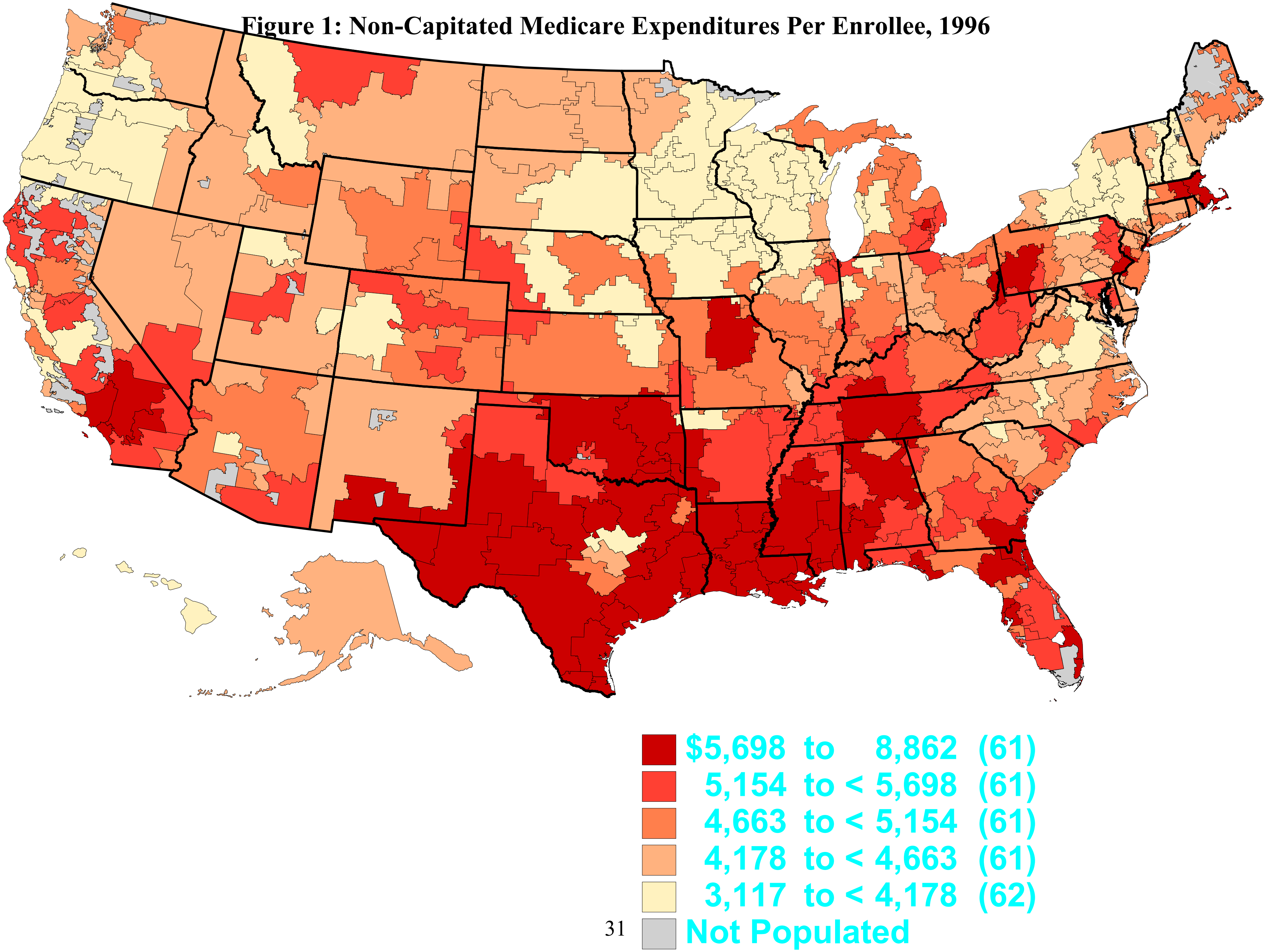


Figure 2: Per Capita Medicare Expenditures and Mortality Rates in the Medicare Population, 1996

(Note: Medicare expenditures and mortality rates are age-sex-race adjusted, both apply to the fee-for-service population)

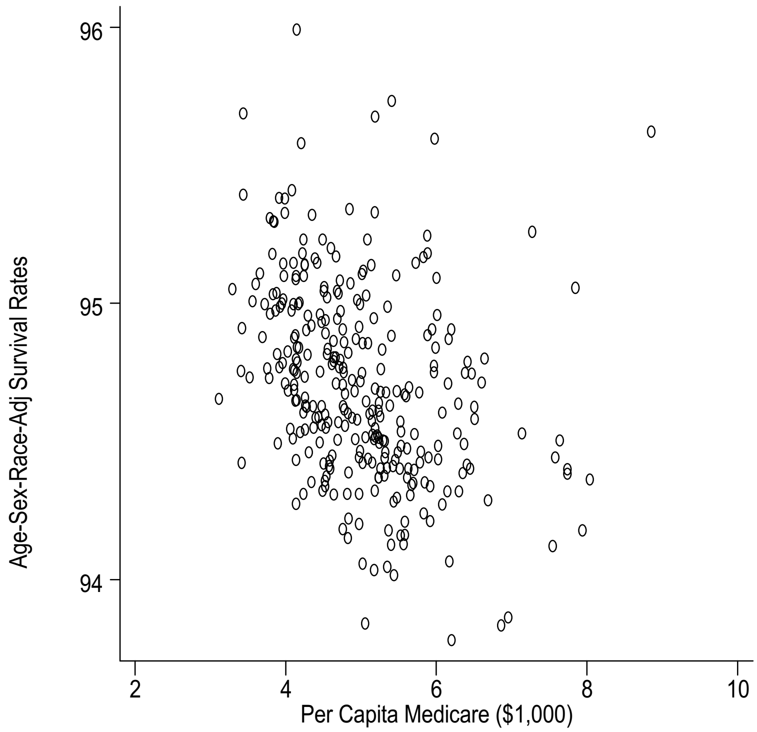
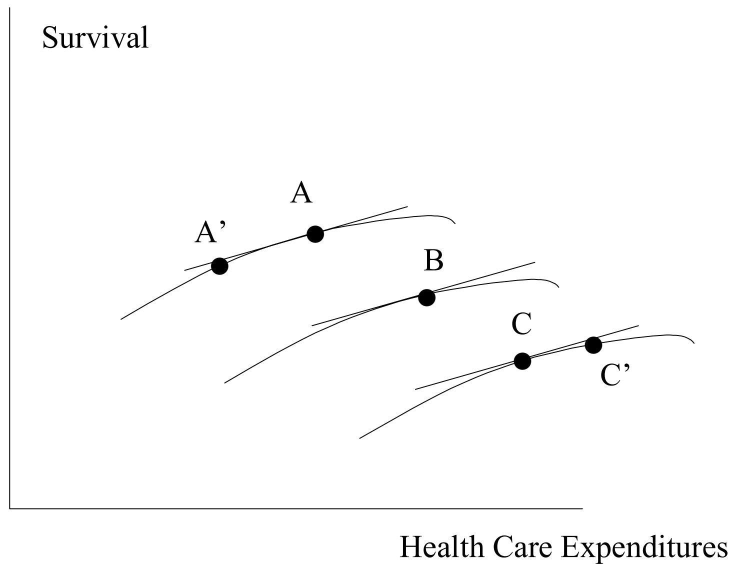


Figure 3: Efficiency in Health Care



Note: A, B, and C represent regions. At each point, the slope of the health care productivity curve is equal (and shown by the straight line passing through points A, B, and C)

Figure 4: Baseline Survival Curve and Counterfactual Survival Curve After 25% decline in Mortality Rate

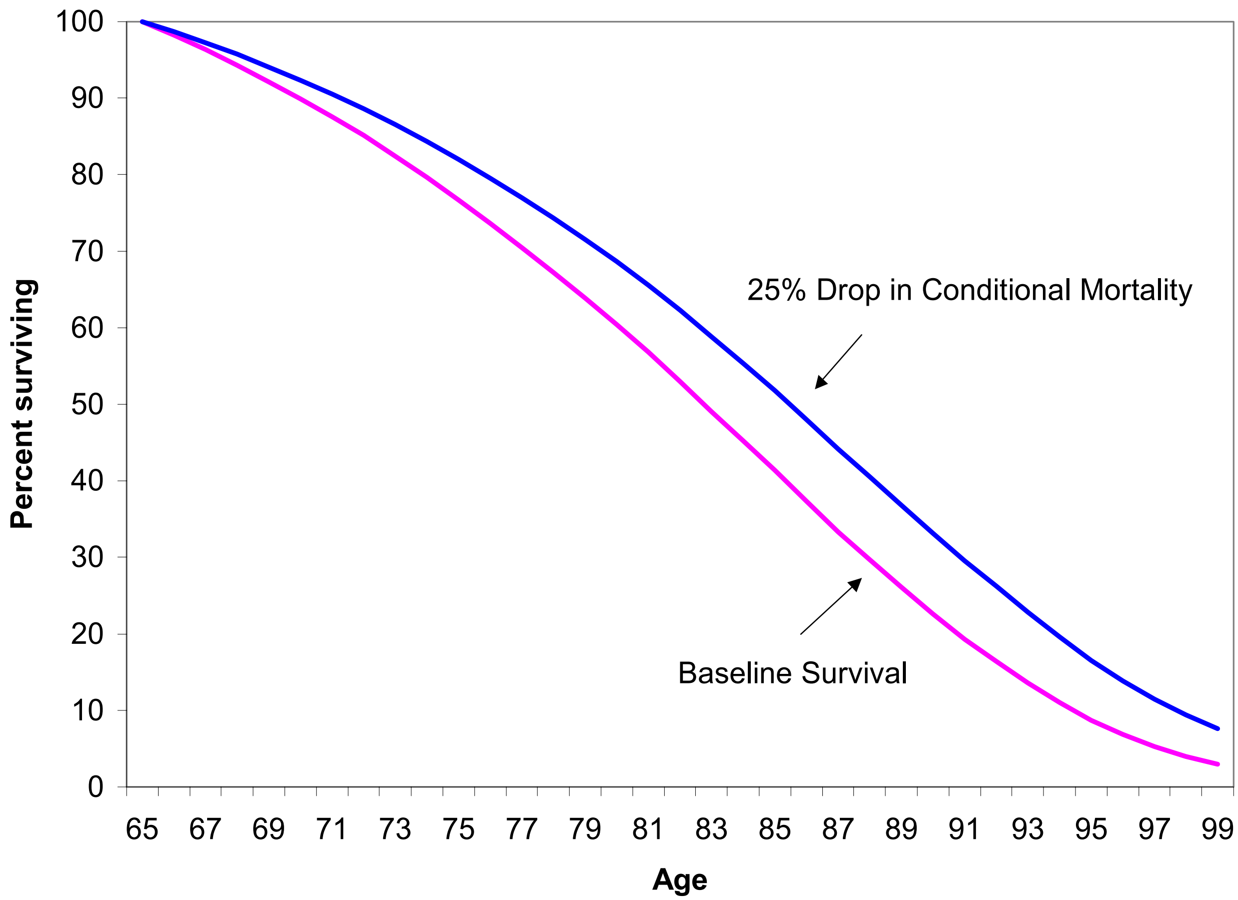
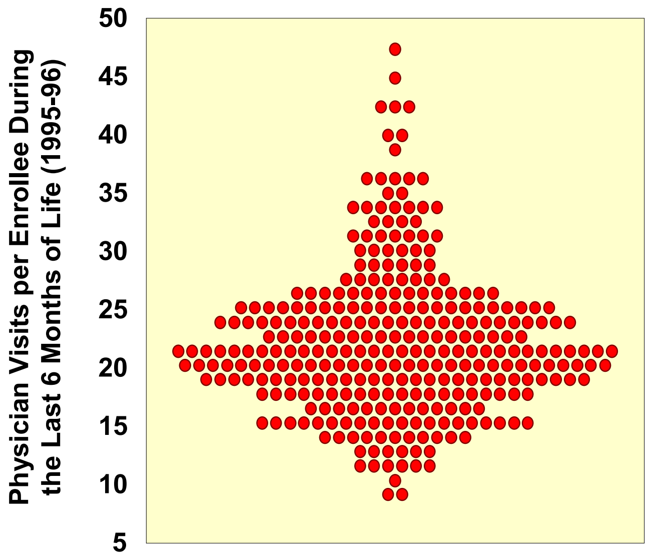


Figure 5: Average Number of Physician Visits per Decedent During the Last Six Months of Life (1995-96)



Source: www.dartmouthatlas.org, Wennberg and Cooper, 1999.

Figure 6: Average Per Capita Medicare Spending 1995/96, by Decile of Physician Visits in the Last Six Months

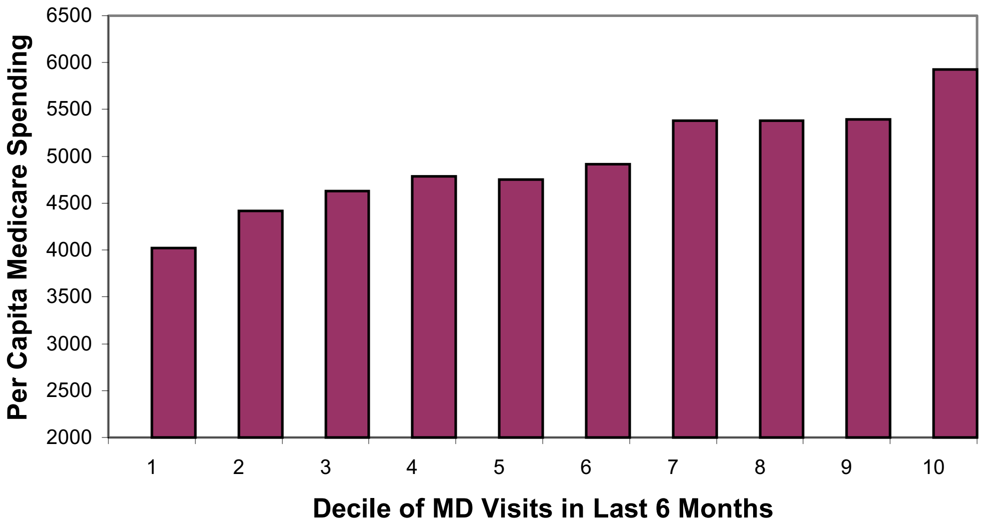
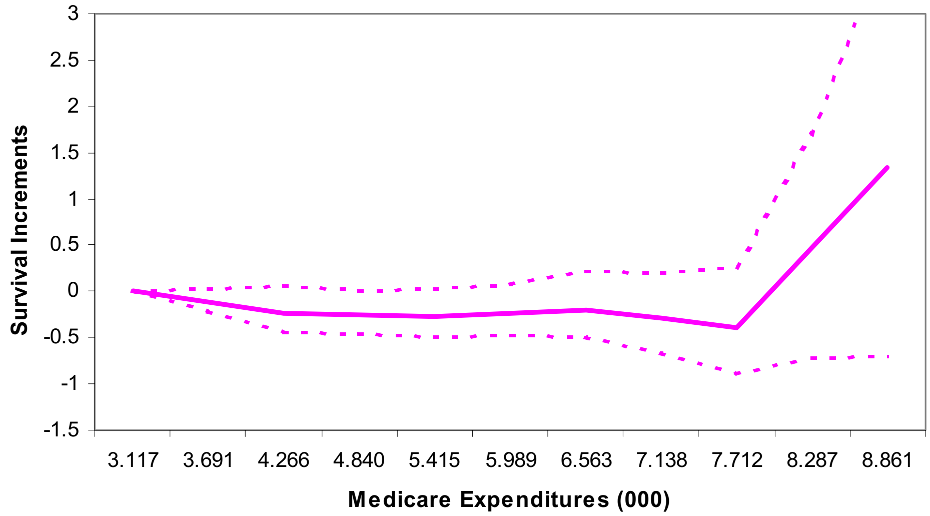


Figure 7: Semiparametric Estimates of the Marginal Effectiveness of Medicare Expenditures

a. Weighted Data



b. Unweighted Data