

© Health Research and Educational Trust  
DOI: 10.1111/j.1475-6773.2011.01276.x  
RESEARCH ARTICLE

# Medical Spending and the Health of the Elderly

*Jack Hadley, Timothy Waidmann, Stephen Zuckerman, and Robert A. Berenson*

---

**Objective.** To estimate the relationship between variations in medical spending and health outcomes of the elderly.

**Data Sources.** 1992–2002 Medicare Current Beneficiary Surveys.

**Study Design.** We used instrumental variable (IV) estimation to identify the relationships between alternative measures of elderly Medicare beneficiaries' medical spending over a 3-year observation period and health status, measured by the Health and Activity Limitation Index (HALex) and survival status at the end of the 3 years. We used the Dartmouth Atlas End-of-Life Expenditure Index defined for hospital referral regions in 1996 as the exogenous identifying variable to construct the IVs for medical spending.

**Data Collection/Extraction Methods.** The analysis sample includes 17,438 elderly (age > 64) beneficiaries who entered the Medicare Current Beneficiary Survey in the fall of each year from 1991 to 1999, were not institutionalized at baseline, stayed in fee-for-service Medicare for the entire observation period, and survived for at least 2 years. Measures of baseline health were constructed from information obtained in the fall of the year the person entered the survey, and changes in health were from subsequent interviews over the entire observation period. Medicare and total medical spending were constructed from Medicare claims and self-reports of other spending over the entire observation period.

**Principal Findings.** IV estimation results in a positive and statistically significant relationship between medical spending and better health: 10 percent greater medical spending over the prior 3 years (mean = U.S.\$2,709) is associated with a 1.9 percent larger HALex value ( $p = .045$ ; range 1.2–2.2 percent depending on medical spending measure) and a 1.5 percent greater survival probability ( $p = .039$ ; range 1.2–1.7 percent).

**Conclusions.** On average, greater medical spending is associated with better health status of Medicare beneficiaries, implying that across-the-board reductions in Medicare spending may result in poorer health for some beneficiaries.

**Key Words.** Medicare efficiency, medical spending, health outcomes

---

A significant body of recent research reports wide variations in Medicare spending per beneficiary across geographic areas, but with little apparent variation in the quality of care or health outcomes (Skinner, Fisher, and

Wennberg 2005; Wennberg et al. 2008; Fisher et al. 2009). Based on this research, policy makers have considered proposals that would limit Medicare payments in high-cost areas or pay bonuses in low-cost areas (Congressional Budget Office [CBO] 2008; U.S. Senate Committee on Finance 2009; Abelson and Harris 2010).

Although prior research has studied people who age into Medicare and found that greater medical care use following Medicare coverage improved health (Lichtenberg 2002; McWilliams et al. 2007a, b, 2003; Card, Dobkin, and Maestas 2009), none of these studies is directly relevant to the question of whether variations in medical spending for people already covered by Medicare affect their health. Earlier studies that found a positive relationship between medical spending and health at the geographic level may be out of date (Hadley 1982; Hadley 1988). More recently, studies by Doyle (2008) and Martin, Rice, and Smith (2008) used instrumental variable (IV) analysis and found that people treated in higher spending areas had better health outcomes, but neither focused only on the Medicare population. Kaestner and Silber (2010) estimated a positive relationship between Medicare spending and health outcomes, but only for hospitalized patients with particular medical conditions.

Like these recent studies, we use IV analysis to investigate the relationship between medical spending and health using data on individual Medicare beneficiaries. If a person-level analysis also finds no relationship between medical spending and health of the elderly, it will reinforce the finding of Medicare inefficiencies drawn from geographic analyses (Fisher et al. 2009).

## METHODS

### *Data*

The analysis sample includes 17,438 Medicare beneficiaries drawn from the Medicare Current Beneficiary Survey (MCBS) (Adler 1998), which uses a rotating panel design and conducts multiple interviews over approximately 3.5 years. Baseline health information is obtained in the fall of the year the person enters the survey. Medical spending information is collected over the

---

Address correspondence to Jack Hadley, Ph.D., Department of Health Administration and Policy, College of Health and Human Services, George Mason University, 4400 University Dr., MSN 2G7, Fairfax VA 22030; e-mail: jhadley1@gmu.edu. Timothy Waidmann, Ph.D., Stephen Zuckerman, Ph.D., and Robert A. Berenson, M.D., are with the Health Policy Center, The Urban Institute, Washington, DC.

next 3 calendar years from Medicare claims and self-reports of other spending. Health information is updated periodically over the 3.5 years.

The sample consists of elderly (age > 64) beneficiaries who entered the MCBS in the fall of each year from 1991 to 1999, were not institutionalized at baseline, stayed in fee-for-service Medicare for the entire observation period, and survived for at least 2 years.<sup>1</sup> (Note that this last criterion does not create a selected sample based on prior health, because every new random sample requires that people be alive at the beginning of the survey period.) Final health status is measured at the end of each person's third year in the MCBS (from 1994 to 2002).

### *Conceptual Framework*

Statistical estimation of the effect of medical spending on health faces significant challenges. Foremost is the endogeneity or observational data bias problem that typically characterizes nonexperimental data (Newhouse and McClellan 1998; McClellan and Newhouse 2000). Poor initial health generally leads to high medical spending, which confounds the ability to detect the effect of medical spending on health because people in poor initial health often have worse health outcomes than people in good initial health. Consequently, absent randomly assigning people to receive different amounts of medical care, empirical estimation with observational data typically finds that people who use more medical care have worse health outcomes (Hadley 2003). Another confounding factor, especially among the elderly, is that medical care's primary impact may often be to slow the rate of health deterioration relative to what it would have been without medical intervention, rather than reestablishing or improving their initial health state. Thus, inadequate risk adjustment as a control for the expected health outcome can also lead to biased conclusions.

Based on these considerations and the nature of our data, we specify the empirical analysis as the IV estimation of a first-stage medical spending equation and a second-stage health outcome equation. We define the health outcome as final health at the *end of a 3-year observation period* ( $FH_3$ ) and note that we observe a vector of health conditions ( $H_t$ ) at various times over the prior 3 years ( $t = 1, 2, 3$ ). Medical spending is observed annually over the 3 years and is measured in alternative specifications as either spending in year 3 ( $M_3$ ) or cumulative spending over the 3 years ( $M = \sum M_t$ ).

Equation (1) posits that final health at the end of year 3 ( $FH_3$ ) is a function of a vector of health measures from the three prior time periods ( $H_t$ ), medical

spending (either contemporaneous,  $M_3$ , or cumulative,  $M$ ), other exogenous factors ( $X$ ) that affect health (and may be measured at baseline and/or during the observation period), and a random error term  $v$ . Equation (2) posits that medical spending (represented here by cumulative medical spending  $M$ ) is a function of the expected final health outcome ( $FH_3$ ), contemporaneous and prior health conditions, exogenous factors that affect health, and exogenous factors that influence medical spending but not health ( $Z$ ):

$$FH_3 = f(M, H_t, X, v), t = 1, 2, 3. \quad (1)$$

$$M = g(FH_3, H_t, X, Z, u), t = 1, 2, 3. \quad (2)$$

Because expected final health is not observed over the time period during which  $M$  is measured, the error term  $u$  in an observational analysis is replaced by  $u^* = h(u, FH_3)$ , which results in the estimating equation

$$M = g(H_t, X, Z, u^*), t = 1, 2, 3. \quad (2^*)$$

Comparing equations (1) and (2\*) illustrates the source of the observational data bias. A person with a poor expected final health outcome, presumably as a function of prior health conditions, some of which are not observed, may be likely to use more medical care than someone with a good expected final health outcome. Thus,  $M$  in equation (1) depends on  $FH_3$  through the error term  $u^*$  in equation (2\*).

IV estimation addresses the potential bias in estimating equation (1) using the variation in  $M$  caused by  $Z$ , which is independent of health, to create an “instrument”  $M^*$  generated from equation (2\*) and substituting  $M^*$  for actual medical spending in equation (1). A positive coefficient on  $M^*$  in equation (1) implies that the person’s health improves with increased medical spending.

Measuring medical spending by  $M$ , cumulative spending over 3 years, reflects the assumption that health at any point in time depends on medical expenditures over multiple time periods, not just the current time period. Ideally, one might try to specify the model as an explicit multiperiod model with repeated annual observations on each person and a dynamic relationship between health at the end of each period and current and lagged medical spending over multiple prior periods. Estimating such a model would be complex because of potentially correlated errors over time and the probable need to develop IV estimates of each year’s medical spending. Moreover, because the MCBS sample structure limits the observation period to 3 years, the number of observations per person would be very small.

Defining  $M$  as cumulative spending over the 3 observation years in the MCBS represents a simplification of this more complicated underlying approach to thinking about the relationship between health and medical spending. However, in order to parallel most previous research and to assess the sensitivity of the results to alternative definitions of the medical spending variable, we estimate alternative versions of the model. One measures medical spending during year 3 only,  $M_3$ . In a second set of alternative specifications, we replace  $M$ , medical spending from all sources, by Medicare spending, either in year 3 ( $MCR_3$ ) or cumulative over the entire 3-year observation period (MCR).

### *Variables and Empirical Specification*

*Dependent Variables.* We use two measures of the final health outcome ( $FH_3$ ). One is a dichotomous measure of whether the person survived to the end of the third year in the MCBS. The other is the Health and Activity Limitations Index (HALex) value at end of the observation period. The HALex is “. . . a generic measure of health that consists of two sets of attributes: perceived health and activity limitation. Using a multiattribute utility scoring system, information from these attributes was combined to form a single score that represents health-related quality of life on a 0 to 1 continuum” (Erickson, Wilson, and Shannon 1995; Erickson 1998). Perceived health is measured by self-reported general health status (excellent, very good, good, fair, poor), and activity limitations are measured by responses to questions about instrumental activities of daily living (IADLs) and activities of daily living (ADLs), for example, ability to walk several blocks, do heavy or light housework, dress, bathe, get into and out of chairs and beds, and feed oneself. HALex values range from 0.995 for someone in excellent health with no limitations to 0.1 for someone in poor health with ADL limitations. People who die are assigned a HALex value of 0.

*Medical Spending.* Medical spending is defined alternatively as total medical spending (Medicare program spending plus spending from all other sources) and Medicare program spending, both over the entire observation period and during the third year only. The Medicare portion of total spending is based on Medicare claims data and captures Medicare program payments. Spending from other sources, such as out-of-pocket or supplementary insurance, is self-reported. The spending data are expressed in 2002 dollars after adjusting for cross-sectional price differences using the Medicare

inpatient hospital wage index, and secular price inflation using the medical care component of the consumer price index.<sup>2</sup>

*Exogenous Identifying Variable.* Specification of the exogenous identifying instrument,  $Z$ , is guided by three criteria: the instrument should be strongly correlated with the endogenous spending variable; it should be uncorrelated with health, other than through its impact on spending, or with unobserved factors that affect health; and there should be a plausible and convincing rationale for its hypothesized relationships with spending and health. The first criterion, instrument strength, is generally satisfied by an  $F$ -statistic  $\geq 10$  from testing the hypothesis that the instrument coefficients are jointly equal to 0 in the first-stage spending equation, and by a meaningful partial  $R^2$  (Staiger and Stock 1997). A “weak” instrument ( $F < 5$ ) signals that the estimated coefficient in the second-stage equation may be biased toward the ordinary least squares (OLS) result. The second criterion, the excludability restriction, cannot be tested when the model is just identified by a single exogenous instrument. Ultimately, the third criterion is the most important because it justifies the relationships assessed by the statistical tests.

Given the considerable controversy over instrument selection (Chandra, Fisher, and Skinner 2007), we specified the instrument as the Dartmouth Atlas End-of-Life Expenditure Index (EOLEI) (Fisher et al. 2003a, b), which was constructed by measuring Medicare spending in the last 6 months of life for beneficiaries grouped by geographic area—hospital referral regions (HRRs) constructed by the Dartmouth Atlas (Wennberg and Cooper 1998). EOLEI values for 1996, the middle of our observation period, were assigned to individual beneficiaries based on their residential zip code, which was matched to a list of zip codes by HRR.

Although the use of the EOLEI as a risk adjuster has been criticized on the grounds that decedents do not all have the same health status over the 6 months preceding death (Bach, Schrag, and Begg 2004; Neuberger 2009; Ong et al. 2009; Bach 2010), the EOLEI may still be an adequate instrument for this analysis because (1) it is defined at the area level for a single year, while our data are for individuals over a 12-year period and (2) it is constructed from data from an essentially nonoverlapping population. To be credible as an instrument, the EOLEI must be significantly related to our measures of individuals’ spending without being directly related to their health outcomes. This will be the case if exogenous geographic factors, such as the number and mix of medical specialists, the presence of major centers of medical education

and health research, the extent of competition among physicians and hospitals, and the local regulatory environment, all of which arguably influence practice patterns and, therefore, medical spending, also influence variations in the EOLEI, even if the end-of-life population used to construct the index does not control perfectly for variations in populations' health across geographic areas.

Ultimately, the plausibility of the EOLEI as an instrument depends on the extent to which it captures variations in practice patterns rather than variations in the health of people at the end-of-life. While this question cannot be resolved fully, numerous studies have shown that the EOLEI varies considerably with local medical capacity and the mix of medical specialties and teaching hospitals, suggesting that it is a plausible indicator of geographic differences in practice patterns (Center for Clinical Evaluative Studies 1999; Sirovich et al. 2008; Wennberg et al. 2008; Baicker and Chandra 2009). We also conduct sensitivity tests that use transformations of the continuous EOLEI values into dummy categorical variables representing deciles of the EOLEI (which is similar to the approach used by Fisher et al. 2003a, b), mean values of the EOLEI by decile, and a decile rank variable that takes values from 1 to 10 corresponding the decile grouping of HRRs by the value of the EOLEI.<sup>3</sup> These alternatives should be less strongly correlated with individuals' health outcomes than the continuous EOLEI.

*Prior Health and Other Exogenous Control Variables.* The vector  $H_t$ ,  $t = 1, 2, 3$  represents initial health at entry to the MCBS and intermediate health changes, which are measured by 43 variables collected at baseline and at the ends of survey years 1 and 2 (see Table 1). These variables include baseline measures of smoking status and body mass index, and baseline and annual measures of self-reported health status, activity limitations, and specific health conditions (diabetes, heart disease, skin cancer, other cancers, stroke, high blood pressure, arteriosclerosis, and arthritis). Because of the large number of variables measuring prior health, we summarize for convenience the annual changes in self-reported health status and activity limitations between baseline, the end of year 1, and the end of year 2 by calculating the changes in the corresponding HALex values at each of those observation points.

Other factors that influence both medical spending and health outcome (the  $X$  variables) include age (dummy variables for ages 68–70, 71–74, 75–79, 80–84, and 85+, relative to age 65–67), gender, race, and ethnicity (dummy

Table 1: Mean Values of Study Variables

<i>Variable (N= 17,438)</i>	<i>Mean</i>	<i>SD</i>
Health outcome measures		
Final HALex value	0.610	0.313
Survival (%)	95.211	21.354
Medical spending measures (U.S.\$1,000s)		
Total spending, all years	27.092	37.927
Total spending, year 3	10.387	18.469
Medicare spending, all years	16.203	28.294
Medicare spending, year 3	6.324	14.506
Prior health status		
Self-reported general health status (reference = excellent)		
Very good	0.285	0.451
Good	0.302	0.459
Fair	0.159	0.366
Poor	0.056	0.230
Any ADL	0.276	0.447
Any IADL	0.339	0.473
Change in HALex, baseline to year 1	0.007	0.221
Change in HALex, year 1–year 2	–0.012	0.217
Conditions		
Had hardening of the arteries at baseline	0.097	0.296
Developed hardening of the arteries, year 1	0.021	0.144
Developed hardening of the arteries, year 2	0.021	0.144
Had hypertension at baseline	0.484	0.500
Developed hypertension, year 1	0.045	0.207
Developed hypertension, year 2	0.038	0.191
Had heart attack at baseline	0.122	0.328
Developed heart attack, year 1	0.015	0.122
Developed heart attack, year 2	0.015	0.122
Had angina pectoris or other coronary heart disease at baseline	0.115	0.319
Developed angina or coronary heart disease, year 1	0.022	0.147
Developed angina or coronary heart disease, year 2	0.020	0.141
Had other heart conditions at baseline	0.217	0.412
Developed other heart condition, year 1	0.048	0.214
Developed other heart condition, year 2	0.044	0.205
Had stroke or brain hemorrhage at baseline	0.082	0.274
Developed stroke or brain hemorrhage, year 1	0.020	0.139
Developed stroke or brain hemorrhage, year 2	0.034	0.182
Had skin cancer at baseline	0.157	0.364
Developed skin cancer, year 1	0.026	0.160
Developed skin cancer, year 2	0.032	0.176
Had other cancer at baseline	0.152	0.359
Developed cancer, year 1	0.019	0.138
Developed cancer, year 2	0.020	0.139
Had diabetes at baseline	0.133	0.339
Developed diabetes, year 1	0.038	0.192
Developed diabetes, year 2	0.067	0.249

*continued*



Table 1. *Continued*

<i>Variable (N = 17,438)</i>	<i>Mean</i>	<i>SD</i>
Additional health variables		
Proxy respondent	0.057	0.231
Current smoker	0.138	0.345
Former smoker	0.418	0.493
LT 10th percentile of body mass index (reference = 25th–75th percentile)	0.085	0.279
10th–25th percentile of body mass index	0.151	0.358
GT 75th percentile of body mass index	0.248	0.432
Upper body limitation (some or lot of difficulty, or unable)	0.275	0.447
Lower body limitation (some or lot of difficulty, or unable)	0.479	0.500
Education (reference = LT high school)		
High school graduate	0.376	0.484
Some college	0.145	0.352
College graduate	0.144	0.351
Family income (reference = LT U.S.\$10,000)		
U.S.\$10,000–20,000	0.290	0.454
U.S.\$20,000–30,000	0.210	0.408
U.S.\$30,000–40,000	0.123	0.328
U.S.\$40,000 or more	0.201	0.401
Age (reference = 65–67)		
68–70	0.168	0.374
71–74	0.200	0.400
75–79	0.217	0.412
80–84	0.125	0.331
85 or older	0.068	0.251
Female (reference = male)	0.596	0.491
Race/ethnicity (reference = white non-Hispanic)		
Asian	0.010	0.099
African American	0.073	0.260
White Hispanic	0.028	0.165
Other race	0.016	0.124
Year	4.560	2.824
Exogenous identifying variables		
EOLEI	11.544	1.925
EOLEI decile mean	11.538	1.895
EOLEI decile rank	5.366	2.882
EOLEI decile dummies (reference = decile 1)		
Decile 2	0.105	0.306
Decile 3	0.114	0.318
Decile 4	0.098	0.297
Decile 5	0.102	0.302
Decile 6	0.094	0.292
Decile 7	0.094	0.292
Decile 8	0.095	0.294
Decile 9	0.097	0.295
Decile 10	0.095	0.293

ADL, activities of daily living; EOLEI, End-of-Life Expenditure Index; HALex, Health and Activity Limitation Index; IADL, instrumental activities of daily living.

variables for African American, Asian, white Hispanic, and other race, relative to white non-Hispanic), education (dummy variables for 12, 13–15, and 16+ years of education, relative to <12 years of education), and family income (dummy variables for U.S.\$10,000–20,000, U.S.\$20,000–30,000, U.S.\$30,000–40,000, and U.S.\$40,000+, relative to <U.S.\$10,000).

All models also include a continuous measure of time<sup>4</sup> (Year) as a control for the average annual impact of unmeasured or unobservable changes in technology, government policies, personal health behaviors, and environmental factors (weather, air quality, infectious diseases) that vary over time and influence both spending and the health outcomes.

All linear IV models were estimated using the *STATA IVREG2* routine (Baum, Schaffer, and Stillman 2007). The IV survival models were estimated using the two-stage residual inclusion method (Terza, Basu, and Rathouz 2008), because the health outcome model is specified as a logistic function. As a special case of a conventional two-stage optimization estimator, the 2SRI estimates have the same desirable asymptotic properties of the class including unbiased standard errors (Terza, Basu, and Rathouz 2008).<sup>5</sup> MCBS cross-sectional survey sampling weights were used in the estimation, and the reported standard errors in both types of analysis are robust to heteroskedasticity. We also tested our results for sensitivity to sample design effects using MCBS's replicate weights. We found no reduction in *p*-values on our key coefficient estimates, but because our sample consisted of pooled observations across multiple years of the survey, it is not clear that the replicate weights were any more appropriate than the base weights for standard error calculations.

## RESULTS

Table 1 reports the mean values of the study variables. Average total medical spending over the 3-year observation period was U.S.\$27,092. Total Medicare spending was 40 percent lower, U.S.\$16,203. Total spending in the third year accounted for nearly 40 percent of all spending, essentially because the third year includes end-of-life spending for beneficiaries who died during that year. Medicare and total spending are highly correlated within the two observational time periods ( $r = 0.92$ ), but less highly correlated across time periods: the correlations between cumulative total spending over the 3 years and year 3 total or Medicare spending are 0.78 and 0.67.

At baseline, 15.9 percent of the sample reported fair health and 5.6 percent poor health, and about 30 percent reported an ADL or an IADL limitation. Almost half the sample had hypertension, and between 12–15 percent reported having had a heart attack, a stroke, or a nonskin cancer. Just over 95 percent of the sample survived to the end of the observation period and the average final HALex value was 0.61.

Table 2 reports the results of the various sensitivity tests of the alternative specifications of the EOLEI instrument in the IV estimation, as well as the OLS estimate of the medical spending coefficient in the health outcome model. Medical spending is measured as total medical spending from all sources over the 3-year observation period in these models. We limit these sensitivity tests to the HALex outcome because both the first and second-stage models are linear.

As expected, the OLS estimate is negative and highly significant. It implies that a 10 percent higher level of medical spending is associated with a 0.6 percent smaller final HALex value, that is, beneficiaries with higher spending have poorer health outcomes.

In contrast, the first IV estimate, based on the continuous EOLEI instrument, suggests that greater medical spending is associated with a significantly higher (better) final HALex value: a 10 percent higher level of medical spending is associated with a 1.93 percent ( $p = .045$ ) larger final

Table 2: Estimates of Medical Spending’s Relationship to Health Outcome (Final HALex Value), by Alternative Estimation Methods and Instrumental Variables

<i>Estimation Method and Instruments (N = 17,438)</i>	<i>Medical Spending Coefficient</i>		<i>Instrument Tests</i>	
			<i>First-Stage F-Test of Excluded Instrument</i>	<i>Partial R<sup>2</sup> of First-Stage Instrument</i>
OLS	– 0.0013	< .001		NA
IV (by instruments)				
a. Continuous EOLEI*	0.0043	.045	16.45	0.0009
b. EOLEI decile means	0.0042	.056	15.36	0.0009
c. EOLEI decile rank	0.0040	.053	16.86	0.0010
d. EOLEI decile dummies	0.0016	.231	2.93	0.0016

\*EOLEI (by hospital referral region; from the 1996 Dartmouth Atlas).  
EOLEI, End-of-Life Expenditure Index; HALex, Health and Activity Limitation Index; IV, instrumental variable; OLS, ordinary least squares.

HALex value. Moreover, the instrument test statistic indicates that the continuous EOLEI exceeds the rule-of-thumb  $F$ -value of 10 for significance in the first-stage spending equation. The partial  $R^2$  value in the first-stage is relatively small, but this reflects the substantial explanatory power of the lagged health variables in explaining medical spending. (If the prior health variables are excluded, then the instrument accounts for 10 percent of the explainable variation in medical spending.)

Specifying the instrument as the mean EOLEI value by decile of the EOLEI or as the decile rank of the beneficiary's HRR has essentially no effect on the magnitude or quality of the IV estimate of the medical spending coefficient. When the EOLEI is transformed into a set of dummy variables representing deciles of its values, the specification does not satisfy the weak instrument test ( $F=2.93$  for the test of excluded instruments in the first-stage model) and the medical spending coefficient falls in magnitude by more than 60 percent to 0.0016, suggesting that the weak instrument biases the coefficient toward the OLS value. Given the weakness of the dummy variable specification and the similarity of the results for the other options, we used the continuous EOLEI instrument in all subsequent model specifications.

Tables 3 and 4 report the complete sets of coefficient estimates for the OLS and IV models for the two health outcome measures.<sup>6</sup> In Table 3 the first-stage spending model shows that spending increases with the value of the EOLEI and with time (Year). The mean value of the EOLEI differs by U.S.\$6,593 between the 1st and 10th deciles of HRRs sorted by EOLEI (from U.S.\$8,644 in the 1st decile to U.S.\$15,237 in the 10th decile). With detailed health and demographic factors held constant, the coefficient of the EOLEI implies that a U.S.\$1,000 increase in the value of the EOLEI is associated with a U.S.\$545 increase in total medical spending. In effect, the instrument accounts for a difference in spending of about U.S.\$3,600 (13.3 percent of average total spending over the 3 years) between beneficiaries in the least and most costly deciles of the HRRs.

Other variables in the spending model indicate that total medical spending increases with education, family income, and age. Women spend less than men, and racial/ethnic minorities spend less than whites, although the differences are statistically significant only for Asians and for other races. The health variables show that spending increases as health worsens and that the implications for spending vary with both the type and timing of the disease or condition. For example, beneficiaries who reported at baseline that they had had a heart attack or cancer had U.S.\$3,500–U.S.\$4,300 higher spending than someone without either condition at baseline. However,

Table 3: Instrumental Variables (IV) and Ordinary Least Squares (OLS) Estimates of Medical Spending (Cumulative Total) and Health Outcomes (Final HALex) Models

Variable (N = 17,438)	IV				OLS	
	First-Stage Spending Model		Second-Stage Health Outcome Model		Health Outcome Model	
	Coefficient	t	Coefficient	z	Coefficient	t
Medical spending (U.S.\$1,000s)	0.545 <sup>a</sup>	4.06	0.0043 <sup>b</sup>	2.00	-0.0013 <sup>a</sup>	-18.00
EOLEI (exogenous identifying variable)						
Prior health status						
Self-reported general health status (reference = excellent)						
Very good	3.427 <sup>a</sup>	6.01	-0.090 <sup>a</sup>	-9.42	-0.071 <sup>a</sup>	-14.53
Good	6.523 <sup>a</sup>	9.58	-0.180 <sup>a</sup>	-11.33	-0.143 <sup>a</sup>	-27.29
Fair	15.312 <sup>a</sup>	13.64	-0.350 <sup>a</sup>	-10.11	-0.263 <sup>a</sup>	-37.64
Poor	24.801 <sup>a</sup>	11.31	-0.450 <sup>a</sup>	-8.04	-0.309 <sup>a</sup>	-32.54
Any ADL	10.721 <sup>a</sup>	11.96	-0.208 <sup>a</sup>	-8.70	-0.148 <sup>a</sup>	-27.04
Any IADL	8.022	9.76	-0.196 <sup>a</sup>	-10.40	-0.151 <sup>a</sup>	-28.17
Change in HALex, baseline to year 1	-37.462 <sup>a</sup>	-20.87	0.668 <sup>a</sup>	8.14	0.456 <sup>a</sup>	40.91
Change in HALex, year 1-year 2	-24.753 <sup>a</sup>	-15.02	0.462 <sup>a</sup>	8.39	0.322 <sup>a</sup>	29.86
Conditions						
Had hardening of the arteries at baseline	0.682	0.57	-0.007	-0.73	-0.002	-0.41
Developed hardening of the arteries, year 1	6.896 <sup>b</sup>	2.31	-0.025	-0.99	0.015	1.19
Developed hardening of the arteries, year 2	9.450 <sup>a</sup>	3.59	-0.059 <sup>b</sup>	-2.14	-0.005	-0.39
Had hypertension at baseline	1.000 <sup>c</sup>	1.72	-0.018 <sup>a</sup>	-3.20	-0.012 <sup>a</sup>	-3.13
Developed hypertension, year 1	2.512 <sup>c</sup>	1.71	-0.028 <sup>b</sup>	-2.14	-0.014	-1.55
Developed hypertension, year 2	4.759 <sup>a</sup>	2.99	-0.036 <sup>b</sup>	-2.20	-0.008	-0.92
Had heart attack at baseline	4.329 <sup>a</sup>	3.91	-0.027 <sup>b</sup>	-2.14	-0.003	-0.47

continued

Table 3. Continued

	IV			OLS		
	First-Stage Spending Model		Second-Stage Health Outcome Model	Health Outcome Model		
	Coefficient	t	Coefficient	z	t	
Developed heart attack, year 1	16.564 <sup>a</sup>	5.17	-0.094 <sup>b</sup>	-2.21	0.000	0.01
Developed heart attack, year 2	14.276 <sup>a</sup>	4.28	-0.056	-1.47	0.025 <sup>c</sup>	1.72
Had angina pectoris or other coronary heart disease at baseline	4.493 <sup>a</sup>	3.95	-0.023 <sup>c</sup>	-1.72	0.003	0.50
Developed angina or coronary heart disease, year 1	1.956	0.93	0.005	0.30	0.017	1.37
Developed angina or coronary heart disease, year 2	9.792 <sup>a</sup>	3.70	-0.039	-1.39	0.017	1.28
Had other heart conditions at baseline	4.071 <sup>a</sup>	5.44	-0.028 <sup>b</sup>	-2.55	-0.005	-1.11
Developed other heart condition, year 1	11.992 <sup>a</sup>	7.02	-0.070 <sup>b</sup>	-2.40	-0.002	-0.28
Developed other heart condition, year 2	5.319 <sup>a</sup>	3.57	-0.056 <sup>a</sup>	-3.32	-0.025 <sup>a</sup>	-3.01
Had stroke or brain hemorrhage at baseline	0.912	0.70	-0.021 <sup>b</sup>	-2.13	-0.016 <sup>b</sup>	-2.55
Developed stroke or brain hemorrhage, year 1	13.669 <sup>a</sup>	4.10	-0.081 <sup>b</sup>	-2.32	-0.005	-0.44
Developed stroke or brain hemorrhage, year 2	8.405 <sup>a</sup>	4.11	-0.071 <sup>a</sup>	-2.99	-0.024 <sup>b</sup>	-2.29
Had skin cancer at baseline	-0.098	-0.11	0.001	0.18	0.001	0.19
Developed skin cancer, year 1	-3.237 <sup>c</sup>	-1.94	0.015	0.95	-0.003	-0.28
Developed skin cancer, year 2	0.018	0.01	-0.014	-1.07	-0.014	-1.42
Had other cancer at baseline	3.537 <sup>a</sup>	4.29	-0.036 <sup>a</sup>	-3.51	-0.016 <sup>a</sup>	-3.26
Developed cancer, year 1	12.042 <sup>a</sup>	5.17	-0.073 <sup>b</sup>	-2.20	-0.005	-0.32
Developed cancer, year 2	20.494 <sup>a</sup>	8.69	-0.135 <sup>a</sup>	-2.78	-0.019	-1.27
Had diabetes at baseline	10.206 <sup>a</sup>	9.47	-0.087 <sup>a</sup>	-3.71	-0.029 <sup>a</sup>	-5.27
Developed diabetes, year 1	-0.473	-0.38	-0.024 <sup>b</sup>	-2.14	-0.027 <sup>a</sup>	-2.95
Developed diabetes, year 2	-1.437	-1.54	-0.030 <sup>a</sup>	-3.11	-0.038 <sup>a</sup>	-5.13

continued

Table 3. Continued

Variable (N = 17,438)	IV				OLS	
	First-Stage Spending Model		Second-Stage Health Outcome Model		Health Outcome Model	
	Coefficient	t	Coefficient	z	Coefficient	t
Additional health variables						
Proxy respondent	5.807 <sup>a</sup>	3.45	-0.050 <sup>a</sup>	-2.95	-0.016 <sup>b</sup>	-2.17
Current smoker	1.849 <sup>b</sup>	2.00	-0.040 <sup>a</sup>	-4.40	-0.029 <sup>a</sup>	-4.84
Former smoker	2.063 <sup>a</sup>	3.55	-0.016 <sup>b</sup>	-2.33	-0.004	-1.05
LT 10th percentile of body mass index (reference = 25th-75th percentile)	1.151	0.99	-0.025 <sup>a</sup>	-2.65	-0.018 <sup>a</sup>	-2.86
10th-25th percentile of body mass index	-0.726	-1.02	-0.009	-1.32	-0.013 <sup>b</sup>	-2.52
GT 75th percentile of body mass index	-1.422 <sup>b</sup>	-2.00	0.003	0.48	-0.005	-1.07
Upper body limitation (some or lot of difficulty, or unable)	0.996	1.19	-0.025 <sup>a</sup>	-3.49	-0.019 <sup>a</sup>	-3.81
Lower body limitation (some or lot of difficulty, or unable)	1.780 <sup>b</sup>	2.45	-0.039 <sup>a</sup>	-5.23	-0.029 <sup>a</sup>	-6.16
Education (reference = LT high school)						
High school graduate	2.415 <sup>a</sup>	3.35	-0.006	-0.79	0.007 <sup>c</sup>	1.68
Some college	2.890 <sup>a</sup>	3.27	-0.007	-0.68	0.009 <sup>c</sup>	1.62
College graduate	3.620 <sup>a</sup>	3.97	-0.012	-1.06	0.009	1.44
Family income (reference = LT U.S.\$10,000)						
U.S.\$10,000-20,000	2.112 <sup>b</sup>	2.31	-0.007	-0.76	0.005	1.04
U.S.\$20,000-30,000	1.966 <sup>b</sup>	2.07	-0.004	-0.47	0.007	1.20
U.S.\$30,000-40,000	2.767 <sup>b</sup>	2.41	0.007	0.63	0.023 <sup>a</sup>	3.33
U.S.\$40,000 or more	3.607 <sup>a</sup>	3.57	0.006	0.52	0.027 <sup>a</sup>	4.23
Age (reference = 65-67)						
68-70	0.926	0.98	-0.024 <sup>a</sup>	-3.08	-0.019 <sup>a</sup>	-3.38
71-74	1.015	1.13	-0.040 <sup>a</sup>	-5.24	-0.034 <sup>a</sup>	-6.08

continued

Table 3. Continued

Variable (N = 17,438)	IV			OLS		
	First-Stage Spending Model		Second-Stage Health Outcome Model	Health Outcome Model		
	Coefficient	t	Coefficient	z	Coefficient	t
75-79	1.256	1.39	-0.063 <sup>a</sup>	-8.07	-0.056 <sup>a</sup>	-10.01
80-84	2.636 <sup>a</sup>	2.60	-0.107 <sup>a</sup>	-10.80	-0.091 <sup>a</sup>	-14.90
85 or older	4.736 <sup>a</sup>	3.48	-0.164 <sup>a</sup>	-11.32	-0.137 <sup>a</sup>	-18.26
Female (reference = male)	-2.245 <sup>a</sup>	-3.36	0.016 <sup>b</sup>	2.21	0.004	0.95
Race/ethnicity (reference = white non-Hispanic)						
Asian	-7.487 <sup>a</sup>	-3.54	0.045	1.62	0.005	0.26
African American	-1.466	-1.23	0.000	0.04	-0.005	-0.79
White Hispanic	-1.847	-1.07	0.000	0.03	-0.006	-0.52
Other race	-4.818 <sup>b</sup>	-2.21	0.015	0.71	-0.010	-0.74
Year	0.311 <sup>a</sup>	3.21	0.001	0.53	0.002 <sup>a</sup>	3.66
Constant	-6.487 <sup>a</sup>	-3.23	0.929 <sup>a</sup>	81.88	0.927 <sup>a</sup>	110.65

Note. IV test statistics.

Partial R<sup>2</sup> of excluded instruments: 0.0009.

Kleibergen-Paap rk Wald F Statistic: 16.448.

Stock-Yogo weak ID test critical values.

10% maximal IV relative bias: 16.380.

<sup>a</sup>p < .01;

<sup>b</sup>.01 < p ≤ .05;

<sup>c</sup>.05 < p ≤ .10.

ADLs, activities of daily living; EOLEI, End-of-Life Expenditure Index; HALEx, Health and Activity Limitation Index; IADL, instrumental activities of daily living.



Table 4: IV (2SRI) and Observational Estimates of Medical Spending (Cumulative Total) and Health Outcome (Survival) Logistic Models

Variable (N = 17,438)	IV (2SRI)*		Observational	
	Coefficient	z	Coefficient	z
Medical spending (U.S.\$1,000s)	1.081 <sup>b</sup>	2.06	0.987 <sup>a</sup>	-14.07
Residual from first-stage model	0.913 <sup>b</sup>	-2.40		
Prior health status				
Self-reported general health status (reference = excellent)				
Very good	0.619 <sup>b</sup>	-2.46	0.841	-1.22
Good	0.352 <sup>a</sup>	-3.61	0.637 <sup>a</sup>	-3.27
Fair	0.098 <sup>a</sup>	-3.81	0.398 <sup>a</sup>	-5.75
Poor	0.038 <sup>a</sup>	-3.39	0.361 <sup>a</sup>	-5.10
Any ADL	0.248 <sup>a</sup>	-3.31	0.655 <sup>a</sup>	-3.76
Any IADL	0.330 <sup>a</sup>	-3.43	0.683 <sup>a</sup>	-3.27
Change in HALex, baseline to year 1	197.234 <sup>a</sup>	3.66	6.526 <sup>a</sup>	7.68
Change in HALex, year 1-year 2	54.745 <sup>a</sup>	4.12	5.800 <sup>a</sup>	7.67
Conditions				
Had hardening of the arteries at baseline	1.086	0.64	1.167	1.24
Developed hardening of the arteries, year 1	0.444 <sup>b</sup>	-2.45	0.828	-0.85
Developed hardening of the arteries, year 2	0.634	-1.05	1.506 <sup>c</sup>	1.73
Had hypertension at baseline	0.935	-0.67	1.023	0.26
Developed hypertension, year 1	0.747	-1.39	0.946	-0.30
Developed hypertension, year 2	0.942	-0.21	1.451 <sup>c</sup>	1.75
Had heart attack at baseline	0.653 <sup>b</sup>	-2.20	0.961	-0.35
Developed heart attack, year 1	0.192 <sup>b</sup>	-2.49	0.868	-0.55
Developed heart attack, year 2	0.200 <sup>a</sup>	-2.76	0.726	-1.40
Had angina pectoris or other coronary heart disease at baseline	0.733	-1.44	1.106	0.81
Developed angina or coronary heart disease, year 1	0.608 <sup>b</sup>	-2.16	0.736	-1.41

continued

Table 4. *Continued*

Variable (N = 17,438)	IV (2SRJ)*		Observational	
	Coefficient	z	Coefficient	z
Developed angina or coronary heart disease, year 2	0.473 <sup>c</sup>	-1.71	1.153	0.60
Had other heart conditions at baseline	0.545 <sup>a</sup>	-3.33	0.787 <sup>b</sup>	-2.58
Developed other heart condition, year 1	0.364 <sup>b</sup>	-2.08	1.072	0.42
Developed other heart condition, year 2	0.609 <sup>c</sup>	-1.88	0.987	-0.08
Had stroke or brain hemorrhage at baseline	0.722 <sup>a</sup>	-2.85	0.780 <sup>b</sup>	-2.29
Developed stroke or brain hemorrhage, year 1	0.326 <sup>b</sup>	-2.00	1.108	0.45
Developed stroke or brain hemorrhage, year 2	0.388 <sup>b</sup>	-2.53	0.826	-0.98
Had skin cancer at baseline	1.227 <sup>c</sup>	1.82	1.219 <sup>c</sup>	1.76
Developed skin cancer, year 1	1.296	1.09	0.973	-0.13
Developed skin cancer, year 2	0.703 <sup>c</sup>	-1.87	0.698 <sup>c</sup>	-1.90
Had other cancer at baseline	0.588 <sup>a</sup>	-3.09	0.810 <sup>b</sup>	-2.10
Developed cancer, year 1	0.182 <sup>a</sup>	-3.27	0.546 <sup>b</sup>	-2.50
Developed cancer, year 2	0.076 <sup>a</sup>	-3.13	0.493 <sup>a</sup>	-2.99
Had diabetes at baseline	0.282 <sup>a</sup>	-3.17	0.714 <sup>a</sup>	-3.20
Developed diabetes, year 1	1.051	0.25	1.011	0.06
Developed diabetes, year 2	1.086	0.47	0.954	-0.28
Additional health variables				
Proxy respondent	0.364 <sup>a</sup>	-4.00	0.624 <sup>a</sup>	-3.97
Current smoker	0.512 <sup>a</sup>	-4.76	0.609 <sup>a</sup>	-3.99
Former smoker	0.709 <sup>a</sup>	-2.85	0.860	-1.62
LT 10th percentile of body mass index (reference = 25th-75th percentile)	0.528 <sup>a</sup>	-4.98	0.587 <sup>a</sup>	-4.49
10th-25th percentile of body mass index	0.768 <sup>b</sup>	-2.32	0.721 <sup>a</sup>	-2.99
GT 75th percentile of body mass index	1.477 <sup>a</sup>	3.08	1.295 <sup>b</sup>	2.32
Upper body limitation (some or lot of difficulty, or unable)	0.882	-1.19	0.968	-0.33
Lower body limitation (some or lot of difficulty, or unable)	0.886	-0.95	1.031	0.28

*continued*

Table 4. Continued

Variable (N= 17,438)	IV (2SRJ)*		Observational	
	Coefficient	z	Coefficient	z
Education (reference = LT high school)				
High school graduate	0.806	-1.62	1.001	0.01
Some college	0.705 <sup>b</sup>	-2.03	0.913	-0.68
College graduate	0.692 <sup>c</sup>	-1.83	0.963	-0.26
Family income (reference = LT U.S.\$10,000)				
U.S.\$10,000-20,000	0.847	-1.22	1.027	0.24
U.S.\$20,000-30,000	0.740 <sup>b</sup>	-2.09	0.888	-0.95
U.S.\$30,000-40,000	0.944	-0.28	1.219	1.14
U.S.\$40,000 or more	0.823	-0.95	1.158	0.95
Age (reference = 65-67)				
68-70	0.472 <sup>a</sup>	-3.49	0.514 <sup>a</sup>	-3.07
71-74	0.384 <sup>a</sup>	-4.66	0.426 <sup>a</sup>	-4.09
75-79	0.291 <sup>a</sup>	-6.38	0.330 <sup>a</sup>	-5.63
80-84	0.173 <sup>a</sup>	-8.60	0.222 <sup>a</sup>	-7.71
85 or older	0.076 <sup>a</sup>	-10.53	0.118 <sup>a</sup>	-10.77
Female (reference = male)	2.304 <sup>a</sup>	6.61	1.901 <sup>a</sup>	6.91
Race/ethnicity (reference = white non-Hispanic)				
Asian	3.170 <sup>b</sup>	2.15	1.643	1.09
African American	1.105	0.66	1.004	0.03
Hispanic	1.474	1.48	1.331	1.14
Other race	1.293	0.79	0.863	-0.51
Year	0.951 <sup>a</sup>	-2.79	0.977	-1.63

\*Second stage only. See Table 3 for first-stage results.

<sup>a</sup>.  $p < .01$ ;

<sup>b</sup>.  $0.01 < p \leq .05$ ,

<sup>c</sup>.  $0.05 < p \leq .10$ .

ADL, activities of daily living;HALEx, Health and Activity Limitation Index; IADL, instrumental activities of daily living; IV, instrumental variable.

spending was much higher for someone who reported the onset of one of these conditions in the second survey year (U.S.\$20,494 for cancer and U.S.\$14,276 for heart attack). The two variables that measure the changes in the HALex values between baseline and year 1 and between years 1 and 2 indicate that improved general health over time is associated with reduced spending.

Comparing the OLS and IV versions of the health outcome model suggests that the generally positive associations between health and education and income in the OLS model operate primarily through the influence on medical spending, because these variables are not statistically significant in the IV model. Health declines significantly and monotonically with age in both models. The OLS model finds no differences by gender and race/ethnicity, while the IV model indicates that women and Asians have better health outcomes. However, differences in health outcomes between whites, African Americans, and Hispanics are not statistically significant, given that the model controls for medical spending, education, income, and detailed health characteristics.

The OLS model also generally finds smaller effects associated with prior health conditions. For example, in the OLS model prior heart attacks or congestive heart disease have positive, though statistically insignificant associations with the final HALex value, suggesting that these people have similar or slightly better health than people without a prior heart condition. In the IV model, five of the six coefficients for these conditions are negative and statistically significant. Similarly, the OLS model results indicate much smaller health effects associated with stroke or cancer compared with the IV model. These apparently anomalous OLS results reinforce the validity of the IV approach.

Comparing selected coefficients from the IV health outcome model (Table 3) can help interpret the magnitude of a particular change in the HALex value associated with an increase in medical spending. Because this is a linear model, each coefficient represents the difference in the HALex value for a person who has a particular condition compared with someone who does not, but has otherwise identical characteristics. The linear coefficient of the medical spending variable is 0.0043, which implies that a 10 percent increase in total medical spending over the 3-year observation period (U.S.\$2,709) is associated with a change of 0.012 in the final HALex value. (Medical spending is measured in thousands of dollars.) This estimated effect is comparable to the positive effect of having an income level  $>$  U.S.\$30,000 relative to an income of  $<$  U.S.\$10,000 (0.016), or the negative effect ( $-$  0.016) of aging from 68–70 years old to 71–74 years old with no other changes in health conditions, or of

having reported hypertension at baseline ( $-0.015$ ). These comparisons suggest that the estimated magnitude of the relationship between medical spending and the HALex is nontrivial and clinically meaningful.

Table 4 reports the OLS and IV estimates of odds ratios from the logistic model for surviving to the end of year 3. (The first-stage spending model is the same as reported in Table 3.) As in the HALex analysis, the observational model suggests that greater medical spending has a highly significant negative effect on the relative odds of survival. In contrast, the IV estimate implies that greater medical spending has a statistically significant and positive effect on the relative odds of survival. Moreover, the residual from the first-stage medical spending model is associated with a lower relative odds of survival, suggesting that unobserved factors that influence medical spending, such as unmeasured disease severity or other health limitations not captured by the observable health factors, have a negative association with survival. When these indirect effects are accounted for, the direct effects of observed prior health conditions on reducing the relative odds of survival are generally larger than indicated by the OLS estimates. This result is similar to the relationships between prior health and the HALex health outcome measure reported in Table 3.

Table 5 reports the results of varying the specification of the measure of medical spending. Rows 1 and 3 compare the coefficients of cumulative total medical spending and Medicare-only spending over the 3-year observation period, and rows 2 and 4 show the results based on measures of medical spending observed over only a single year (year 3), as is typical of most cross-sectional studies of the relationship between medical spending and health. The table also reports the coefficient estimates of the EOLEI identifying variable from the first-stage spending equations and the corresponding  $F$ -statistic for the strength of the instrument.

The estimates associated with Medicare spending, whether measured as cumulative spending over 3 years or spending during year 3 only, are very similar for both health outcomes. A 10 percent higher level of Medicare spending is associated with a health or probability of survival improvement of 1.2–1.4 percent. The percentage changes in the outcomes associated with a 10 percent increase in medical spending from all sources are somewhat larger than those derived from the Medicare spending estimates. However, this primarily reflects the differences in mean values. The coefficients from the linear HALex models indicate that a dollar of additional medical spending has the same effect on health, regardless of whether it is a Medicare or non-Medicare dollar. Moreover, the coefficients of the cumulative spending variables, which

Table 5: IV Estimates of Relationship between Medical Spending and Health Outcomes (Final HALex Value and Survival at End of Year 3), by Alternative Measures of Medical Spending

	Health Outcome Measure				
	HALex	Medical Spending Coefficient (p-Value)	Pct. Change in HALex (10% Change in Spending)	Medical Spending Odds Ratio (p-Value)	Survival Odds
<i>EOLEI First-Stage Coefficient (F-Statistic)*</i>					
<i>Alternative Medical Spending Measures (N = 17,438)</i>					
All sources, cumulative spending over 3 years	0.545 (16.45)	0.0043 (0.045)	1.9%	1.081 (0.039)	1.5%
All sources, year 3 spending only	0.184 (7.21)	0.013 (0.096)	2.2	1.266 (0.034)	1.7
Cumulative Medicare spending over 3 years	0.536 (26.61)	0.0044 (0.044)	1.2	1.092 (0.022)	1.2
Medicare spending, year 3 only	0.184 (10.86)	0.013 (0.066)	1.4	1.29 (0.023)	1.3

\*F-statistic for test of instrument strength; first-stage equations for spending measures are identical for HALex and survival second-stage models. EOLEI, End-of-Life Expenditure Index; HALex, Health and Activity Limitation Index.

measure spending over the prior 3 years, are essentially one-third the magnitude of the coefficients of the annual spending measures for year 3. In effect, an increase of U.S.\$3 in cumulative spending over the 3 years has the same impact as an increase of U.S.\$1 in spending in year 3. Overall, the results reported in Table 5 indicate that the general conclusion that medical spending has a positive association with better health is quantitatively and statistically robust with respect to the measures of both medical spending and health outcome.

## DISCUSSION

IV estimation results in a positive and statistically significant relationship between greater medical spending and better health. Depending on the specific measure of medical spending, 10 percent greater medical spending is associated with a 1.2–2.2 percent larger final HALex value, which is a weighted combination of self-reported health status and measures of ADL and IADL limitations, and a 1.2–1.7 percent higher probability of surviving to the end of the year. These estimates represent averages over the entire range of medical spending that may differ at various levels of spending and/or health if the true underlying relationship is nonlinear or if the relationship between spending and health varies across alternative populations. Thus, the marginal effect of increased medical spending might be quite different at higher and lower levels of spending, or for people in initially very good or very poor health.

One question that inevitably arises in considering these findings is why they differ from those based on analyses conducted at the geographic level using data generated by the Dartmouth Atlas of Health Care. Fisher et al. (2003b) reported that Medicare beneficiaries' mortality rates did not vary significantly across areas grouped into quintiles based on their EOLEI values. Dartmouth research has also investigated the HALex index as a health measure, analyzing the change in the HALex value for a large sample of beneficiaries surveyed by the MCBS between 1992 and 1996 (Fisher et al. 2003b) and found no significant difference in the change in the HALex across Dartmouth Atlas spending quintiles.

Our analyses differ from the Dartmouth analyses in several important dimensions. First, we use IV analysis to estimate the relationship between beneficiaries' own medical spending and their health outcomes, rather than average medical spending across groups (quintiles) of geographic areas in

which they live. Our sensitivity analyses suggest that dichotomous measures of the EOLEI may not be a strong instrument, which biases estimated effects toward the negative OLS results. (We also estimated an alternative IV specification of the survival model that used EOLEI decile dummy variables as instruments rather than the continuous EOLEI measure. Similar to the result reported in Table 2 [row d], the estimated coefficient for the medical spending variable was smaller in magnitude, 1.039, and not statistically significant [ $p = .19$ ].) A finding of no relationship between spending and health across geographic areas grouped by the EOLEI may reflect this weak-instrument bias.

Second, we use a more extensive set of measures of prior health status, activity limitations, and health conditions, as well as education and income, to control for their effects on both medical spending and health outcomes. Thus, some of the differences in findings may be due to differences in the extent of risk adjustment in the analyses. Controlling for differences in individuals' health and socioeconomic characteristics more directly may reduce possible bias from population heterogeneity within geographic areas.

Third, our data span an 11-year period from 1992 to 2002. It is possible that the Dartmouth results, which refer primarily to relationships at a point in time or over a relatively short-time period, are accurate for the time periods covered, but do not capture the nature of the relationship between medical spending and health over a longer time period. For example, changes in medical technology may improve health outcomes over time through their effects on spending, even if the spending-health relationship is "flat" at a point in time.

To the extent that our results indicate that on average, across the entire range of medical spending and health conditions, the relationship between medical spending and health is positive, then across-the-board reductions in or limits on Medicare spending may result in poorer health for some Medicare beneficiaries. Moreover, because the variation in medical spending in our analyses was driven by variations in practice patterns across geographic areas, as represented by the EOLEI, then targeting Medicare spending reductions on geographic areas with high average Medicare costs per beneficiary would also appear to be unwarranted.

Conversely, as noted above, even if our primary result is accurate on average, it does not imply that all medical spending for all medical conditions and situations necessarily has the same positive effects on health. For example, Kaestner and Silber (2010) found significant variations in the magnitudes of the estimated positive effect of spending on health across several medical conditions. These results suggest that the search for inefficiencies in Medicare



spending needs to be conducted on a more disaggregated basis, focusing on the effectiveness of specific treatments and care patterns for particular conditions and combinations of conditions. If research finds that medical spending inefficiencies are more common in care provided to beneficiaries who are high cost because of being in relatively poor health, such as those with multiple chronic conditions or terminal illnesses near the end of life, then policy should focus on improving the efficiency of care provided to those people, rather than on targeting geographic areas with high Medicare costs. Along the same lines, recent analysis by the Medicare Payment Advisory Commission (Glass 2010) has highlighted possible fraudulent spending for durable medical equipment, home health, and hospice care as a contributor to unwarranted geographic variations in total Medicare spending.

Overall, our analysis suggests that the search for inefficiency in Medicare spending should progress on a more disaggregated basis and that the use of IV methods applied to data on individual Medicare beneficiaries may be a fruitful way to pursue this objective. In other words, future research should focus on beneficiaries with particular conditions, combinations of conditions, or types of service use. Comparative effectiveness analyses of alternative treatments for specific conditions represent a logical extension of this suggestion (Pizer 2009). Moreover, it would be desirable that this research be comprehensive in terms of disease and health status groupings. Generalizing from analyses based on a small number of conditions could miss significant intergroup differences in medical care efficiency (Kaestner and Silber 2010; Rothberg et al. 2010). Lastly, research might focus on comparing patterns of care for beneficiaries with particularly costly health conditions across high- and low-spending areas (Lieberman et al. 2003). Because these beneficiaries account for a disproportionate share of Medicare spending, detailed analyses of both demand-side and supply-side financial and organizational incentives affecting their care are important for developing appropriate policy responses.

## ACKNOWLEDGMENTS

*Joint Acknowledgment/Disclosure Statement:* This study was supported by a grant from the Robert Wood Johnson Foundation's program on Health Care Financing and Organization to George Mason University. Joel Ruhter and Matthew Cravens provided excellent research and programming assistance while employed as research assistants at The Urban Institute. We also thank

Thomas Stratmann, Bryan Dowd, Jose Escarce, and two anonymous referees for their very helpful comments on earlier drafts of this manuscript.

*Disclosures:* None.

*Disclaimers:* None.

## NOTES

1. The data collection and management methods were reviewed and approved the George Mason University Human Subjects Review Board—protocol no. 5409, August 14, 2007.
2. A second advantage of instrumental variable estimation is that it also reduces bias due to measurement error, which might occur because of inaccurate self-reporting of non-Medicare health spending and/or incomplete adjustment for intertemporal and cross-sectional price variations.
3. For examples of studies that use a rank variable as an exogenous instrument, see Stratmann (2009) and Koenker and Bassett (1978).
4. As a sensitivity test, we also estimated all models using a set of dummy variables to represent the first year a person entered the MCBS survey. Coefficient estimates and standard errors were virtually identical to those reported below.
5. We also used STATA's bootstrap procedure to calculate standard errors in order to correct for the two-step procedure used by the 2SRI method. Based on 1,000 iterations, the estimated 95% confidence interval for the second-stage parameter estimate of 1.082 was (1.0098, 1.2446). Although slightly larger than the confidence interval generated by the 2SRI procedure of (1.0045, 1.1646), we are still able to reject the hypothesis of no effect on survival.
6. Complete results for all other model specifications are available on request.

## REFERENCES

- Abelson, R., and G. Harris. 2010. "Critics Question Study Cited in Health Debate." *The New York Times*, June 3, p. A1.
- Adler, G. 1998. "Concept and Development of the Medicare Current Beneficiary Survey." *Proceedings of the Survey Research Methods Section*, American Statistical Association [accessed on August 19, 2009]. Available at [http://www.amstat.org/sections/srms/Proceedings/papers/1998\\_021.pdf](http://www.amstat.org/sections/srms/Proceedings/papers/1998_021.pdf).
- Bach, P. 2010. "A Map to Bad Policy—Hospital Efficiency Measures in the Dartmouth Atlas." *New England Journal of Medicine* 362 (7): 569–74.
- Bach, P., D. Schrag, and C. Begg. 2004. "Resurrecting Treatment Histories of Dead Patients: A Study Design That Should Be Laid to Rest." *Journal of the American Medical Association* 292 (22): 2765–70.
- Baicker, K., and A. Chandra. 2009. "A Trillion Dollar Geography Lesson." *Health Affairs* 28 (5): 1448–51.

- Baum, C. F., M. E. Schaffer, and S. Stillman. 2007. "ivreg2: Stata Module for Extended Instrumental Variables/2SLS, GMM and AC/HAC, LIML and k-class Regression" [accessed on May 3, 2011]. Available at <http://ideas.repec.org/c/boc/bocode/s425401.html>
- Card, D., C. Dobkin, and N. Maestas. 2009. "Does Medicare Save Lives?" *Quarterly Journal of Economics* 124 (2): 597–636.
- Center for Clinical Evaluative Studies. 1999. "The Quality of Medical Care in the United States: A Report on the Medicare Program" [accessed on May 3, 2011]. Available at <http://www.dartmouthatlas.org/downloads/atlas/99Atlas.pdf>
- Chandra, A., E. Fisher, and J. Skinner. 2007. "Pitfalls in the Analysis of Regional Variation in Health Care: A Response to Hadley, Berenson, Waidmann, and Zuckerman." Unpublished manuscript. Hanover NH: The Dartmouth Institute for Health Policy and Clinical Practice, September 21 [accessed on August 19, 2009]. Available at <http://intensity.dartmouth.edu/?q=node/92>
- Congressional Budget Office (CBO). 2008. *Budget Options: Volume I Health Care, Chapter 5*. Washington, DC: CBO.
- Doyle, J. 2008. *Returns to Local-Area Emergency Health Care Spending: Using Health Shocks to Patients Far from Home*. NBER Working Paper #13301. Cambridge, MA.
- Erickson, P. 1998. "Evaluation of a Population-Based Measure of Quality of Life: The Health and Activity Limitation Index (HALex)." *Quality of Life Research* 7 (2): 101–14.
- Erickson, P., R. Wilson, and I. Shannon. 1995. "Years of Healthy Life." *Statistical Note No. 7*. Hyattsville, MD: National Center for Health Statistics.
- Fisher, E., D. Goodman, J. Skinner, and K. Bonner. 2009. *Health Care Spending, Quality, and Outcomes*. A Dartmouth Atlas Project Topic Brief. Hanover, NH: The Dartmouth Institute for Health Policy and Clinical Practice, February 27.
- Fisher, E., D. Wennberg, T. Stukel, D. Gottlieb, F. Lucas, and E. Pindar. 2003a. "The Implications of Regional Variations in Medicare Spending, Part 1: The Content, Quality, and Accessibility of Care." *Annals of Internal Medicine* 138 (4): 273–87.
- . 2003b. "The Implications of Regional Variations in Medicare Spending, Part 2: Health Outcomes and Satisfaction with Care." *Annals of Internal Medicine* 138 (4): 288–98.
- Glass, D. 2010. "Accountability for DME, Home Health, and Hospice Use. Presentation to the Medicare Payment Advisory Commission" [accessed May 3, 2011]. Available at <http://www.medpac.gov/transcripts/Accountability%20Sept%202010%20public.pdf>
- Hadley, J. 2003. "Sicker and Poorer: The Consequences of Being Uninsured." *Medical Care Research and Review* 60 (2, suppl): 3S–75S.
- . 1982. *More Medical Care, Better Health? An Economic Analysis of Mortality Rates*. Washington, DC: The Urban Institute Press.
- . 1988. "Medicare Spending and Mortality Rates of the Elderly." *Inquiry* 25: 486–93.
- Kaestner, R., and J. H. Silber. 2010. "Evidence on the Efficacy of Inpatient Spending on Medicare Patients." *Milbank Quarterly* 88 (4): 560–94.
- Koenker, R., and G. Bassett. 1978. "Regression Quantiles." *Econometrica* 46 (1): 107–12.

- Lichtenberg, F. 2002. "The Effects of Medicare on Health Care Utilization and Outcomes." In *Frontiers in Health Policy Research*, edited by A. Garber, Vol. 5, pp. 27–52. Cambridge, MA: MIT Press.
- Lieberman, S. M., J. Lee, T. Anderson, and D. Crippen. 2003. "Reducing the Growth of Medicare Spending: Geographic Versus Patient-Based Strategies." *Health Affairs* 22: w3-603–13.
- Martin, S., N. Rice, and P. Smith. 2008. "Does Health Care Spending Improve Health Outcomes? Evidence from English Programme Budgeting Data." *Journal of Health Economics* 27 (4): 826–42.
- McClellan, M., and J. Newhouse. 2000. "Overview of Special Supplemental Issue on Instrumental Variable Analysis Applications in Health Services Research." *Health Services Research* 35: 1061–9.
- McWilliams, J., E. Meara, A. Zaslafsky, and J. Ayanian. 2007a. "Health of Previously Uninsured Adults after Acquiring Medicare Coverage." *Journal of American Medical Association* 298 (24): 2886–94.
- . 2007b. "Use of Health Services by Previously Uninsured Medicare Beneficiaries." *New England Journal of Medicine* 357 (2): 143–53.
- . 2003. "Impact of Medicare Coverage on Basic Clinical Services for Previously Uninsured Adults." *Journal of American Medical Association* 290 (6): 757–64.
- Neuberg, G. 2009. "The Cost of End-of-Life Care: A New Efficiency Measure Falls Short of AHA/ACC Standards." *Circulation: Cardiovascular Quality and Outcomes* 2 (2): 127–33.
- Newhouse, J., and M. McClellan. 1998. "Econometrics in Outcomes Research: The Use of Instrumental Variables." *Annual Review of Public Health* 19: 17–34.
- Ong, M., C. Mangione, P. Romano, Q. Zhou, D. Auerbach, A. Chun, B. Davidson, G. Ganiats, S. Greenfield, M. Gropper, S. Malik, J. Rosenthal, and J. Escarce. 2009. "Looking Forward, Looking Back: Assessing Variations in Hospital Resource Use and Outcomes for Elderly Patients with Heart Failure." *Circulation: Cardiovascular Quality and Outcomes* 2 (6): 548–557.
- Pizer, S. D. 2009. "An Intuitive Review of Methods for Observational Studies of Comparative Effectiveness." *Health Services Outcomes Research Methods* 9: 54–68.
- Rothberg, M. B., J. Cohen, P. Lindenauer, J. Maselli, and A. Auerbach. 2010. "Little Evidence of Correlation between Growth in Health Care Spending and Reduced Mortality." *Health Affairs* 29 (8): 1523–31.
- Sirovich, B., P. Gallagher, D. Wennberg, and E. Fisher. 2008. "Discretionary Decision Making by Primary Care Physicians and the Cost of U.S. Health Care." *Health Affairs* 27 (3): 813–23.
- Skinner, J., E. Fisher, and J. Wennberg. 2005. "The Efficiency of Medicare." In *Analyses in the Economics of Aging*, edited by D. Wise, pp. 129–60. Chicago, IL: University of Chicago Press.
- Staiger, D., and J. Stock. 1997. "Instrumental Variable Regression with Weak Instruments." *Econometrica* 65: 557–86.
- Stratmann, T. 2009. "How Prices Matter in Politics: The Returns to Campaign Advertising." *Public Choice* 140: 357–77.

- Terza, J., A. Basu, and P. Rathouz. 2008. "Two-Stage Residual Inclusion Estimation: Addressing Endogeneity in Health Econometric Modeling." *Journal of Health Economics* 27: 531–43.
- U. S. Senate Committee on Finance. 2009. "Financing Comprehensive Health Care Reform: Proposed Health System Savings and Revenue Options," May 20 [accessed on August 19, 2009]. Available at <http://finance.senate.gov/sitepages/leg/LEG%202009/051809%20Health%20Care%20Description%20of%20Policy%20Options.pdf>
- Wennberg, J., and M. Cooper (eds). 1998. *The Dartmouth Atlas of Health Care*. Chicago, IL: American Hospital Publishing.
- Wennberg, J., E. Fisher, D. Goodman, and J. Skinner. 2008. *Executive Summary—Tracking the Care of Patients with Severe Chronic Illness: The Dartmouth Atlas of Health Care 2008*. Hanover NH: The Dartmouth Institute for Health Policy and Clinical Practice, April.

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Appendix SA1: Author Matrix.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.