TOMAS J. PHILIPSON
University of Chicago

DANA P. GOLDMAN
University of Southern California

SETH A. SEABURY
RAND Corporation

DARIUS N. LAKDAWALLA
University of Southern California

LEE M. LOCKWOOD
University of Chicago

Geographic Variation in Health Care: The Role of Private Markets

ABSTRACT The Dartmouth Atlas of Health Care has documented substantial regional variation in health care utilization and spending, beyond what would be expected from such observable factors as demographics and disease severity. However, since these data are specific to Medicare, it is unclear to what extent this finding generalizes to the private sector. Economic theory suggests that private insurers have stronger incentives to restrain utilization and costs, while public insurers have greater monopsony power to restrain prices. We argue that these two differences alone should lead to greater regional variation in utilization for the public sector, but either more or less variation in spending. We provide evidence that variation in utilization in the public sector is about 2.8 times as great for outpatient visits ($p < 0.01$) and 3.9 times as great for hospital days ($p = 0.09$) as in the private sector. Variation in spending appears to be greater in the private sector, consistent with the importance of public sector price restraints.

There is considerable variation in health care utilization and spending across geographic areas in the United States, but little evidence of corresponding differences in health outcomes or satisfaction with care.[1] This variability is often cited as evidence that current levels of health care spending reflect "flat-of-the-curve" medicine, that is, treatment for which the marginal benefit of an additional unit of care is approximately zero.

1. The main data source used to document regional variations is the Dartmouth Atlas of Health Care, which can be found at www.dartmouthatlas.org/ (accessed January 15, 2010).

Interpreted this way, these findings have dramatic implications for the potential to increase the productivity of health care spending, and for this reason they have figured prominently in the policy debate.

However, the evidence on regional variation is almost exclusively limited to the public sector, because it relies on Medicare data. Less is known about the corresponding patterns in the private sector. A venerable literature in economics has argued that private firms and their managers have stronger incentives to restrain costs and boost efficiency than their public counterparts.[2] In the health insurance context, Medicare does not face competition over premiums that might otherwise restrain its costs, and unlike private sector firms, Medicare does not have direct residual claimants whose standard of living improves with the efficiency of the enterprise.

To develop the implications of these incentive differences, this paper provides a theoretical and empirical analysis of how regional variation in health care differs across the public and the private sectors. We first examine conceptually how private efforts to control costs *within* a region, through selection of providers, might translate into differences in care *across* regions. In particular, our analysis implies that utilization controls within regions in the private sector should lead to lower regional variation in the private sector than in Medicare. However, the implications for variation in spending are less clear, because Medicare may also be able to better control prices through its greater monopsony power. If the private sector controls utilization while the public sector controls prices, the result is an ambiguous prediction for variation in spending.

We examine these implications empirically using individual-level data on patients with heart disease, comparing utilization and spending on patients who have private insurance with that on similar patients within Medicare. Data on the former come from a large database of private sector medical claims, and on the latter from the Medicare Current Beneficiary Survey. Both datasets include patient-level demographics and co-morbidities, which allow us to identify regional variation distinct from individual characteristics such as health. The focus on heart disease helps

2. Alchian and Demsetz (1972) showed a greater incentive for shirking and inefficiency in public enterprise, where managers' and employees' own standards of living are unaffected by poor performance. De Alessi (1974a, 1974b) observed that inefficient private firms disappear, whereas inefficient public firms can last for long periods. Spann (1977) argued that private firms typically produce similar goods and services at much lower cost than their public counterparts.

mitigate the confounding impact of regional differences in health status on our analysis.

Our main object of interest is the regional variation in utilization and spending across sectors that cannot be explained by variation in patient characteristics. Our data suggest greater variation in utilization in the public sector: our main analysis suggests that variation in the public sector is about 2.8 times as great for outpatient visits ($p < 0.01$) and 3.9 times as great for hospital days ($p = 0.09$) as in the private sector. There is some evidence of greater variation for the number of hospitalizations in the public sector, but this evidence is less robust. Prescription drug utilization serves as our "placebo" case of insurance that was privately provided in *both* samples during the period investigated. Significantly, and unlike other types of medical care, drug utilization exhibits less variation among Medicare patients. On the other hand, there is greater spending variation in the private sector, suggesting the potential importance of monopsony power in the public sector.

The paper proceeds as follows. Section I provides the conceptual analysis of how differing cost-control measures within a region might lead to differences in regional variation in utilization and spending. Section II reports our empirical analysis comparing regional variation in the public and the private sectors. Section III discusses how our findings relate to the existing literature on health care variation and the resulting policy implications. Section IV discusses some limitations of our analysis and presents several robustness tests. Section V concludes.

I. A Simple Analysis of Regional Variation in Utilization and Spending

This section presents a simple analysis of how private and public incentives interact to create different degrees of regional variation in health care utilization and spending between the public and the private sector.[3] A key assumption is that private insurers have stronger incentives to restrain costs and utilization than a public insurer such as Medicare. This assumption is based on the literature demonstrating that, unlike public enterprises,

3. This analysis is general enough to include several possible sources of regional differences, and in particular it allows for such differences to be efficient. However, differences in liability (Kessler and McClellan 1996, 2002a, 2002b, Baicker and Chandra 2007) or productivity (Chandra and Staiger 2007), for example, may imply differences in efficient levels of care.

private firms have to restrain costs in order to compete on price, and private firms' inefficiencies have direct impacts on the welfare of their owners and employees (Alchian and Demsetz 1972, De Alessi 1974a). For example, private payers may explicitly manage care and exert pressure on providers through utilization review and case management. They can also selectively contract with lower-cost providers, steer patients to preferred providers, and exclude inefficient doctors or hospitals. In addition, prior authorization of large expenditures is prevalent in the private health insurance sector, a practice that allocates major spending decisions to the payer rather than the provider. Finally, private payers can steer patients toward efficient care through benefits management—for example, by not covering certain services unless certain clinical criteria are met. In what follows, we use the shorthand of "utilization restrictions" (UR) to refer to all these practices.

We interpret UR as a limit on the provision of treatments whose costs exceed their benefits. This may still lead to regional variation in utilization, because there is substantial heterogeneity among apparently similar patients in the efficacy of different treatments. Excessive care for one patient may be cost-effective for another.

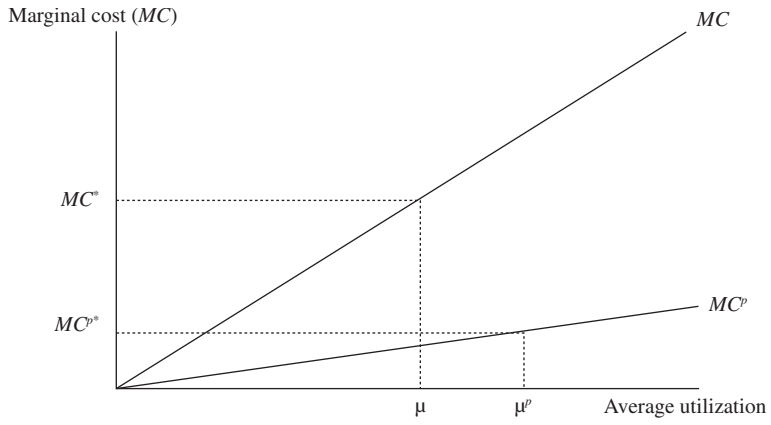
I.A. Causes of Sectoral Differences within Regions

We first consider the level of utilization in both the private and the public sectors. Define y^* as the efficient utilization level, that at which marginal benefit equals marginal cost. Following the earlier literature, we assume that private insurers have stronger incentives to limit utilization that rises above this level. They do this through UR, which we assume places an upper bound on utilization, $y_{UR} \geq y^*$, and perfectly eliminates inefficient utilization above that level.[4] The assumption of full efficiency is an analytical simplification; the positive predictions do not depend on it, and we do not emphasize the normative predictions.

Within any region there is a distribution of providers, who vary in the level of care they would provide to an identical patient. We characterize this distribution using the cumulative distribution function $F(y)$ for the random utilization variable Y. Private payers' UR procedures limit utilization and thereby truncate the support of the providers participating in their plans. This results in the private mean utilization level, $\mu = E(Y|Y \leq y_{UR})$.

4. Imperfect UR has qualitatively similar theoretical implications. The difference is one of degree rather than nature.

Figure 1. Variation in Spending Given Public Monopsony Power



Source: Authors' model described in the text.

This constrained private sector mean is thus lower than the unconstrained public sector mean, $\mu^p = E(Y)$.

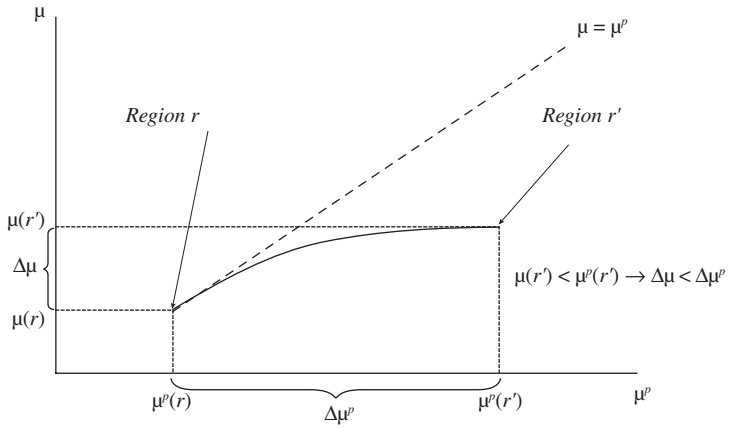
Now consider a pure increase in utilization, holding health status fixed. This can be represented as a rightward shift in the function $F(y)$. Assuming the efficient level of utilization remains fixed, the result is a greater difference in mean utilization across the two sectors, $\mu^p - \mu$. In other words, in regions with providers who have greater tendencies toward inefficiency, the difference in utilization between sectors will be larger.

The second key assumption is the presence of greater monopsony power in the public sector. The result is greater restraint of *prices,* as opposed to utilization, in the public sector. This affects the analysis of variation in spending, which combines the utilization effect and the price effect. If the government pays below-market prices through the exercise of either monopsony power or direct price regulation, the cost curves will differ across sectors. The result is depicted in figure 1. Average spending per patient in the private sector may exceed that in the public sector, if equilibrium marginal cost in the public sector, MC^{p*}, is less than equilibrium marginal cost in the private sector, MC^*.

I.B. Causes of Sectoral Differences across Regions

Next consider how mean utilization for each sector might vary across regions. Define the joint distribution $G(\mu^p, \mu)$ of mean utilization levels across regions. Specifically, suppose that the underlying distribution $F(y)$

Figure 2. Comparing Differences in Utilization across Sectors



Source: Authors' model described in the text.

differs across regions. Figure 2 illustrates how one might then characterize the relationship between changes in the public mean and the mean difference between sectors:

$$\frac{d\left(\mu^p - \mu\right)}{d\mu^p} = 1 - \frac{d\mu}{d\mu^p}.$$

For example, consider the case of normally distributed public sector utilization, $Y \sim N(\mu^p, \sigma^2)$. In this case, mean utilization in the private sector follows from the formula for the mean of a truncated normal random variable, $\mu = \mu^p + \sigma\lambda(\alpha)$, where $\lambda(\alpha) \equiv \frac{-\varphi(\alpha)}{\Phi(\alpha)}$ is the inverse Mills ratio and $\alpha = \frac{y_{UR} - \mu^p}{\sigma}$. This implies that the slope of the private mean as a function of the public mean is less than unity or, equivalently, that the between-sector difference rises with the public mean:[5]

$$0 \leq \frac{d\mu}{d\mu^p} \leq 1$$

$$\frac{d\left(\mu^p - \mu\right)}{d\mu^p} \geq 0.$$

5. We use the fact that the derivative of the inverse Mills ratio with respect to α is strictly between zero and 1, $\lambda'(\alpha) \in (0,1)$. (Sampford 1953).

When the public sector provides more care above the efficient level, this raises the between-sector difference. This in turn implies that the variance in the regional means in the public sector will exceed the variance in the regional means in the private sector: $V(\mu^p) > V(\mu)$.

This simple framework leads to several testable empirical predictions: Private provision should lead to lower mean utilization and less variance in mean utilization across regions, but not necessarily lower mean spending. In addition, the difference in utilization between sectors is likely to rise with the mean level of public utilization. Note that all these predictions hold patient health status constant.

II. Empirical Analysis of Regional Variation across Sectors

In this section we describe our empirical analysis of regional variation in the public and private sectors aimed at testing the implications discussed above.

II.A. Data and Empirical Specification

We compare regional variation between a sample of privately insured patients and a sample of Medicare patients. The private data come from a large database of health insurance claims. The data capture all health care claims, including prescription drugs and inpatient, emergency, and ambulatory services, by employees and retirees while they are enrolled in the health plans of 35 Fortune 500 firms. The analytical database integrates component datasets of medical claims, pharmacy claims, and enrollment records. This allows us to calculate spending and utilization for all services provided to the patients over our study period. The enrollment records allow us to identify basic demographics of the patients, including age, sex, and some information on income.[6] Importantly for our purposes, the data also include information on area of residence, coded by metropolitan statistical area (MSA) and 3-digit zip code. This allows us to analyze health care spending and utilization patterns at different levels of geographic aggregation.

Our Medicare sample is taken from the Medicare Current Beneficiary Survey (MCBS), which is administered to a nationally representative sample of aged, disabled, or institutionalized Medicare beneficiaries.

6. Our proxy for income is median household income at the 3-digit zip code level; this is taken from the 2000 Census.

Respondents, whether living in the community or residing in health care facilities, are interviewed up to 12 times over a 4-year period. Institutionalized respondents are interviewed by proxy. There is oversampling of the disabled under 65 years of age and of the oldest old (85 years of age or older). The MCBS uses a rotating panel design with limited periods of participation. Each fall a new panel is introduced with a target sample size of 12,000 respondents, and each summer a panel is retired. The MCBS data include detailed information on self-reported health status, health care use and expenditure, insurance coverage, and demographic characteristics. Additional Medicare claims data for beneficiaries enrolled in fee-for-service plans are also incorporated to provide more accurate information on health care use and expenditure. The MCBS data do not include actual claims data on prescription drugs; all information on prescription drug spending and utilization in the MCBS is self-reported. This leads to a known undercount of drug spending in the MCBS.[7]

Both datasets include information on medical claims that is used to compile utilization, spending, and baseline health information. That is, although the MCBS contains a survey component, all data on spending and utilization are compared with Medicare's administrative claims data (Eppig and Chulis 1997). However, since Medicare does not cover prescription drugs over our sample period, this validation procedure applies to medical care but not drugs. Finally, for both datasets we use information from 2000 to 2006. The one exception is prescription drug utilization and spending: to abstract from the complexities of Medicare Part D's introduction, we eliminate the 2006 data for these variables.

To mitigate differences in health status across sectors and regions, we condition inclusion in the sample on a diagnosis of ischemic heart disease (IHD).[8] We also use the diagnosis codes on medical claims to identify whether patients were treated for any of 30 different conditions in a

7. When estimating the cost of Medicare Part D (for example), the Congressional Budget Office scaled the reported MCBS prescription drug spending up by 33 percent for the noninstitutionalized population (Christensen and Wagner 2000).

8. Also called myocardial ischemia, IHD is characterized as reduced blood flow to the heart. In the private data we identify patients with IHD as those with at least one inpatient or outpatient claim with a primary diagnosis ICD-9 code of 410.xx, 411.xx, 412.xx, 413.xx, or 414.xx. In the public data we identify patients based on self-reports of ever diagnosed with heart disease. See the online data appendix (available on the *Brookings Papers* webpage at www.brookings.edu/economics/bpea.aspx, under "Conferences and Papers") for more information.

calendar year.[9] The claims-based measures of the number of diseases are available in both the MCBS and the private health insurance data.[10] This is important because unmeasured differences in severity across regions could lead to spurious positive correlation between sectors.

The primary geographic unit of analysis for our study is the MSA. An alternative candidate would be the hospital referral region (HRR), used by the Dartmouth Atlas. However, HRRs are not reported in either of our datasets, and the private sector data do not contain 5-digit zip codes, which are required to construct an individual's HRR. We restrict our sample to the 99 MSAs where we have the largest samples. MSAs are somewhat larger than HRRs, and this may compress the variation for both sectors in our data.

Our final sample contains 240,028 private patients and 24,800 public patients.[11] Since there are many fewer public patients, it is important to correct for the effects of sample size on our estimates. We derive and report these corrections in detail below.

Table 1 reports some summary statistics comparing demographic characteristics in the public and the private samples. As one would expect, the average age in the private sample is lower than in the sample of Medicare patients, most of whom are older than 65. The private sample contains a greater fraction of males, in part because it is influenced by current or past employment status. (The private sample contains both active workers and retirees receiving benefits from their current or past employers.) Average income is also higher in the private sample. The greater variance in income for the public sample is likely due to the fact that income is reported individually in the MCBS, but imputed at the local level in the private sample.

9. The specific conditions considered are essential hypertension, congestive heart failure, diabetes, asthma, hypercholesterolemia, ulcer, depression, chronic obstructive pulmonary disease, allergic rhinitis, migraine, arthritis, chronic sinusitis, anxiety disorder, cardiac disease, vascular disease, epilepsy, gastric acid disorder, glaucoma, gout, hyperlipidemia, irritable bowel syndrome, malignancy, psychotic illness, thyroid disorder, rheumatoid arthritis, tuberculosis, angina, human immunodeficiency virus infection, anemia, and stroke. Most co-morbidities are relatively uncommon, except for the ones involving heart disease (or risk factors for heart disease).

10. The MCBS also contains self-reports of a number of distinct health conditions, as well as the individual's self-reported general health status (coded 1 to 5, with 1 indicating poor and 5 indicating excellent). Our regression analysis relies on the claims-based, rather than self-reported, disease measures for both the public and the private samples. More details appear in our online data appendix.

11. See the footnotes to tables 1 and 2 for a few sample size issues specific to certain variables.

Table 1. Selected Patient Demographic and Health Characteristics[a]

	Private		*Public*	
Patient characteristic	*Mean*	*Standard deviation*	*Mean*	*Standard deviation*
Age	55.4	7.1	78.2	7.9
Percent male	65	48	46	50
Income (thousands of 2004 dollars)	42.8	10.8	28.5	46.8
Percent with heart disease in year	32	47	37	48
No. of adverse health conditions	1.4	1.4	2.9	2.4

Sources: Data on private patients come from a modified version of the Ingenix Touchstone product. Data on public patients come from the MCBS.

a. History of heart disease is self-reported in the public sample and identified using medical claims in the private sample. The private sample has 240,028 observations and the public sample 24,800 observations.

The table also compares the health of individuals in the two samples. Since both samples are limited to individuals with a history of heart disease, we include a variable indicating the fraction of individuals who are diagnosed with heart disease in a particular year. In all cases, the presence of disease is taken from claims rather than from self-reported data. The incidence of heart disease is similar in the two samples: 0.32 in the private sample and 0.37 in the public sample.

In addition, the table reports the average number of adverse health conditions (out of the total of 30, including heart disease) per patient. As with heart disease, the health conditions are determined using the ICD-9 diagnosis codes from medical claims in both the public and the private samples. Unsurprisingly, the elderly individuals in the public sample are much sicker on average, with 2.9 adverse health conditions in the year compared with 1.4 in the private sample.

As a matter of course, the public and the private samples are drawn from different populations. We include a number of controls and analyses designed to mitigate and test for the impact of these differences, but heterogeneity across samples remains a possibility. Later we discuss the sources of heterogeneity, the methods we have employed to address them, and their possible implications for the analysis.

II.B. Descriptive Statistics

Table 2 presents some descriptive statistics for health care spending and utilization in the public and the private samples aggregated over all regions and patient characteristics. We present not only the mean and the standard deviation but also the 25th-percentile, median, and 75th-percentile values. Our utilization measures (all measured as yearly averages per patient)

Table 2. Distributions of Spending and Utilization Measures[a]

Measure	*Sample*	*Mean*	*Standard deviation*	*25th percentile*	*Median*	*75th percentile*
Utilization (number per patient per year)						
Hospitalizations	Private	0.36	1.15	0	0	0
	Public	0.57	1.14	0	0	1
Hospital days	Private	1.23	7.13	0	0	0
	Public	2.93	8.59	0	0	1
Outpatient visits	Private	5.56	5.86	1	4	8
	Public[b]	8.59	11.05	1	5	12
Drug prescriptions[c]	Private	45.78	42.20	13	36	66
	Public	35.45	29.93	14	29	50
Spending (thousands of 2004 dollars)						
Total spending	Private	8.40	22.98	0.56	2.10	6.88
	Public	10.25	18.8	1.25	3.91	11.4
Inpatient spending	Private	4.02	18.36	0	0	0
	Public	4.94	13.21	0	0	4.65
Outpatient spending	Private	4.38	9.83	0.54	1.85	4.86
	Public	5.30	9.14	1.13	2.94	6.44
Prescription drug spending[c]	Private	2.80	5.78	0.53	1.67	3.42
	Public	1.92	2.05	0.58	1.39	2.63

Source: Authors' calculations.

a. Figures are yearly averages during 2000–06 (2000–05 for drug prescriptions) for patients with a history of heart disease, which is self-reported in the public sample and identified using medical claims in the private sample. Except where noted otherwise, the private sample has 240,028 observations and the public sample 24,800 observations.

b. Survey responses (used to cross-validate the claims data) were incomplete in 3,769 cases, so that the public sample has 21,031 observations.

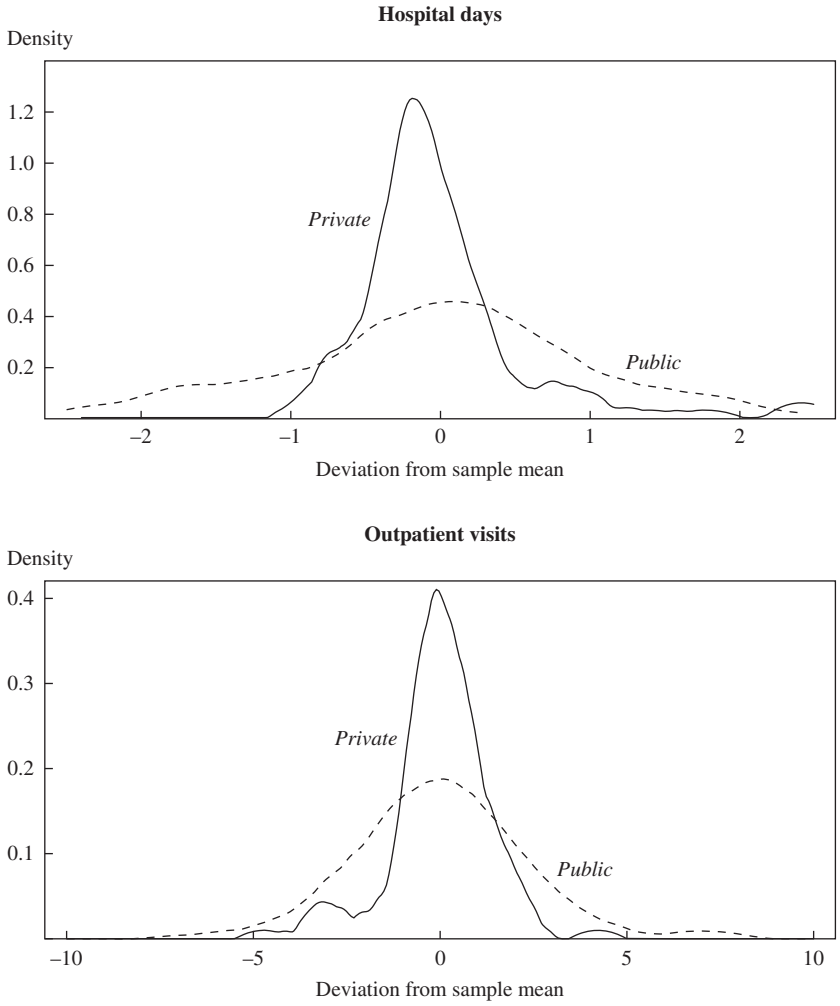
c. Because observations from 2006 are omitted, the private sample has 231,802 observations and the public sample 21,140 observations. Number of prescriptions is in terms of 30-day-equivalents.

include the number of hospitalizations, total hospital days across all hospitalizations, the number of outpatient visits, and the number of 30-day-equivalent prescriptions in both samples. For spending, we record total (inpatient plus outpatient), inpatient, and outpatient spending, as well as spending on prescription drugs.

Utilization, in terms of hospitalizations, hospital days, and outpatient visits, is lower for the private patients. Spending for this group also tends to be lower. Total medical spending for individuals in the private plans is $8,401 per year, compared with $10,245 for the Medicare patients—about a 20 percent difference. The exception to the pattern is prescription drugs, for which both utilization and spending are greater among private patients.

Figures 3 and 4 provide a broad sense of the variation present in our samples. Figure 3 reports for both samples the estimated kernel densities of MSA-level deviations from the mean for both hospital days and

Figure 3. Kernel Density Estimates of Regional Fixed Effects for Selected Utilization Measures[a]

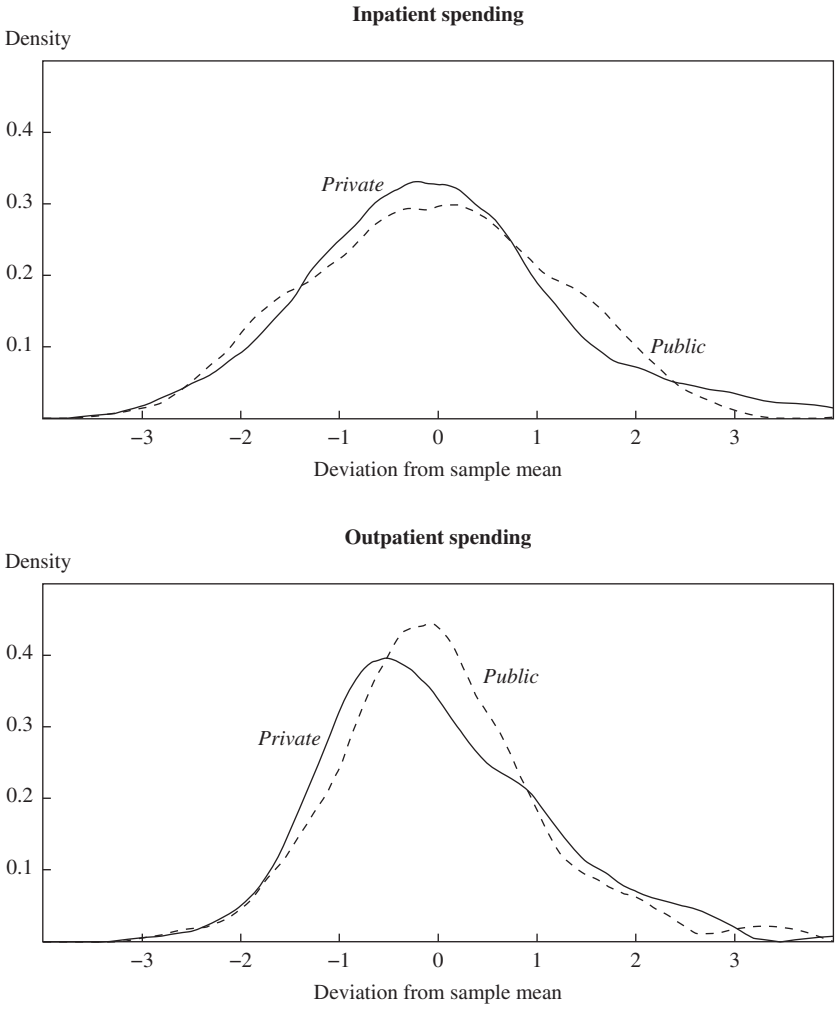


Source: Authors' calculations.

a. Estimated kernel densities of the deviation of mean hospital days and outpatient visits per patient per year across MSAs.

outpatient visits. Each data point underlying the kernel estimate is the difference between an MSA-level mean and the overall sample mean. For both variables, the distributions appear to be tighter for the private than for the public sample. However, these distributions are based on raw, unadjusted numbers that do not account for disease or other covariates.

Figure 4. Kernel Density Estimates of Regional Fixed Effects for Selected Spending Measures[a]

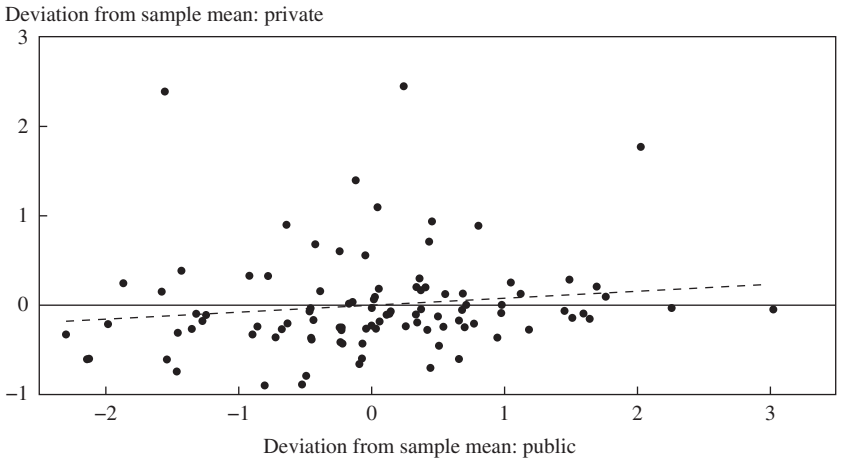


Source: Authors' calculations.

a. Estimated kernel densities of the deviation of mean inpatient and outpatient spending per patient per year across MSAs.

Figure 4 repeats this exercise for inpatient and outpatient spending. Here the findings are decidedly more mixed. For outpatient spending the distribution appears to be slightly tighter for the public sample. The figure for inpatient spending is harder to interpret visually, as the differences in the densities are small and asymmetric. In any event, the differences

Figure 5. Hospital Days: Relationship between Public and Private Deviations from the Mean[a]



Source: Authors' calculations.

a. Each observation pairs the deviation for a single MSA from mean hospital days per patient per year in the public sample with that for the same MSA for the private sample. The line represents the fitted ordinary least squares regression line.

observable visually between the spending and the utilization distributions suggest the possible importance of public sector price restraints, which would lower spending variation even with greater variation in utilization.

Finally, figure 5 plots the relationship between deviations from the MSA-level means for public and private hospital days. This is the empirical analogue to the theoretical relationship in figure 2. The figure suggests that mean private hospital days increase slightly with mean public hospital days, but much less than one for one. This is consistent with there being less regional variation in the private sector; we test this hypothesis more formally in the following analyses.

II.C. Framework for Estimating Regional Variation

We are particularly interested in the between-MSA variance in spending and utilization for the public and the private samples. We begin with the simplest possible approach that evaluates the variance between MSAs in the sample means. We then move to estimating the variance in regression-adjusted means, which we estimate from regressions that control for various factors that might also influence spending and utilization. In both cases we account for the relative bias that is created by the substantial differences in

sample size across sectors: because the public samples are much smaller than the private samples, there is greater sampling variance in the public sector estimates and thus greater variation in the MSA-level means for Medicare patients. To estimate the true between-sector differences in regional variation, we estimate and remove the variability that is due to sample size differences alone.

Formally, the observed regional variation within a sector is due to the true variation and the sampling variance in estimating that variation. Denote by μ_r the true mean for region r and by $\hat{\mu}_r$ the corresponding sample estimate, whether unconditional or regression-adjusted. The sample mean is equal to the true mean plus sampling error, according to

$$\hat{\mu}_r = \mu_r + z_r.$$

The sampling error z_r has zero mean, and the covariance of the sampling error across regions is $E(z, z_s) = \sigma_{rs}$. Define $\bar{\mu} \equiv \frac{1}{R}\sum_{r=1}^{R}\mu_r$, the "grand mean" across regions. Similarly, define the corresponding sample analogue, $\bar{\hat{\mu}} \equiv \frac{1}{R}\sum_{r=1}^{R}\hat{\mu}_r$. Finally, define the average sampling error across regions, $\bar{z} \equiv \frac{1}{R}\sum_{r=1}^{R} z_r$. The object of interest is the degree of regional variation in the true MSA means, $RV \equiv \frac{1}{R}\sum_{r=1}^{R}\left(\mu_r - \bar{\mu}\right)^2$, which has the sample analogue $\widehat{RV} \equiv \frac{1}{R}\sum_{r=1}^{R}\left(\hat{\mu}_r - \bar{\hat{\mu}}\right)^2$.

The observed variation $\widehat{RV}$ is a biased estimator of *RV,* as a result of sampling error in the estimates. Moreover, this bias is likely to be larger for our public sector estimates because of the smaller public sector sample size, which yields noisier estimates of public sector utilization. However, we can recover a consistent estimate of the bias and correct for it, according to

$$E\left(\widehat{RV}\right) = RV + \frac{1}{R}\sum_{r=1}^{R} E\left(z_r - \bar{z}\right)^2.$$

In the appendix we show how to estimate this bias from sample variances and covariances. Our formula works for both the case of unconditional sample means and the case of regression-adjusted means. In the simple case without covariance across regions or zero average error across regions, this expression simply states that the observed variation is the true

variation plus the average squared standard error. More generally, the more precisely the sample means are estimated, the smaller is the bias correction.

In sum, the object of interest in our analysis is RV, which we estimate as $\widehat{RV} - \widehat{Bias}$ for both the public and the private sector. Using these estimates of regional variation, we report both the ratio of public to private variation and the difference between public and private variation. We construct standard errors around these by means of a bootstrap procedure, which samples individuals with replacement within each MSA, so that each bootstrap sample contains exactly the same number of individuals in each MSA as the original sample.[12] The bootstrap procedure reflects the nature of our sample design. We regard the set of MSAs as fixed but each sample within an MSA as a random sample of that MSA's population. Statistically, our set of MSAs approximates a population, but we have samples within each MSA.

Our regression-adjusted estimates employ a model with regional fixed effects that controls for disease severity and demographics.[13] For each sector s we estimate

$$Y_{irts} = \alpha_{0s} + \mathbf{X}_{its}\beta_s + \delta_{ts} + \delta_{rs} + \epsilon_{irts}.$$

Here Y_{irts} represents some measure of utilization or spending by patient i in region r at time t and in sector s. The vector $\mathbf{X}$ includes the following demographic characteristics for each patient: age, age squared, sex, income, income squared, age and age squared interacted with sex, as well as dummy variables for each of the adverse health conditions listed above. The terms δ_{ts} and δ_{rs} are sector-specific fixed effects for year and MSA, respectively. The sector-specific variance in the fixed effect δ_{rs} is the regression-adjusted analogue to the variance in the MSA-level sample means.

As a general matter, the covariates have relatively little predictive power within MSAs but a fair amount between MSAs. Across all specifications, for instance, the MSA means of the covariates explain about 50 to 70 percent of the between-MSA variation in utilization and spending, in the sense of R^2.

12. The alternative block-bootstrap that samples MSAs with replacement generates nearly identical inferences for statistical significance in our analysis, and so does a "flat" bootstrap.

13. A possible alternative is a random-effects model, but the Hausman test rejected this more efficient model in favor of the fixed-effects model in the majority of cases we analyzed.

Table 3. Regional Variation in Mean Utilization[a]

	Observed variation[b]		*Corrected variation*[c]		*Difference, public minus private*	*Ratio of public to private*
Utilization measure	*Private*	*Public*	*Private*	*Public*		
Unconditional means						
Hospitalizations	0.013	0.015	0.012	0.009	–0.003	0.728
					(0.003)	(0.204)
Hospital days	0.322	1.016	0.230	0.659	0.429**	2.870*
					(0.199)	(1.017)
Outpatient visits	1.736	5.154	1.676	4.585	2.909***	2.735***
					(0.502)	(0.323)
Drug prescriptions[d]	72.746	32.758	70.090	28.403	–41.687***	0.405***
					(3.896)	(0.043)
Regression-adjusted means[e]						
Hospitalizations	0.006	0.011	0.005	0.006	0.001	1.266
					(0.002)	(0.430)
Hospital days	0.169	0.610	0.080	0.313	0.233*	3.907*
					(0.124)	(1.684)
Outpatient visits	0.988	3.255	0.942	2.677	1.735***	2.841***
					(0.322)	(0.379)
Drug prescriptions[d]	29.190	25.856	27.086	21.131	–5.955*	0.780**
					(3.130)	(0.106)

Source: Authors' calculations.

a. Numbers in parentheses are standard errors on the difference between public and private variation or the ratio of public to private variation and are bootstrapped within MSAs, separately for public and private patients, with 200 bootstrap draws. For both sectors, then, the number of patients in each region in each bootstrapped sample is the same as the number of patients in the original sample. Asterisks indicate differences statistically significantly different from zero or ratios statistically significantly different from 1 at the ***1 percent, **5 percent, and *10 percent level.

b. Variance in the regional means or fixed effects of the utilization variables.

c. Unbiased measure of the true variance in the regional means or fixed effects corrected for sampling error, as described in the text and the appendix. All differences and ratios are based on these numbers.

d. 30-day-equivalents.

e. Estimates of regional fixed effects on each utilization variable from a regression that includes as other independent variables year fixed effects, quadratic specifications of patient age and income, patient sex, sex interacted with age, and dummy variables for 30 separate types of disease.

II.D. Regional Variance in Utilization and Spending

Table 3 reports the estimated regional variance in four utilization measures: number of hospitalizations, number of hospital days, number of outpatient visits, and number of prescription drugs (in terms of 30-day-equivalents). Again, prescription drug coverage is provided by the private sector in both populations throughout the sample period, and therefore we do not expect to see similar differences for prescription drugs as for the other measures.

The table shows overall between-MSA variation in the public and the private sectors. The observed variation (first two columns) is computed as the average MSA-level deviation from the overall mean. The top panel reports the variation based on unconditional means; in the bottom panel, both the overall mean and each MSA-level mean are regression-adjusted, as described above. The corrected variation (second two columns) is computed by subtracting the expected bias due to sampling error, as described above. The next column shows the absolute difference between the public and private variances, and the last column the ratio of the variances. Asterisks indicate statistically significant differences from zero for the differences, and from unity for the ratios.

Variation in hospital days is about three times, and variation in outpatient visits about two times, higher in the public sector. These differences are statistically significant at the 10 percent level or higher and appear regardless of whether we adjust for covariates (although the magnitudes differ somewhat). On the other hand, prescription drug utilization exhibits statistically less variation in the Medicare population; this is important because, again, even Medicare patients obtain their prescription drug insurance privately in our sample. Finally, there is no statistically significant difference in the variation for hospitalizations. It is likely that more statistical power is needed to pin down this variance, in one direction or the other. Overall, these results provide evidence suggesting higher variance in the public sector, but for a few of the outcomes our statistical tests lack the power to generate definitive results.

Table 4 reports the estimated regional variance in four spending measures: total spending, inpatient spending, outpatient spending, and prescription drug spending. The regression-adjusted estimates indicate that outpatient spending exhibits only about 35 percent as much variation in the public sector as in the private sector. Inpatient spending exhibits roughly equal variation in the two sectors. Finally, prescription drug spending varies less for Medicare patients. With that exception, these results are quite different from the utilization results and suggest that price restraints play a role in the public sector. In spite of greater variation in utilization, the public sector exhibits less variation in spending.

III. Comparisons with Existing Literature

Regional variation in spending and utilization in the public sector has been well documented in a literature that is almost 40 years old and well accepted by the academic community. In that sense, our contribution is to

Table 4. Regional Variation in Mean Spending[a]

	Observed variation[b]		*Corrected variation*[c]		*Difference, public minus private*	*Ratio of public to private*
Spending measure	*Private*	*Public*	*Private*	*Public*		
Unconditional means						
Total medical spending	4.443	3.907	3.571	2.352	–1.219 (0.981)	0.659* (0.207)
Inpatient spending	1.634	1.587	1.109	0.842	–0.266 (0.416)	0.760 (0.247)
Outpatient spending	1.263	1.048	1.082	0.634	–0.447 (0.355)	0.586 (0.266)
Prescription drug spending	0.418	0.144	0.377	0.124	–0.253*** (0.049)	0.329*** (0.060)
Regression-adjusted means[d]						
Total medical spending	2.890	2.782	2.111	1.463	–0.647 (0.728)	0.693 (0.258)
Inpatient spending	1.186	1.357	0.698	0.695	–0.004 (0.324)	0.995 (0.298)
Outpatient spending	0.924	0.628	0.758	0.265	–0.494* (0.261)	0.349** (0.272)
Prescription drug spending	0.251	0.086	0.214	0.064	–0.150*** (0.043)	0.300*** (0.082)

Source: Authors' calculations.

a. Spending is measured in 2004 dollars. Numbers in parentheses are standard errors on the difference between public and private variation or the ratio of public to private variation and are bootstrapped within MSAs, separately for public and private patients, with 200 bootstrap draws. For both sectors, then, the number of patients in each region in each bootstrapped sample is the same as the number of patients in the original sample. Asterisks indicate differences statistically significantly different from zero or ratios statistically significantly different from 1 at the ***1 percent, **5 percent, and *10 percent level.

b. Variance in the regional means or fixed effects of the spending variables.

c. Unbiased measure of the true variance in the regional means or fixed effects corrected for sampling error, as detailed in the text and the appendix. All differences and ratios are based on these numbers.

d. Estimates of regional fixed effects on each spending variable from a regression that includes as other independent variables year fixed effects, quadratic specifications of patient age and income, patient sex, sex interacted with age, and dummy variables for 30 separate types of disease.

compare this with variation in the private sector, rather than to establish the existence of public sector variation.

Table 5 summarizes a few representative papers from this vast literature.[14] John Wennberg and Alan Gittelsohn (1973) provide an early example. Their study analyzed variation across "hospital service areas," a precursor to the HRRs typically analyzed in the modern Dartmouth Atlas

14. For useful summaries from both the economic and the clinical literatures on geographic variation in health care, see Wennberg and Cooper (1998), Phelps (2000), Fisher and others (2003a, 2003b), Chandra and Staiger (2007) and Sutherland, Fisher, and Skinner (2009).

Table 5. Key Findings on Variation in Regional Health Care Spending Using Medicare Data[a]

Study	*Geographic aggregation*	*Summary*	*Key findings*
Wennberg and Gittelsohn (1973)	Hospital service area	Studied geographic variation in utilization and spending in Vermont.	Found wide variations apparently due to differences in practice style rather than in population health. Hospital days in highest-use area were 1.5 times that in lowest.
Cutler and Sheiner (1999)	HRR	Calculated share of regional variation in Medicare spending attributable to regional differences in health and demographics.	Regional differences in demographics can explain about 70 percent of regional differences in Medicare spending, but significant differences remain unexplained.
Fisher and others (2003a)	HRR	Compared patients across regions holding other characteristics constant.	Patients in higher-spending regions received approximately 60 percent more care. The increased utilization mostly arose from more frequent physician visits.
Sirovich and others (2006)	MSA	Compared variation in intensity of treatment with physician perceptions of quality of care.	Medicare spending per capita in highest intensity quintile was 1.58 times that in lowest.
Chandra and Staiger (2007)	HRR	Specified a model of patient treatment choice with productivity spillovers and tested the model using treatment choices and health outcomes of heart attack patients.	Patterns of which patients benefit and which lose from intensive medical care are consistent with productivity spillover model.
Fowler and others (2008)	HRR	Used a patient survey to compare local variation in spending and utilization with patient perceptions of quality.	Regional differences in spending were not systematically related to differences in patient perceptions of care quality.

Wennberg and others (2008)	HRR	Summarized Dartmouth Atlas findings on geographic variation in Medicare spending and their implications.	Three states spent more than 20 percent above the national average of $46,412. Conversely, three states spent 25 percent below the national average or less. Interquartile ratio (75th percentile over 25th) is 1.26 for HRRs.
Rettenmaier and Saving (2009)	State	Studied how state rankings in medical spending per capita change when different definitions of spending are used, such as Medicare only or total spending by all payers.	Found a state-level correlation between Medicare spending and total spending of 0.21, and that variation in Medicare spending exceeds variation in private spending.
Sutherland, Fisher, and Skinner (2009)	HRR	Updated Dartmouth Atlas findings on geographic variation in Medicare spending and their implications.	Inpatient days per beneficiary in highest cost quintile were 1.50 times that in lowest; physician visits in highest cost quintile were 1.36 times that in lowest.
Chernew and others (2010)	HRR	Compared HRR-level variation in medical spending between Medicare and large firms.	Found substantial regional variation in spending, greater for large firms than for Medicare (coefficient of variation 0.21 v. 0.16). Correlation between private and public inpatient utilization was 0.59.
Gottlieb and others (2010)	HRR	Examined role of Medicare prices in driving geographic variations in health care.	Prices explain a small portion of variation in spending. The 80th percentile of price-adjusted Medicare Part B spending was 1.37 times the 20th percentile.

Sources: Literature cited.
a. HRR = hospital referral region; MSA = metropolitan statistical area.

of Health Care studies. The table also lists a couple of important studies that use states or MSAs. It is important to recognize this difference when comparing our MSA-level analysis with HRR-level analyses elsewhere, and it is important for future work to assess the potential implications of this difference.

The 2008 Dartmouth Atlas of Health Care reports that average spending on health care in the last 2 years of life (for deaths occurring from 2001 to 2005) ranged from a high of $59,379 in New Jersey to $32,523 in North Dakota (Wennberg and others 2008). This range, from 28 percent above to 30 percent below the national average, is similar to the range of quantity utilization reported across MSAs by MedPac: from 39 percent above the national average in Miami to 25 percent below in rural Hawaii (Medicare Payment Advisory Commission 2009).

These variations are not fully explained by factors such as age, insurance coverage, average income, and rates of illness or disease. David Cutler and Louise Sheiner (1999) investigate the extent to which variation in spending across HRRs can be explained by regional differences in illness, in the demand for health (for example, as measured by income and race), or in "exogenous differences in the structure of medical care markets" (for example, in the ratio of generalists to specialists). They find that regional demographics can explain about 70 percent of the variation in medical spending across regions, but the unexplained variation remains large. For example, when differences in demographics and the illness of the population are accounted for, bringing Medicare spending down to the level of the 10th-percentile region would reduce total spending by 15 percent.

Perhaps the existing study most closely related to ours is that of Michael Chernew and others (2010), who compare HRR-level variations in Medicare against those in a sample of large firms in the Thomson Reuters (Medstat) MarketScan Commercial Claims and Encounters Database. They estimate that the geographic variation in private sector spending is greater than that in Medicare spending (coefficient of variation of 0.21 versus 0.16). This is consistent with our findings for spending. They focus less on variation in utilization, although they do report a positive correlation between Medicare and non-Medicare inpatient days.

IV. Limitations of Our Analysis

There are several empirical questions that our data cannot address but that should be addressed in future work. The populations of privately insured and publicly insured patients differ, because the latter have often opted out

of private health insurance options. The empirical implications of this are not clear a priori. Fee-for-service Medicare patients are likely to be sicker than their counterparts in private Medicare health maintenance organizations (HMOs), because HMOs attempt to select healthier patients (Morgan and others 1997). On the other hand, the privately insured nonelderly may also be healthier than the nonelderly overall, if private health insurers select against the sickest patients for similar reasons. The link between health insurance and employment in the nonelderly population adds further complexity, as those who are eligible for employment-based health insurance may be richer or healthier, or both, than their peers. Finally, the fact that our private sector data are based only on employees of large (Fortune 500) firms adds a further dimension of selection.

We ran several supplementary analyses to investigate some of these issues, but our data lack the power to reach definitive conclusions across the board. First, we narrowed the age range of our comparisons, to mitigate some of the differences in health status. We compared 60- to 64-year-olds in the commercially insured population with 66- to 70-year-olds in the fee-for-service Medicare population. As this restriction further reduces the sample, we limit our analysis to the 70 MSAs for which we have at least 25 observations in both samples.

Table 6 reports the result for the samples with the narrow age ranges. Generally, the point estimates based on these restricted age ranges are similar to those based on the full sample, but the precision of the estimates declines enough to eliminate statistical significance. The point estimates indicate that variation in the public sector is about 5.1, 3.4, and 1.2 times that in the private sector for hospital days, outpatient visits, and hospitalizations, respectively. As in the analysis based on the full sample, variation in prescription drug use is smaller in the public sector, about 53 percent as large as variation in the private sector.

Next we investigated the issue of selection based on employment by comparing our privately insured sample with Medicare patients who also have coverage from an employer. If an individual has such coverage, we know that he or she was employed and privately insured at one point. Roughly 35 percent of Medicare enrollees in our sample also have employer-provided coverage. They are slightly younger (averaging 77 years, compared with 79 years for those without such coverage), richer (average income is 58 percent higher), and more likely to be male (52 percent versus 40 percent) than the average Medicare enrollee. Having employer coverage is associated with very small differences in the fraction of total expenses paid for by Medicare: Medicare pays 39 percent

Table 6. Regional Variation in Regression-Adjusted Mean Utilization, Patients Aged 60 to 70[a]

Utilization measure	*Observed variation*[b]		*Corrected variation*[c]		*Difference, public minus private*	*Ratio of public to private*
	Private	*Public*	*Private*	*Public*		
Hospitalizations	0.015	0.035	0.011	0.014	0.003	1.244
					(0.010)	(0.921)
Hospital days	0.258	2.038	0.130	0.656	0.526	5.060
					(0.886)	(4.134)
Outpatient visits	1.460	7.207	1.335	4.524	3.190	3.390
					(2.316)	(1.695)
Prescriptions	52.869	47.198	46.741	24.765	–21.975**	0.530
					(11.090)	(0.209)

Source: Authors' calculations.

a. The private sample is restricted to patients aged 60 to 64 and the public sample to patients aged 66 to 70. Both samples are restricted to include only the 70 MSAs with at least 25 observations in both samples. The private sample has 67,414 observations and the public sample 3,568 observations. Numbers in parentheses are standard errors on the difference between public and private variation or the ratio of public to private variation and are bootstrapped within MSAs, and separately for public and private patients, with 200 bootstrap draws. For both sectors, then, the number of patients in each region in each bootstrapped sample is the same as the number of patients in the original sample. Asterisks indicate differences statistically significantly different from zero or ratios statistically significantly different from 1 at the ***1 percent, **5 percent, and *10 percent level.

b. Variance in the regional means or fixed effects of the utilization variables.

c. Unbiased measure of the true variance in the fixed effects corrected for sampling error, as detailed in the text and the appendix. All differences and ratios are based on these numbers.

of the expenses of those without employer coverage and 38 percent of those with such coverage. The lack of a disparity is due to the fact that once an elderly Medicare beneficiary retires, the employer-provided coverage becomes secondary to Medicare. In our data just 9 percent of individuals in the Medicare sample with employer coverage are working, so for the vast majority Medicare is the primary payer. It thus seems reasonable to assume that Medicare is the primary driver of resource allocation for these individuals. A number of MSAs are left with very small samples after this restriction, so we limit our analysis to the 77 MSAs where we have at least 50 observations in both samples.

These results are presented in table 7. Again, the point estimates are similar to those based on the full sample, but the precision of the estimates declines enough to eliminate much of the statistical significance. The point estimates indicate that variation in the public sector is about 4.1, 3.8, and 1.6 times that in the private sector for hospital days, outpatient visits, and hospitalizations, respectively. The greater variation in outpatient visits in the public sample is statistically significant at the 1 percent level. The other

Table 7. Regional Variation in Regression-Adjusted Mean Utilization, Patients with Some Private, Employer-Provided Coverage[a]

Utilization measure	*Observed variation*[b]		*Corrected variation*[c]		*Difference, public minus private*	*Ratio of public to private*
	Private	*Public*	*Private*	*Public*		
Hospitalizations	0.006	0.017	0.005	0.008	0.003	1.572
					(0.004)	(0.984)
Hospital days	0.144	0.709	0.049	0.198	0.149	4.058
					(0.240)	(3.861)
Outpatient visits	3.774	4.462	0.911	3.438	2.527***	3.774***
					(0.518)	(0.614)
Prescriptions	28.982	34.778	26.885	25.107	–1.778	0.934
					(6.032)	(0.213)

Source: Authors' calculations.

a. The public sample is restricted to patients who report at least some form of private, employer-provided insurance coverage. Both samples are restricted to include only the 77 MSAs that have at least 50 observations in both samples. The private sample has 202,202 observations and the public sample 8,416 observations. Numbers in parentheses are standard errors on the difference between public and private variation or the ratio of public to private variation and are bootstrapped within MSAs, and separately for public and private patients. For both sectors, then, the number of patients in each region in each bootstrapped sample is the same as the number of patients in the original sample. Asterisks indicate differences statistically significantly different from zero or ratios statistically significantly different from 1 at the ***1 percent, **5 percent, and *10 percent level.

b. Variance in the regional means or fixed effects of the utilization variables.

c. Unbiased measure of the true variance in the fixed effects corrected for sampling error, as detailed in the text and the appendix. All differences and ratios are based on these numbers.

differences are not significant at the 10 percent level. Variation in prescription drug use is slightly lower in the public sector, about 93 percent as large as in the private sector.

V. Concluding Remarks

It has long been recognized that public and private enterprises face different incentives to control costs. This paper has analyzed these differences in the health insurance context, along with their implications for variation in care. Public payers are likely able to restrain prices better than private payers but have weaker incentives to control costs through utilization controls. As a result, one might expect greater variation in utilization for the public sector, but the effects on total spending are ambiguous. Using samples of heart disease patients, we presented empirical evidence consistent with these implications.

Further research should focus more closely on the issue of whether and to what extent variations across sectors are the result of differences in the

baseline health of the publicly and privately insured populations. Work is also needed to assess whether our basic findings can be generalized across other disease categories and geographical classifications. In addition, the analysis of health outcomes must be integrated into the analyses of utilization and of spending. As a related point, although we have focused on estimated variations, further research should be conducted into the sources of the mean differences in utilization and spending. Finally, and perhaps most important, research is needed to draw out the normative implications of variations in care both within and between sectors.

The normative implications of variation in care are not straightforward, in spite of the conventional wisdom that greater variation implies inefficiency. On the one hand, the literature has consistently found that areas exhibiting higher utilization of health care services do not exhibit demonstrably better outcomes for patients (Fisher and others 2003a). This has led many to conclude that these areas are practicing "flat-of-the-curve" medicine, where the marginal benefit approaches zero. However, Amitabh Chandra and Douglas Staiger (2007) demonstrate that productivity spillovers and specialization can explain regional variation in the utilization of intensive procedures, without resorting to inefficiency. Most notably, their model can reconcile the seemingly contradictory evidence that intensive treatments such as most surgery are often highly effective at the individual level, but that regions using these treatments more intensively do not have better average health outcomes. Chandra and Staiger observe that regions specializing in intensive treatment will find it optimal to provide that treatment to more patients; therefore, the marginal patient in such regions will be less suited to it than the marginal patient elsewhere. This mitigates the greater benefits of intensive treatment.

For this and other reasons, the efficiency implications of variation in care require a more careful analysis. Even in our simple framework, the normative impact of variation is unclear. For instance, if the private sector is pricing and producing efficiently, then the theory suggests that the public sector is engaging in inefficiently high utilization and inefficiently low pricing. On the other hand, if private sector prices are too high or if utilization is too low, then the effects of public insurance may actually represent second-best improvements to welfare. Evidently, it is important to investigate the baseline efficiency properties of the private health insurance market and to characterize how these are affected by the presence of publicly financed health insurance.

Regardless of the conclusions, normative analysis of this issue will likely generate a number of important policy implications. Many have noted that

Medicare has lower administrative costs than the private sector. This is often interpreted as part of the value generated by centralized insurance. This is a typical finding when one is comparing a centralized with a decentralized model, but it could also be explained by the cost of administering utilization controls in the private sector. If so, any efficiency benefits of utilization controls would need to be weighed against these administrative costs. The benefits generated by administrative costs are often neglected in the policy debate, as are related issues such as the deadweight costs of the tax revenue required to fund public enterprise, the efficiency gains of marketing activities by private firms, and higher rates of fraud in the public Medicare system. The last of these is directly related to lax utilization controls. A fuller analysis of the costs and benefits of public versus private health insurance is needed.

The relative merits of public enterprise have a number of policy implications. The first concerns the appropriate size of Medicare Advantage, which operates through publicly provided premium subsidies to private HMOs. Medicare Advantage plans are not directly comparable to private payers, because they compete on quality rather than price, as long as there is no price competition through competitive bidding for plan members. Thus differences in incentives for utilization control operate through the need to enhance quality, subject to available premium resources, or result from residual claims on profits. Future research needs to investigate more carefully the differences and similarities in cost-control measures from this type of coverage and their effects on regional variations and efficiency.

The second implication regards the timely issue of comparative effectiveness research (CER), which has been offered as a means of raising health care quality and reducing costs. The rationale for CER is to generate better evidence, and to disseminate it to patients, payers, and providers, about what works and does not work in health care. Indeed, a common motivation for the use of CER is to reduce cost inefficiencies due to regional differences in care. Awareness of CER has been heightened recently by its significant public subsidization through the American Recovery and Reinvestment Act of 2009.[15] An overriding question raised by our analysis is whether regional variation in care occurs because of a lack of information or a lack of incentives for utilization control in the public sector.

15. The explicit use of comparative effectiveness assessments is much more common outside the United States, particularly in the European Union. However, this is a relatively recent trend: no European countries formally required economic assessments for pricing and reimbursement decisions as of 1993, but a majority had such a policy either in place or in development by 1999 (Drummond and others 1993, 1999).

Health economists have not yet paid sufficient attention to the differences in incentives across the public and the private sectors or to the corresponding implications for health care variation. The regional variations documented in the Dartmouth Atlas of Health Care have led several prominent researchers to conclude that high-use regions ought to model themselves after their low-use peers (Fisher, Bynum, and Skinner 2009). Our study suggests the importance of research focusing on another, different question: whether or not public sector health insurers ought to model themselves after their peers in the private sector.

APPENDIX

Sampling Error in Estimation of Regional Variation

In the text we outlined our approach for obtaining a consistent estimate of regional variation, defined as

$$RV \equiv \frac{1}{R}\sum_{r=1}^{R}\left(\mu_r - \bar{\mu}\right)^2.$$

In this appendix we show how we solve for the bias in the sample analogue, $\widehat{RV} \equiv \frac{1}{R}\sum_{r=1}^{R}\left(\hat{\mu}_r - \bar{\hat{\mu}}\right)^2$, and estimate it consistently using the variance-covariance matrix of the estimates. Recall the definitions from the text: μ_r is the true population fixed-effect parameter for region r, $\hat{\mu}_r$ is the corresponding sample estimate, and z_r is a mean-zero sampling error with covariance across regions $E(z_r z_s) = \sigma_{rs}$. The sample estimate is the true value plus sampling error,

$$\hat{\mu}_r = \mu_r + z_r.$$

Define $\bar{\mu} \equiv \frac{1}{R}\sum_{r=1}^{R}\mu_r$, the mean regional fixed effect across regions; $\bar{\hat{\mu}} \equiv \frac{1}{R}\sum_{r=1}^{R}\hat{\mu}_r$, its sample analogue; and $\bar{z} \equiv \frac{1}{R}\sum_{r=1}^{R} z_r$, the average sampling error across regions.

Using the definitions above, we can write

$$E\left(\widehat{RV}\right) \equiv \frac{1}{R}\sum_{r=1}^{R}E\left(\hat{\mu}_r - \bar{\hat{\mu}}\right)^2 = \frac{1}{R}\sum_{r=1}^{R}E\left(\mu_r + z_r - \bar{\mu} - \bar{z}\right)^2,$$

where we rely on the fact that we are dealing with regional fixed effects, rather than random effects, to move the expectations operator inside the summation. Expanding the right-hand side of the expression results in

$$E\left(\widehat{RV}\right) = \frac{1}{R}\sum_{r=1}^{R}\left[E\left(\mu_r - \bar{\mu}\right)^2 - 2E\left(\mu_r - \bar{\mu}\right)\left(z_r - \bar{z}\right) + E\left(z_r - \bar{z}\right)^2\right].$$

Since μ_r and $\bar{\mu}$ are both scalars, this simplifies to

$$E\left(\widehat{RV}\right) = \frac{1}{R}\sum_{r=1}^{R}\left[\left(\mu_r - \bar{\mu}\right)^2 - 2\left(\mu_r - \bar{\mu}\right)E\left(z_r - \bar{z}\right) + E\left(z_r - \bar{z}\right)^2\right].$$

The distributional assumptions on z imply that $E(z_r) = E(\bar{z}) = 0$. Therefore, we can write

$$E\left(\widehat{RV}\right) = RV + \frac{1}{R}\sum_{r=1}^{R}E\left(z_r - \bar{z}\right)^2.$$

To characterize the bias, note that

$$E\left(z_r - \bar{z}\right)^2 = E\left(z_r^2\right) - 2E\left(z_r\frac{1}{R}\sum_{s=1}^{R}z_s\right) + E\left(\frac{1}{R}\sum_{s=1}^{R}z_s\right)^2,$$

which we can write in terms of the variance and covariance parameters as

$$E\left(z_r - \bar{z}\right)^2 = \sigma_{rr} - \frac{2}{R}\sum_{s=1}^{R}\sigma_{rs} + \frac{1}{R^2}\sum_{s=1}^{R}\sum_{t=1}^{R}\sigma_{st}.$$

The bias due to sampling variance is equal to the above expression, averaged across all regions. A consistent estimate of the bias can be calculated by summing up and taking the appropriate averages of estimated variances of and covariances between the estimated regional fixed effects. The more precisely the regional fixed effects are estimated, the smaller is the bias correction.

ACKNOWLEDGMENTS We thank the editors, David Cutler, Mark McClellan, and seminar participants at the University of Chicago, RAND, and the Brookings Institution for comments. Financial support from the George Stigler Center for the Economy and the State at the University of Chicago, the Leonard Schaeffer Center for Health Economics and Policy at the University of Southern California, and Pfizer Inc. is gratefully acknowledged.

References

Alchian, Armen A., and Harold Demsetz. 1972. "Production, Information Costs, and Economic Organization." *American Economic Review* 62, no. 5: 777–95.

Baicker, Katherine, Elliott S. Fisher, and Amitabh Chandra. 2007. "Malpractice Liability and the Practice of Medicine in the Medicare Program." *Health Affairs* 26, no. 3: 841–52.

Chandra, Amitabh, and Douglas O. Staiger. 2007. "Productivity Spillovers in Health Care: Evidence from the Treatment of Heart Attacks." *Journal of Political Economy* 115, no. 1: 103–40.

Chernew, Michael E., and others. 2010. "Geographic Correlation between Large-Firm Commercial Spending and Medicare Spending." *American Journal of Managed Care* 16, no. 2: 131–38.

Christensen, Sandra, and Judith Wagner. 2000. "The Costs of a Medicare Prescription Drug Benefit." *Health Affairs (Millwood)* 19, no. 2: 212–18.

Cutler, David M., and Louise Sheiner. 1999. "The Geography of Medicare." *American Economic Review* 89, no. 2: 228–233.

De Alessi, Louis. 1974a. "An Economic Analysis of Government Ownership and Regulation: Theory and the Evidence from the Electric Power Industry." *Public Choice* 19: 1–42.

———. 1974b. "Managerial Tenure under Private and Government Ownership in the Electric Power Industry." *Journal of Political Economy* 82, no. 3: 645–53.

Drummond, Michael, and others. 1993. "Economic Evaluation of Pharmaceuticals: A European Perspective." *Pharmacoeconomics* 4, no. 3: 173–86.

———. 1999. "Current Trends in the Use of Pharmacoeconomics and Outcomes Research in Europe." *Value in Health* 2, no. 5: 323–32.

Eppig, Franklin J., and George S. Chulis. 1997. "Matching MCBS (Medicare Current Beneficiary Survey) and Medicare Data: The Best of Both Worlds." *Health Care Financing Review* 18, no. 3: 211–29.

Fisher, Elliott S., Julie P. Bynum, and Jonthan S. Skinner. 2009. "Slowing the Growth of Health Care Costs—Lessons from Regional Variation." *New England Journal of Medicine* 360, no. 9: 849–52.

Fisher, E. S., and others. 2003a. "The Implications of Regional Variations in Medicare Spending. Part 1: The Content, Quality, and Accessibility of Care." *Annals of Internal Medicine* 138, no. 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, no. 4: 288–99.

Fowler, Floyd J., and others. 2008. "Relationship between Regional per Capita Medicare Expenditures and Patient Perceptions of Quality of Care." *JAMA* 299, no. 20: 2406–12.

Gottlieb, Daniel J., and others. 2010. "Prices Don't Drive Regional Medicare Spending Variations." *Health Affairs* 29, no. 3: 537–43.

Kessler, Daniel, and Mark McClellan. 1996. "Do Doctors Practice Defensive Medicine?" *Quarterly Journal of Economics* 111, no. 2: 353–90.

———. 2002a. "Malpractice Law and Health Care Reform: Optimal Liability Policy in an Era of Managed Care." *Journal of Public Economics* 84, no. 2: 175–97.

———. 2002b. "How Liability Law Affects Medical Productivity." *Journal of Health Economics* 21, no. 6: 931–55.

Medicare Payment Advisory Commission. 2009. "Measuring Regional Variation in Service Use." Washington.

Morgan, Robert O., and others. 1997. "The Medicare-HMO Revolving Door—The Healthy Go In and the Sick Go Out." *New England Journal of Medicine* 337, no. 3: 169–75.

Phelps, Charles E. 2000. "Information Diffusion and Best Practice Adoption." In *Handbook of Health Economics,* vol. 1A, edited by Anthony J. Culyer and Joseph P. Newhouse. Amsterdam: Elsevier Science.

Rettenmaier, Andrew J., and Thomas R. Saving. 2009. "Perspectives on the Geographic Variation in Health Care Spending." Special Publication. Washington: National Center for Policy Analysis (June).

Sampford, M. R. 1953. "Some Inequalities on Mill's Ratio and Related Functions." *Annals of Mathematical Statistics* 24, no. 1: 130–32.

Sirovich, Brenda E., and others. 2006. "Regional Variations in Health Care Intensity and Physician Perceptions of Quality of Care." *Annals of Internal Medicine* 144, no. 9: 641–49.

Spann, Robert M. 1977. "The Macroeconomics of Unbalanced Growth and the Expanding Public Sector: Some Simple Tests of a Model of Government Growth." *Journal of Public Economics* 8, no. 3: 397–404.

Sutherland, Jason M., Elliott S. Fisher, and Jonathan S. Skinner. 2009. "Getting Past Denial—The High Cost of Health Care in the United States." *New England Journal of Medicine* 361, no. 13: 1227–30.

Wennberg, John, and Alan Gittelsohn. 1973. "Small Area Variations in Health Care Delivery." *Science* 182, no. 4117: 1102–08.

Wennberg, John E., and Megan McAndrew Cooper, eds. 1998. *Dartmouth Atlas of Health Care, 1998.* Chicago: Dartmouth Medical School and American Hospital Association.

Wennberg, John E., and others. 2008. *Tracking the Care of Patients with Severe Chronic Illness: The Dartmouth Atlas of Health Care 2008.* Lebanon, N.H.: Dartmouth Institute for Health Policy and Clinical Practice.