

The Association between Medical Care Utilization and Health Outcomes: A Spatial Analysis

Francesco Moscone
Brunel University London
Francesco.Moscone@brunel.ac.uk

Jonathan Skinner
Dartmouth College, Hanover NH 03755
jon.skinner@dartmouth.edu

Elisa Tosetti
European Commission
Elisa.Tosetti@ec.europa.eu

Laura Yasaitis
Dartmouth College
yasaitis@gmail.com

03 April 2019

Abstract

As health care spending continues to strain government and household budgets, there is increasing interest in measuring whether the incremental dollar spent on health care is worth it. In studying this question, researchers often make two key assumptions: that health care intensity can be summarized by a single index such as average spending, and that samples of hospitals or regions are spatially independent: Manhattan and the Bronx are no more alike than are Manhattan and San Diego, for example. In this paper we relax both assumptions. Using detailed data on 897,008 elderly Medicare enrollees with acute myocardial infarction (or a heart attack) during 2007-11, we find first that the total level of health care spending has little impact on health outcomes; more important is *how* the money is spent. Same-day stenting, a treatment with proven effectiveness, positively predicts survival, while home health care spending does not. Second, there is strong evidence of spatial autocorrelation; without corrections this can lead to inefficient estimates and standard errors that are biased downward. Spatial autocorrelation in outcomes appears to be the consequence both of unmeasured health status and spatial correlation in new and effective technology.

We are grateful for comments from 2 anonymous referees, the editor, Amitabh Chandra, and conference participants at the First Spatial Health Economics Workshop, Girton College, University of Cambridge. Financial support from the National Institute on Aging (PO1-AG19783 and U01-AG046830) is gratefully acknowledged.

I. Introduction

With health care spending in all OECD countries projected to rise over the next several decades (OECD, 2018), there is a growing interest in understanding whether spending more on health care yields health benefits, or the converse – will cutbacks in health spending have an adverse impact on health? Many studies use a single measure of intensity – such as end-of-life spending – to test the hypothesis that more health care spending yields better health outcomes. The results vary widely, depending on the cohort considered, the method of risk-adjustment, and the type of disease.¹ As well, nearly all studies using regional variation in spending and outcomes make statistical inferences under the assumption of spatial independence; that the unexplained residuals in San Francisco and San Jose (two adjacent regions) are no more likely to be correlated than those between San Francisco and Miami, Florida, 2585 miles distant.²

Yet there is increasing evidence that behavioural health and economic factors play an important role in explaining spatial patterns of back surgery, quality of life, health utilization, and physician practice patterns (e.g., Joines et al., 2003; Eibich and Ziebarth, 2014; Bhattacharjee et al., 2014; Chaix et al., 2005; Sriubaite, 2018). Assuming independence in the presence of spatial autocorrelation can lead both to inefficient estimators and standard errors that are biased downward, leading to falsely rejecting the null hypothesis (Anselin, 1988).

¹ For example, see Hussey et al., 2013, Fisher et al., 2003a,b, Skinner, et al. 2005, Skinner and Staiger, 2015, Chandra and Staiger, 2007, Hadley, et al., 2011, Romley, et al., 2011, Silber, et al., 2010, Doyle, 2011, Doyle, et al., 2015, 2017, Yasaitis et al., 2014, and Kibria et al., 2013.

² An exception is Rickett and Holmes (2007), discussed below. By implicitly comparing outcomes in adjacent hospitals, the Doyle et al. 2015, 2017 studies sidestep these issues; also see Gravelle et al., 2014, for evidence on spatial correlation of quality for local hospitals.

In this study, we relax both the assumption of a single “intensity” measure for health care, and the assumption of independence across regions. We first specify a model that allows for multiple types of health care inputs, and for the presence of spatial autocorrelation across regions. We then draw on data from 897,008 Medicare fee-for-service patients age 65 or older hospitalized for acute myocardial infarction (AMI) or heart attack, during 2007-11, aggregated by Hospital Referral Region (HRR) and year, and analyse these data using a general econometrics framework that allows for correlation both in the dependent variable and the error term (Anselin, 1988; Moscone and Tosetti, 2014).³

We first demonstrate that even when using highly risk-adjusted AMI survival rates, there remains considerable spatial autocorrelation across HRRs in the U.S. After adjusting for spatial autocorrelation, we find that the marginal productivity of different health inputs are quite different; indeed for the specific inputs we consider, they are of opposite sign; what matters most for health outcomes is *how* the money is spent. Like Likosky et al. (2018), we find that nearly all the survival benefits for this cohort arose from primary or same-day stenting for AMI patients, an inpatient treatment with proven effectiveness and minimal incremental costs. Raising the fraction of patients receiving such treatments by 10 percentage points is associated with a 0.7 percentage point increase in survival. And like Doyle et al. (2015, 2017), we find that the intensity of acute care is more important for health outcomes than is post-acute home health care, which in our model provides negative benefits; a doubling of home health care spending is predicted to reduce survival by 0.5 percentage points (also see McKnight, 2006).

What are the sources of this spatial autocorrelation? We find that including at least two measures associated with health outcomes – regional income and smoking rates --

³ While the likelihood of an AMI in a given year is less than 2% for these Medicare enrollees, the lifetime risk of a cardiovascular event is roughly one-third (Berry et al., 2012).

reduces the degree of autocorrelation, which suggests the importance of unmeasured health even for risk-adjusted AMI patients.⁴ Another factor that can explain spatial autocorrelation is that the diffusion of new and effective technologies such as same-day PCI are themselves spatially correlated; we find that including these measures directly in a regression attenuates the degree of measured spatial autocorrelation by as much as one-half. In sum, empirical studies relying on regional patterns of treatments and outcomes should consider addressing the problem of spatial autocorrelation as well as capturing more granular health inputs that go beyond a single measure of overall expenditures.

II. A Model of Health Expenditures and Health Outcomes

Numerous studies have tried to understand the value of money spent on health care, motivated by the question of whether more spending “buys” better health outcomes. Several studies examining the association between overall population health and total spending found a null or negative association (Fisher et al., 2003a,b; Skinner, et al. 2005; Yasaitis et al., 2009; Doyle et al., 2017, Kibria et al., 2013), while others demonstrated a positive association (Hadley, et al., 2011; Romley, et al., 2011; Silber, et al., 2010, Doyle, 2011; Doyle et al., 2015). Yet, total spending, even for specific health events, may not be the best measure of the care provided to patients. A dollar spent on highly effective care may buy more health than a dollar spent on another, less effective, type of care.

We therefore hypothesize that the marginal value of the incremental dollar of health care spending depends critically on how the money is spent (Chandra and Skinner, 2012; Chandra et al., 2019). For example, percutaneous coronary interventions (PCI), or stents, are

⁴ Since these are risk-adjusted measures of survival conditional on having a heart attack, researchers have often assumed that these are reliable measures of hospital or regional productivity; see Chandra et al. (2016). And as noted below, since the risk-adjusted measures of survival already adjust for ZIP code level income, HRR-level income may also proxy for supply-side effects.

a good example of this type of treatment, but where treatment within 12 hours of the AMI is far more beneficial than PCI performed at a later stage (Hochman et al., 2006; Weintraub et al., 2008). To capture these effects, we include variables capturing the likelihood of AMI patients to receive same-day PCI. As well, we also hypothesize that quality of care in academic medical centers is better than in non-academic hospitals (Burke et al., 2017). By contrast, other treatments are those with little documented benefit for patients. We hypothesize that spending on home health care services represent care in this category, as there is little evidence that these excess dollars are purchasing better health outcomes (McKnight, 2006) and may even lead to harm (Doyle et al., 2017).

Why might such differences in treatment patterns arise? While the objective of this study is less about the exogenous sources of treatment patterns, we can draw on other evidence that suggests physician beliefs or capacity for PCI, rather than the underlying illness in regions or patient demand, plays a central role to explaining regional variations. For example, Cutler et al. (2019) have shown that physician beliefs about the efficacy of a variety of cardiovascular procedures – as measured by a national survey including detailed patient vignettes along with options for treatment – explained as much as 60 percent of variations in end-of-life regional spending, with only a modest role for patient preferences (see also Baker et al., 2014). Molitor (2017) similarly found that when cardiologists move, they tend to adjust to the prevailing norms for PCI in their new institutions.

A real concern with observational data is the problem of reverse causation; that regions or hospitals with a higher level of unobserved illness will lead to higher levels of spending, thus biasing conventional regression estimates against finding positive effects of spending. Restricting attention to AMI patients is one approach to reducing the potential for endogeneity, since regions with worse health may have more people who have a myocardial infarction, but conditional on having an AMI of a specific type, and with specific comorbidities, there's much

less chance of unobservable factors unrelated to the health system biasing the estimates. Other approaches to avoiding endogeneity include using tourists with acute conditions (Doyle, 2011) and ambulance services loyal to specific hospitals (Doyle et al., 2015, 2017).⁵

We used a dataset of 897,008 heart attack (AMI, or acute myocardial infarction) patients during 2007-11 with follow-up through 2012. We follow previous research (such as Chandra and Staiger, 2007, Chandra et al., 2016, Cutler et al., 2019, Doyle, 2011, and Cutler et al., 1998) in using heart attack patients because any patient with an AMI is taken immediately to a nearby hospital, thus minimizing endogenous hospital or regional selection.

There are 306 HRRs across the United States, and while some of the spatial boundaries cross state lines, we place the HRR “city” in its corresponding state. For each beneficiary in this cohort, demographic data included age, sex, and race or ethnicity. The median household income in a beneficiary’s ZIP code of residence was used as a proxy for his or her income. We also recorded diagnoses present on the beneficiaries’ claims from their inpatient admission, as well as creating an HCC risk adjustment measure based on Medicare claims 6 months prior to the AMI.

An AMI is based on the first diagnosis code (410.x1 or 410.x2), and not on the diagnostic related group (DRG), which can often vary depending on how the patient is subsequently treated. Risk adjustment measures at the individual level include median ZIP code income (averaged across AMI patients), age (65-69, 70-74, 75-79, 80-84, 85+), fully interacted with sex, vascular disease, pulmonary disease, dementia, diabetes, liver failure, renal failure, cancer, plegia (stroke), rheumatologic disease, HIV, race (African-American, Hispanic and other) and location of the AMI: Anterolateral, inferolateral, inferoposterior, all other

⁵ That is, there is generally more than one ambulance sent out for a medical emergency, yet some ambulances are loyal to specific hospitals; this creates a natural randomization as to which hospital the patient is admitted. Regarding the use of tourists as being unrelated to “place,” Chandra et al. (2019) used an AMI cohort similar to this one and did not find differences in estimates between samples of tourists and non-tourists, although the tourist sample estimates were much less precise.

inferior, true posterior walls, or subendocardial, other site, or not otherwise specified. In most cases, we also use Hierarchical Condition Categories (HCC), which, while leading to over-adjustment and potential biases (Song et al., 2010; Wennberg et al., 2014; Finkelstein et al., 2017), counts the number of different diagnoses that patients have received in the 6 months prior to the index admission, and weights them for severity. We created HRR-level variables by aggregating individual-level regression models with HRR-year fixed effects for risk-adjusted survival and spending measures from the beneficiaries living in those HRRs.⁶ Initially, we estimate models without HCCs, but then adopt them as the default risk adjustment. Descriptive statistics are provided in Table 1.

As noted above, treatment variables were created from the individual-level Medicare claims data for the AMI patients, and aggregated up to the year-HRR level. Primary or early PCI was defined as angioplasty or stenting within one day of admission to the hospital, while home health care expenditures was average per capita spending for home health care (over the entire population of Medicare enrollees, even those who did not receive any home health care), by year and HRR. Our data are limited to the fee-for-service population, since we cannot measure utilization among those in managed care plans (Medicare Advantage).

To adjust further for regional health risk, we created year-specific estimates of HRR-level smoking rates by combining county-level small area estimates of smoking derived from the Behavioural Risk Factor Surveillance System, or BRFSS (Dwyer-Lindgren, 2014). Finally, we also included average HRR-level income for the AMI patients, aggregated up from the patient-specific ZIP code level income. Note that in our risk adjustment, we have already adjusted for the patient's ZIP code, so this aggregate zip code could capture an agglomeration

⁶ Recall that when one regresses both the dependent variable Y , and a specific independent variable X , against a vector of risk adjusters Z (as we do), the coefficient from the bivariate coefficient of the regression of Y on X (as reported in the first column of Table 2) is equivalent to the corresponding coefficient on X for a fully risk-adjusted regression of Y on X and Z . However, variables such as smoking and PCI rates are not risk-adjusted.

effect if (for example) higher-quality physicians locate in regions with higher incomes and more amenities.

III. Estimating Spatial Models

A wide literature points to the existence of geographical concentration in population health and health care services (Rushton, 2003; Lorant, 2001; James et al., 2004; www.dartmouthatlas.org). Yet nearly all studies assume a zero correlation with regard to shocks affecting nearby regions, whether hospitals, states, or HRRs. One source of dependence, spatial correlation, is related to the location and distance among statistical units, with respect to the geographical, economic or social space in which they are embedded. Neighbouring units may share common general population characteristics or underlying socio-economic features that may affect health outcomes. For example, environmental stressors such as air pollution could be linked to regional rather than local factors, influencing prevalence and health needs across a wide area. Similarly, diet and health behaviors not already measured in our analysis could vary across broad regions of the United States.

An alternative hypothesis is that unmeasured factors related to treatments are correlated across regions. One simple example is if there is spatial correlation in the use of a highly effective treatment such as same-day PCI, a pattern we find in the data. In the absence of measuring PCI rates directly, the dependent variable would exhibit spatial dependence because of the unmeasured spatial correlation in the X variable; if this hypothesis were true, including PCI rates (or other measures such as academic medical centers) directly would then be expected to attenuate measured spatial autocorrelation in the dependent variable.

Another potential explanation would be that the spatial diffusion of new knowledge and practice patterns not measured in the data would cause nearby providers to deliver similar types

of care.⁷ For example, if physicians or hospital administrators migrate to, or interact with, health centers in nearby areas, their skills conditional on inputs (e.g., they trained at either a high-quality or low-quality residency, and move nearby), could lead to spatial autocorrelation; indeed our measure of HRR-level income could capture the fact that skilled physicians are attracted to higher-income regions, whether because of higher relative wages or the presence of amenities in such regions.⁸

For a variety of reasons then, spatial spillovers are likely to induce correlations across regions in health care measures, whether of inputs or outputs. When data contain cross section dependence, conventional estimators such as ordinary least squares are inefficient, and the estimated standard errors are biased. In an earlier study, Ricketts and Holmes (2007) allow for spatial autocorrelation, and importantly, heterogeneity across regions in the association between physician supply and mortality.⁹ While we do not allow for heterogeneity, we do account for potential spatial autocorrelation both in the dependent variable and in the error term by estimating a spatial autoregressive model with spatially correlated errors (SAR-SEM). For the spatial weights matrix, we use contiguity information, and assigned weights $w_{ij}=1$ when HRR i and j are contiguous according to the queen contiguity criterion and 0 otherwise. These spatial weights were also used to compute a set of diagnostics.

In some regressions we have included state dummies to account for possible heterogeneity in health outcomes across US states. Such heterogeneity may be due, for

⁷ Spatial correlation might also be generated by cross-state border migration of health services beneficiaries, although our focus here on heart attack patients – emergencies in which ambulances take patients quickly to a nearby hospital.

⁸ Recall that we control for ZIP-code-level income in our risk-adjusted patient mortality level, so we suspect that income in this case could capture other supply-side factors.

⁹ The spatial heterogeneity in the association between physician supply and mortality may also reflect differences across regions in the importance of physician location; in areas where there is insurance coverage and generosity, physicians may tend to locate in healthier regions, rather than in those with the greatest demand for their services.

example, to the common reaction of units belonging to the same state to external forces and unanticipated events such as technological advances, health shocks, the implementation of new health policies and sociological structural changes. Finally, as sensitivity analysis, we also consider the model using county level. This requires that we drop many of the smaller counties where sample sizes are insufficient for analysis.¹⁰

IV. Results

Figure 1 shows a map of risk-adjusted survival, while Figure 2 displays a map of risk-adjusted expenditures. In both cases, there is considerable spatial correlation; for survival, the spatial error coefficient is 0.39 and highly significant. Figure 3 shows the association between one-year Medicare expenditures (on the horizontal axis) and one-year survival (on the vertical axis), both adjusted using HCCs. There is considerable variance across regions of the U.S., both with regard to risk-adjusted mortality, and risk-adjusted expenditures.

In Table 2 we report results of conventional OLS and spatial autoregressive models (with spatially lagged dependent variable and spatially lagged error term) of one-year survival. Because of concerns about risk-adjustment noted above (e.g., Song et al., 2010; Finkelstein et al., 2017), the upper panel of the table the dependent variable is a non-HCC measure of survival, while in the lower panel of the table, the dependent variable is a HCC-adjusted measure of survival (these models all use HCC-adjusted spending measures), which we use in subsequent regression analysis. We also report the Lagrange Multiplier (LM) tests for error dependence and for a missing spatially lagged dependent variable (see Bera et al. 1996 for details).

¹⁰ The Centers for Medicare and Medicaid research requires that all reported “cells” have counts of at least 11 individuals. For this reason, we drop counties with fewer than 11 observations in any of the cells, with the critical measure typically being the number of patients with same-day PCI.

Looking at regressions (I) in the upper panel of Column 1 in Table 2, the variable health expenditure is statistically significant with a negative sign (-0.04), suggesting that a 10 percent increase in spending leads to a 0.4 percentage point decline in one-year survival. The corresponding regression in Column 4 that includes spatial autocorrelation adjustment exhibits a coefficient of -0.024, suggesting that the lack of adjustment leads to an upward bias (in absolute terms) of roughly two-thirds, but only a slightly elevated standard error. Including additional covariates (Columns 3 and 6) imply that the coefficient on overall spending is essentially zero in both cases, with a similar standard error.

We next turn to the lower panel of Table 2 that includes the HCC risk-adjustment in creating HRR-level spending and survival measures for AMI patients. Focusing on the fully specified model (Columns 3 and 6) suggest that the spatial adjustment essentially erases a significance “star” from the estimate; while the model implies a positive coefficient on spending (conditional on PCI and home health care spending) of 0.018 (implying a 10 percentage point increase in overall spending should increase survival by 0.18 percentage points), the coefficient is smaller in magnitude (0.014) and insignificant at conventional levels after adjustment for spatial autocorrelation (Column 6).

The coefficients on income and smoking are not as sensitive to adjustment for spatial autocorrelation; in both cases they exhibit coefficients with the expected magnitude, and in the case of smoking, is associated with large and significant reductions in survival even after adjusting for other covariates. While the magnitude of the income coefficient is not affected by spatial autocorrelation, the coefficient on smoking is nearly one-third lower (again, in absolute terms) with adjustment for spatial autocorrelation; the results imply that a 10 percentage point increase in smoking rates would reduce survival by 0.6 percentage points.

Finally, we consider hospital input measures such as primary PCI, home health care, and the fraction of patients treated in a teaching hospital. While teaching hospital status might

appear to provide significant health benefits in the absence of adjustments for spatial autocorrelation, once the adjustment is made, the coefficient shrinks in magnitude and is no longer significant. However, both primary PCI and home health care spending are highly significant, although – like the findings in Doyle et al. (2015, 2017), the quality of acute care (e.g., whether a PCI is provided quickly) is associated with higher survival (a 10 percent increase in PCI implies a 0.67 percentage point greater likelihood of survival), while the magnitude of post-acute home health care is associated with a decline in survival.

As well, the introduction of these “supply-side” factors reduces the degree of estimated spatial autocorrelation from .32 (Column 5 in Table 2) to .18 (Column 6), suggesting that spatial correlation in PCI rates – which are consistently higher in mountain and Midwest regions, and lower in the South – can explain nearly half of the observed spatial autocorrelation even after adjusting for health measures. Figure 4 provides a map of the adjusted errors, which do not appear to exhibit spatial autocorrelation.

We also include for sensitivity analysis the regression analysis with an inverse distance weighted approach in Table 3; the coefficient estimates are similar, although now smoking is no longer significant. What is different is the estimate of spatial autocorrelation, which appear to be larger when inverse distance weights are used compared to simple contiguity. This seems to suggest that absolute distance between HRRs is more important than whether they share a border; Boston and New York (for example) may share practice styles and patient characteristics, even if they are not adjacent.¹¹

Our LM tests point to a spatial error model over a spatial lag model as most suitable for describing the process underlying our data. A spatial error model is consistent with the hypothesis that spatial correlation arises from the geographical concentration of unobservables,

¹¹ An alternative approach is to create “neighbors” not spatially, but with regard to similarities in economic and demographic characteristics, as in Case et al. (1993).

such as environmental risks that are difficult to measure like air pollution, or unmeasured life style factors, which are known to be spatially concentrated (Baltagi et al. 2018). Another potential explanation for spatial error correlation is, as noted above, technology or productivity spillover arising from the local diffusion of certain technological standards, due, for example, to professional interaction. In the light of results pointed by LM tests, we have run a SEM for each regression, but we have not reported the results since they do not add further information when compared to our spatial regressions. Similarly, we have not reported the estimation of the Durbin spatial model: the spatially lagged regressors turn out to be statistically insignificant.

In Table 4 we introduce state fixed effects, so that all estimates are “within state” by HRR. (Recall that there are 306 HRRs, so some of the larger states, like Texas, have more than 20 HRRs.) The use of state-level effects effectively knocks out spatial autocorrelation – perhaps not surprisingly – and the coefficient estimates are largely consistent with Columns 3-6 in Table 2, except that here home health care spending is positively associated with survival. State fixed effects are therefore a somewhat blunt instrument to adjust for spatial autocorrelation, because state borders are somewhat arbitrary from the viewpoint of (e.g.) smoking rates; large states have many HRRs, while smaller states in New England contribute little to the regression estimates because there is so little within-state variation.

To check the robustness of our results the same empirical analysis has been conducted at the county level. In Table 5 we report results of conventional OLS and spatial autoregressive models (with spatially lagged dependent variable and spatially lagged error term) of one-year survival. Results from both the conventional OLS and spatial model, confirm that overall spending does not appear to have a strong association with survival when controlling for variables such as teaching hospital, PCI within 1 day, and average home health spending. Although these results are likely to be weaker than for HRRs because of missing counties, there

is evidence of spatial autocorrelation in the error term, though this effect does not seem to impact on the standard errors of the estimated coefficients.

V. Discussion

Policy-makers are increasingly concerned with whether continued growth in health care spending is delivering a reasonable return in terms of improved health outcomes. Economic studies have found mixed results, with some studies suggesting large health gains to more spending, while others suggest the opposite. In this paper, we have relaxed two common assumptions typically made in these studies, that (1) health care inputs can be summarized by a single index of “intensity,” and (2) that in studies using regional variation in utilization and outcomes, each region or hospital is an independent draw. We reject both assumptions. First, we find that spending alone is a poor predictor of survival for our sample of heart attack (AMI) patients. What appears to be most important is how the money is spent; hospitals providing higher rates of same-day stenting (PCI), for example, exhibit substantially higher rates of one-year survival, while hospitals whose patients receive large quantities of post-acute home health care generally do worse. One could be concerned here of reverse causality; that patients who are sicker require more home health care, but Doyle et al. (2017) have found much the same result even with the strong natural randomization of their ambulance assignment approach.

We also found strong evidence of spatial correlation in the residuals, even for AMI survival rates, which have often been used as measures of hospital productivity (e.g., Chandra et al., 2016; Skinner and Staiger, 2015). Using our time-series cross-section data structure, we estimated our model with several types of adjustments for spatial autocorrelation; some (but not all) coefficients shrank in magnitudes; these typically lost a “star” in statistical significance. As well, the use of state-level fixed effects attenuated spatial autocorrelation, although this approach also throws out a considerable amount of statistical power.

We also found that including regional measures of health behaviors, income, and most importantly, health inputs such as PCI reduce the degree of spatial autocorrelation. This latter characteristic is consistent with models of diffusion through professional contacts (Coleman, Katz and Menzel, 1966); Papageorgiou et al. (2007) for example found interdependence in the adoption of medical technology that arises when one country strategically mimics neighbouring health policies, for example by adopting the same vaccine to prevent the diffusion of a contagious disease (also see Birke, 2009 and Moen et al., 2016).

There are several limitations to the study. First, our analysis is performed at the HRR or county level; a substantially more complicated model might consider spatial effects at the hospital or even patient level. Second, we acknowledge the potential for reverse correlation; that patients who are sicker on an unobservable basis may require more spending, thus leading to a positive bias on the association between spending and health outcomes. However, given that our results are consistent for the fully-specified model whether one includes HCCs or excludes them is reassuring that the input measures (e.g., PCI) are not directly associated with unobservable health status (Chandra et al., 2019). Finally, our measures are specific only to Medicare fee-for-service patients; it could be that the association between medical inputs and survival differ for the under-65 population, or for those in a managed care plan. However, there is a close association between the under-65 and over-65 population for specific treatments such as joint surgery (Cooper et al., 2018) and end-of-life care (Baker et al., 2008).

A final more general question is: How generalizable are these results to other studies using regional variation outside of health care? For example, to what extent are regional studies that compare state-level policies biased by the general problem of spatial autocorrelation, as in Betz et al. (2019) who question the use of instrumental variables in the presence of spatial autocorrelation? Adjusting for autocorrelation in our application to AMI patients does not change the basic implications of our model that the cross-sectional association between

spending and survival is very small and difficult to sign especially when conditioning on key inputs with proven effectiveness (or key inputs that have been shown to be ineffective). We don't yet understand entirely why some regions are so much more effective in adopting new and effective treatments, (Skinner and Staiger, 2007), but a careful modelling of spatial factors affecting health care and health status could provide additional insights into the question of whether we're getting our "money's worth" in health care.

REFERENCES

- Anselin, L. 1988. *Spatial econometrics: methods and models*. (Dordrecht: Kluwer).
- Baker, L.C., Fisher, E.S. and Wennberg, J.E. 2008. Variations in hospital resource use for Medicare and privately insured populations in California. *Health Affairs*. 27, w123-w134.
- Baker, L.C., Bundorf, M.K. and Kessler, D.P. 2014. Patients' preferences explain a small but significant share of regional variation in Medicare spending. *Health Affairs*, 33, 957-963.
- Baltagi, B.H., Moscone, F., and Santos, R. 2018. *Spatial health econometrics*. In Baltagi B.H. and Moscone, F. (eds) *Health econometrics, contributions to economic analysis*. Emerald Publishing. b
- Bera, A. K., Florax, R. and Yoon, M. J. 1996. Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics*. 26, 77-104.
- Berry, J.D., Dyer, A., Cai, X., Garside, D.B., Ning, H., Thomas, A., Greenland, P., Van Horn, L., Tracy, R.P. and Lloyd-Jones, D.M. 2012. Lifetime risks of cardiovascular disease. *New England Journal of Medicine*. 366, 321-329.
- Betz, T., Cook, S.J., and Hollenbach, F.M. 2019. "Spatial interdependence and instrumental variable models." *Political Science Research and Methods*, 1-16.
- Bhattacharjee, A., Maiti, T., Petrie, D. 2014. General equilibrium effects of spatial structure: Health outcomes and health behaviours in Scotland. *Regional Science and Urban Economics*, 49, pp.286-297.
- Birke D. 2009. The economics of networks: A survey of the empirical literature. *Journal of Economic Surveys*, 23, 762-793.
- Burke, L.G., Frakt, A.B., Khullar, D., Orav, E.J. and Jha, A.K. 2017. Association between teaching status and mortality in US hospitals. *JAMA*. 317, 2105-2113.
- Case, Anne C., Harvey S. Rosen and James R. Hines, Jr., 1993. Budget spillovers and fiscal policy interdependence," *Journal of Public Economics*. 52, 285-307.
- Chaix, B., Merlo, J., Chauvin, P. 2005. Comparison of a spatial approach with the multilevel approach for investigating place effects on health: the example of healthcare utilisation in France. *Journal of Epidemiology & Community Health*. 59, 517-526.
- Chandra, A., Finkelstein, A., Sacarny, A., Syverson, C. 2016. Health care exceptionalism? Performance and allocation in the US health care sector. *American Economic Review*. 106, 2110-2144.
- Chandra, A. and Staiger, D.O. 2007. Productivity spillovers in health care: Evidence from the treatment of heart attacks. *Journal of Political Economy*. 115, 103-140.
- Chandra, A., Colla, C., Skinner, J. 2019. The elusive link between health care spending and health outcomes. Working paper, Dartmouth.
- Coleman J.S., Katz E. and Menzel H. 1966. *Medical innovation: A diffusion study*. Indianapolis: The Bobbs-Merrill Company, Inc.
- Cooper, Z., Craig, S.V., Gaynor, M. and Van Reenen, J. 2018. The price ain't right? Hospital prices and health spending on the privately insured. *Quarterly Journal of Economics*, forthcoming.
- Cutler, D.M., McClellan, M., Newhouse, J.P., Remler, D. 1998. Are medical prices declining? Evidence from heart attack treatments." *Quarterly Journal of Economics*. 113, 991-1024.
- Cutler, D., Skinner, J., Stern, A.D., Wennberg, D. 2019. Physician beliefs and patient preferences: A new look at regional variation in health care spending. *American Economic Journals: Economic Policy*. (forthcoming)

- Doyle Jr. J. 2011. Returns to Local-Area Healthcare Spending: Using Health Shocks to Patients Far from Home.” *American Economic Journal: Applied Economics* 3(3):221-243.
- Doyle Jr., J., Graves, J., Gruber, J., Kleiner, J. 2015. Measuring returns to hospital care: Evidence from ambulance referral patterns. *Journal of Political Economy*. 123, 170-214.
- Doyle Jr, J.J., Graves, J.A., Gruber, J. 2017. Uncovering waste in US healthcare: Evidence from ambulance referral patterns. *Journal of Health Economics*. 54, 25-39.
- Dwyer-Lindgren, L., Mokdad, A.H., Srebotnjak, T., Flaxman, A.D., Hansen, G.M., Murray, C.J. 2014. Cigarette smoking prevalence in US counties: 1996-2012. *Population Health Metrics*. 12, 5.
- Eibich, P. and Ziebarth, N.R. 2014. Examining the structure of spatial health effects in Germany using Hierarchical Bayes Models. *Regional Science and Urban Economics*. 49, 305-320.
- Finkelstein, A., Gentzkow, M., Hull, P. and Williams, H. 2017. Adjusting risk adjustment—accounting for variation in diagnostic intensity. *The New England Journal of Medicine*. 376, 608-09.
- Finkelstein, A., Gentzkow, M., Williams, H. 2016. Sources of geographic variation in health care: Evidence from patient migration. *The Quarterly Journal of Economics*. 131, 1681-1726.
- Fisher, E.S., Wennberg, D.E., Stukel, T.A., Gottlieb, D.J., Lucas, F.L., Pinder, E.L. 2003a. The implications of regional variations in Medicare spending. Part 1: The content, quality, and accessibility of care. *Annals of Internal Medicine*. 138, 273-287.
- Fisher, E.S., Wennberg, D.E., Stukel, T.A., Gottlieb, D.J., Lucas, F.L., Pinder, E.L. 2003b. The implications of regional variations in Medicare spending. Part 2: Health outcomes and satisfaction with care. *Annals of Internal Medicine*. 138, 288-298.
- Gravelle, H., Santos, R., Siciliani, L. 2014. Does a hospital's quality depend on the quality of other hospitals? A spatial econometrics approach. *Regional science and urban economics* 49, 203-216.
- Hochman, J.S., Lamas, G.A., Buller, C.E., Dzavik, V., Reynolds, H.R., Abramsky, S.J., Forman, S., Ruzyllo, W., Maggioni, A.P., White, H. and Sadowski, Z. 2006. Coronary intervention for persistent occlusion after myocardial infarction. *New England Journal of Medicine*. 355, 2395-2407.
- James, W. L., Cossman, R. E., Cossman, J. S., Campbell, C., Blanchard, T. 2004. “A brief visual primer for the mapping of mortality trend data,” *International Journal of Health Geographics*, 3, 7.
- Joines, J.D., Hertz-Picciotto, I., Carey, T.S., Gesler, W., Suchindran, C. 2003. A spatial analysis of county-level variation in hospitalization rates for low back problems in North Carolina. *Social Science & Medicine*. 56, 2541-2553.
- Hussey, P.S., Wertheimer, S. and Mehrotra, A. 2013. The association between health care quality and cost: a systematic review. *Annals of Internal Medicine*. 158, 27-34.
- Keller W. 2004. Geographic localization of international technology diffusion. *American Economic Review*. 92, 120-142.
- Kibria, A., Mancher, M., McCoy, M.A., Graham, R.P., Garber, A.M., Newhouse, J.P. eds. 2013. *Variation in health care spending: target decision making, not geography*. National Academies Press.
- Likosky, D.S., Van Parys, J., Zhou, W., Borden, W.B., Weinstein, M.C., Skinner, J.S. 2018. Association between Medicare expenditure growth and mortality rates in patients with acute myocardial infarction: a comparison from 1999 through 2014. *JAMA Cardiology*. 3, 114-122.

- Lorant, V., Thomas, I., Delière, D., Tonglet, R. 2001. Deprivation and mortality: the implications of spatial autocorrelation for health resources allocation. *Social Science & Medicine*, 53, 1711-1719.
- McKnight, R. 2006. Home health care reimbursement, long-term care utilization, and health outcomes. *Journal of Public Economics* 90, 293-323.
- Moen, E.L., Austin, A.M., Bynum, J.P., Skinner, J.S. and O'Malley, A.J. 2016. An analysis of patient-sharing physician networks and implantable cardioverter defibrillator therapy. *Health Services and Outcomes Research Methodology*. 16, 132-153.
- Molitor, D. 2018. The evolution of physician practice styles: evidence from cardiologist migration. *American Economic Journal: Economic Policy*. 10, 326-56.
- Moscone, F. and Tosetti, E. 2014. Spatial econometrics: Theory and applications in health economics. In T. Culyer (Ed.), *The Encyclopedia of Health Economics*, pp. 329–334. Elsevier.
- OECD. 2018.. Public spending on health and long-term care: a new set of projections. <http://www.oecd.org/els/health-systems/public-spending-on-health-and-long-term-care.htm>
- Papageorgiou C., Savvides A., Zachariadis M. 2007. International medical technology diffusion. *Journal of International Economics*, 72, 409-427.
- Ricketts, T.C. and Holmes, G.M. 2007. Mortality and physician supply: Does region hold the key to the paradox? *Health Services Research* 42, 2233-2251.
- Romley, J.A., Jena, A.B., Goldman, D.P. 2011. Hospital spending and inpatient mortality: evidence from California: an observational study. *Annals of Internal Medicine*. 154, 160-167.
- Rothberg, M.B., Cohen, J., Lindenauer, P., Maselli, J., Auerbach, A. 2010. Little evidence of correlation between growth in health care spending and reduced mortality. *Health Affairs*, 29, 1523-1531.
- Rushton, G. 2003. "Public health, GIS, and spatial analytic tools." *Annual Review of Public Health*, 24(1), 43-56.
- Silber, J. H., R. Kaestner, O. Even-Shoshan, Y. Wang, and L. J. Bressler. 2010. "Aggressive treatment style and surgical outcomes," *Health Services Research* 45, 1872-92.
- Skinner, J. 2011. "Causes and consequences of geographic variations in health care," in T. McGuire, M. Pauly, and P. Pita Baros (eds.) Vol. 2, North Holland; 1st edition.
- Skinner, J., Fisher, E.S. and Wennberg, J. 2005. The efficiency of Medicare. In *Analyses in the Economics of Aging* (pp. 129-160). University of Chicago Press.
- Skinner, J. and Staiger, D. 2007. Technological diffusion from hybrid corn to beta blockers. *Hard-to-Measure Goods and Services: Essays in Honor of Zvi Griliches*. University of Chicago Press and NBER.
- Skinner, J., Staiger, D. 2015. Technology diffusion and productivity growth in health care. *Review of Economics and Statistics*. 97, 951-964.
- Song, Y., Skinner, J., Bynum, J., Sutherland, J., Wennberg, J.E. and Fisher, E.S. 2010. Regional variations in diagnostic practices. *New England Journal of Medicine*. 363, 45-53.
- Sriubaite, I. 2018. "My way or the highway? The influence of peers in the formation and consequences of physician practice styles," Paper presented at the Austrian Economic Association.
- Weintraub, W.S., Spertus, J.A., Kolm, P., Maron, D.J., Zhang, Z., Jurkovitz, C., Zhang, W., Hartigan, P.M., Lewis, C., Veledar, E. and Bowen, J. 2008. Effect of PCI on quality of life in patients with stable coronary disease. *New England Journal of Medicine*. 359, 677-687.

- Welch, H.G., Sharp, S.M., Gottlieb, D.J., Skinner, J.S. and Wennberg, J.E. 2011. Geographic variation in diagnosis frequency and risk of death among Medicare beneficiaries. *JAMA*, 305, 1113-1118.
- Wennberg, D.E., Sharp, S.M., Bevan, G., Skinner, J.S., Gottlieb, D.J. and Wennberg, J.E. 2014. A population health approach to reducing observational intensity bias in health risk adjustment: Cross sectional analysis of insurance claims. *BMJ*. 348, 2392.
- Yasaitis, L., Fisher, E.S., Skinner, J., Chandra, A. 2009. Hospital quality and intensity of spending: Is there an association?. *Health Affairs*. 28, w566-w572.

Figure 1: Map of One-Year Survival after Acute Myocardial Infarction, HCC Risk-Adjusted, 2007-2011

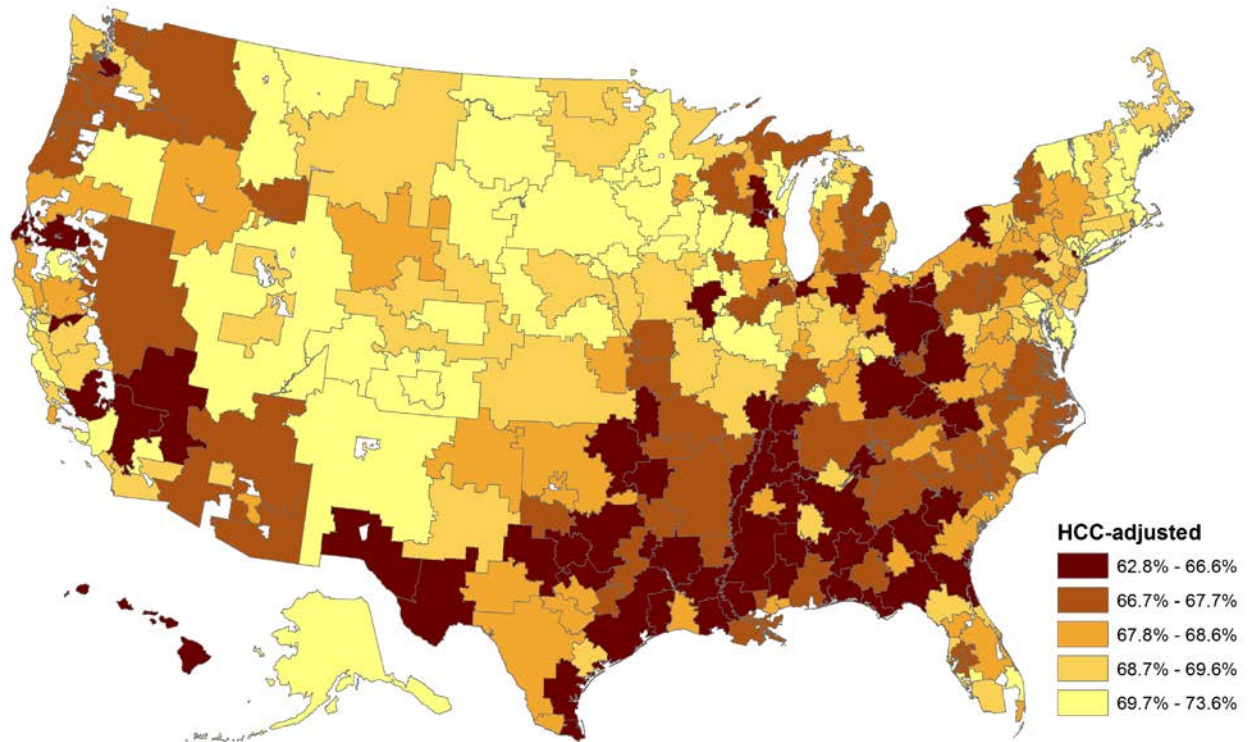


Figure 2: Map of One-Year Health Care Expenditures after Acute Myocardial Infarction, HCC Risk-Price Adjusted, 2007-2011

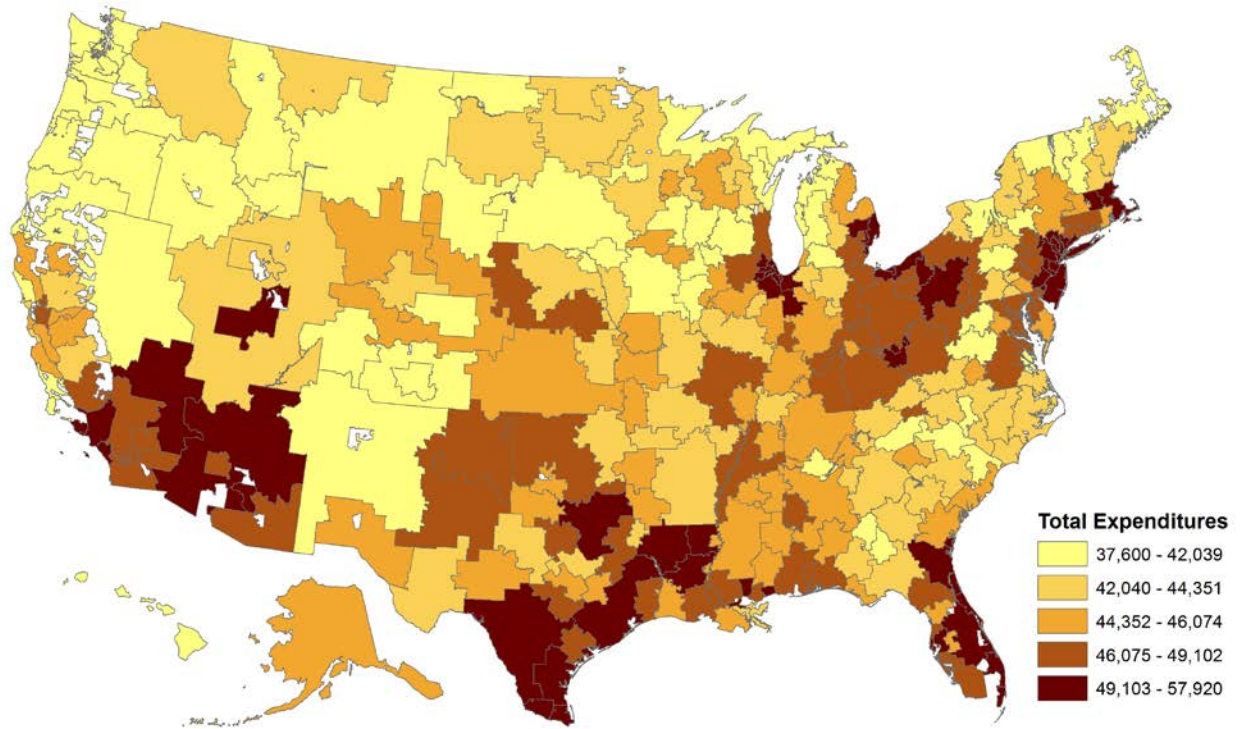
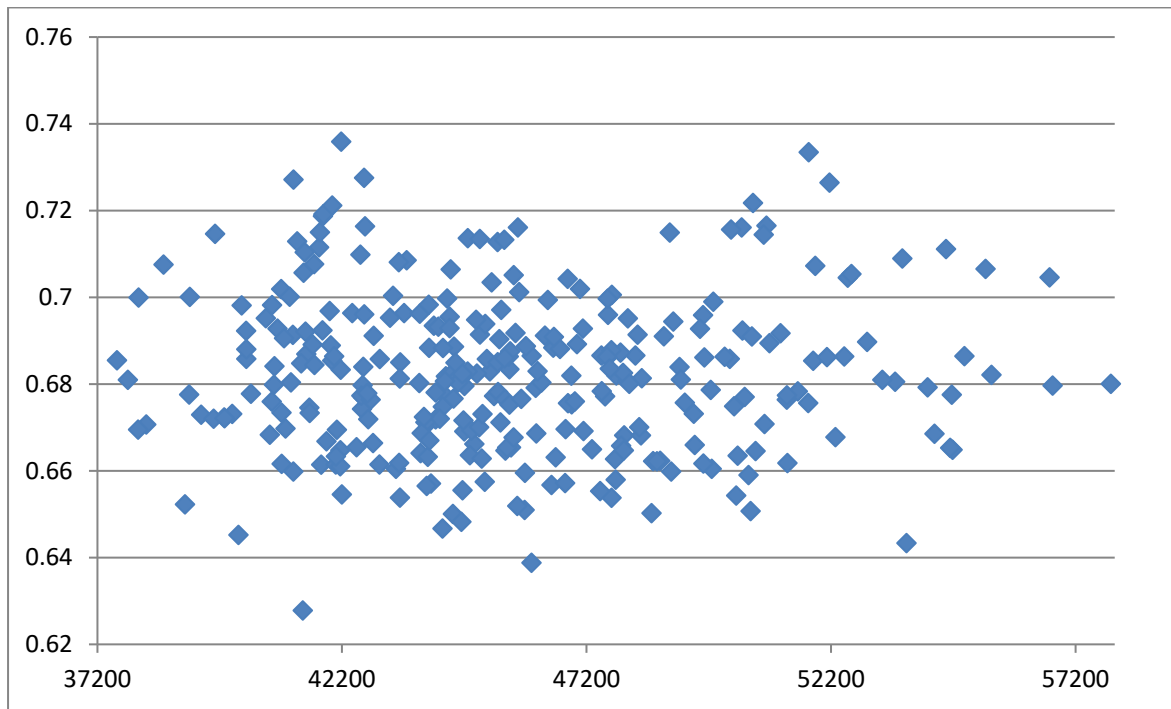
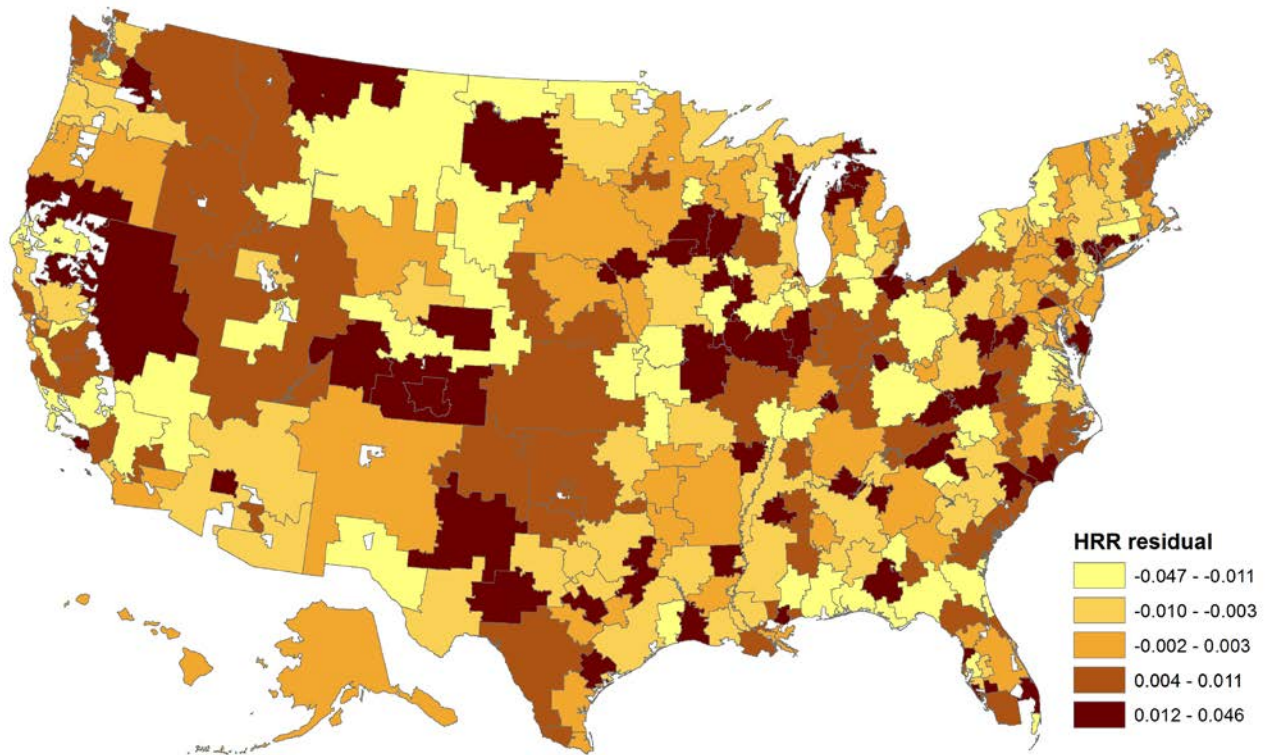


Figure 3: Plot of One-Year HCC-adjusted survival (Y) and One-Year HCC-adjusted health expenditure (X), time average 2007-2011



Note: Each dot represents a hospital referral region.

Figure 4: Survival Residuals with Spatial Adjustment



**Table 1: Descriptive Characteristics For Aggregated HHR-Level Variables
2007-2011**

	Mean	Std. dev.	Min	Max
Fraction Survival (1 Year)	0.68	0.03	0.57	0.81
Fraction HCC-adjusted Survival	0.68	0.03	0.57	0.80
Mean Health Expenditure	\$45,685	\$4,548	\$34,836	\$61,477
Fraction Teaching Hospital	0.14	0.18	0.00	0.86
Fraction PCI within 1 day	0.29	0.07	0.10	0.61
Average Home Health Spending	\$1,675	\$892	\$202	\$9,345
Average Income	\$51,904	\$12,978	\$29,969	\$112,442
Fraction Smoking	0.23	0.04	0.08	0.31

Note: N = 1,530 (306 HRRs over 5 years)

Table 2: HRR level regressions.

Variable	(I) CONVENTIONAL OLS						(II) SPATIAL REGRESSIONS					
	Par.	Std.err.	Par.	Std.err.	Par.	Std.err.	Par.	Std.err.	Par.	Std.err.	Par.	Std.err.
Dependent variable Survival non HCC-adjusted												
Health exp.	-0.0400***	0.0078	-0.0559***	0.0079	-0.0043	0.0093	-0.0242**	0.0088	-0.0412***	0.0087	-0.0068	0.0096
Income			0.0218***	0.0049	0.0195***	0.0048			0.0229***	0.0054	0.0215***	0.0051
Smoke			-0.0997***	0.0278	-0.0522	0.0269			-0.0687*	0.0317	-0.0413	0.0290
Teaching hospital					0.0056	0.0042					0.0015	0.0043
PCI within 1 day					0.1040***	0.0106					0.0980***	0.0112
HH care exp.					-0.0122***	0.0018					-0.0109***	0.0019
Spatial lag coeff.							-0.0020	0.0131	0.0158	0.0130	0.0271*	0.0125
Spatial error coeff.							0.3918***	0.0330	0.3162***	0.0350	0.1772***	0.0377
LM spatial lag							0.182	[0.67]	3.102	[0.08]	5.528*	[0.02]
LM spatial error							141.7***	[0.00]	73.927***	[0.00]	18.005***	[0.00]
Dependent variable Survival HCC-adjusted												
Health exp.	0.0087	0.0054	-0.0104	0.0074	0.0180*	0.0090	0.0078	0.0083	-0.0075	0.0080	0.0140	0.0093
Income			0.0210***	0.0046	0.0189***	0.0046			0.0214***	0.0050	0.0205***	0.0049
Smoke			-0.1049***	0.0260	-0.0761**	0.0261			-0.0808**	0.0288	-0.0591*	0.0280
Teaching hospital					0.0097*	0.0041					0.0054	0.0041
PCI within 1 day					0.0679***	0.0103					0.0668***	0.0108
HH care exp.					-0.0065***	0.0017					-0.0055**	0.0019
Spatial lag coeff.							0.0000	0.0126	0.0179	0.0124	0.0211	0.0123
Spatial error coeff.							0.3348***	0.0344	0.2259***	0.0368	0.1732***	0.0377
LM spatial lag							0.081	[0.78]	2.914	[0.09]	3.633	[0.06]
LM spatial error							104.2***	[0.00]	35.421***	[0.00]	17.025***	[0.00]

Notes: * <= .05, ** <= .01, and *** <= .001.

Table 3: HRR-level regressions. Use of the inverse of distance as spatial weights matrix

Variable	Par.	Std.err.	Par.	Std.err.	Par.	Std.err.
Health exp.	0.0022	0.0082	-0.0108	0.0073	0.0115	0.0087
Income			0.0204***	0.0046	0.0194***	0.0045
Smoke			-0.0521	0.0271	-0.0311	0.0263
Teaching hospital					0.0065	0.0040
PCI within 1 day					0.0558***	0.0101
HH care exp.					-0.0049**	0.0017
Sp. lag coeff.	0.6575	12.9998	0.6986***	0.1977	0.6903***	0.1222
Sp. error coeff.	0.6565	13.0305	0.2311	0.3963	-0.0578	0.3155
LM spatial lag	5.3770*	[0.02]	21.3138***	[0.00]	44.3145***	[0.00]
LM spatial error	2.5414	[0.11]	7.8449**	[0.01]	0.0200	[0.89]

Notes: * $\leq .05$, ** $\leq .01$, and *** $\leq .001$. HCCs included in constructing the HRR/state level data.

Table 4: HRR level regressions. Controlling for State dummies.

Variable	(I) CONVENTIONAL OLS						(II) SPATIAL REGRESSIONS					
	Par.	Std.err.	Par.	Std.err.	Par.	Std.err.	Par.	Std.err.	Par.	Std.err.	Par.	Std.err.
Dependent variable Survival non HCC-adjusted												
Health exp.	-0.0077	0.0088	-0.0220*	0.0090	-0.0174	0.0109	-0.0071	0.0088	-0.0214*	0.0089	-0.0174	0.0106
Income			0.0189***	0.0057	0.0160**	0.0056			0.0196***	0.0057	0.0159**	0.0055
Smoke			-0.0987*	0.0384	-0.0783*	0.0383			-0.0920*	0.0381	-0.0805*	0.0374
Teaching hospital					-0.0032	0.0047					-0.0032	0.0046
PCI within 1 day					0.1053***	0.0132					0.1061***	0.0129
HH care exp.					0.0008	0.0032					0.0006	0.0031
Spatial lag par.	-		-		-		-0.0094	0.0228	0.0029	0.0225	0.0114	0.0219
Spatial error par.	-		-		-		0.0770	0.0434	0.0413	0.0438	-0.0279	0.0443
LM spatial lag							0.769	[0.38]	0.082	[0.77]	0.001	[0.97]
LM spatial error							3.705	[0.05]	0.546	[0.46]	0.088	[0.77]
Dependent variable Survival HCC-adjusted												
Health exp.	0.0233**	0.0085	0.0103	0.0087	0.0018	0.0105	0.0226**	0.0085	0.0104	0.0086	0.0019	0.0103
Income			0.0117*	0.0055	0.0091	0.0055			0.0123*	0.0055	0.0088	0.0053
Smoke			-0.1343***	0.0372	-0.1053**	0.0372			-0.1269***	0.0369	-0.1081**	0.0364
Teaching hospital					-0.0005	0.0046					-0.0004	0.0045
PCI within 1 day					0.0932***	0.0128					0.0938***	0.0126
HH care exp.					0.0068*	0.0031					0.0068*	0.0030
Spatial lag par.	-		-		-		-0.0197	0.0224	-0.0066	0.0222	-0.0007	0.0217
Spatial error par.	-		-		-		0.0918*	0.0429	0.0386	0.0436	-0.0172	0.0441
LM spatial lag							0.187	[0.66]	0.017	[0.90]	0.274	[0.60]
LM spatial error							2.494	[0.11]	0.594	[0.44]	0.301	[0.58]

Notes: * <= .05, ** <= .01, and *** <= .001.

Table 5: County-level regressions.

	(I) CONVENTIONAL OLS						(II) SPATIAL REGRESSIONS					
	Par.	Std.err.	Par.	Std.err.	Par.	Std.err.	Par.	Std.er.	Par.	Std.err.	Par.	Std.err.
Health exp.	0.0087	0.0054	0.0098	0.0054	0.0078	0.0059	0.0097	0.0056	0.0106	0.0055	0.0074	0.0060
Income			-0.0016	0.0041	0.0011	0.0041			-0.0019	0.0043	0.0009	0.0043
Smoke			-0.0018***	0.0002	-0.0015***	0.0002			-0.0018***	0.0002	-0.0015***	0.0002
Teaching hosp.					0.0137***	0.0037					0.0124**	0.0038
PCI within 1 day					0.0812***	0.0119					0.0823***	0.0125
HH care expend.					0.0031*	0.0013					0.0037**	0.0014
Spatial lag par.							-0.0001	0.0042	-0.0009	0.0042	-0.0008	0.0042
Spatial error par.							0.0893***	0.0164	0.0705***	0.0165	0.0651***	0.0165
LM spatial lag							0.885	[0.35]	1.613	[0.20]	2.194	[0.14]
LM spatial error							35.081***	[0.00]	22.340***	[0.00]	18.498***	[0.00]

Notes: * <= .05, ** <= .01, and *** <= .001. HCCs included in constructing the HRR/state level data.