Spatial heterogeneity of quality, use and spending on medicare for the elderly

Felipa de Mello-Sampayo

Department of Economics, University of Lisboa (ISCTE-IUL), Lisbon, Portugal

Abstract

The spatial variation of the relations between Medicare spending (MS), use and quality in the United States was investigated employing spatial regression. A focus of the study was whether, and to what extent, MS and use vary by service type. Employing different spatial regression designs based on Medicare regional data, the impact of the heterogeneous spatial effects of hospital readmissions on MS for the elderly at the aggregate level was examined. The results were followed up by investigation whether the effects of hospital readmissions are heterogeneous with regard to service type. It was found that poor quality indicators lead to increased MS at the aggregate level and thus higher costs per beneficiary, and that the quality effects are heterogeneous with variable impacts, both spatially and by type of medical service. The results shed new light on the relationship between quality and MS highlighting the pitfalls of global averaging models that hide the reality of a highly diversified and spatially stratified country. Reducing payments to high-spending areas and increasing payments to low-spending areas should reduce spending variability but the quality indicators of care become ambiguous and not easy to interpret.

Correspondence: Felipa de Mello-Sampayo, Department of Economics, University of Lisboa (ISCTE-IUL), cacifo 187, Av. Forças Armadas, 1649-026, Lisbon, Portugal.
Tel: +351.918729720.
E-mail: fdmso@iscte.pt

Key words: Medicare; Quality; Elderly; Geographically weighted regression; Spatial econometric models; USA.

Conflict of interest: the author declares no potential conflict of interest.
Funding: funding from Fundação para a Ciência e Tecnologia, under UID/GES/00315/2013 grant is gratefully acknowledged. The funding body had no influence in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Received for publication: 30 November 2017.
Revision received: 4 December 2017.
Accepted for publication: 18 February 2018.

©Copyright F. de Mello-Sampayo, 2018
Licensee PAGEPress, Italy
Geospatial Health 2018; 13:655
doi:10.4081/gh.2018.655

This article is distributed under the terms of the Creative Commons Attribution Noncommercial License (CC BY-NC 4.0) which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

Introduction

Public health and primary health care are common elements in all health systems and nations must design/develop functioning health systems in accordance with their needs and resources. Healthcare planning can be seen as a concerted effort to deliver planned services targeted to the populations served. Generations currently reach a higher age than before with the consequence that the elderly today account for the majority of medical spending in both industrialized and developing countries. At the same time, governments are under increasing pressure to contain costs, which is done turning to an avenue of restraining the growth of medical expenditures by reducing costly complications and unnecessary procedures. Measuring quality is a non-trivial concept in healthcare. Hospital readmission (HR) is indicative of the interaction and coordination between hospital care and primary healthcare (PC). Elderly patients frequently require hospitalisation for treatment of complications such as pneumonia. Provided that patients are discharged in adequate functional status, and with appropriate follow-up instructions, there should be no need for repeated admissions within a short interval. As discussed by the Organization for Economic Co-operation and Development (OECD), the high cost of hospital care (combined with the high prevalence of disease) underscores the relevance of measuring readmission rates (OECD, 2017).

The Dartmouth Atlas research (http://www. dartmouthatlas.org/) found that most of the regional differences in healthcare spending result less from differences in technology, prices, payment levels, illness levels or patient preferences, than from the quantity and use of certain types of services in the local area. These services tend to be supply-sensitive, with no specific clinical theories of benefit governing their relative frequency of use, including, for example, when to schedule a revisit to the doctor, perform a diagnostic test, hospitalize or admit a patient to intensive care. For these types of services, utilization, and thus costs, vary greatly from region to region depending on the per capita supply of resources, such as hospital beds, medical specialists and diagnostic equipment (Fisher *et al*., 2009).

Compared to those in lower-spending regions, Medicare beneficiaries in higher-spending regions are hospitalized more frequently for conditions that could be treated outside the hospital. In addition, beneficiaries with serious chronic illness visit the physician twice as often, has a greater proportion of care provided by specialists compared to PC physicians and are more likely to see ten or more physicians a year (Sutherland *et al*., 2009). This variation in provider-practice patterns, particularly for supply-sensitive care where diagnosis and treatment are uncertain and the physician's decisions more discretionary, means more use of services for the same type of patient and thus greater healthcare spending. Areas with higher spending tend to score no better, sometimes worse, than other areas on measures of quality and

health outcomes. The additional services provided in higher healthcare spending regions, even for serious conditions such as heart attack or hip fracture, do not produce better quality or outcomes of care when looking at the technical quality and reliability of hospital or ambulatory care (Fisher *et al*., 2009). Arriving at an understanding of the relationship between healthcare spending and the quality of care is a critical part of interpreting geographic variation. Patterns of treatment in high-spending areas tend to be more intensive than those in low-spending areas, improving health outcomes for some types of patient but worsening outcomes for others. Looking at the state level, Baicker and Chandra (2004) found that higher Medicare spending (MS) per beneficiary, due to high concentrations of specialists, is sometimes associated with lower quality of care or has no relationship with the quality of care. Davis and Schoen (2007) found that health spending per capita by state is not correlated with mortality rates. Indeed, an inverse relationship exists between states that spend more on personal healthcare and state rankings on quality of care, while state MS per capita is highly correlated with preventable hospitalization (Davis and Schoen, 2007). Cooper (2008a, b), on the other hand, finds that states with more per enrolee have poorer state quality rankings, while states with more total health spending per capita have better quality rankings. He therefore argues that Medicare cannot be used as a proxy for overall healthcare spending or for the performance of the health system as a whole, and that commonly used quality standards alone are not valid measures of healthcare spending because socio-demographic factors (*e.g*., race, ethnicity, poverty, insurance coverage) also have an effect. His overall finding is that quality is better in states with more physicians per capita, both specialists and family physicians. Baicker and Chandra (2008) and Skinner *et al*., (2008) have, however, challenged his findings based on statistical methods and interpretation reasons. The relationships between quality, use, and cost of healthcare are important elements needed to consider when analyzing the geographic variation in healthcare spending. More importantly, patients in higher-cost areas are not necessarily receiving better care. Rather, the spending variations have been explained by the availability and volume of services used by patients in similar conditions (Fisher *et al*., 2009). An understanding of regional spending differences could answer important questions whether Medicare includes waste or inappropriate treatment practices. It also deserves to be asked what can be done to address these problems; what is the impact on providers and patients; how much money can be saved; and what role would changes play in healthcare reform. Although an extensive literature exists on health spending variation, surprisingly few studies have examined the elderly population in this respect despite their rapidly increasing share of the total medical spending. Also, relatively little is known about whether MS responds heterogeneously with regard to quality for the elderly population, both spatially and by service type. Such information should help policy makers implement better healthcare for the elderly. This paper aims to filling this gap by examining the Medicare market in the United States (US) whose healthcare spending drastically varies across the country.

Materials and Methods

Background and data selection

Helping chronically ill patients to cope with their diseases is an important goal for any healthcare system, while the HR indicator cannot be regarded as a comprehensive description of the quality of PC (Fischer *et al*., 2014). Thus, admission rates for congestive heart failure and respiratory disease were also included in the analysis. However, it has to be kept in mind that they are not necessarily indicators for PC only. The risk of readmission may be related to the type of drugs prescribed at discharge, patient compliance with post-discharge therapy, the quality of follow-up care in the community or the availability of appropriate diagnostic or therapeutic technologies during the initial hospital stay. Consequently, these indicators reflect quality of hospital care as well as PC and community-based care (Lee *et al*., 2003).

The data used in this study cover the year 2014 and are from the Geographic Variation Public Use File (GV PUF) of the Centers for Medicare & Medicaid Services (CMS) in the US. This file is for public use and primarily based on information from CMS's Chronic Conditions Data Warehouse (CCW), which contains 100% of Medicare claims for beneficiaries enrolled in the fee-for-service (FFS) program as well as enrolment and eligibility data. The study focus was on people aged 65 and older (nearly 28.0 million in 2014) because beneficiaries over 65 differ in numerous respects from other age groups and could have different health service needs that are difficult to adjust across geographic regions. Since the primary study goal was to analyze differences in healthcare utilization and spending for elderly Medicare beneficiaries living in different parts of the US, the sample comprises of beneficiaries over 65 who have both Part A and Part B coverage (see below) and are enrolled in Medicare's FFS program, excluding beneficiaries enrolled at any point during the year in a Medicare Advantage (MA) plan.

Medicare Part A is an insurance program providing basic protection against the costs of hospital and related post-hospital services for individuals at age 65 or over and eligible for retirement benefits under Social Security or the Railroad Retirement Board system. Part A covers inpatient hospital care, skilled nursing facility (SNF) care, hospice care and some home health care. Medicare Part B is a voluntary insurance program that provides insurance benefits for physicians, outpatient hospital services, ambulatory services and other medical supplies and services. FFS reimbursement is the payment system under which physicians and other providers receive a payment for each unit of service they provide.

One concern that arises when focusing on those who are 65 and above is that some might have died during the year of study (2014), which could inflate medical spending because it typically increases immediately before someone dies. One approach would be to exclude these subjects, but this is likely to bias estimates because it leaves out the most seriously ill from the sample as pointed out by Fukushima *et al*. (2016). For this reason, data on beneficiaries who died during the calendar year were included in this study.

The geographical unit of analysis to delineate regional healthcare markets in the US is the Hospital referral region (HRR) developed by the Dartmouth Atlas of Health Care. County data could have been utilized, but this would have led to loss of several variables because of missing and suppressed data, leading to the choice of HRRs for this assessment. The Dartmouth Atlas constructed the HRRs by grouping zip codes together based on the referral patterns for tertiary care for Medicare beneficiaries and requiring an overall population of 120,000 who must receive at least 65% of their hospitalizations within the HRR as discussed in detail by Wennberg and Cooper (1999). Continental US, excluding Alaska and Hawaii,

comprises 304 HRRs which were used in this study. Boundaries often cross state lines, *e.g.*, the Memphis, Tennessee HRR includes parts of south-eastern Missouri, eastern Arkansas and northern Mississippi. Although the HRRs are smaller than states, they are large enough to encompass most of the care received by beneficiaries even if they obtain care in multiple localities or counties. Furthermore, HRRs generally have populations large enough to generate stable averages for comparisons of cost and utilization even for narrowly defined combinations of conditions and services. MS can vary for reasons that are not attributable to patterns of practice or willingness to seek care. Two are particularly important, *i.e.* i) Medicare often pays different amounts for the same service in different areas (for example, to reflect variation in local wages or input prices). These payment rates are standardized to account for local wages or input prices and extra payments that Medicare makes, such as compensating certain hospitals for the cost of training doctors. Finally, CMS reduce all payment amounts to reflect any cost sharing that Medicare beneficiaries paid through deductible copayments or coinsurance; ii) the health of Medicare beneficiaries also varies geographically and those differences clearly affect spending and utilization. CMS has developed a risk-adjustment model based on Hierarchical Condition Categories (HCCs) to assign risk scores. Those scores estimate how beneficiaries' FFS spending compares to the overall average for the entire Medicare population. The average risk score is set at 1.0 and beneficiaries with scores greater than that are expected to receive above-average spending and vice versa. The risk scores are based on the beneficiary's age and gender, whether eligible for Medicaid, first qualified for Medicare on the basis of disability, living in an institution (usually a nursing home) and the beneficiary's diagnoses from the previous year. CMS uses the risk scores to adjust spending data at the beneficiary level rather than in aggregate. As a result, the aggregate standardized, risk-adjusted spending of a region does not equal the aggregate standardized costs divided by the average HCC risk score. By standardizing payment amounts and adjusting for differences in beneficiaries' health status, these data provide a more accurate picture of how resource use varies for Medicare beneficiaries across the country.

The three maps in Figure 1 show the distribution of MS and use across the US at the beneficiary level in 2014. The top map shows the regional variation of the original MS, the middle the standardized MS (SMS) and the bottom one the standardized and risk adjusted MS, *i.e.* the Medicare Service Use (MU). Throughout this paper, the term MU refers to MS adjusted for Medicare payment rates and the health status of the patient with data adjusted for differences in beneficiaries' health status so that MU was compared between groups of people with comparable needs for clinical resources.

MU has less regional variation than MS and SMS but substantial variation remains. Looking at the extremes, there is nearly a twofold difference between the area with the greatest MU (Monroe, Los Angeles) and that with the least (San Francisco, California). In terms of MS, there is more than a twofold difference between the area with the highest spending per beneficiary (Bronx, New York) and that with the lowest (Grand Junction, Colorado). There are also clusters of high MS regions which are largely concentrated in Florida, the deep South and urban areas on the East and West coasts, whereas the regions of high MU are the South, the East coast and the central West.

Many factors drive MU, such as differences in physician practice patterns and care decisions and differences in beneficiaries' preferences for seeking care. Looking at the extremes, the low MU in San Francisco may reflect unique characteristics and preferences that result in beneficiaries having similar levels of physician outpatient visits but lower levels of institutional (hospital, skilled nursing facility, hospice) care than in other parts of the country. When looking at regions with very high levels of MU, factors such as physician practice patterns and beneficiaries' predilection for care may drive MU above average, but different factors may account for the most extreme reported MU in areas like Monroe, Los Angeles. Reported MU there was almost 23% higher than the national average in 2014. Looking more closely at the data, it was noted that per beneficiary spending on home MU s and clinics in Monroe were more than three times and five times the national average, respectively. These types of pattern in the data raise concerns about abuse of Medicare use.

Figure 1 raises some immediate policy issues. If the higher spending in some regions actually leads to better health, then the Medicare program may be inequitable to the extent that taxpayers in the low-expenditure regions are paying for the better health of those in the high-spending regions (Feenberg and Skinner, 2000). Conversely, if the higher spending yields nothing in health benefits, then Medicare would represent tremendous waste that could be remedied by reducing spending in the high-spending areas thereby extending the solvency of the Medicare trust fund (Skinner and Wennberg, 2000). It could also be the case that people in the high-spending areas simply prefer more intensive care. One might then ask why other regions should be subsidizing their preferences (Skinner and Wennberg, 2005).

Table 1 reports summary statistics of the SMS at i) the aggregate level; ii) by service type; and iii) medication. We have a total of 304 observations, representing 28.0 million enrolees in 2014. The aggregated results panel reports that the average total SMS was approximately USD8,708.00 per beneficiary-year and the average total standardized risk-adjusted MS was approximately USD9,405.00. Spending for inpatient (outpatient) services, which included prescription drugs, was approximately USD4,019 (4,655) per beneficiary, accounting for 46% (54%) of the total MS.

We also analyzed spending by type of service to adjust for regional service differences. However, the HCC model was not designed to adjust spending risk for individual services and therefore not applied to service-level spending. Hence, we analyze the SMS by service type: inpatient and outpatient services and prescription drugs. In the outpatient setting, MS was highest for hospital outpatient services (HOP) with USD1,261.14 per beneficiary-year, followed by evaluation and management services (E&M) and physician procedures (PP). Among the inpatient treatments, that are paid by fee for service, MS for inpatient hospital care (IPHC) was the highest, which accounted for approximately 28% of the total MS. The medication panel summarizes outpatient prescription drug spending with USD326.03 per beneficiary-year, which accounts for approximately 7% of total outpatient MS.

Statistics (econometric methodology)

Earlier research has shown that various demographic, economic, and institutional factors are important in explaining MS. To control for these factors, the following model was used:

$$SMS_i = \alpha_k X_{ik} + \varepsilon_i \qquad \text{Eq. 1}$$

where SMS_i is the SMS of HRR_i, X an observation matrix of k conditional variables, and ε_i the error term. The focus was on empirical

work on the determinants of the cross-region variation in MS. The observed MS level of a given region could not only depend on its own determinants, but also on the determinants of the MS of its neighbours. A high level of MS in a particular region could increase the MS in all surrounding regions, with a resulting positive impact on their MS. Spatial econometric models are needed that take into account such proximity effects (Moscone *et al*., 2007a). The geographically weighted regression (GWR) approach estimates local rather than global parameters. Instead of calibrating a single regression equation, GWR generates a separate regression

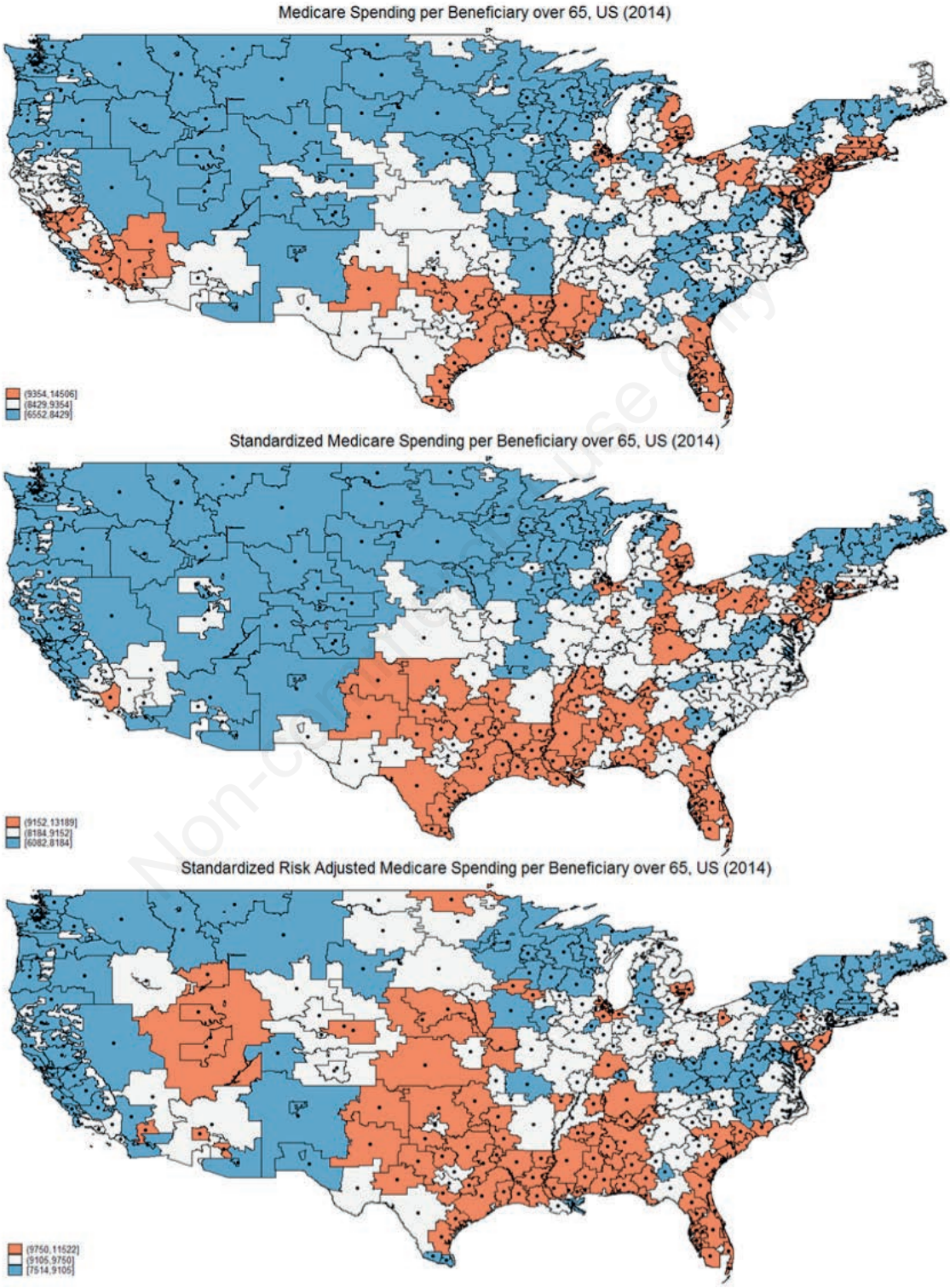

Figure 1. Total Medicare spending (MS) per beneficiary, US 2014.

equation for each observation and each equation is calibrated using a different weighting of the observations contained in the data (Graif and Sampson, 2009). At each regression point (which is represented by the centroid of HRR) the equation above (Eq. 1) can be reformulated as follows:

$$SMS_i = \Sigma_k \alpha_k (u_i, v_i)\, x_{ik} + \varepsilon_i \qquad \text{Eq. 2}$$

where (u_i, v_i) represents the coordinate location of the data point i, in this case the HRR centroid, and $\alpha_k (u_i, v_i)$ realizes the continuous function $\alpha_k (u, v)$ at point i. We estimated α as:

$$\hat{\alpha}_k(u_i, v_i) = (X^T W(u_i, v_i) X)^{-1} X^T W(u_i, v_i) SMS \qquad \text{Eq. 3}$$

where X is an observations matrix of k conditional variables, $W(u_i, v_i)$ an n-by-n matrix whose off-diagonal elements are zero and whose diagonal elements denote the geographical weighting of observed data for point i (Fotheringham *et al*., 2002). Each observation is weighted according to its proximity to i. When moving across space, the weight of an observation is no longer constant in the calibration as in ordinary least squares (OLS), but varies with i (see Eq. 5 below).

Spatial regression models likewise relax the assumption of spatial independence and ad- just for spatially autocorrelated processes by incorporating local relationships in the error covariance structure (Anselin, 1988). In the spatial error model (SARE), spatial dependence is modeled as a spatial autoregressive process in the error term, *i.e* there is spatial clustering in the unobservables. Ignoring spatial dependence in the error term does not yield biased least squares estimates, although their variance will be biased, thereby resulting in misleading inferences (Anselin, 1988, 1990). The source of spatial dependence may be due to spillover effects in the outcomes. In this case, spatial dependence is modelled as a spatially lagged regressor that is correlated with the error term, and the least square (OLS) estimation turns out to be biased and inconsistent due to the simultaneity bias.

We estimated the MS as given by Eq. 1 using the combined spatial-autoregressive model with spatial autoregressive disturbances (SARAR) model (Anselin and Florax, 1995). By modelling the outcome for each observation as related to a weighted average of the outcomes of other units, this model determines the outcomes simultaneously (Drukker *et al*., 2013). According to this approach, Eq. 1 can be rewritten as follows:

$$SMS_i = \lambda VSMS_i + X\alpha_i + u_i \qquad \text{Eq. 4}$$
$$u_i = \rho Vu_i + \varepsilon_i$$

where V is an n-by-n spatial link matrix with zero diagonal elements whose non-zero off- elements, v_{ij} express the degree of potential spatial interaction between each possible i^{th} and j^{th} pair of locations, λ the spatial dependence parameter, the spatial error parameter. The spatial-weighting matrix, V, is employed to compute weighted averages in which more weight is placed on nearby observations than on distant observations parameterizing Tobler's law of geography, *i.e. Everything is related to everything else, but*

Table 1. Medicare spending in USD: summary statistics.

Variable	HRR*	Mean	SD**	Min.	Max.
Aggregated results					
Total	304	9,049.94	1,264.52	6,552.04	14,505.5
Total standardized	304	8,674.51	1,201.55	6,082.43	13,188.76
Total standardized and risk-adjusted	304	9,404.58	753.48	7,513.74	11,521.53
By treatment spending					
Inpatient setting					
Inpatient hospital care (IPHC)	304	2,427.31	301.88	1,639.69	3,545.15
Skilled nursing facilities (SNFs)	304	884.2	207.58	332.31	1,518.32
Outpatient setting					
Hospital outpatient services (HOPs)	304	1,261.14	271.66	679.69	2,249.85
Clinics	304	44.94	50.2	0.96	334.28
Outpatient dialysis facilities (OPDs)	304	136.01	52.44	49.69	393.62
Home health services (HHs)	304	519.52	321.13	86.55	2,617.15
Ambulatory surgical centres (ASCs)	304	90.67	37.95	3.17	247.62
Evaluation & management services (E&Ms)	304	976.28	217.36	520.86	1,647.65
Physician procedures (PPs)	304	575.2	114.73	333.61	1,054.4
Imaging	304	202.64	67.39	77.53	440.33
Durable medical equipment (DME)	304	184.53	37.71	93.41	350.96
Tests	304	214.79	82.96	60.53	489.13
Ambulance	304	124.05	43.96	41.27	336.58
Medication					
Prescription drugs	304	326.03	116.06	63.11	777.61

*Number of; **Standard deviation.

Non-commercial use only

near things are more related than distant things (Tobler, 1970). In inverse distance spatial-weighting matrices, the weights are inversely related to distance between the locations ($v_{ij} = 1/\tau_{ij}$ where τ_{ij} is the distance between regions *i* and *j*).

In contrast to GWR, the spatial regression models produce global parameter estimates and the influence of neighbours is stipulated by a spatial weight function that is not calibrated at each observation point (Fotheringham *et al*., 2002). When the spatial weighting function is fixed or applied equally at each calibration point, one assumes that the weight-distance relationship is globally applicable at all calibration points across space, which can be problematic for several reasons. Firstly, the global statement may not be true, for instance in situations where physical (built or natural) buffers such as highways or parks between two neighbourhoods radically affect their impact on one another; secondly if data are sparse in parts of the larger area the local regressions may be based on too few data points. To account for these possibilities in the GWR models a spatially adaptive weighting function was used instead. This function allows for smaller bandwidths in which the data are dense and for larger bandwidths in which the data are sparse. Specifically, a bandwidth represents how far out from a focal neighbourhood, *i*, the other neighbourhoods will count in the calibration of parameters at point *i*.

The following bi-square function allows for such spatially adaptive bandwidths:

$$v_{ij} = \begin{cases} \left[1 - \left(\tau_{ij}/b_i\right)^2\right]^2, & \text{if } \tau_{ij}/b_i \\ 0, & \text{otherwise} \end{cases} \qquad \text{Eq. 5}$$

where b_i denotes the bandwidth and the distance of the nth nearest neighbour from *i*, while τ_{ij} represents the distance between places *i* and *j*. Instead of fixing the distance, the number of nearest neighbours was fixed allowing the kernel to go as far in space as needed in order to find that number of neighbours. This is a continuous, near-Gaussian weighting function up to distance *b* from the regression point, becoming zero at any data point beyond b_i (Fotheringham *et al*., 2002).

The estimated parameters depend in part on the weighting function and bandwidth selected, while the selection of the weighting function did not appear to be as consequential for the results as the selection of the bandwidth. When the bandwidth tended to infinity the weights become uniformly close to 1 and the spatial variance of the estimated parameters tends to zero rendering GWR equivalent to OLS. The weight function was calibrated using the Akaike Information Criterion (AIC) minimization procedure of Hurvich *et al*. (1998), which provides a trade-off between goodness-of-fit and degrees of freedom. The AIC indices were also compared to assess if GWR provided a better fit than a global model while adjusting for the different degrees of freedom in the two models.

Some of the local spatial variability may result from sampling variation. As a check we used the computationally intensive Monte Carlo method (Metropolis and Ulam, 1949), which tests if the observed variation in a parameter is sufficient to reject the null hypothesis of a globally fixed parameter. When no real spatial pattern in the parameter exists, any permutation of the regression variables against their locations should be equally likely, providing a model for the null distribution of the variance. The Monte Carlo approach was used to randomly permute the geographical coordinates of the observations against the variables a certain number of times. This produces n values of the variance of the coefficient of interest, which were used as an experimental distribution. By comparing the actual variance against this distribution an experimental significance level for the spatial variability of each individual parameter was obtained.

Results

Equations (2) - GWR Model and (4) - Spatial Regression Model were estimated to analyze the spatial MS across the US using HRR data from the study year (2014). Data for the spatial analysis, *i.e*. the geographic coordinates of HRR centroids and respective boundaries, were created using the shapefile for the HRR from the topologically integrated geographic encoding and referencing system (TIGER) data 2008 from the US Census Bureau. The measure used in the weighting matrices, *W* and *V* of Equations (2) - GWR Model and (4) - Spatial Regression Model, respectively, was the Euclidean distance between the *i* and *j* centroids. Though not reported here, common boundaries were tested in the weighting matrices and found to produce generally consistent results.

Our dependent variable was the SMS, standardized at the aggregate level and by service type across the US comprising the 304 US continental geo-referenced HRRs. The spatial effects on MS with respect to specific quality variables from the Prevention Quality Indicators' measure set, a publicly available software developed by the Agency for Healthcare Research and Quality based on administrative data, could be used to analyse hospital admission rates for conditions sensitive for ambulatory care. After controlling for the health risk status (HS) of Medicare beneficiaries proxied by the HCC score, hospital admission (HA) and readmissions (HR) rates due to congestive heart failure (HACHF) and respiratory disease (HARD) were included in the analysis. Focusing on these two disorders were felt to be useful because they represent two of the leading causes of morbidity and mortality among the elderly according to the National Center for Health Statistics. Rigorous outpatient management has been shown to reduce disease progression and the need for hospital care. All other things being equal, we expected higher levels of HR, HACHF and HARD to have a positive effect on SMS.

In Table 2 we estimated Eq. 2 with regard to HS using Monte Carlo testing. The coefficients reported in column 1 represent the median of all local coefficients estimated across space. The probability markers reflect results from the Monte Carlo tests, not t-tests. The optimal bandwidth, calculated based on the AIC criterion, varied across the models between the three and five nearest HRRs. The test of the bandwidth in column 2 suggests that the GWR model is a much better model for our data than the global model.

Table 2 indicates that the estimated HS coefficients vary significantly across the US in predicting SMS. As expected, the results say that the HS affects the SMSs positively for almost every Medicare type of service. However, SMS decreases with HS for HOP, clinics, ambulatory surgical centres (ASCs), and Durable Medical Equipment (DME), which may suggest that lower risk patients are encouraged to use these supply-sensitive types of Medicare services.

Non-commercial use only

In Table 3 Eq. 2 is estimated adding HR as a quality indicator after controlling for HS. The MC tests of spatial variability on the coefficients reported in parentheses indicate that both HS and HR vary significantly across the US. The test of the bandwidth suggests that the GWR model is a significantly better model for these data than the global model. When controlling for HR, it was found that the HS risk effect maintains the same pattern when compared with those in Table 2. With regard to the effect of HR on SMS, we noted a positive effect for almost every type of Medicare service, with the exception of outpatient dialysis facilities (OPD), physician procedures (PP), tests and drugs, *i.e.* the higher the quality indicator (which means a lower HR), the more often the beneficiaries use these supply-sensitive types of Medicare services.

The spatial distribution of GWR estimates for the SMS at the aggregate level shown in Table 3 is also presented in maps (Figure 2). It can be seen that the relationship between SMS and HS is always positive and particularly strong in the mountainous region, the centre West and southern Atlantic areas. In relation to the mapping of the HR coefficients, it was observed that the positive effects were accentuated in counties of the Southwest region, with the exception of some HRRs in the West and East North centre, the New England region and also in Texas. In these regions there is evidence that lower quality increases the use of Medicare services. However, in the West and the South Atlantic regions, the effect of HR becomes negative, which may give support to the hypothesis that the higher SMS is due to higher quality.

Table 4 shows estimates of Eq. 2 after adding HACHF and HARD as quality indicators. The first column refers to local estimates of the HS, and the second and third columns to HACHF and HARD, respectively. The MC tests of the spatial coefficient variability indicate that all variables vary significantly across the US. The test of the

Table 2. Geographically weighted regression (GWR) estimates of the standardized Medicare spending (SMS) ($SMSi = \alpha (u_i, v_i) HS_i + \varepsilon_i$).

Dependent variable	HS[a] and (Monte Carlo test)[b]	Test for bandwidth
Aggregated results		
Total standardized	1.259*** (0.205)	4.983***
By treatment spending		
Inpatient setting		
Inpatient hospital care (IPHC)	1.281*** (0.333)	3.961***
Skilled nursing facilities (SNFs)	1.252*** (1.239)	3.088***
Outpatient setting		
Hospital outpatient services (HOPs)	–0.185* (0.738)	3.961***
Clinics	–4.925*** (12.384)	1.917***
Outpatient dialysis facilities (OPDs)	3.283* (2.103)	1.917***
Home health services (HHs)	3.622*** (2.061)	4.983***
Ambulatory surgical centres (ASCs)	–0.204* (1.722)	3.088***
Evaluation & management services (E&Ms)	1.491* (1.116)	1.917***
Physician procedures (PPs)	0.532** (1.121)	1.917***
Imaging	1.528* (0.929)	3.961***
Durable medical equipment (DME)	–0.113* (0.526)	4.983***
Tests	1.770* (2.750)	1.917***
Ambulance	1.791*** (1.213)	3.088***
Medication		
Prescription drugs	0.498*** (1.581)	3.961***

[a]Health risk status; [b]These tests of spatial variability indicate the extent at which coefficient variation across space significantly differs from random distribution; *Null rejection at the 10% level; **Null rejection at the 5% level; ***Null rejection at the 1% level.

Table 3. Geographically weighted regression (GWR) estimates of the standardized Medicare spending (SMS) ($SMS_i = \alpha_1 (u_i, v_i) HS_i + \alpha_2(u_i, v_i) HR_i + \varepsilon_i$).

Dependent variable	HS[a] (Monte Carlo test)[b]	HR[c] (Monte Carlo test)[b]	Test for bandwidth
Aggregated results			
Total standardized	0.867* (0.243)	0.179** (0.100)	4.983***
By treatment spending			
Inpatient setting			
Inpatient hospital care (IPHC)	0.465** (0.175)	0.375** (0.043)	3.961***
Skilled nursing facilities (SNFs)	0.360*** (1.416)	0.434* (0.351)	3.088***
Outpatient setting			
Hospital outpatient services (HOPs)	–0.322* (0.990)	0.095*** (0.403)	3.961***
Clinics	–5.699* (13.964)	0.852* (3.775)	1.917***
Outpatient dialysis facilities (OPDs)	3.349* (2.210)	–0.112* (0.567)	3.088***
Home health services (HHs)	3.309*** (1.971)	0.138*** (0.687)	4.983***
Ambulatory surgical centres (ASCs)	–0.382* (2.934)	0.941* (1.100)	3.088***
Evaluation & management services (E&Ms)	1.299* (0.653)	0.126*** (0.347)	3.088***
Physician procedures (PPs)	0.663* (0.744)	–0.057* (0.279)	3.088***
Imaging	1.345** (1.043)	0.055* (0.396)	3.961***
Durable medical equipment (DME)	–0.525* (1.115)	0.225* (0.330)	3.961***
Tests	1.474** (1.887)	–0.003*** (0.694)	3.088***
Ambulance	0.587* (0.888)	0.593* (0.325)	4.983***
Medication			
Prescription drugs	0.814** (1.298)	–0.157* (0.433)	4.983***

[a]Health risk status; [b]These tests of spatial variability indicate the extent at which coefficient variation across space significantly differs from random distribution; [c]hospital readmission; *Null rejection at the 10% level; **Null rejection at the 5% level; ***Null rejection at the 1% level.

Non-commercial use only

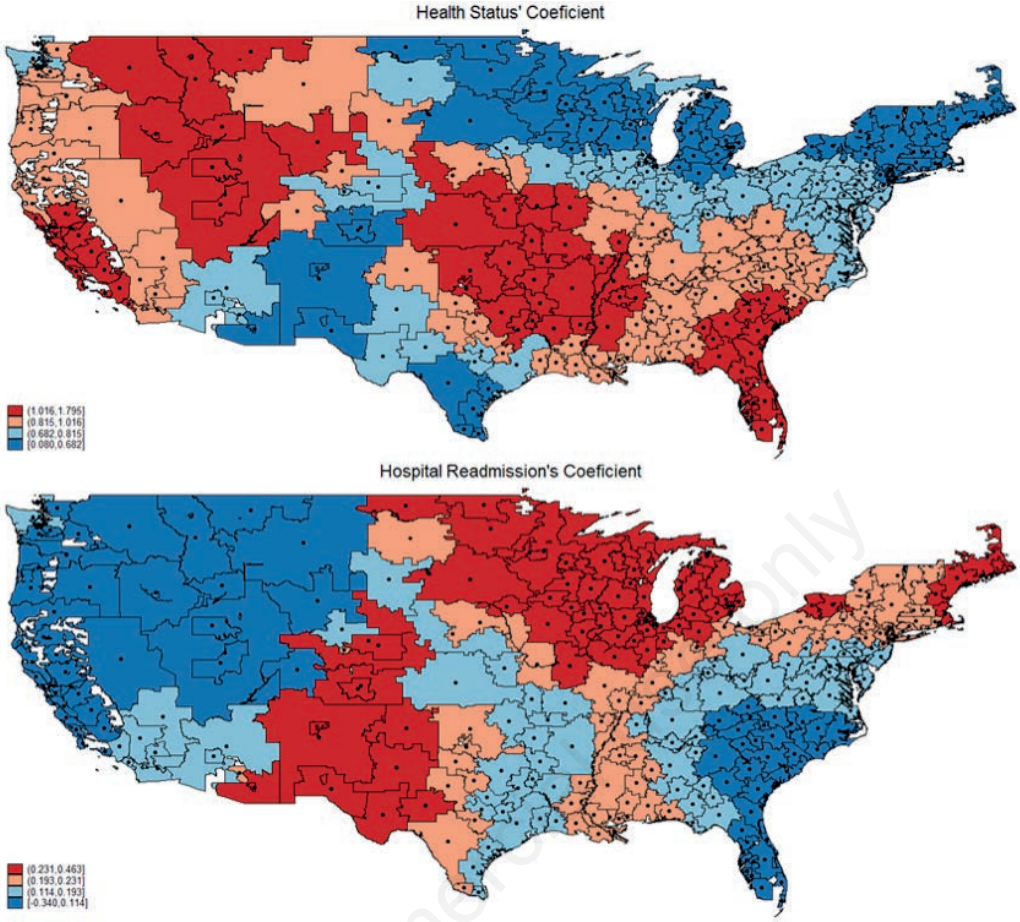

Figure 2. Estimated local parameters relative to Table 3.

Table 4. Geographically weighted regression (GWR) estimates of the standardized Medicare spending (SMS) ($SMS_i = \alpha_1 (u_i, v_i) HSi + \alpha_2(u_i, v_i) HACHF_i + \alpha_3(u_i, v_i) HARD_i + \varepsilon_i$).

Dependent variable	HSa (Monte Carlo test)[b]	HRc (Monte Carlo test)[b]	Test for bandwidth	Dependent variable
Aggregated results				
Total standardized	1.236* (0.205)	0.009* (0.057)	0.026*** (0.046)	4.983***
By treatment spending				
Inpatient setting				
Inpatient hospital care (IPHC)	0.990*** (0.232)	0.107* (0.031)	0.045* (0.025)	8.048***
Skilled nursing facilities (SNFs)	1.238* (1.279)	–0.102* (0.204)	0.127* (0.181)	3.088***
Outpatient setting				
Hospital outpatient services (HOPs)	–0.374*** (0.619)	0.040*** (0.191)	0.149* (0.108)	4.983***
Clinics	–9.206*** (7.019)	1.179*** (1.259)	1.178* (0.616)	4.983***
Outpatient dialysis facilities (OPDs)	2.469*** (1.550)	0.288*** (0.303)	–0.275* (0.179)	4.983***
Home health services (HHs)	3.418*** (1.713)	0.082* (0.296)	–0.009* (0.264)	4.983***
Ambulatory surgical centres (ASCs)	0.449* (1.349)	–0.115 (0.333)	–0.081 (0.288)	4.983***
Evaluation & management services (E&Ms)	2.097*** (0.562)	–0.103* (0.096)	–0.156* (0.093)	4.983***
Physician procedures (PPs)	1.171** (0.424)	–0.150* (0.084)	–0.124* (0.118)	4.983***
Imaging	1.760** (0.825)	–0.058* (0.194)	–0.163* (0.143)	4.983***
Durable medical equipment (DME)	–0.528* (0.449)	0.104* (0.157)	0.269* (0.158)	4.983***
Tests	2.111* (1.294)	–0.222* (0.235)	–0.091*** (0.276)	3.961***
Ambulance	0.916* (0.740)	0.308* (0.175)	0.113* (0.126)	4.983***
Medication				
Prescription drugs	1.274* (0.369)	–0.252* (0.204)	–0.214*(0.136)	8.048***

[a]Health risk status; [b]These tests of spatial variability indicate the extent at which coefficient variation across space significantly differs from random distribution; [c]hospital admission due to congestive heart failure; [d]hospital admission due to respiratory disease; *Null rejection at the 10% level; **Null rejection at the 5% level; ***Null rejection at the 1% level.

Non-commercial use only

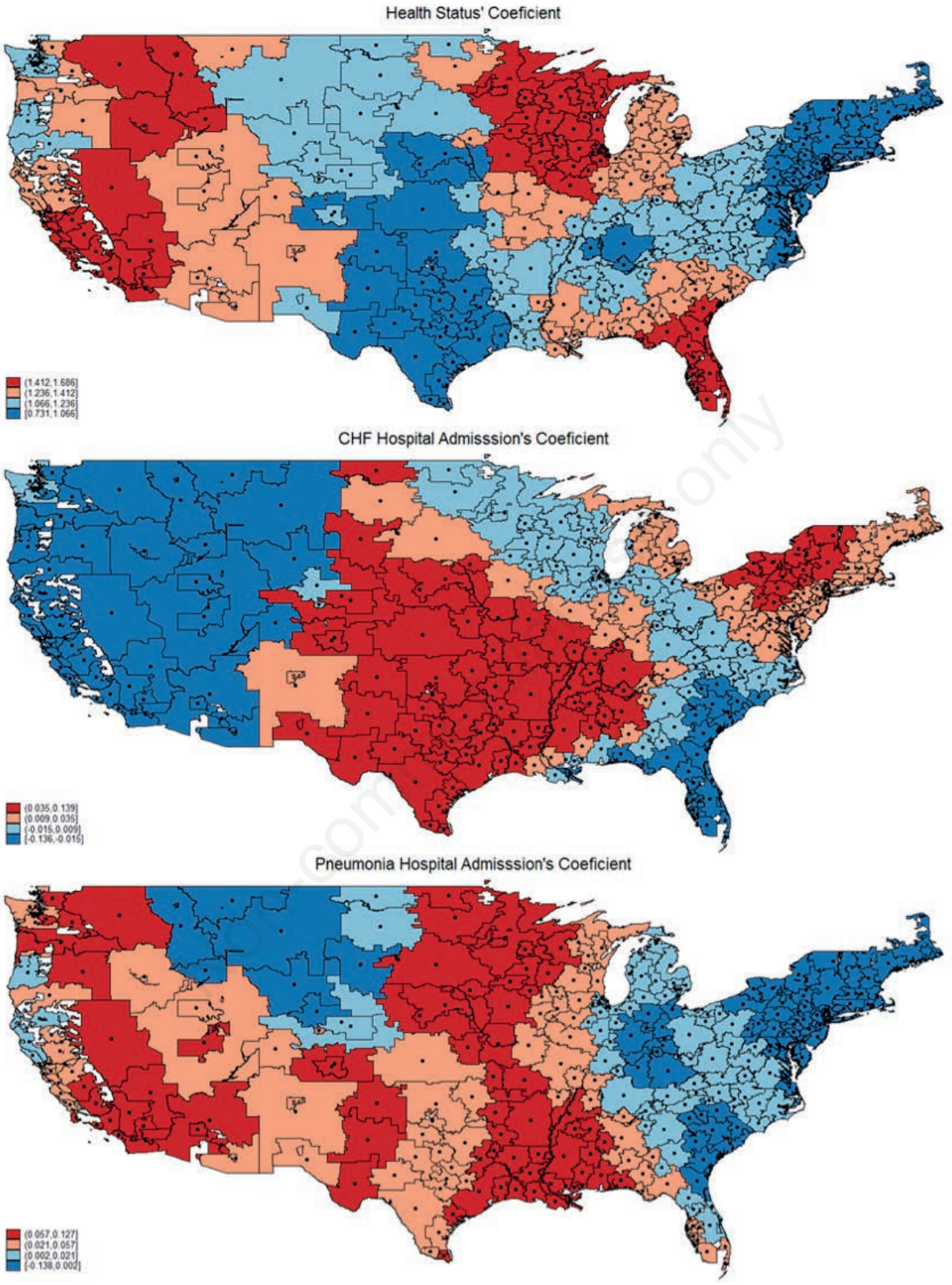

Figure 3. Estimated local parameters relative to Table 4.

bandwidths suggests that the GWR model is significantly better for these data than the global model. Mixed results in relation to the effect of these quality indicators on SMS were found here. At the aggregated level and with reference to inpatient settings, the results support the hypothesis that the lack of quality aggravates the SMS. However, in outpatient settings, the evidence points to a positive relationship between higher quality (*i.e.* lower HA) and SMS of these services.

The spatial distribution of GWR estimates for the SMS at the aggregate level shown in Table 3 is also presented as maps (Figure 3). The mapping of the HACHF estimates confirms that the low-quality indicators in centre West South up to centre West North and centre East South regions have a positive impact on SMS. The HARD effect is positive and stronger in some parts of the mountainous region and in the centre East South up to the centre East North. The West region is the HRR where a negative relationship between SMS and HA rate was observed suggesting that the good quality indicator leads to SMS reduction. The heterogeneity of the quality effects on SMS shown in Figures 2 and 3 may become aggravated if geographic variation in transportation options and natural resources are considered.

Large differences in spending for the care of elderly patients across the US were noted. It was found that Medicare spent an average USD9,050 per beneficiary over 65 in 2014. However, considerable variation in spending occurred among the 304 US continent HRRs, with the highest-cost regions spending more than twice the amounts spent in the lowest-cost regions. The highest-spending regions were Bronx, New York (USD 14,506) and Miami, Florida (USD14,436), while the lowest-spending were Grand Junction, Colorado (USD6,552) and Missoula, Montana (USD6,791).

Robustness checks

To assess the robustness of the spatial MS analysis, three checks were conducted. The GWR models discussed above were first estimated but now using four socio-demographic factors correlated to HS: age, male, African descent and Medicaid. As shown in Table 5, there was a negative effect associated with male gender, while positive effects were recorded for age, minorities (as proxied by the number of beneficiaries of African descent), and low-income beneficiaries (as proxied by the number of Medicaid beneficiaries). The quality indicators reported in columns 2 and 3 positively affected Medicare use and varied significantly across the US. These results confirm the heterogeneity in the effects of quality on SMS with regard to the elderly. However, at the aggregate level, higher MS per beneficiary was associated with lower quality, or no relationship, with the quality indicators. Secondly, Eq. 4 was estimated using maximum likelihood (ML) spatial regression, which corrects for both spatial error and spatial lag (the SARAR model), at the aggregated level. The Wald tests (Engle, 1983) reported in Table 6 do not reject the overall significance of the regressors, all variables have the expected sign and are significant, with the exception of the parameter λ, which is not significant in columns 1, 3, and 5. In other words, the inclusion of spatial corre-

Non-commercial use only

Table 5. Geographically weighted regression (GWR) estimates of Medicare spending (SMS) - Equation (2) at the aggregate level.

Independent variable	i) Monte Carlo test[a]	ii) Monte Carlo test[a]	iii) Monte Carlo test[a]
Age	0.355* (3.556)	1.313* (2.553)	1.152* (3.219)
Male	–1.175** (1.237)	–0.882* (0.912)	–1.335* (1.123)
African descent	0.016*** (0.034)	0.001** (0.026)	0.008*** (0.036)
Income	0.042* (0.079)	0.006* (0.071)	–0.030* (0.084)
Hospital readmission (HR)		0.306*** (0.133)	
Hospital admission HACHF[b]			0.125* (0.115)
Hospital admission HARD[c]			0.043*(0.084)
Test for bandwidth	3.088***	3.088***	3.088***

[a]These tests of spatial variability indicate the extent at which coefficient variation across space significantly differs from random distribution; [b]congestive heart failure; [c]respiratory disease; *Null rejection at the 10% level; **Null rejection at the 5% level; ***Null rejection at the 1% level.

Table 6. Spatial autoregressive model with autoregressive disturbances (SARAR) and spatial error model (SARE) estimates of Medicare spending (SMS) - Equation (4) at the aggregate level.

Independent variable	i) SARAR (Standard error)	ii) SARE (Standard error)	iii) SARAR (Standard error)	iv) SARE (Standard error)	v) SARAR (Standard error)	vi) SARE (Standard error)
HS	1.047*** (0.054)	1.080*** (0.051)	0.786*** (0.067)	0.812*** (0.069)	0.946***(0.063)	0.991*** (0.065)
HR			0.151*** (0.024)	0.149***(0.023)		
HACHF					0.011 (0.016)	0.006 (0.016)
HARD					0.056*** (0.012)	0.049*** (0.012)
Lambda	–0.008 (0.006)		–0.008 (0.007)		–0.007 (0.005)	
Rho	8.781*** (0.329)	7.909*** (0.149)	7.942*** (0.139)	7.266*** (0.261)	10.409*** (0.311)	8.979*** (0.339)
Sigma^2	0.003*** (0.000)	0.003*** (0.000)	0.003*** (0.000)	0.002 (0.000)	0.002 (0.000)	0.002 (0.000)
Wald test	380.1***	453.1***	524.5***	556.9***	421.5***	441.6***

*Null rejection at the 10% level; **Null rejection at the 5% level; *** Null rejection at the 1% level.

lation from the disturbances completely eliminates the spatial dependence of the lag variable. Thus, columns 2, 4, and 6 represent the results of the spatial error model (SARE). The results further assure the importance of accounting for spatial effects when analyzing MS across the US. Finally, Eq. 4 was estimated using ML spatial regression and the four socio-demographic factors. The Wald tests reported in Table 7 did not reject the overall significance of the regressors and all variables had the expected sign and were significant, with the exception of parameter λ, which was not significant in column 1. Thus, column 2 represents the results of the SARE model. Medicaid appeared not significant and had the incorrect sign in columns 3 and 4, *i.e.* when controlling for quality indicators. This may suggest that factors other than low income should be considered when trying to understand the quality effects on SMS.

Discussion

The spatial variation of MS on the elderly in the US was investigated by different spatial regression designs A focus of the study was whether, and to what extent, MS varies by service type. This question was addressed by using Medicare data that contain accurate and detailed spending information for nearly all medical services provided. The present study builds on a large body of literature pointing out that an important element explaining cross-sectional variation of per-capita health and social care spending is represented by spill-over or spatial effect. The spending on health and social services in a region can thus have beneficial or harmful effects across a widening geographical area (Revelli, 2002, 2006; Brueckner 2003; Baicker, 2005; Moscone and Knapp, 2005; Moscone *et al*., 2007a; Moscone *et al.*, 2007b; Costa-Font and Pons-Novell, 2007, Rettenmaier and Wang, 2012). Moscone *et al*. (2007b) posited a number of reasons why public authorities are concerned about how their spending decisions compare with those of their neighbours. First, one local authority's good performance may encourage its neighbours to mimic the activities and expenditure patterns associated with such performance. For instance, an authority that has not previously faced high suicide rates may follow the policies of neighbour authorities that have more experience. Then, changes in the organizational-financial structure of any given health or social care system might significantly affect surrounding systems. For instance, the closure of a large psychiatric hospital that accommodates people from a number of geographical areas might impact on the social care sector across regions. Finally, the regionally organized regulatory, inspection and auditing function of central government might offer common guidance across a region that may have influenced certain patterns of activity or expenditure across regions.

Geographical proximity plays an important role in the activities of regions. Recalling the first law of geography, the transfer of knowledge and information translating into a local spending decision, occurs more naturally between neighbouring regions than between those which are far apart, even if this runs counter to the argument that knowledge is a public good which freely diffuses in the economy, as suggested by Arrow (1962). The latter has led to calls for governments to employ centralized incentive mechanisms to diminish territorial inequality in delivering health and social services. Rather than identifying these regions *a priori*, advances in GWR were employed with regional data from the CMS claims data for 2014 estimating what can be thought of as spatially moving clusters of structural covariations (Fotheringham *et al*., 2002). Spatial regression models, that likewise relax the assumption of spatial independence (Anselin, 1988), were also employed. MS was found to be spatially dependent and it was noted that one region's MS does not only depend on its own characteristics, but also on the value of its neighbours' spending.

It was also found that the HR effect is spatially heterogeneous and not always positively correlated with the level of spending or use. In other words, some low-use regions have high-quality indicators and some high-use regions have low-quality indicators. Examining whether HR effects are heterogeneous by service type,

Non-commercial use only

Table 7. Spatial autoregressive model with autoregressive disturbances (SARAR) and spatial error model (SARE) estimates of Medicare spending (SMS) - Equation (4) at the aggregate level, controlling for socio-demographic factors.

Independent variable	i) SARAR (Standard error)	ii) SARE (Standard error)	iii) SARAR (Standard error)	iv) SARE (Standard error)
Age	0.088 (0.679)	0.069 (0.665)	0.556 (0.526)	0.728 (0.616)
Male	–1.029*** (0.223)	–0.898*** (0.201)	–0.869*** (0.177)	–1.108*** (0.206)
African descent	0.019*** (0.005)	0.019*** (0.005)	0.007** (0.004)	0.010** (0.005)
Income	0.036*** (0.012)	0.040*** (0.012)	–0.019* (0.011)	–0.013 (0.013)
HR			0.310*** (0.023)	
HACHF				0.123*** (0.020)
HARD				0.040*** (0.158)
Lambda	–0.012 (0.009)		–0.029*** (0.008)	–0.019** (0.010)
Rho	9.244*** (0.267)	9.229*** (0.279)	8.006*** (0.104)	8.027*** (0.068)
$Sigma^2$	0.005*** (0.000)	0.005*** (0.000)	0.003*** (0.000)	0.004*** (0.000)
Wald Test	106.6***	117.6***	371.3***	196.4***

*Null rejection at the 10% level; **Null rejection at the 5% level; ***Null rejection at the 1% level.

strong heterogeneity of the effects by service type at the aggregate level was found. This result lends support to the view that higher MS is inversely related to higher quality PC services. Although wide regional variation in MS per beneficiary over 65 was found, there was less regional variation in Medicare per beneficiary utilization of services. The data we used were first adjusted for differences in Medicare payment rates (due to regional wages, special payments to teaching hospitals, rural add-on payments). Removing these differences in payment rates is a necessary step to isolate differences in MU.

The methodology utilized here solves some, but not all of the problems associated with the estimation of spatial interdependence among the US regions. The distance measure was not adjusted to account for transportation options and natural barriers, which may vary considerably among the HRRs. Finally, there was the problem of delineation of the bounded areas *i* and *j*. First, the delineation of the bounded area strongly influenced the results, as any change in the area definition would change population characteristics and the quality of care. Secondly, since we used aggregated data and HRRs are geographically large, these factors are potentially misleading since they ignore internal variations of spatial dependence, especially differences between rural and urban areas. It would also be worthwhile to conduct research to determine the most appropriate bandwidth to use in the weighting matrices. This would ideally be the distance beyond which HRRs would not maintain consistent influence upon others. A representative survey indicating the consistency of HRR influences could reveal a better bandwidth. Also, given that burden of transportation probably varies with socioeconomic status and neighbourhood characteristics, it might be that the bandwidth should vary with ecological circumstances.

A number of limitations of this study should be mentioned. Firstly, with the data originating from Medicare, the results may or may not be possible to generalize to other elderly populations. Caution is needed in interpreting the results. Secondly, although it was found that poor quality results in an increase in MS at the aggregate level, it is difficult to conclude that additional spending would not improve quality. To reach such a conclusion, one has to carefully analyze the long-term benefits of additional medical treatment. Regardless though, our results provide valuable information for policy makers on the heterogeneous effects of quality on MS for the elderly.

The discussion above identifies several avenues for methodological improvement. Although readmission rates are promising quality indicators, future research should include other measures of quality, the search for a satisfactory adjustment for MS measure, an adjustment for travel mode options, assessment of the improvements achievable through use of travel time or travel distance, and collection and incorporation of mid-level data.

Conclusions

Analysis of supply-sensitive type of services shows has shown that the effect of hospital readmissions (HRs) are heterogeneous and that their impact varies both spatially and by type of Medicare service. High HRs increase MS and thus the cost per beneficiary at the aggregate level, in particular the use of services rendered by clinics and ambulatory surgical centres and decrease that of outpatients dialysis facilities, physician procedures and tests. Lower quality in terms of admission rates due to congestive heart failure and respiratory disease decreases the spending in ambulatory surgical centres, evaluation and management services, physician procedures, imaging and also drugs. In addition, the effect of health risk status on MS differs by treatment type; specifically, healthier elderly use more hospital outpatient services, clinics, ambulatory surgical centres and durable medical equipment, *i.e.* supply-sensitive types of services. Reduction of the geographic variation in healthcare spending would not necessarily improve the overall quality of healthcare. Although reducing payments to high-spending areas and increasing payments to low-spending areas reduce spending variation, the impact on quality would be ambiguous, since the quality effects tend to be heterogeneous, both spatially and by service. Another important question is whether appropriate treatment for individual patients is compromised when payment-spending is reduced locally by instituting average payment levels.

References

Anselin L, 1988. Spatial econometrics: methods and models. Boston: Kluwer Academic Publishers.
Anselin L, 1990. Some robust approaches to testing and estimation in spatial econometrics. Reg Sci Urban Econ 20:141-63.
Anselin L, Florax R, 1995. Small sample properties of tests for spatial dependence in regression models: some further results in new directions. In: Anselin L, Florax R, eds. Spatial econometrics. New York: Springer. pp 75-95.
Arrow K, 1962. Economic welfare and the allocation of resources to invention. In: Nelson R, ed. The rate and direction of inventive activity: economic and social factors. Cambridge, MA: NBER. pp 609-626.
Brueckner JK, 2003. Strategic interaction among governments: an overview of empirical studies. Intern Reg Scie Rev 26:175-88.
Baicker K, 2005. The spillover effects of state spending. J Pub Econ 89:529-44.
Baicker K, Chandra M, 2004. Medicare spending, the physician workforce, and beneficiaries' quality of care. Health Aff W4:184-97.
Baicker K, Chandra A, 2008. Cooper's analysis is incorrect. Health Aff 28: 116-18.
Cooper RA, 2008a. States with more health care spending have better-quality health care: Lessons about Medicare. Health Aff 28:103-15.
Cooper RA, 2008b. States with more physicians have better-quality health care. Health Aff 28:91-102.
Costa-Font J, Pons-Novell J, 2007. Public health expenditure and spatial interactions in a decentralized health system. Health Econ 16:291-306.
Davis K, Schoen C, 2007. State health system performance and state health reform. Health Aff 26:664-6.
Drukker DM, Egger P, Prucha IR, 2013. On two-step estimation of a spatial autoregressive model with autoregressive disturbances and endogenous regressors. Econometr Rev 32:686-733.
Engle RF, 1983. Wald, likelihood ratio, and lagrange multiplier tests in econometrics. In: Intriligator MD, Griliches Z, eds. Handbook of econometrics. New York, Elsevier. pp 796-801.
Feenberg D, Skinner J, 2000. Medicare transfers across states: winners and losers. Nation Tax J 53:713-32.
Fischer C, Lingsma HF, Marang-van de Mheen P, Kringos DD,

Non-commercial use only

Klazinga NS, Steyerberg EW, 2014. Is the readmission rate a valid quality indicator? a review of the evidence. PLoS ONE 9:1-9.

Fisher E, Goodman D, Skinner J, Bronner K, 2009. Health care spending, quality, and outcomes - more isn't always better. Technical report. Dartmouth Institute for Health Policy & Clinical Practice. Available from: www.dartmouthatlas.org.

Fotheringham AS, Brunsdon C, Charlton M, Geographically weighted regression: The analysis of spatially varying relationships. Chichester, UK: Wiley; 2002.

Fukushima K, Yamamoto S, Lizuka T, 2016. Patient cost sharing and medical expenditures for the elderly. J Health Econ 45:115-30.

Graif C, Sampson RJ, 2009. Spatial heterogeneity in the effects of immigration and diversity on neighborhood homicide rates. Homicide Stu 13:242-60.

Hurvich CM, Simonoff JS, Tsai C-L, 1998. Smoothing parameter selection in nonparametric regression using an improved Akaike information criterion. J Royal Stat Soc 60:271-93.

Lee DS, Tran C, Flintof V, Grant FC, Liu PP, Tu JV, 2003. CCORT/CCS Canadian quality indicators for congestive heart failure care. Canad J Cardiol 19:357-64.

Metropolis N, Ulam S, 1949. The Monte Carlo method. J Amer Stat Assoc 44:335-41.

Moscone F, Knapp M, 2005. Exploring the spatial pattern of mental health expenditure. J Ment Health Pol Econ 8:205-17.

Moscone F, Knapp M, Tosetti E, 2007a. SUR model with spatial effects: an application to mental health expenditure. Health Econ 16:1403-08.

Moscone F, Tosetti E, Knapp M, 2007b. Mental health expenditure in England: a spatial panel approach. J Health Econ 26:842-64.

OECD, 2017. Health care quality indicators - primary care. Technical report. Available from: www.oecd.org/els/health-systems/hcqi-primary-care.htm. Accessed: November 2017.

Rettenmaier AJ, Wang Z, 2012. Regional variations in medical spending and utilization: a longitudinal analysis of US Medicare population. Health Econ 21:67-82.

Revelli F, 2002. Testing the tax mimicking versus expenditure spillover hypotheses using English data. Appl Econ 34:1723-31.

Revelli F, 2006. Performance rating and yardstick competition in social service provision. J Pub Econ 90:459-75.

Skinner J, Chandra A, Goodman G, Fisher EF, 2008. The elusive connection between health care spending and quality. Health Aff 28:W119-23.

Skinner J, Wennberg J, 2000. Regional inequality in Medicare spending: The key to Medicare reform? In: Garber A, eds. Frontiers of health economics. Cambridge, MA: MIT Press. pp. 69-90.

Skinner J, Wennberg J, 2005. The efficiency of Medicare. In: Wise DA, eds. Analyses in the economics of aging. Chicago, IL: University of Chicago Press. pp. 129-160.

Sutherland JM, Fisher ES, Skinner JS, 2009. The high cost of health care in the United States. N Eng J Med 361:1227-30.

Tobler W, 1970. A computer movie simulating urban growth in the Detroit region. Econ Geogr 47:234-40.

Wennberg JE, Cooper MM, 1999. The quality of medical care in the United States: A report on the Medicare program. Technical

Non-commercial use only