Rural isolation and the availability of hospital services

James W. Henderson, Beck A. Taylor*

Department of Economics, Baylor University, Waco, TX 76798-8003 USA

Abstract

This study empirically examines some important factors affecting the geographic distribution of hospitals. Using 1996 cross-sectional data on hospital locations in Texas, we employ a count-data methodology to estimate the impact of demand and rural isolation on the frequency of hospital service in a given geographical area. Using the model's estimates, we calculate the population thresholds needed to support a given frequency. Results suggest that population, population density, per capita income, and rural isolation are important factors in determining the number of hospitals in an area. Our methodology and results provide a means to inform and evaluate policy decisions regarding the designation of medically underserved areas.

Keywords: Rural hospitals; Demand thresholds

1. Introduction

Access to quality health care is a continuing challenge for most rural communities. The availability and accessibility of health services largely determine their adequacy in rural areas (Comer and Mueller, 1995; Edelman and Menz, 1996). The presence of a hospital is an important factor in determining the adequacy of health-related services in any community, rural or urban. A community without a hospital not only lacks local access to vital health services, but the presence of a hospital is an important factor in the location decision of many physicians, especially those specialists who rely on the facility for the provision of key inputs. Recent changes in US health care delivery—improved transportation services, the phenomenal growth of outpatient procedures, and the increased use of technology—have served to increase the already sizeable gap in the availability of health services between rural and urban areas.

This study examines the geographic distribution of Texas hospitals in 1996. Our objectives are to isolate critical determinants of the geographical distribution of hospitals and measure important tradeoffs among them. Of particular interest are the measurable tradeoffs between rural isolation (or, conversely, urban proximity) and populations needed to support a given frequency of hospitals. Another important tradeoff is the potential substitution between rural isolation and a population’s per capita income in determining the spatial distribution of hospitals. To this end, we relate the frequency of hospitals across all 254 Texas counties to county-level demographics, such as population, population density, and per capita income. We also estimate the effect of rural isolation on the likelihood that a hospital is located in a specific county. Using these estimates, we then calculate the population thresholds needed to support a given frequency of hospitals. Our results indicate that counties more geographically isolated from urban centers require larger populations to support a hospital, other things equal, and that higher per capita income can indeed serve as an important substitute for urban proximity when considering the availability of hospital services.

The paper is organized as follows. In the second section we identify and discuss the challenges of providing health care in rural settings. We present the theory and methods used in the study in Section 3. In Section 4, we describe the data used in the study, and our empirical results are presented in Section 5. A discussion of the results and their implications are provided in Section 6, followed by conclusions in the final section. An appendix developing the statistical model is also included.
2. Challenges in rural health care

Promoting access to medically necessary care is just one of the many challenges facing health care delivery systems. The logistics of access are particularly challenging in the US due to the geographic dispersion of the population. Using the US Census Bureau’s definition, more than 70 percent of all US counties are rural. These rural areas comprise one-fourth of the US population and are home to only 10 percent of all active physicians (Medicare Payment Advisory Commission, 2002). Rural residents are more likely to be poor, elderly, and in poorer health than their urban counterparts. Rural areas have fewer hospitals, fewer physicians, and fewer higher-ordered health services than urban areas (Henderson et al., 2000). As a result, rural residents have fewer physician contacts per year and experience more difficulty in gaining access to care than urban residents.

Maintaining adequate health care services in isolated areas has not been particularly easy. The evidence seems to indicate that rural hospitals are more vulnerable to competitive pressures. The risk of closure is much higher for rural hospitals since they tend to be smaller in both size and capitalization, their facilities on average are older, and they have limited opportunities to form strategic alliances with other hospitals (Morrissey et al., 1991; Succi et al., 1997). Lower occupancy rates and a case-mix skewed toward less complex medical conditions also contribute to the higher risk of closure (Lillie-Blanton et al., 1992).

In the 1980s, more than 330 rural US hospitals closed their doors. California and Texas had the largest number of closures during that decade with almost one-fourth of all US closures occurring in Texas. While the number of closures slowed during the 1990s, rural areas still lost 186 hospitals during the decade. By 2000, many states had lost more than 10 percent of their rural hospitals that were in existence in 1990 (Poley and Ricketts, 2001). The recent and relatively slower rate of rural closures is encouraging but may actually mask a crisis that still exists in rural areas. Data from the Medicare Payment Advisory Commission (2002) indicate that over one-third of all rural hospitals reported negative operating margins in 1999.

Americans spent more than $412 billion on hospital services in 2000, representing one-third of all US health care spending and 4.2 percent of US gross domestic product. Community hospitals providing short-stay acute care were by far the most common classification with approximately 5900 nationwide. Primarily a market-driven system, the hospital industry in the US is nonetheless dependent upon government sources for the majority of its revenues. Government payments for inpatient and outpatient services are made prospectively and adjusted for both disproportionate share payments that provide extra funding for those hospitals with a disproportionate share of low income and indigent patients, and medical education payments intended to cover Medicare’s share of the cost of running medical residency programs. These adjustments in part contribute to a differential in inpatient payment rates between rural and urban hospitals, resulting in Medicare operating margins of 5 percent for rural hospitals and 14 percent for large urban hospitals in 2002.

Efforts to improve the prospects for the survival of rural facilities are vitally important since access to a nearby hospital is a critical factor in determining the health status of a community (Robst and Graham, 1997). Rural communities must overcome a number of competitive disadvantages. Differences in socio-economic, demographic, and environmental characteristics influence the nature and severity of the health problems experienced in rural areas. Longer commuting distances lead to slower access to critical care, particularly emergency and obstetric care, and may be a prime contributor to higher infant mortality rates and death rates from accidents in rural communities.

Surprisingly, very little research has examined the factors that influence the geographic distribution of hospitals—that is, the hospital location decision. The methodology and results presented in this paper will better inform both hospital administrators and policy makers in a number of ways. Most importantly, the proposed methodology can assist federal and local policy makers in their efforts to improve access to, and availability of, hospital services, and provide an evaluation tool that can be used to assess the effectiveness and appropriateness of existing policies. In addition, a hospital is a business, and a growing hospital adds to the economic health of a community. Quite often the hospital located in a rural area is the first or second largest employer in the community. Its very existence enters into business location decisions and the loss of a hospital can have a profound negative impact on the employment and income in a rural community for years after its closure. In this sense, local community developers and health care administrators may use population threshold estimates like those presented here to coordinate public and managerial policy regarding hospital location decisions.

3. Theory and methods

The underlying methodology for our study is drawn largely from central place theory (CPT) (Christaller, 1933; Lösch, 1940; Eaton and Lipsey, 1982; Berry and Parr, 1988). CPT predicts that the size of the market for a good or service depends on the geographic distribution of consumers, the transportation costs borne by consumers, and any cost savings associated with the production and distribution of the good or service.
In our context, geographic variations in demand give rise to geographic market areas for hospitals. For example, if it is assumed that potential hospital customers are uniformly distributed across geographical space and that transportation costs are proportional to the distance traveled, the geographic distribution of hospitals would depend solely on the per patient cost savings due to hospital size (i.e., economies of scale). Service providers with lower average costs of production could offer a total price (including transportation costs) that would induce more of the population to buy their product. Thus, these low-cost providers would have larger market areas.

In order for a community to support an additional hospital, traditional CPT states that an increase in demand in the local market area is required. An increase in demand can result from an increase in population, higher population density within the existing geographic service area, changes in consumer income, and/or an increase in a facility’s potential service area due to favorable changes in transportation costs. In order for a hospital to be provided, demand must be at least sufficient to cover the average cost of production. This requirement is known as the profitability constraint. Even in the world of not-for-profit hospitals, revenues must exceed expenditures for the operation to remain solvent. Finally, the minimum market size required to support a particular service is called the demand threshold. Demand thresholds are most often measured by estimating the population needed to support a given level of service (Berry and Garrison, 1958a,b; Brunn, 1968; Murray and Harris, 1978; Salyards and Leitner, 1981; Schuler and Leistritz, 1990; Deller and Harris, 1993; Harris and Shonkwiler, 1993; Harris et al., 1996; Wensley and Stabler, 1998; Henderson et al., 2000).

Thus, throughout the remainder of this paper, the terms demand threshold and population threshold can be used interchangeably.

The theory states, therefore, that the frequency of hospitals observed in a geographic area is a function of demand patterns and factors that affect transportation costs, particularly rural isolation. Population, population density, and per capita income are generally used to measure demand. Distance to the closest urban center is used to measure rural isolation. Such a measure is used almost exclusively in the extant literature when estimation of demand thresholds is desired. Certainly other proxies for “ruralness” or urban proximity can be devised, but a community’s distance from the closest major population center is a reasonable and tractable measure. Limiting the analysis to a dichotomous distinction between rural and non-rural, for instance, will also limit the power of the analyses to follow.

Other things equal, the observed number of hospitals is expected to increase as population increases. Holding population constant, an increase in population density is expected to result in a decrease in the number of hospitals. As per capita income rises, the number of hospitals observed is expected to rise. The impact of rural isolation on the number of hospitals is less clear. CPT predicts that more geographically isolated counties require a smaller population to support a given number of hospitals. Taking into consideration high transportation costs to distant urban areas, the convenience of local services offsets any cost disadvantage due to smaller size. Thus, rural areas should have a higher frequency of hospitals than urban areas, other things equal.

The relationship between population and rural isolation may be more complex, however, than the predictions of CPT. Reductions in the average cost of providing service, because of the clustering of services in an urban area, can play a significant role (Henderson et al., 2000). These agglomeration economies may be due to shared resources, lower information costs, and proximity of suppliers, customers, and transportation facilities, and are expected to be present to a greater extent as a geographic area becomes more urbanized. If these cost savings are sufficiently large, we can expect urban areas to have a higher frequency of service providers after controlling for other factors.

4. Data

In order to examine the geographic distribution of hospitals, we utilize the detailed hospital data collected by the Texas Department of Health (TDH) in its 1996 Annual Survey of Hospitals (Texas Department of Health, 1996). The 1996 TDH survey includes detailed information from each of the 484 non-federal hospitals in Texas. The survey was mandatory and there was a 100 percent response rate. Data include information on hospital characteristics such as location, organizational structure, utilization, financing, and staffing. The TDH data recognize only those facilities that are designated as hospitals by the TDH. The basic (lowest-ordered) level of service is defined as general medical and surgical care, though hospitals in Texas are differentiated on a range of services including pediatrics, obstetrics, intensive care, emergency and trauma services, cardiac care, neonatal care, and oncological services, with higher-ordered, more specialized services offered only at hospitals in major urban centers. The analysis to follow generalizes from this level of differentiation and concentrates on only the basic level of service, though other researchers have examined the determinants of higher-ordered hospital services (see, e.g., Henderson et al., 2000).

Because hospitals draw customers from a potentially large geographic area, we use the 254 counties in Texas as our units of analysis. Aggregating within each of the
254 counties results in a hospital frequency variable that is county specific. Counties in Texas are diverse. Table 1 reports descriptive statistics on civilian population, population density (population per square mile), per capita income, and distance from each county to the closest standard metropolitan statistical area (SMSA) (US Census Bureau, 1996). For instance, civilian population ranges from 97 persons in Loving County to approximately 3.1 million persons in Harris County (where Houston is located). A similar dispersion characterizes population density. Loving County, for instance, has only 0.14 persons per square mile, while Dallas County contains 2273 persons per square mile. Finally, the average distance between a county and the closest SMSA is approximately 47 miles.

Table 1 also presents county demographics across counties with the same number of hospitals. For example, 61 counties in Texas do not have a hospital, while 129 counties have only one. Only 16 counties have more than four hospitals, and only five have more than six. The average population for a county without a hospital is 7264. As Table 1 reports, the average county population increases as the number of hospitals increases. Likewise, population density increases as the number of hospitals increases. Counties with more hospitals have higher per capita incomes and counties that are farther from urban areas have fewer hospitals.

5. Results

5.1. Parameter estimates

In our empirical model, we estimate the relationship between the observed frequency of hospitals in a county to the county’s civilian population, population density (per square mile), per capita income, and distance to the closest SMSA. Since the dependent variable takes on only nonnegative integer values, a Poisson distribution is assumed. Using maximum-likelihood procedures, we relate the dependent variable to the exogenous variables through a logarithmic link function. Additionally, we control for the overdispersion caused by a prevalence of zeros in our count data using a zero-inflated Poisson regression technique. A detailed description of our empirical model is available in the Appendix.

Parameter estimates and robust standard errors are reported in Table 2. The Pearson (goodness-of-fit) $\chi^2$ statistic is statistically insignificant, thus failing to reject the null hypothesis that the data are distributed Poisson. Additionally, a Wald $\chi^2$ statistic, measuring the joint significance of the right-hand side variables, is statistically significant at better than the 1-percent level.

As Table 2 reports, county population is a significant predictor of the number of hospitals across counties.
The estimated population elasticity of 0.858 indicates that when all other control variables are held at their mean values, a 10-percent increase in population contributes to an 8.58 percent increase in the number of hospitals.

Population is not the only determining factor in hospital location decisions. Population density, per capita income, and distance from the nearest SMSA also play important roles in determining the number of hospitals in a given county. The estimated coefficient on population density indicates that, other things equal, a 10-percent increase in the number of persons per square mile decreases the number of hospitals 4.04 percent. Recall that, holding population constant, an increase in the size of the service area decreases population density. An increase in service area should lead to a higher frequency of service providers. Finally, a 10-percent increase in per capita income leads to a 10.03 percent increase in the number of hospitals, other things equal.

A closer examination of the estimated coefficients capturing the relationship between the number of hospitals and the distance to the closest SMSA is warranted. Recall that rural isolation will increase the demand for hospitals because of increased transportation costs to consumers. This, alone, leads to the (traditional) prediction of lower population thresholds and higher frequencies of hospitals in rural areas, other things equal. This is not what our results reveal, however. The negative coefficient on the distance variable indicates that more geographically isolated counties have fewer hospitals, other things equal. This result is consistent with the existence of substantial agglomeration economies, that is, cost-savings due to providers clustering in more urban areas.

5.2. Population thresholds

Using the estimation results from Table 2, it is relatively straightforward to calculate population thresholds. These thresholds are calculated by substituting any number of hospitals for the dependent variable and solving for the population that satisfies the equation, holding all other variables at their mean values. Table 3 presents the population thresholds required for the existence of one hospital for counties located at various distances from the closest SMSA. The distances chosen are the mean distance across all counties from Table 1 and distances that are ±1 standard deviation from the mean. The population density and per capita income from the average county are assumed. As Table 3 reports, the population needed to support one hospital increases as the distance to the nearest SMSA increases. In other words, counties located farther from metropolitan areas have higher population thresholds. Note, however, the small trade-off between population and distance. In fact, approximately the same population is needed to support one hospital for counties located 50 or 150 miles from a SMSA. The small tradeoff between the population and urban proximity required to support one hospital is better illustrated in Fig. 1.

Another interesting tradeoff occurs between the population and per capita income required to support a hospital. This tradeoff is illustrated in Fig. 2. Note that as a given county’s population increases, the per capita income required to support one hospital decreases. Additionally, note that there is a greater rate of substitution between population and per capita income than between population and urban proximity.

Table 4 presents population thresholds needed to support a particular number of hospitals for the average county. The theory pertaining to establishment multiplication has traditionally focused on the incremental population required to support an increasing number of establishments. It is generally assumed that average costs decrease as the level of demand increases at a given location (Parr and Denike, 1970). Given these expected cost-savings, doubling the number of hospitals would require more than a doubling of population. That is, the

| Table 2 | Model estimation results
| Parameter estimates and (robust std. errors) |
|---|---|
| Constant | −16.727*(2.504) |
| Population | 0.858* (0.125) |
| Population density | −0.404* (0.118) |
| Per capita income | 1.003* (0.214) |
| Distance to closest SMSA | −0.083* (0.029) |

| Table 3 | Population required to support one hospital at various distances
<table>
<thead>
<tr>
<th>Distance from closest SMSA (miles)</th>
<th>Population threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>19.1</td>
<td>32,696</td>
</tr>
<tr>
<td>46.9</td>
<td>35,675</td>
</tr>
<tr>
<td>74.7</td>
<td>37,326</td>
</tr>
</tbody>
</table>

*These population threshold estimates assume demographics from an average Texas county in 1996 (see Table 1). The chosen distances are the mean distance across all counties and distances which are ±1 standard deviation from the mean.
elasticity of the frequency of hospitals with respect to population is expected to be less than unity. Our estimates indicate a decreasing rate of multiplication (larger population increments are required to support an additional hospital).

In summary, our results suggest that (i) population is the most significant determinant of the number of hospitals in a county; (ii) as a county’s distance from an urban center increases, population thresholds increase resulting in a lower frequency of hospitals in outlying areas, other things equal, though this tradeoff is not large; (iii) the tradeoff between income and the population needed to support a given frequency of hospital service is much larger than the tradeoff between urban proximity and population; and (iv) there exists a decreasing rate of establishment multiplication implying that ever larger population increments are required to support additional hospitals.

**Table 4**

<table>
<thead>
<tr>
<th>Number of hospitals</th>
<th>Population thresholds</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>35,675</td>
</tr>
<tr>
<td>2</td>
<td>80,027</td>
</tr>
<tr>
<td>3</td>
<td>128,374</td>
</tr>
<tr>
<td>4</td>
<td>179,514</td>
</tr>
<tr>
<td>5</td>
<td>232,837</td>
</tr>
</tbody>
</table>

*aThese population threshold estimates assume demographics and location of an average Texas county in 1996 (see Table 1).*
6. Discussion

There is a widely held belief that certain groups of Americans are being underserved by the US health care system. These underserved groups include the poor, the elderly, and people living in rural and inner-city areas. Ironically, as evidenced by a shortage of service providers and facilities in these areas, the adequacy of medical care for these groups may have more to do with the proximity and availability of medical care services than, say, access to health insurance. A growing concern for policy makers is the problem of providing medical care in isolated areas. Almost one-third of the US population lives in market areas with less than 180,000 inhabitants. In fact, a majority of the population in 19 US states lives in these small-market areas (Kronick et al., 1993). Inner cities in the US face the same challenges in attracting medical care services and maintaining access to those services.

Recall that 61 Texas counties did not have a hospital in 1996. Of this number 43 counties had less than 10,000 inhabitants, well below the estimated population threshold needed for one hospital reported in Table 4. However, of the 129 counties with one hospital, 45 had less than 10,000 inhabitants. While population is not the only determinant in the hospital location decision, it is critical to understand why some of these less inhabited counties have hospitals and others do not. Of primary interest is why a county that appears to have the necessary demographic characteristics to support a hospital does not have one (call this Type-I error), and why a county that does not seem to merit a hospital, given its demographic characteristics, has one (call this Type-II error).

Table 5 reports information on the actual and predicted values for the 39 counties in Texas where the difference between the actual number of hospitals and the predicted number was greater than 1. Using the estimates from the model in Table 2, counties where the actual number of hospitals exceeded the predicted number by more than one were designated as overserved. Those where the predicted number exceeded the actual number by more than one were designated as underserved.

Only three of the 61 counties without a hospital have predicted values greater than one, implying Type-I error of less than 5 percent. Dallam county, for example, with a population of 5765, does not meet the population threshold for one hospital, but a high per capita income of $29,861 places it in the underserved category. Seventy-eight of the 129 counties with one hospital have predicted values of less than one, indicating a Type-II error rate of approximately 60 percent. Populations in these overserved counties range from 2505 in

<table>
<thead>
<tr>
<th>Underserved counties</th>
<th>Actual number of hospitals</th>
<th>Predicted number</th>
<th>Overserved counties</th>
<th>Actual number of hospitals</th>
<th>Predicted number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bell</td>
<td>3</td>
<td>4.98</td>
<td>Cameron</td>
<td>5</td>
<td>3.61</td>
</tr>
<tr>
<td>Collin</td>
<td>3</td>
<td>6.88</td>
<td>Cass</td>
<td>3</td>
<td>1.21</td>
</tr>
<tr>
<td>Comal</td>
<td>1</td>
<td>2.04</td>
<td>Colorado</td>
<td>3</td>
<td>1.14</td>
</tr>
<tr>
<td>Dallam</td>
<td>0</td>
<td>1.16</td>
<td>Comanche</td>
<td>2</td>
<td>0.84</td>
</tr>
<tr>
<td>Denton</td>
<td>4</td>
<td>5.09</td>
<td>Frio</td>
<td>2</td>
<td>0.60</td>
</tr>
<tr>
<td>Ector</td>
<td>2</td>
<td>3.28</td>
<td>Gregg</td>
<td>3</td>
<td>1.86</td>
</tr>
<tr>
<td>Ellis</td>
<td>1</td>
<td>2.57</td>
<td>Hardeman</td>
<td>2</td>
<td>0.48</td>
</tr>
<tr>
<td>Fort Bend</td>
<td>3</td>
<td>4.82</td>
<td>Harris</td>
<td>40</td>
<td>30.24</td>
</tr>
<tr>
<td>Hardin</td>
<td>1</td>
<td>2.21</td>
<td>Jefferson</td>
<td>7</td>
<td>3.91</td>
</tr>
<tr>
<td>Johnson</td>
<td>1</td>
<td>2.25</td>
<td>Jones</td>
<td>3</td>
<td>0.80</td>
</tr>
<tr>
<td>Kendall</td>
<td>0</td>
<td>1.37</td>
<td>Montague</td>
<td>2</td>
<td>0.97</td>
</tr>
<tr>
<td>McLennan</td>
<td>3</td>
<td>4.97</td>
<td>Potter</td>
<td>5</td>
<td>3.64</td>
</tr>
<tr>
<td>Midland</td>
<td>2</td>
<td>5.21</td>
<td>Runnels</td>
<td>2</td>
<td>0.88</td>
</tr>
<tr>
<td>Montgomery</td>
<td>3</td>
<td>4.41</td>
<td>Tarrant</td>
<td>18</td>
<td>13.27</td>
</tr>
<tr>
<td>Parker</td>
<td>1</td>
<td>2.15</td>
<td>Upton</td>
<td>2</td>
<td>0.56</td>
</tr>
<tr>
<td>Randall</td>
<td>1</td>
<td>2.79</td>
<td>Wheeler</td>
<td>2</td>
<td>0.79</td>
</tr>
<tr>
<td>Taylor</td>
<td>2</td>
<td>3.93</td>
<td>Williamson</td>
<td>3</td>
<td>1.10</td>
</tr>
<tr>
<td>Tom Green</td>
<td>3</td>
<td>4.37</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Travis</td>
<td>6</td>
<td>11.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upshur</td>
<td>0</td>
<td>1.50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Webb</td>
<td>2</td>
<td>4.65</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wilson</td>
<td>1</td>
<td>2.92</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Fitted values for each county were obtained by estimating a model similar to the model estimated in Table 2 that allows for differential intercepts and slopes for each federal MUA designation. Results are available upon request.

* A county is designated underserved if the county’s predicted frequency of hospitals is greater than the county’s observed frequency by more than 1.

* A county is designated overserved if the county’s predicted frequency of hospitals is less than the county’s observed frequency by more than 1.
Calhoun County to 49,206 in Starr County. (Starr County most likely receives the overserved designation due to a low per capita income.)

It is not our intention to imply that a county must have a predicted value of one before it merits a hospital. A critical value of 0.5 instead of 1.0 would change the implications of the model by increasing Type-I error and reducing Type-II error. For example, 23 of the 61 counties without a hospital have a predicted value greater than 0.50. This translates into a 38 percent Type-I error rate. Only 20 of the counties with one hospital have a fitted value less than 0.5, accounting for a 15 percent Type-II error.

The model accurately predicts the actual number of hospitals for most counties. The predicted number of hospitals is within ±1 of the actual number in 215 of the 254 counties, or almost 85 percent. Of the 39 other counties, the predicted value is greater than the actual value in 22 cases resulting in the county being designated as underserved (see left-hand side of Table 5). For underserved counties the differential ranges from 1.04 in Comal County to 5.00 in Travis County. In 17 cases the actual number of hospitals is greater than the predicted number leading to the county being classified as overserved (see right-hand side of Table 5). The differential in overserved counties ranges from 1.03 in Montague County to 9.75 in Harris County.

In results not presented here for brevity, logit analysis reveals that the most important predictors determining whether a county receives the underserved designation by our standards include population, per capita income, and elderly population. In particular, counties with a higher population, lower per capita income, and a smaller percentage of elderly persons were significantly more likely to be underserved, other things equal. With respect to the overserved designation, counties with a larger population of elderly persons were significantly more likely to have a greater number of hospitals than the model predicted, other things equal. These results point to the potentially important influence of the age demographic in determining the level of hospital services rendered. Potential explanations for this include both the greater level and increased intensity of health care services demanded by the elderly.

While we argue that this methodology can shed some light on the determinants of hospital location, we caution that care must be taken in interpreting the size and sign of the estimated residuals. When the model's predictions do not accurately reflect the actual number of hospitals in a county, this does not indicate that planners and policy makers "got it wrong." Part of the differential may be due to differences in the market share enjoyed by managed care in the various counties. Given the large percentage of counties in Texas where the predicted number of service providers closely approximates the actual number, we feel confident in the model's quantitative results.

7. Conclusion

The availability of health services in rural communities is an important policy issue. The increased emphasis on managed care in the US and the popularity of a network model that utilizes contracts with a number of different providers (including individual physicians, physicians' practices, hospitals, and pharmacies) to make a full range of medical services available to its enrollees has made it increasingly difficult for rural health systems to adapt to the changing environment. A combination of rural isolation and low overall demand for health services challenge rural providers to offer an attractive package of services to complement existing networks.

Economic developers and local, state, and federal policy makers all look to the configuration of health care delivery systems when making decisions regarding the provision of medical services. This paper examines the parameters that define the geographic distribution of hospitals. Using 1996 cross-sectional data on the location of hospitals in Texas, several key results emerge. First, population, population density, per capita income, and rural isolation all play a significant role in the observed differences in the observed frequency of hospitals across counties. Second, to support a given frequency of hospitals, the tradeoff between the needed level of population in a county and that county's distance to an urban center is small. That is, estimated population thresholds do not change substantially across counties that vary in their proximity to urban areas. On the other hand, our results suggest that rural counties with high per capita income can overcome their rural isolation with respect to the existence of hospital services. Finally, population thresholds are smaller for urban areas, leading to an increased frequency of hospitals, even after controlling for differences in population.

The empirical model can be used to address a number of issues. One application is the examination of federally designated medically underserved areas (MUAs). A lot is at stake. With billions of dollars in public and private funds to be distributed, the health care shortage designation is particularly important for rural communities. Because a significant contributor to the federal designation of MUAs is the existence of, and access to, medical services, the actual frequency of hospitals can be compared with the model's predicted frequency and policy inferences can be drawn. We point out, however, that access to medical services is not the only determinant of MUA status. Policymakers and rural health care advocates also consider equity and medical
efficacy issues when designating areas for underserved status. Such analysis is beyond the scope of this paper. A more thorough analysis of these determinants is warranted and remains a topic for further research.

Acknowledgements

We are grateful to Steve Green, Tom Kelly, Chuck North, and two anonymous referees for their helpful comments on earlier drafts. A portion of Taylor’s work was completed while a visiting scholar at Harvard University. Any remaining errors are our sole responsibility.

Appendix

Empirical model

In specifying the empirical model, we must accommodate the nature of our dependent variable (the frequency of hospitals within a particular county). Given the nonnegative discrete nature of the frequency data, a count-data technique was adopted. Assuming thenonnegative discretenatureofthefrequency frequency of hospitals within a particular county).

The influence that county population, population density, per capita income, and distance from closest SMSA have on the expected frequency of hospitals was specified in a generalized linear model. We specify a double-log link function that recognizes the necessity of a nonnegative \( \lambda \) estimate and takes into account the traditional logarithmic relationship between demand parameters (specifically, population) and the observed frequency of service providers (Wensley and Stabler, 1998). The dependent variable (\( \lambda \)) is related to the exogenous variables through a logarithmic link function because parameter estimates of a linear combination of independent variables can result in negative values even though \( \lambda \) is known to take on only nonnegative values (Congdon, 1989). Thus, the relationship between the expected frequency of service in county \( i \) and the exogenous variables takes the form

\[
\ln \lambda_i = \alpha + \beta \text{Population}_i + \gamma \text{Density}_i + \delta \text{PCI}_i + \phi \text{Distance}_i,
\]

where \( \text{Population} \) is the civilian population in county \( i \), \( \text{Density} \) is the population density (per square mile) in county \( i \), \( \text{PCI} \) is the per capita income in county \( i \), and \( \text{Distance} \) is the distance between county \( i \) and its closest SMSA.

Two econometric issues must be addressed before estimating the model. First, one potential problem when using the Poisson distribution is the possibility of overdispersion (underdispersion). The Poisson distribution assumes equality between the conditional mean and variance of the dependent variable. Overdispersion (underdispersion) exists when the conditional variance of the dependent variable is greater than (less than) the conditional mean. Overdispersion (underdispersion) is essentially a form of heteroskedasticity. While parameter estimates are not biased under either situation, the standard errors of the estimates may be biased resulting in unreliable inference. The existence of overdispersion (underdispersion) and the appropriateness of using the Poisson distribution can be tested using Pearson’s generalized \( \chi^2 \) test. If overdispersion (underdispersion) exists, the most common practice is to change the distributional assumption to the negative binomial distribution that effectively controls for the heteroskedasticity in the data.

Second, it may be the case that the zero outcomes in our dependent variable are qualitatively different from the positive outcomes. In other words, there may be two regimes at work in the data generating process, and we would like to control for the prevalence of zero counts in the data. Mullahy (1986) argued that this regime-splitting process constituted a shortcoming of the Poisson model and suggested the use of a “hurdle” model as an alternative. Following Lambert (1992), we employ a zero-inflated Poisson (ZIP) model. The ZIP model allows for overdispersion via the introduction of a splitting process which models the outcomes as zero or nonzero. In this formulation, a binary probability (logit) model determines whether a zero or nonzero outcome occurs, then, in the latter case, a (truncated) Poisson distribution describes the positive outcomes. A likelihood-ratio test comparing the ZIP Poisson model with the (nested) Poisson model will indicate whether the zero or nonzero outcome occurs.
the log-likelihood function is written as
\[
\ln L = \sum_{i \in S} \ln \left[ F(\zeta_i^S) + (1 - F(\zeta_i^S)) \exp(-\lambda_i) \right] \\
+ \sum_{i \notin S} \left[ \ln(1 - F(\zeta_i^S)) - \lambda_i + y_i \gamma_i^S - \ln(y_i!) \right],
\]
where
\[
\zeta_i^S = \beta X_i \text{ and } \zeta_i^S = \gamma Z_i,
\]
\( F \) is the logit link function, and \( S \) is the set of observations for which the outcome \( y_i = 0 \). The vector \( X \) contains the covariates listed above that determine the expected frequency of services, and the vector \( Z \) contains the covariates that will be included in the logit model determining the zero and nonzero outcomes in the dependent variable. For our purposes, \( \text{Population} \) and \( \text{Distance} \) were included in \( Z \).

References