



Use of Elective Surgery in Public hospitals: Modeling Access-Cost Quality Trade-offs in a Spatial Framework

Dr Anurag Sharma

Research Fellow, Centre for Health Economics,
Monash University

Associate Professor Anthony Harris

Deputy Director, Centre for Health Economics,
Monash University

Professor Jeff Richardson

Foundation Director, Centre for Health Economics,
Monash University

March, 2008

Centre for Health Economics

ISSN 1833-1173

ISBN 1 921187 25 5

Use of Elective Surgery in Public Hospitals: Modeling Access-Cost Quality Trade-offs in a Spatial Framework*

Anurag Sharma[†] Anthony Harris[‡] Jeff Richardson[§]

March 2008

All correspondence to:

Anurag Sharma

Centre for Health Economics,

Monash University,

Melbourne, VIC 3800,

Australia

Fax: + 61 3 9905 8344

Phone: +61 3 9905 0519

Email: anurag.sharma@buseco.monash.edu.au

*This paper is part of a project supported by NHMRC Grant (ID:334114). We are thankful to participants of CHE internal seminar, AHES06 conference, iHEA 2007 Congress and the 1st World Conference of the Spatial Econometrics Association, Cambridge 2007 for useful comments. The usual disclaimer applies.

[†]Anurag Sharma, Centre for Health Economics, Monash University, Clayton Campus, VIC 3800. Email: Anurag.Sharma@Buseco.monash.edu.au

[‡]Anthony Harris, Center for Health Economics, Monash University, Clayton Campus, VIC 3800. Email: Anthony.harris@Buseco.monash.edu.au

[§]Jeff Richardson, Centre for Health Economics, Monash University, Clayton Campus, VIC 3800. Email: Jeff.Richardson@Buseco.monash.edu.au

Abstract

The overall objective of the paper is to model and econometrically analyze the impact of access costs (travel time to hospital) and quality of health care on the utilization of elective health services in public hospitals. We argue that patients might face a trade-off between better perceived quality of care and access costs. The extant literature has not yet developed a common framework which explicitly incorporates the quality and access trade-off. The first aim of this paper is to help fill this gap. We propose a stylized model where GPs and Specialists act as gatekeepers, waiting times act as a rationing device; and perception of quality and access costs contributes to the choice of hospital for treatment. A secondary objective of the paper is to explore econometric approaches to simultaneously deal with access cost quality tradeoff and its effect on patient flows across regions. The geographic access costs lead to interaction between regions which is termed as spatial dependence. This is econometrically tested by applying spatial regression techniques focussing on spatial panel models recently proposed but not yet been widely applied to health economics. The results show that spatial effects especially the geographic neighborhood effects significantly affect the hospital utilization rates at a regional level. Travel time is found to have significant and negative effect on hospital utilization for some Diagnostic categories. The effect of quality of care (measured by the rate of adverse events) is negative and significant for one category of separations. However the effect is quantitatively small. We do not find any evidence of trade-off between quality and travel time for all category of separations. Policy implications are discussed.

JEL Classification: *I10, R53, C21, C33*

Keywords: *Quality, Access Costs, Spatial Panel Regression, Patient Choice, Elective Care, Hospital Utilization*

1 Introduction

The empirical analysis of quality of care and travel time on hospital utilization has not received widespread attention in the literature and the work that has been undertaken has not been based upon explicit formulation of individual behavior. The modeling of individual utility maximizing individuals will be influenced by their perception of quality of care, financial affordability and costs associated with waiting times and access to hospital. These factors are likely to interact. For example, an individual might prefer a hospital solely on the basis of her perception of quality even if this involves longer waiting times or higher access costs. Thus the individuals might face trade-off between perceived quality of care and these costs.

The extant literature does not seem to have yet developed a common framework which explicitly incorporates quality and access cost trade-offs. The first aim of this paper is to partially fill this gap. We propose a stylized model of the interaction between individuals seeking health care, GPs and Specialists, Hospitals and Government. The main properties of the model are that GPs and Specialists act as gatekeepers (Scott, 2000; Dusheiko et al., 2005), waiting times act as a rationing device (Lindsay and Feigenbaum, 1984; Cullis and Jones, 1986; Martin and Smith, 1999; Ackere and Smith, 1999) and that waiting times reconcile the demand and supply of hospital care; and that perception of quality and access costs contribute to the choice of hospital for treatment. The advantage of such a formalization is that it allows the derivation of testable restrictions on observable behavior (utilization of hospital care) and the use of reduced form equations to estimate parameters which may subsequently be used for empirical and policy analysis.

A second aim of the paper, is to examine the “*spatial distribution*”, of attributes and determine how quality and accessibility costs affect hospital utilization rates. The issue is important for policy. For example, attempts to improve the quality of care are more likely to be effective if the “target populations” are responsive to quality in their choice of provider. Similarly, the response to access costs is important for the effectiveness of policies attempting to achieve similar service use by individuals with similar needs.

Duckett (2000) argues that the quest for equity has been a major issue in the Australian

health care system and one of the main aims of Medicare was, and still is, to ensure fair access to services for the entire population and, in particular, by the provision of a universal, free, public hospital service. This, nevertheless, still varies significantly by location, socio-economic status (SES) and by type of procedure. This results in significant variation in the often long wait lists for elective surgery. Small area studies of variation in the standardized use of particular services indicate that very significant inequities still exist in access to services (Richardson, 1998). In particular there are significant differences between urban and rural areas in access to primary and hospital health care (Duckett, 2000). These geographic inequities are largely a result of the highly variable distances people travel to access specialist and hospital services, which is one of the main themes of this paper.

Geographic access costs generate a “spatial pattern” of utilization of elective care. Partly to minimize these inequities state governments have recently been encouraging patients to move to other hospitals in the service area which have shorter waiting lists: *“For example, the Central Northern Adelaide Health Service is giving patients the choice to have their elective surgery done at other hospitals within the service’s area, if the hospital where they are booked has longer waiting times”* (DHS, 2003). In principle this right has always existed but, in practice, it has not been exercised to a significant extent because of a lack of information by consumers (and doctors) about the availability of beds. With policies to increase effective choice of hospital, a patient might decide to seek treatment in a hospital with a shorter waiting list or lower access costs despite the perception that there is a lower quality of care than in some other hospitals. If the opposite occurs and patients are prepared to substitute quality for higher access costs then the government’s policy of reducing such costs is less likely to be effective. To test for the existence of this substitution effect we carried out an econometric spatial analysis of the reduced form equations derived from our theoretical model in section 2 below.

The econometric analysis uses hospital utilization rates from 114 hospitals in financial years 1998-99 to 2003-04. These were obtained from the annual Victorian Admitted Episodes Dataset (VAED), a statewide patient-level dataset which also classifies patients according to their Main Diagnostic Categories (MDCs). The unit of analysis is Statistical Subdivision

(SSD) in Victorian state of Australia. An advantage of this unit is its policy relevance as results may be related to the planning regions used by the Victorian government and the effectiveness of policy assessed by repetition of the analysis through time. SSD boundaries are based on census tracts and may or may not capture the regional variations in hospital utilization. Hospitalization rates of one SSD may influence rates in neighboring SSDs. This might happen because of the patients' expectation of better quality care in neighboring SSDs, proximity of hospitals in neighboring SSD, social networks (Deri, 2005), clustering of a disease pattern (Neidell, 2004) over multiple SSDs, shorter waiting times, competition between hospitals or as a result of referrals by specialists. Such a phenomenon between regions is termed as "spatial dependence" by Anselin (1988).

The relevant models that do exist in the economic literature and potentially explain individual behavior leading to spatial dependence between regions are discussed by Manski (2000). The author argues that an action chosen by one agent may influence the constraints, expectations, and/or preferences of her reference group. Such an interaction could be due to three effects: i) exogenous: agent's action are dependent on observed attributes that define group membership, ii) correlated: agent's in the same neighborhood tend to behave similarly because they have similar characteristics or face similar opportunities or constraints and iii) endogenous: an agent's behavior is causally influenced by other members of the group. In the context of our framework, patients choices could result in regional interactions which could lead to an emergent collective behavior or aggregate pattern (Anselin and Center, 1999) resulting in correlation of data.

The empirical analysis in this paper tests and incorporates these spatial interaction. Specifically, we test if there exists an aggregate pattern of hospital utilization within a finite subset of regions defined by economic status (regions having similar SES) or location (regions sharing same border). The existing literature in health economics incorporating spatial effects is limited. Joines et al. (2003) apply cross-section spatial regression techniques on county level hospitalization rates for low back problems in North Carolina. The authors conclude that controlling for county level characteristics, a separate spatial process produces geographic clustering of high rate counties. They therefore recommend use of spatial effect modeling in small area analyses. Crighton et al. (2007) examine spatial characteris-

tics of pneumonia and influenza hospitalizations with county level data for Ontario. The authors find evidence of hospitalization ratios clustering across Ontario counties and also conclude that common spatial processes may be at play in determining hospitalization rates in these areas. Semaan et al. (2007) conducted a spatial regression analysis to account for clustering of sexually transmitted diseases using data from 48 contiguous states of the USA. The spatial regression results showed that spatial effects were significant and higher social capital was associated with lower STD rates.

Our analysis contributes to the existing empirical literature in two ways. Firstly, we explicitly control for quality of care and travel time in spatial analyses. The theoretical model proposed below allows quality and travel time to lead to spatial clustering of hospitalization rates. Previous studies have noted the role of quality of care or patient choices in a spatial framework but have assumed these effects part of the unobservables. Secondly, previous studies have examined spatial and temporal dimensions of hospitalization rates separately by conducting cross section spatial regression analysis for each year. Our analysis incorporates both spatial and temporal analysis by using spatial panel data models. To our knowledge these have only been applied once in the health literature in a very interesting paper by Moscone et al. (2007) where the authors empirically investigate the determinants of local authority mental health expenditure in England. Our econometric modeling below draws upon the insights of this work.

The plan of paper is as follows: Section 2 presents the model. Section 3 discusses estimation methodology and data and section 4 reports the main results which are discussed in Section 5. Section 6 concludes.

2 Model

2.1 Government

Let us assume that there is a continuum of patients with severity of illness $z \in [0, z^{max}]$. Health policy results in a criterion for admittance to a waiting list based upon the severity of illness of the patients z in time period $t - 1$. Formally, health authority chooses a cut-off value $0 < \bar{z} < z^{max}$ such that all the patients with severity of illness $z > \bar{z}$ are admitted

to the public hospital waiting list (Barros and Olivella, 2005). Based on this admittance criteria, the interaction between demand and supply of elective health care leads to a waiting list $Wait_t$ in time period t . We further assume that waiting list $Wait_t$ corresponds to waiting time of Q_t . Thus the variation in waiting times is attributed to two factors in the model: demand supply interaction and admittance criteria \bar{z} . Thus for example the government's decision to increase the health budget (leading to increase in hospital supply) or change in admission criteria will affect waiting times.

2.2 General Practitioners (GPs)

GPs act as gatekeepers (Scott, 2000; Dusheiko et al., 2005). The role of a GP is to access the severity of illness of patients and refer them to specialists or suggest alternative treatment. GPs referrals are determined by their utility maximization behavior which is described by a neoclassical labor supply function which includes a variable for ethical, professional behavior (Folmer et al., 1997). The function is given by:

$$V^{GP} = \log(c_t) + \log(x_t) - eth^{GP} \quad (1)$$

where c_t is consumption and x_t leisure of GP and eth^{GP} is a variable for ethical/professional behavior. Utility falls if the GP provides an inadequate service to the patient which is possible because of the asymmetry of information available to the GP and patient. For example, a GP would derive disutility if he provides inadequate information about quality of specialists or hospitals care to a patient. Assuming a time endowment of unity, the total consultation time by GP is given by $1 - x_t$. This consultation time can be divided into two parts: (i) time taken to consult patients which are referred to specialists ($\nu_1 f(z')$) and (ii) time taken to consult patients which are referred for alternative treatment ($\nu_2 f(z'')$), where ν_1 and ν_2 are time taken for each consultation for two types of patients and $f(z')$ and $f(z'')$ are number of consulting patients in first and second category respectively. Thus:

$$1 - x_t = (\nu_1 f(z') + \nu_2 f(z'')) \quad (2)$$

Assuming the GP charges a price of t_p for per unit time spent on consultation and incurs expenditure τ_t , the budget constraint facing the GP is:

$$c_t + \tau_t = t_p(\nu_1 f(z') + \nu_2 f(z'')) = B_{GP} \quad (3)$$

where B_{GP} is the total earnings accrued by the GP. The GP maximizes her utility with respect to c_t and x_t subject to the budget constraint Eq. 3. Based on the optimal values of c_t and x_t , the total referrals to specialists by a GP ($f(z')$) is given by¹:

$$f(z') = \frac{t_p + \tau_t}{2\nu_1 t_p} - \frac{\nu_2}{\nu_1} f(z'') = \frac{B_{GP}}{\nu_1 t_p} - \frac{\nu_2}{\nu_1} f(z'') \quad (4)$$

Assuming there are l GPs in a specific region, the total referrals to specialists in that region is given by aggregating the referrals over all GPs in that region:

$$f_{GP}(z') = \sum_{i=1}^l f_i(z') \quad (5)$$

2.3 Specialists

The role of specialists is to assess the severity of patients (which are all referred by GPs) and divide them into two groups based on current capacity and hospitals rationing policies: The group consists of patients whose severity of illness is greater than \bar{z} and they are placed on the waiting list of a public hospital. The second group consists of patients who are referred to alternative treatment including specialist out-of-hospital management which is therefore a substitute service. Let us assume that the number of patients referred to waiting lists (by specialists) are $f(z^*)$ (where $z^* > \bar{z}$) and number of patients sent for alternative treatment are $f(z^{**})$ (where $z^{**} < \bar{z}$). Their consultation time can be divided into two parts: (i) time taken to consult patients who are referred to waiting lists ($\mu_1 f(z^*)$) and (ii) time taken to consult patients which are referred for alternative treatment ($\mu_2 f(z^{**})$) where μ_1 and μ_2 are time taken for each consultation for the two types of patient. Like GPs, specialists maximize utility which is affected by their ethical/professional behavior eth^{SP} . Assuming the specialists charge a price t_s per unit time of consultation and incur expenditure τ_s , total referrals to public hospital waiting list are given by:

$$f(z^*) = \frac{t_s + \tau_s}{2 \mu_1 t_s} - \frac{\mu_2}{\mu_1} f(z^{**}) = \frac{B_{SP}}{\mu_1 t_s} - \frac{\mu_2}{\mu_1} f(z^{**}) \quad (6)$$

where B_{SP} is the total earnings of a specialist from consultation. Assuming there are m specialists in a specific region, total referrals to waiting list from that region is an

¹See Appendix A.1 for detailed derivation.

aggregation over all specialists referrals ($f(z^*)$) and is given by:

$$f_{SP}(z^*) = \sum_{i=1}^m f_i(z^*) \quad (7)$$

Similarly, total referrals to alternative treatment are given by:

$$f_{SP}(z^{**}) = \sum_{i=1}^m f_i(z^{**}) \quad (8)$$

Since all the patients consulting the specialists are referred by GPs we have:

$$f_{SP}(z^*) = f_{GP}(z') - f_{SP}(z^{**}) \quad (9)$$

Let us call the patients who are referred to waiting list as “eligible patients”. We argue that some of these eligible patients will be reluctant to choose public care as an option of treatment because of, inter alia, high access costs, longer waiting times, or perception of poor quality of care. The next section formally models the demand for hospital care by eligible patients, incorporating access costs, perception of quality and waiting times.

2.4 Demand for Hospital Care

The individual derives a utility, U , from the expectation of improved health and therefore from the perception of the quality of health care. Let us assume that patient has two types of expectations about quality of care: good or bad. The health gains in these two scenarios are given by: For Public care:

$$U = \begin{cases} U_1 & \text{with probability } p \text{ (Good quality care)} \\ U_2 & \text{with probability } (1 - p) \text{ (Bad quality care)} \end{cases}$$

We make the following assumptions on above utility specification.

Assumption 1 $U_1 > 0$, $U_2 > 0$, $U_1 > U_2$ or $\frac{\partial U}{\partial p} > 0$. We require that an individual’s health gain increases with an increase in the probability of better quality care, utility derived from good quality care is higher than from bad quality care and that health gains are always positive.

Thus the perception of quality is a source of uncertainty. Formally, the expected health gain from treatment without any wait is:

for Public Care:

$$E(U) = pU_1 + (1 - p)U_2 \quad (10)$$

for Private Care:

$$E(U^{pvt}) = p_1U_1 + (1 - p_1)U_2 \quad (11)$$

The value of p is assumed to be exogenous to the system, i.e. an individual's perception of quality of care for treatment of his illness in a particular hospital, is based on informal communications with GP and specialist or upon general knowledge and it is not affected by the access cost of the treatment. Expected health gains from public and private care ($E(U)$ and $E(U)^{pvt}$ respectively) differ because of different perceptions of quality of care (p and p_1). Let us denote this difference in expected health gains by η :

$$E(U) - E(U^{pvt}) = \eta \quad \text{where } \eta \begin{matrix} \leq \\ \geq \end{matrix} 0 \quad (12)$$

The expected health gain from private care might be greater, equal or less than that from public care, depending on the value of η . We assume η to be the same for all individuals. We also assume that an individual derives disutility from waiting for treatment. If she waits for a period Q her health gain is depressed by an exponential decay function e^{-zQ} where z is perceived severity of illness (Cullis and Jones, 1986). Thus individual with more severe illness will derive more disutility from waiting. The accessibility cost (kA_c) negatively affects utility by a scale factor of k . An individual also faces a fixed cost of treatment F irrespective of the type of care (i.e. public or private care), and has to make a copayment C_p for private care. Thus the net expected utility from the treatment is given by:

For Public Care:

$$E(\xi) = \{pU_1 + (1 - p)U_2\}e^{-zQ} - kA_c - F \quad (13)$$

For Private Care:

$$E(\xi^{pvt}) = \{p_1U_1 + (1 - p_1)U_2\} - C_p - kA_c^{pvt} - F \quad (14)$$

Since we consider only non-urgent conditions in our framework, patient faces three options (Goddard et al., 1995; Martin and Smith, 1999):

- Receiving treatment in a public hospital as a public patient. In this scenario patient will be put on a waiting list and faces a positive waiting time ($Q > 0$) but no cost of copayments ($C_p = 0$).
- Receiving treatment in a private hospital or as a private patient in a public hospital: In this scenario the patient faces almost no waiting time, we assume it to be zero ($Q = 0$) and positive copayments ($C_p > 0$)
- Receiving no treatment: This scenario arises if the costs of treatment in public or private hospitals exceed the expected health gains from the treatment.

The remainder of the paper focusses on the patterns of utilization of public hospital services by public patients. We assume that a single patient is unable to affect the perception of quality of care in hospitals, waiting times in public hospitals, co-payment rates, or the fixed cost of the treatment. Thus all these variables are exogenous and uniform across all patients in the population group. The patients can make two choices: (i) Whether to seek a hospital treatment; and (ii) if so, in which hospital. The only choice variable for the patient is the access cost A_c (proxied by travel time to hospital). Thus a patient can choose to travel for a longer time for better quality health care. We now derive the conditions under which a patient will prefer care in public hospitals.

Proposition 1 *The patient will prefer public care to no care if:*

$$E(\xi) > 0 \Rightarrow E(U)e^{-zQ} > (kA_c + F) \quad (15)$$

or the expected health gains (net of waiting times) from the treatment are greater than the sum of access costs and fixed costs of treatment.

Proof 1 *The patient will prefer public care to no care if her net utility from public care is greater than the utility from no care (assumed to be zero) or her net utility is positive. Thus solving $E(\xi) > 0$ using (10), (13), and taking $C_p = 0$ (for public care) we derive the above result.*

Lemma 1 Access-Cost Quality Trade-off: The patient is willing to face higher access costs for better quality of care:

$$\frac{\partial A_c}{\partial p} = \frac{(U_1 - U_2)e^{-zQ}}{k} > 0 \quad (16)$$

Proof 2 Substituting (13) in (15), solving for A_c , first order differentiating with respect to p and using Assumption 1 ($U_1 > U_2$) we get above result.²

Proposition 2 The patient will prefer private care to public care if:

$$E(U)e^{-zQ} - E(U^{pvt}) < kA_c - (kA_c^{pvt} + C_p) \quad (17)$$

or the incremental health gain from public care is less than the additional cost faced by the patient.

Proof 3 The patient will prefer private care to public care if the expected net utility from public care is less than expected net utility from private care or $E(\xi) < E(\xi^{pvt})$. Substituting the values from Eq. 13, Eq. 14 and using Eq. 10 and 11 we get Eq. 17.

It should be noted that the difference in expected health gains between public and private care is solely attributed to differences in the perception of quality whereas the difference in net utility from public and private care is attributed additionally to differences in access costs ($kA_c - kA_c^{pvt}$) and copayment rates (C_p). The total demand for public care from a specific region is determined by aggregating over a distribution of patients with varying expected health gains ($E(U)$) and severity of illness (z). Let us denote it by $\phi(E(U), z)$. The value of z among eligible patients will vary between \bar{z} and z^{max} where we assume that z^{max} is the maximum value of z at which public care is sought. The expected health gains ($E(U)$) among eligible patients will vary between following two limits³:

Lower Limit: This is the limit above which the patient will prefer public care to no care. Rearranging Eq. 15 the lower limit is: $(kA_c + F)e^{zQ}$

Upper Limit: This is the limit below which the patient will prefer public care to private care. Substituting Eq. 12 in 17 and rearranging the upper limit is: $\frac{(kA_c^{pvt} + C_p) - kA_c + \eta}{(1 - e^{-zQ})}$. Total

²See Appendix A.2 for detailed derivation.

³See Appendix A.3 for detailed derivation.

demand for public care in a region is derived by integrating the population density over the region:

$$DD = \int_{\bar{z}}^{z^{max}} \int \frac{(kA_c^p + C_p) - kA_c + \eta}{(1 - e^{-zQ})} \phi(E(U), z) dU dz \quad (18)$$

Each patient's preference and choice of hospital will be based on his preferences of perception of quality of care and accessibility costs. The preferred hospital may be located outside the region. Let us assume that the demand of public hospital care DD_r from a specific region $r \in [1, R]$ is satisfied by H public hospitals. Thus the aggregate demand from a region r (DD_r) can be decomposed into individual demand for each of these H hospitals. Let us denote the demand for hospital h from region r patients as DD_r^h . Thus:

$$DD_r = \sum_{h=1}^H DD_r^h \quad (19)$$

DD_r is a net demand adjusted for the eligible patients who have opted for care in other public hospitals, private hospital or for no care. Let us denote number of such patients by χ_r (where $\chi_r \geq 0$). Thus net demand in terms of GP and specialist referrals (Using (8) and (9)) can be written as:

$$DD_r = f_{SP}(z^*) - \chi_r = f_{GP}(z') - f_{SP}(z^{**}) \quad (20)$$

2.5 Supply of Health care

This section derives the supply of care by a public hospital which is defined as the number of treatments by a hospital in a financial year. The supply is derived by assuming the hospital manager to be a utility maximizing agent (Gravelle et al., 2003; Martin and Smith, 1999; Propper, 1995) facing constraints imposed by the health authority by hospital efficiency, the exogenous budget and by available resources. These include FTE Staff, beds, operating theaters, diagnostic and other equipment. Since the focus of our paper is demand and the choice of hospital, an explicit modeling of these factors is outside the scope of the paper and they are treated as exogenous supply shifters (\bar{X}^h). The hospital manager maximizes following utility function:

$$M_t^h = f(S_t^h, N_t^h, Wait_t) \quad (21)$$

where: M_t^h is hospital manager's utility function S^h is the supply of elective care by hospital h , N_t^h is the supply of non elective care, $Wait_t$ is the waiting list in the current period. The waiting list in the current period is dependent on the past period waiting list ($Wait_{t-1}$) adjusted for current period demand-supply mismatch ($D_t^h - S_t^h$) (Windmeijer et al., 2004):

$$Wait_t = Wait_{t-1} + D_t^h - S_t^h \quad (22)$$

where D_t^h is the demand for hospital h in period t . Further, the expected waiting time can be defined as:

$$Q_t = \frac{\lambda Wait_{t-1}}{S_t^h} \quad (23)$$

where Q_t is the expected waiting time λ is the bed days required for each episode of elective health care. With supply measured by active beds, the expected waiting time is determined by the ratio of potential bed days demand to the supply of beds (where λ is the length of stay (Lindsay and Feigenbaum, 1984; Mulligan, 1985)). Equation 23 defines the intertemporal evolution of waiting times and is the main link between the supply of hospital beds in any two consecutive periods. The manager faces a budget constraint determined by government funding policy. We assume a case-mix funding approach where hospitals are paid according to case complexity (Duckett, 1995) . Assuming the government pays $\bar{G}(z)$ dollars for each episode with severity of illness z the total payment to hospitals B^h is given by integrating the payments of non urgent and urgent conditions over the distribution of patients with varying z :

$$B^h(z) = \int_{\bar{z}}^{z^{max}} (S^h(z) \bar{G}(z))dz + \int_0^{\infty} (N^h(z) \bar{G}(z))dz \quad (24)$$

where: $B^h(z)$ is the total funding to the hospital $\bar{G}(z)$ is the dollar payment for each episode with severity of illness z . The manager also faces a resource constraint:

$$S^h(z) + N^h(z) = \bar{X}^h \quad (25)$$

where: \bar{X}^h : Vector of variables treated as exogenous supply shifters. The maximization of manager's utility subject to resource and budget constraint corresponds to optimal values of S^h , N^h and $Wait^h$. The reduced form supply equation for elective procedures can be expressed as:

$$S^h = f(Q_t, \bar{X}^h, B^h) \quad (26)$$

2.6 Equilibrium

We assume that an equilibrium is achieved when the supply of hospital beds is equal to the demand for hospital beds (rationed by waiting times) for elective care in public hospitals:

$$D_t^h = S_t^h \quad (27)$$

The main characteristic of equilibrium is that the waiting list is stable. The marginal benefit equates to the marginal cost of entering the waiting list. The corresponding waiting time will thus be Q_h^* which becomes time invariant. ($Q_{h_{t+1}}^e = Q_{h_t}^e$) and there will be no addition or deletion from the waiting list. Thus the waiting times acts as a price variable or the rationing variable. However, the price variable in a standard demand and supply model ensures an equilibrium is reached by raising the cost of good whereas here the waiting times clear the health market by making the perceived benefits of health less valuable.

Corollary 1 *Given that each hospital is at equilibrium characterized by Eq. 27, the demand for elective public care from a region r is equal to the aggregate supply of beds to the patients of region r .*

Corollary 1 alternatively specifies the equilibrium condition at the regional level. The aggregate supply of beds to the region r is calculated by adding the supply of all hospital beds which were utilized by the patients from region r . Thus if the patients in region r utilized services from H hospitals (some of which might be located outside region r), the equilibrium condition can be specified as:

$$\sum_{h=1}^H DD_r^h \equiv \sum_{h=1}^H S^h \quad \text{for each } DD_r^h > 0 \quad (28)$$

3 Analytic Framework

We now derive estimable equations on observable behavior (utilization of hospital care) in the form of reduced form equations. These can be subsequently used for empirical analysis.

3.1 Empirical Specification

The demand for elective care in public hospitals is given by Eq. 19 which is a function of referrals by specialists and GPs (Eq. (20)). Thus the reduced form demand for elective care in public hospitals can be specified as:

$$DD = f(A_c^p, p, Q, C_p, F, A_c, f_{SP}, f_{GP}) \quad (29)$$

The reduced form equation for Supply is given by equation 26 reproduced here:

$$S^h = f(Q, \bar{X}_h, B^h)$$

Both reduced form equations can be solved to derive utilization of health care and the corresponding waiting time. We focus upon the possible impact of quality-access cost trade-offs on the utilization of health care and thus the empirical analysis of the waiting time equation is not considered here.⁴ The regional utilization of elective care in public hospitals can be derived by solving the reduced form demand and supply equations. Thus the utilization can be specified as:

$$UT_r = f(A_c^p, p, C_p, F, A_c, f_{SP}, f_{GP}, \bar{X}_h, B^h) \quad (30)$$

Variables are defined in Table 1.

—————Insert Table 1 about here—————

3.2 Data

Data were obtained from various sources. Patient level information from 1998-99 to 2003-04 was taken from Victorian Admitted Episode dataset (VAED). The data span from 1998-99 to 2003-04. Patients' residence postcodes are used to determine the SSD in which patient resides. There are 45 SSDs in Victoria based on 2001 ASGC boundaries. The dependent variable hospital utilization is proxied by surgical separation rates per ten thousand population and are standardized for age and gender using the direct standardization method.

⁴Moreover, the data for waiting times is not yet available for elective surgery patients making it impossible to estimate the waiting times equation.

The average quality of care expected by public patients is proxied by the actual quality of care received by public patients in region r *in the previous period*:

$$p_t = qual_{t-1}$$

where p_t is the current period perception of quality and $qual_{t-1}$ is the actual quality of care reflecting the time lag between reality and its effects upon perceptions. Quality is measured by the proportion of patients who experienced an adverse event while in hospital. They are identified by the methodology suggested in Jackson et al. (2006) and unplanned readmission rates. The detailed method is discussed in Appendix A.4.1. Use of lagged quality variable also avoids the problem of endogeneity between quality of care and hospital volume discussed extensively in literature by Gaynor et al. (2002) among others.

PHI coverage at SSD level is imputed using National Health Survey (NHS) and Australian Bureau of Statistics (ABS) data.⁵ The GP and Specialist Benefits per thousand population are derived from Australian Institute of Health and Welfare (AIHW) data. The public and private bed density is derived from Social Health Atlas (Glover et al., 1999). Some patients receive public treatment outside their region. This will be influenced by public bed density in the corresponding region and by other unobserved influences. For example, the appointment of a reputed surgeon to a hospital in region A might attract patients from neighboring regions. Similarly, a patient residing in region A might opt for elective surgery in region B because there is less travel time. The unobservable and variable referral norms of specialists and GPs could also lead to inter-regional interactions. This interdependence is explicitly incorporated in the econometric modeling using spatial regression techniques.

Utilization is disaggregated by Main Diagnostic Categories to account for the effect of intra and inter regional heterogeneity in severity of illness (z). The data consistently uses ICD-10 AM Codes to categorize episode type. These codes are used to divide the data into Main Diagnostic Categories (MDCs) for surgical procedures. The empirical analysis in this paper uses elective surgery data for three main diagnostic categories: Diseases and Disorders of Female Reproductive System (FRS), Diseases and disorders of the musculoskeletal system

⁵The detailed imputation methodology is skipped here due to space constraints and is available on request from the authors. Note that PHI Coverage in a region is used as a proxy for copayments in that region. Alternative specification with private separations in a region as a proxy for copayments led to qualitatively similar results.

and connective tissue (Hip & Knee) and Diseases and Disorders of the Digestive System (Digestive). These three MDCs fall in the top five list of MDCs having highest surgical waiting times.

Ideally, the empirical analysis of patient choices should be carried out at patient, not regional level, as done here. However VAED reports only three patient characteristics: age, sex and postcode of residence which are insufficient for a patient level analysis. Regional analysis could suffer from ecological fallacy if the region is an administrative area (Carr-Hill et al., 1994; Martin and Smith, 1999). In order to minimize this risk, the present paper uses Statistical Sub Divisions (SSDs) as unit of analysis which is an Australian Standard Geographical Classification (ASGC) defined area which represents an intermediate level, general purpose, regional type geographic unit which is not an administrative area.

The SSD characteristics based on census data (SEIFA indices and Proportion of Indigenous population) are extracted from the 1995 and 2000 census for pre and post-2000 data respectively. Summary statistics are reported in Table 2. Table 3 reports the descriptive statistics for panel data classified by the different diagnostic categories.

—————Insert Table 2, Table 3 about here—————

Table 3 reports means and deviations for the panel data by decomposing the variable into “between” and “within” variation. For example, for FRS, the rate of adverse events for each SSD varied between 1.45 % to 16.67% and travel time varied between 4 minutes to 231 minutes. The “within” number refers to the deviation from each SSD’s average and can be negative. The within numbers for FRS show that for some SSD rate of adverse events deviated from its average by around 4.7%. The deviations are even higher for travel time. The variation in adverse events across and within SSDs is significantly lower compared to travel time. Thus SSDs differ more in terms of access costs than in terms of quality of care. The trends are qualitatively the same for other diagnostic categories.

3.3 Econometric Methodology

Spatial panel models allow spatial heterogeneity and autocorrelation over space and time (Elhorst, 2003; Baltagi and Li, 1999).

Spatial Error Specification:

$$UT_{it} = x'_{it}\beta + \epsilon_{it} \quad i = 1, \dots, n; t = 1, \dots, T \quad (31)$$

where UT_{it} denotes the hospital separations per ten thousand population in SSD i and x the vector of explanatory variables. The disturbance term ϵ_{it} is assumed to follow an error component model with spatially autocorrelated residuals. The disturbance vector for time t is can be further split into SSD level effects and remainder disturbances:

$$\epsilon_t = v + \varphi_t$$

where $\epsilon_t = (\epsilon_{1t}, \epsilon_{2t}, \dots, \epsilon_{nt})'$; $v = (v_1, \dots, v_n)'$ denotes the vector of SSD effects and $\varphi_t = (\varphi_{1t}, \varphi_{2t}, \dots, \varphi_{nt})'$ are the remainder disturbances which are independent of v . The φ_t 's follow the spatial autoregressive error dependence model:

$$\varphi_t = \lambda W \varphi_t + \psi_t$$

where W is a matrix of known spatial weights of dimension n by n and λ is the spatial autoregressive coefficient. The error vector $\psi_t \sim iid(0, \sigma_\psi^2)$ and independent of φ_t and v . The spatial weights matrix W is a matrix of zeros and ones where SSDs sharing common criteria take a value 1, otherwise it is zero. The v_i 's are the unobserved SSD specific effects which can be fixed or random. Since the error terms are correlated, the OLS estimates of above model are unbiased and inefficient (Anselin, 1988; Elhorst, 2003).

Spatial Lag Specification:

$$UT_{it} = \rho W y_{it} + x'_{it}\beta + v + \varphi_t \quad i = 1, \dots, n; t = 1, \dots, T \quad (32)$$

where ρ is a spatial autoregressive coefficient. The v_i 's are the unobserved SSD specific effects which can be fixed or random. In a spatial lag model assumptions of independent observations and uncorrelated error terms are violated and hence OLS estimates are biased and inefficient. Maximum likelihood estimation techniques are suggested to overcome the problem of bias and inefficiency of estimates (Anselin, 1988). The econometric estimation in this paper uses random effects spatial panel model which allows us to incorporate time-invariant heterogeneity across SSDs through an individual SSD specific error component (v).

The use of a fixed effects model is not appropriate in our context for two reasons: most of our variables (e.g. bed density) show little variation over time and will be highly correlated with fixed effects and, secondly, the data covers a short period of time because of the lack of consistent data over a longer time horizon. The test for random effects is reported in the results section.

The issue of choice between spatial error and spatial lag model specification is crucial. A Spatial lag specification is more appropriate when there is a direct interaction between spatial units and the focus is on the assessment of the strength of spatial interaction. Spatial error models are appropriate when the concern is with correcting the potential biasing influence of spatial autocorrelation due to the use of spatial data (Anselin and Rey, 2004). Moreover, as Griffith (1992) argues spatial error models which explicitly incorporate spatial autocorrelation in error terms, can be perceived as a surrogate for an omitted or unobserved variables. Spatial error models are also applied when there is a mismatch between the scale of process and scale of analysis. Spatial error models are better suited for our application as our sole purpose to apply such models is to test for any remaining spatial autocorrelation once we account for quality and distance effects. In addition, there are some unobservables like practice style of GPs and Specialists and their interactions which could diffuse the local standards of hospital care across SSDs. Such effects can be best incorporated by using a spatial error specification. Under a random effects specification the random effect associated to the i^{th} SSD (v_i) is IID distributed with zero mean and variance σ_v^2 .

4 Results

The econometric analysis used Matlab code kindly provided by J. P. Elhorst. The results for spatial error random effects are reported in Table 4. The p-values are in parentheses.

—————Insert Table 4 about here—————

The estimation is carried out for three diagnostic categories using two types of weights matrices. The first is based on geographic contiguity of SSDs and the second upon the socioeconomic characteristics of SSDs. Specifically SSDs which fall in same percentile of

socio economic index are assigned value 1 otherwise zero.

The test for spatial autocorrelation (significance of λ) shows that spatial autocorrelation is not significant for model specification with economic weights. However significant remaining spatial error dependence exists for Hip & Knee and Digestive Sys procedures in spatial error models with geographic weights. The value of λ is -0.503 and -0.410 for Digestive Sys. and Hip & Knee Surgeries respectively indicating that the indirect effect of unobservable and environmental factors of neighboring SSDs have a negative impact on the hospital utilization of an SSD. The results for FRS show no residual spatial error dependence and thus the estimation reduces to a classical random effects model. Consequently results are very similar for both geographic and economic weight specifications in the second and third columns of Table 4.

FRS category has highest number of patients on Victorian surgical waiting lists and from Table 4 the main predictors of utilization for this category are travel time, Private Health Insurance, Surgeon Density and Public Bed Density. The coefficient for travel time and private health insurance are negative as expected indicating that, *ceteris paribus*, SSDs having higher travel time to public hospitals or higher private health insurance coverage will have lower public hospital utilization. The negative impact of surgeon density on hospital utilization of FRS implies that surgeons in an SSD are a substitute for elective public hospital care. The positive coefficient on public bed density implies, *ceteris paribus*, that SSDs with more beds will have higher FRS separation rates. This is consistent with the existing literature where hospital bed supply has been positively related to hospital use for several conditions (Folland and Stano, 1989).

The main predictors of hospital utilization of Digestive Sys services in the spatial error model with geographic weight are quality of care, the SES (SEIFA) Index, Public and private bed densities and the spatial dependence between neighboring SSDs. The negative coefficient on quality of care suggests that SSDs with a higher number of patients having adverse events in previous period (in public care) will have lower public hospital utilization in current period. The negative sign on the SEIFA (EO) index for education and occupation implies that higher education levels in an SSD are associated with lower surgery rates, a finding that is also consistent with the literature (Mcmahon et al., 1991). Public bed

supply has a positive effect on utilization whereas an SSD with higher private bed density has lower public surgery rates in this category.

The results for Hip & Knee elective surgery also indicate spatial dependence among neighboring SSDs. Apart from this, only surgeon density and regional area significantly affect hospital utilization having negative and positive effects respectively. The adjusted R^2 for all three categories is around 0.66. The spatial panel models were estimated using MLE and a LR test conducted to determine whether spatial random effect models differ significantly from classical random effect models. The results indicate that for Hip & Knee and Digestive Sys. the results of the two methodologies are significantly different. The results also reveal that θ (where $\theta^2 = \frac{\sigma_\psi^2}{\sigma_\epsilon^2}$) is significant for all cases thereby confirming the existence of significant random effects in the model.

5 Discussion

5.1 Testing for Access-cost Quality Trade-offs

Unlike previous work this study uses both quality and travel time in a single framework which enables us to analyze travel time-quality interactions. One of the potential reasons for significance of either quality or travel time (but not both) on hospital utilizations could be the travel time-quality trade-off discussed earlier. We tested for such a trade-off by estimating the sign and significance of $\frac{\partial A_c}{\partial p}$ defined in (16). According to Lemma 1, for any trade-off to exist this expression should be positive. Rewriting the expression above in terms of the coefficients of the econometric model:

$$\frac{\partial A_c}{\partial p} \equiv \frac{\frac{\partial UT_r}{\partial p}}{\frac{\partial UT_r}{\partial A_c}} = \frac{\beta_{quality}}{\beta_{time}} \quad (33)$$

where $\beta_{quality}$ and β_{time} are the regression coefficients on quality and time variables respectively. Since we have taken the rate of adverse events (negative quality) as a proxy for quality the expression should be negative if a trade-off exists. The ratio for main diagnostic categories are reported in Table 5.

—————Insert Table 5 about here—————

The results from Table 5 show that the although the ratio is negative for all categories

except Digestive Sys none is significant. Thus there is no strong evidence of a trade-off for any of the categories.

5.2 Policy Implications

The main broad policy relevant results concern quality. Like a number of countries Australia has a high level of adverse events associated with hospital care (Wilson et al., 1995). In an efficient market information about this would be expected to influence patient choice of hospital. As indicated above our results are inconsistent with the hypothesis that patients substitute travel time costs for quality. Further, results in 4 indicate that the expected negative effect of quality upon separations is only observed for conditions of the digestive system, with total separations unaffected by quality. The coefficient of -0.470 for the digestive system represents an elasticity of 0.06 at the mean. Putting this in a different perspective, moving from the level of adverse events one standard deviation above to one standard deviation below the mean would reduce admissions by 1.98 or by 3.6 percent of the mean level. *Ceteris paribus*, moving from the lowest Digestive System adverse events rate in an SSD (2.08) to the highest (10.33) - a five-fold difference - would reduce separations by 3.89 or 7.2 percent of the mean separation rate which, would have the secondary effect of reducing adverse events in the worst SSD by 0.72 pa or 6.9 percent. These results indicate that the market does not perform very efficiently as judged by the effect of quality upon admissions and the self adjusting effect this might have. Individual hospitals would have little or nothing to worry about in terms of their loss of market share and budget revenue. The 3.6% and 7.2% reductions in separations rates cited above (for the one significant diagnostic category) assumed all else to be equal in the face of significant differences in quality. But these effects could be offset very easily by a reduction in queuing times implying no incentive via 'the market' - revenues - for quality improvement, and adverse events would be unaffected.

The three sets of results on the effect of private sector on public hospital separations are: i) an insignificant or quantitatively small association between total separations and the percentage of the population with private health insurance ii) an insignificant effect of private hospitals beds in an SSD and iii) a statistically significant negative impact of an

increase in the availability of surgeons. However, their impact is, once again, quantitatively small. An additional surgeon per 1,000 population decreases admissions by 3.6 per 1,000 or by 0.65% representing an elasticity of 0.16 with respect to surgeons at their mean level. These results indicate that the substitution effect between the public and private sectors is relatively small. Although public private sector substitution is not the focus of this paper our results provide a baseline to explore this issue in detail in a fully specified model for the private sector.

6 Concluding Remarks

This paper developed a framework incorporating patient choices for elective surgeries when access acts as a rationing mechanism. This was done by focussing on the quality of care and travel time to care as two main determinants of the demand for hospital services. Unlike previous studies the theoretical framework models the complex interaction between quality and access costs allowing for a trade-off where a patient is willing to incur higher access cost for better quality of care. The second part of the paper constructed an empirical model of supply and demand of elective surgery which is consistent with the theoretical framework developed in this paper. Perhaps the most important finding was that our empirical results confirm that spatial effects, especially those based on geographical neighborhood effects, are important and should be considered in studies involving regional level data.

One unique feature of the empirical analysis is that it uses real data on both patients' travel time and the quality of care they received. In addition, the empirical analysis explicitly incorporated spatial interactions between regions over time by using spatial panel models. Such an interaction is necessary as patients can travel between regions for better quality care or lower access costs. The empirical analysis was carried out separately for three MDCs and the results demonstrate that quality access costs and spatial dependence effects differ for different MDCs. For example, for FRS travel time was one of the determinants of hospital utilization and quality effect was insignificant whereas for Digestive surgeries, quality was significant and travel time insignificant. For Hip & Knee procedures most of the variation was explained by spatial dependence and not by quality or travel time. The existing evidence of the impact of distance on utilization of health care in the literature is

mainly focussed on GP services (Gravelle et al., 2002; Namet and Bailey, 2000; Croxson et al., 2001). However our finding a negative and significant effect of travel time on hospital utilization is consistent with these results. Similarly our results are consistent with the somewhat ambiguous literature on quality. After analyzing a 1983 dataset Luft et al. (1990) concluded that “..... hospitals with poorer than expected outcomes attracted significantly fewer admissions” even in the absence of publicly available data on quality. Our results found a similar effect in one category of separations. However the effect was quantitatively small and for other categories, and for total separations, it was insignificant. This is consistent with the more recent US literature (for example see Chassin (2002)).

Anticipating our results, the prescient new Australian government announced in 2008 that all Australian hospitals will be required to make report cards publicly available, including information on quality! US experience suggests that this will not necessarily create an efficient patient/agent driven market in which patient knowledge forces doctors and hospitals to improve quality. It is possible that, in the face of excess demand, nothing will happen or, as suggested by Chassin, doctors in hospitals will respond directly to the publicising of their relative performance. The present study represents a baseline from which to assess the possible effects.

References

- Ackere, A. and P. C. Smith (1999). Towards a Macro Model of National Health Service Waiting Lists. *System Dynamics Review* 15, 225–252.
- Anselin, L. (1988). *Spatial econometrics*. Kluwer.
- Anselin, L. and B. Center (1999). *Spatial Econometrics*.
- Anselin, L. and S. Rey (2004). *Advances in spatial econometrics*. Springer.
- Baltagi, B. H. and D. Li (1999). *Prediction in the Panel Data Model with Spatial Correlation*, Chapter 13, pp. 283–296. Springer.
- Barros, P. and P. Olivella (2005). Waiting Lists and Patient Selection. *Journal of Economics & Management Strategy* 14 (3), 623–646.
- Carr-Hill, R., G. Hardman, S. Martin, S. Peacock, T. Sheldon, and P. Smith (1994). A Formula for Distributing NHS Revenues Based on Small Area use of Hospital Beds. Occasional paper.
- Chassin, M. (2002). Achieving And Sustaining Improved Quality: Lessons From New York State And Cardiac Surgery. *Health Affairs* 21(4), 40.
- Crichton, E., S. Elliott, R. Moineddin, P. Kanaroglou, and R. Upshur (2007). A spatial analysis of the determinants of pneumonia and influenza hospitalizations in Ontario (1992–2001). *Social Science & Medicine* 64(8), 1636–1650.
- Croxson, B., C. Propper, and A. Perkins (2001). Do doctors respond to financial incentives? UK family doctors and the GP fundholder scheme. *Journal of Public Economics* 79(2), 375–398.
- Cullis, J. and P. Jones (1986). Rationing by Waiting Lists: An Implication. *The American Economic Review* 76(1), 250–256.
- Deri, C. (2005). Social Networks and Health Service Utilization. *Journal of Health Economics* 24, 1076–1107.

- DHS (2003). Policy for the Management of Elective Surgical Admissions in Metropolitan Public Hospitals. Technical report, South Australia Department of Human Resources.
- Duckett, S. (1995). Hospital payment arrangements to encourage efficiency: the case of Victoria, Australia. *Health Policy* 34(2), 113–34.
- Duckett, S. J. (2000). *The Australian Health Care System*. South Melbourne, Victoria: Oxford University Press.
- Dusheiko, M., H. Gravelle, R. Jacobs, and P. Smith (2005). The effect of financial incentives on gatekeeping doctors: Evidence from a natural experiment. *J Health Econ.*
- Elhorst, J. P. (2003). Specification and Estimation of Spatial panel Data Models. *International Regional Science Review* 26, 244–268.
- Folland, S. and M. Stano (1989). Sources of Small Area Variations in the Use of Medical Care. *Journal of Health Economics* 8, 85–107.
- Folmer, K., J. Stevens, F. Van Tulder, and E. Westerhout (1997). Towards an economic model of the Dutch health care sector. *Health Economics* 6(4), 351–363.
- Fuchs, V. R. (2004). Reflections on the Socio-economic Correlates of Health. *Journal of Health Economics* 23, 653–61.
- Gaynor, M., H. Seider, and W. Vogt (2002). Is There A Volume-Outcome Effect, and Does It Matter?(Yes, and Yes). *Abstr Acad Health Serv Res Health Policy Meet* 19, 2.
- Glover, J., K. Harris, and S. Tennant (1999). A Social Health Atlas of Australia, Vol 1: Australia. *Public Health Information Development Unit, South Australia: University of Adelaide, Adelaide*.
- Goddard, J., M. Malek, and M. Tavakoli (1995). An Economic Model of Market for Hospital Treatment for Non-Urgent Conditions. *Health Economics* 4, 41–55.
- Gravelle, H., M. Dusheiko, and M. Sutton (2002). The Demand for Elective Surgery: Time and Money Prices in the UK National Health Service. *Journal of Health Economics* 21, 423–449.

- Gravelle, H., P. Smith, and A. Xavier (2003). Performance Signals in the Public Sector: the Case of Health Care. *Oxford Economic Papers* 55, 81–103.
- Griffith, D. (1992). What is spatial autocorrelation? Reflections on the past 25 years of spatial statistics. *LEspace Geographique* 21(3), 265–280.
- Jackson, T., S. Duckett, J. Shephard, and K. Baxter (2006). Measurement of Adverse Events Using 'Incidence Flagged' Diagnosis Codes. *J Health Serv Res Policy* 11, 21–26.
- Joines, J., I. Hertz-Picciotto, T. Carey, W. Gesler, and C. Suchindran (2003). A spatial analysis of county-level variation in hospitalization rates for low back problems in North Carolina. *Soc Sci Med* 56(12), 2541–53.
- Lindsay, C. M. and B. Feigenbaum (1984). Rationing by Waiting Lists. *American Economic Review* 74(3), 404–417.
- Luft, H., D. Garnick, D. Mark, D. Peltzman, C. Phibbs, E. Lichtenberg, and S. McPhee (1990). Does quality influence choice of hospital? *JAMA* 263(21), 2899–2906.
- Manski, C. (2000). Economic Analysis of Social Interactions. *The Journal of Economic Perspectives* 14(3), 115–136.
- Martin, S. and P. C. Smith (1999). Rationing By Waiting Lists: An Empirical Investigation. *Journal of Public Economics* 71, 141–164.
- McMahon, L., G. Petroni, P. Tedeschi, et al. (1991). Small area analysis of hospital discharges for musculoskeletal diseases in Michigan: the influence of socioeconomic factors. *The American journal of medicine* 91(2), 173–178.
- Moscone, F., M. Knapp, and E. Tosetti (2007). Mental health expenditure in England: A spatial panel approach. *Journal of Health Economics* 26(4), 842–864.
- Mulligan, J. G. (1985). The Stochastic Determinants of Hospital Bed Supply. *Journal of Health Economics* 4, 177–181.
- Namet, G. F. and A. J. Bailey (2000). Distance and Health care utilization among the rural elderly. *Social Science and Medicine* 50, 1197–1208.

- Neidell, M. J. (2004). Air Pollution, Health, and Socio-economic Status: the Effect of Outdoor Air Quality on Childhood Asthama. *Journal of Health Economics* 23, 1209–36.
- Propper, C. (1995). Agency and incentives in the NHS internal market. *Soc Sci Med* 40(12), 1683–90.
- Richardson, J. (1998). The health care financing debate. *Economics and Australian Health Policy*, 192–213.
- Scott, A. (2000). Economics of general practice. *Handbook of Health Economics* 2, 1775–12000.
- Semaan, S., M. Sternberg, A. Zaidi, and S. Aral (2007). Social capital and rates of gonorrhoea and syphilis in the United States: Spatial regression analyses of state-level associations. *Soc Sci Med*.
- Wilson, R., W. Runciman, R. Gibberd, B. Harrison, L. Newby, and J. Hamilton (1995). The Quality in Australian Health Care Study. *Med J Aust* 163(9), 458–71.
- Windmeijer, F., H. Gravelle, and P. Hoonhout (2004). Waiting Lists, Waiting Times and Admissions: an Empirical Analysis at Hospital and General Practice level. *IFS Working Paper WP04/35*.

Table 1: Variables and Definitions

Variable	Definition	Remarks
UT_r	Utilization of public elective care by patients residing in region r .	Measured by hospital separations per thousand population.
A_c^p	Average access costs for private care.	Measured by the private bed density (number of private beds per thousand population) in region r .
p	Average quality of care expected by public patients.	Proxied by the actual quality of care received by public patients in region r <i>in the previous period</i>
C_p	Average copayments for private care in region r	Proxied by the private health insurance (PHI) coverage in the region r .
F	Average fixed cost for health care	Proxied by the population characteristics of a region (Fuchs, 2004). These characteristics include the socio-economic index for areas (SEIFA) for Economic Resources (ER) and proportion of indigenous population in region r .
A_c	Average access costs for public care in region r .	Proxied by the actual travel time taken by patients for hospital care. The detailed method is discussed in Appendix A.4.2.
f_{SP}	Referrals by specialists in region r	Proxied by Medicare benefits (B_{SP}) to specialists per thousand population
f_{GP}	Referrals by GPs in region r	Proxied by Medicare benefits (B_{GP}) to GPs per thousand population
X_h	Resource and efficiency constraints of the hospitals	Proxied by hospital characteristics in region r : Two dummy variables used for teaching and regional and sub-urban hospitals.
B^h	Total budget of hospitals	It is dependent on the severity of illness of patients (z) and hospital size. Proxied by the public bed density (number of public beds per thousand population) in region r .

Table 2: Summary Statistics: Area Level Variables

Variables	N	Mean	SD	Min	Max
1998-1999					
PHI Coverage (%)	45	33.6	1.5	30.2	37.4
SEIFA (EO)	45	993.0	41.4	940.8	1115.6
Indigenous Population (%)	45	0.7	0.6	0.1	2.9
2000-2003					
PHI Coverage (%)	45	51.6	6.2	43.0	68.0
SEIFA (ER)	45	973.5	55.4	901.5	1153.9
Indigenous Population (%)	45	0.8	0.7	0.1	3.2
Time Constant Variables					
Surgeon Density	45	24.6	35.8	0.3	209.9
GP Density	45	115.2	32.7	76.8	250.6
Bed Density (Pub.Hosp)	45	2.79	2.18	0.0	12.5
Bed Density (Pvt. Hosp)	45	1.09	1.43	0.0	9.1

Table 3: Summary Statistics by Diagnostic Category

Variable		Mean	Std. Dev.	Min	Max
Aggregate Utilization					
Separations- per 10000 pop.	overall	547.92	229.90	132.91	1532.45
	between		209.20	153.13	1092.76
	within		99.37	163.12	1029.16
Quality (%)	overall	10.45	1.99	4.70	16.22
	between		0.83	8.19	11.79
	within		1.82	5.83	17.35
Travel Time (mins)	overall	35.47	49.49	2.00	388.00
	between		40.92	4.00	234.80
	within		28.37	-197.33	188.67
FRS					
Separations- per 10000 pop.	overall	234.02	108.80	57.91	620.29
	between		98.62	66.54	447.70
	within		47.80	33.06	406.61
Quality (%)	overall	6.64	2.33	1.45	16.67
	between		1.86	4.13	11.96
	within		1.46	2.05	11.35
Travel Time (mins)	overall	35.15	55.43	2.00	389.00
	between		44.74	4.00	231.60
	within		33.28	-194.45	192.55
Hip & Knee					
Separations- per 10000 pop.	overall	84.86	41.42	18.52	243.45
	between		38.39	21.93	188.97
	within		16.39	5.48	142.56
Quality (%)	overall	12.59	3.31	3.11	19.55
	between		2.70	5.25	16.95
	within		1.91	7.79	18.65
Travel Time (mins)	overall	40.44	54.46	2.00	388.00
	between		46.06	4.00	234.60
	within		29.72	-192.16	193.84
Digestive					
Separations- per 10000 pop.	overall	51.91	22.57	13.64	126.57
	between		20.90	15.10	103.20
	within		8.98	3.01	95.02
Quality (%)	overall	13.77	3.85	5.00	24.76
	between		2.80	8.72	20.88
	within		2.62	6.91	22.80
Travel Time (mins)	overall	31.48	49.12	0.00	389.00
	between		40.06	3.00	234.50
	within		28.92	-201.02	185.98

Table 4: Spatial Error Random Effects Regression

Variables Weight \rightarrow	FRS		Digestive Sys		Hip & Knee		All	
	Geographic	Economic	Geog.	Econ.	Geog.	Econ.	Geog.	Econ.
PHI (%)	-2.040 (0.000)	-2.060 (0.000)	-0.003 (0.950)	-0.030 (0.690)	-0.080 (0.541)	-0.020 (0.864)	0.859 (0.158)	0.385 (0.600)
GP Density	0.369 (0.580)	0.351 (0.598)	0.023 (0.856)	0.028 (0.829)	-0.090 (0.680)	-0.210 (0.399)	1.650 (0.240)	1.420 (0.327)
Surgeon Density	-1.709 (0.012)	-1.620 (0.018)	-0.088 (0.535)	-0.110 (0.430)	-0.310 (0.019)	-0.280 (0.246)	-3.580 (0.001)	-3.290 (0.018)
Travel Time (mins.)	-0.180 (0.052)	-0.186 (0.049)	-0.033 (0.110)	-0.037 (0.083)	-0.046 (0.209)	-0.059 (-1.560)	-0.370 (0.102)	-0.340 (0.149)
Quality	0.445 (0.878)	0.676 (0.819)	-0.470 (0.052)	-0.610 (0.025)	0.257 (0.700)	0.247 (0.730)	-2.850 (0.440)	0.081 (0.983)
SEIFA (ER)	0.295 (0.064)	0.247 (0.145)	-0.104 (0.036)	-0.083 (0.114)	-0.060 (0.220)	-0.167 (0.116)	0.250 (0.640)	-0.040 (0.940)
Public Bed Density	18.630 (0.030)	17.770 (0.030)	3.130 (0.050)	3.260 (0.050)	2.440 (0.410)	2.190 (0.476)	37.160 (0.036)	36.720 (0.039)
Private Bed Density	-11.380 (0.438)	-11.730 (0.422)	-5.590 (0.048)	-5.680 (0.047)	-3.820 (0.450)	-1.850 (0.727)	-47.830 (0.122)	-48.260 (0.118)
Indigenous Pop. (%)	6.830 (0.706)	6.850 (0.700)	0.400 (0.910)	1.260 (0.730)	7.950 (0.210)	6.070 (0.353)	50.750 (0.166)	43.680 (0.250)
Teaching (Dummy)	-39.350 (0.331)	-38.600 (0.337)	-10.080 (0.204)	-8.910 (0.264)	-7.030 (0.621)	-10.200 (0.486)	-66.780 (0.430)	-66.780 (0.420)
Regional (Dummy)	35.150 (0.208)	34.040 (0.220)	6.410 (0.234)	6.650 (0.219)	20.120 (0.040)	19.490 (0.048)	77.830 (0.179)	76.300 (0.176)
Constant	-10.530 (0.955)	39.640 (0.840)	54.500 (0.004)	35.650 (0.017)	52.820 (0.000)	69.190 (0.000)	66.920 (0.910)	90.640 (0.570)
θ	1.508 (0.000)	1.480 (0.000)	1.510 (0.000)	1.440 (0.000)	1.450 (0.000)	1.450 (0.000)	1.480 (0.000)	1.400 (0.000)
λ	-0.184 (0.174)	-0.013 (0.900)	-0.503 (0.000)	-0.060 (0.550)	-0.410 (0.004)	-0.220 (0.080)	-0.630 (0.000)	-0.360 (0.000)
\bar{R}^2	0.670	0.660	0.720	0.700	0.670	0.660	0.720	0.700

Table 5: Test for Access-cost Quality Trade-off

Diagnostic Category	Ratio: $\frac{\beta_{quality}}{\beta_{time}}$	p-Value
FRS	-3.87	0.817
Digestive Sys.	15.39	0.204
Hip& Knee	-3.15	0.604
All	-22.79	0.940

APPENDIX

A.1 Derivation of GP and Specialist Utility Maximization

Utility function of GP:

$$V_t = \log(c_t) + \log(x_t) \quad (\text{A-1})$$

GP Leisure:

$$1 - x_t = (\nu_1 f(z') + \nu_2 f(z'')) \quad (\text{A-2})$$

GP Budget Constraint:

$$c_t + \tau_t = t_p(\nu_1 f(z') + \nu_2 f(z'')) \quad (\text{A-3})$$

Substituting (A-2) in (A-3) we get:

$$c_t = t_p(1 - x_t) - \tau_t \quad (\text{A-4})$$

Substituting (A-4) in (A-1) Utility function can be written in terms of x_t :

$$V_t = \log(x_t) + \log(t_p(1 - x_t) - \tau_t)$$

Maximizing V w.r.t. x_t , the FOC is:

$$\frac{\partial V}{\partial x_t} = 0 \quad \Rightarrow \quad \frac{1}{x_t} + \frac{1}{t_p(1 - x_t) - \tau_t}(-t_p) = 0$$

Rearranging the FOC we get the optimal value for x_t :

$$x_t^* = \frac{t_p - \tau_t}{2t_p} \quad (\text{A-5})$$

Substituting (A-5) into (A-4) the optimal value of consumption is given by:

$$c_t = \frac{t_p - \tau_t}{2} \quad (\text{A-6})$$

Substituting (A-6) into (A-3) and solving for $f(z')$ we get:

$$f(z') = \frac{t_p + \tau_t}{2\nu_1 t_p} - \frac{\nu_2}{\nu_1} f(z'') = \frac{B_{GP}}{\nu_1 t_p} - \frac{\nu_2}{\nu_1} f(z'')$$

The utility maximization problem for specialist can be derived in exactly same way.

A.2 Proof of Lemma 1: Access-Cost Quality Tradeoff

Solving (15) at the threshold gives:

$$E(U)e^{-zQ} = (kA_c + F) \tag{A-7}$$

Substituting (10) in (A-7) and rearranging we get:

$$A_c = \frac{(pU_1 + (1-p)U_2)e^{-zQ}}{k} - F \tag{A-8}$$

First Differentiating (A-7) w.r.t. p we get

$$\frac{\partial A_c}{\partial p} = \frac{(U_1 - U_2)e^{-zQ}}{k}$$

Since by assumption $U_1 > U_2$, $k > 0$ the above expression is positive.

A.3 Derivation of Upper and Lower Limits in Equation

Lower Limit: Lower limit is the limit above which the patient will prefer public care to no care. Rearranging (15) we get

$$E(U) > (kA_c + F)e^{zQ}$$

Upper Limit: Upper limit is the limit below which the patient will prefer public care to private care. Substituting (12) in 17 we get:

$$E(U)e^{-zQ} - E(U) + \eta < kA_c - k(A_c^p + C_p)$$

Rearranging above equation we get:

$$E(U) < \frac{(kA_c^p + C_p) - kA_c + \eta}{(1 - e^{-zQ})}$$

A.4 Construction of Quality and Travel Time Variables

A.4.1 Quality

The quality of care variable is proxied by the rate of adverse event for a particular MDC. The rate of adverse event is calculated by using the method suggested in Jackson et al. (2006). All admissions to Australian Hospitals are have diagnoses coded using the International Classification of Diseases 10th revision Australian Modification (ICD-10-AM) after discharge. The codes are assigned using the Australian Coding Standards and for public hospitals a random sample of records is recoded independently as part of a state-wide coding audit to assure accuracy of the diagnosis and procedure data. Adverse events are identified as follows: i) ICD codes that by definition imply an adverse event are used. These are codes in the range of T80-T88.9 that relate to complications of surgical or medical care; ‘end of chapter codes’ which are used for coding only complications of care, ii) The ‘external cause codes’ (a requirement of Australian coding rules) in the range of Y40-Y84.9 are used to identify the diagnosis that is externally caused and was a complication of medical or surgical care. The external cause code is also accompanied by the ‘place of occurrence code’ where the code value Y92.22 indicates that external cause of any injury or illness arose within a health facility. In addition code Y95 indicates nosocomial conditions specifically documented as hospital-acquired in the medical record and iii) Unplanned readmission rate in the same hospital within 28 days of discharge is also used as an indicator of adverse events. Hospital episode which satisfies any of the above criteria is identified as an adverse event.

A.4.2 Travel Time

Travel time is calculated between the centroid of the patient’s postcode of residence and the hospital street address. This is done by using Roadnet software which classifies roads into seven categories and assigns an average speed to each category. Travel time is subsequently calculated by multiplying the average speed of travel and distance through the shortest route.