Demand Variability, Demand Uncertainty and Hospital Costs: A Selective Survey of the Empirical Literature

Zoe Boutsioli

Athens Institute for Education and Research (ATINER)

8 Valaoritou, street, 10671, Athens, Greece

Tel: 30-210-363-4210   E-mail: zb5@atiner.gr

Abstract

Demand for hospital services varies in an uncertain way and affects both the optimal size and the variable cost of hospitals. From this selective review the following main points for further research are resulted. The theory of hospital costs needs to take into consideration the impact of demand variability. The insurance role of hospital excess capacity should be treated endogenously. Better econometric techniques can improve the efficiency of empirical results. Disaggregated data may be more appropriate, and finally, more countries in order to be able to generalize the conclusions concerning the impact of demand variability on hospital costs are needed.

Keywords: Hospital demand, Variability, Review, Hospital cost

1. Introduction

The demand for hospital services is characterized by considerable variability and is often unpredictable. Demand variability affects both the optimal size and the variable cost of the hospital. Long and Feldstein (1967, p. 119), in an early article of this literature, have stated "…hospitals must build larger facilities and hire more staff than would be necessary if demand were constant over time." Surprisingly, very few empirical studies exist that examine the impact of hospital demand variability on hospital capacity levels and production costs. In an early survey of hospital costs, Cowing *et al.* (1983) concluded that there was a "…complete absence of any significant work in this area." As an exception, they mention the work of Joskow (1980) and Friedman and Pauly (1981). They predicted that "…we will begin to see empirical analyses of hospital cost structures that take account of this most important dimension of hospital services," namely variability in the demand for hospital services. In the 1990s, Gaynor and Anderson (1995) were still emphasizing the dearth of empirical research on this topic. Ten years after there exists very few empirical studies that deal with the issue of demand variability, but they are sufficient to survey them in order to identify common analytical frameworks, applied methodologies, modeling techniques, types of data, covered area (country or region), and major empirical findings.

This review emphasizes the role of demand variability and demand uncertainty on hospital costs. The two concepts are interrelated but they are not identical. Hospital demand can vary in a very predictable and certain way, i.e. during holidays the demand for hospital beds is lower. It can, generally, vary day to day in a way that can be approximated by a distribution density function such as the Poisson distribution or the Normal Density Function. In this case, hospital decisions become one of risk. The parameters of the probability function are known, which enables hospitals to estimate *a priori* the probability of an excess demand. In risk situations, hospitals can plan ahead in choosing optimal capacity and purchase of inputs, such as staffing. However, hospital demand is often unpredictable in the sense that its variation cannot be described by any known probability distribution. Contrary to predicted variations in hospital demand, unpredicted demand is only *ex post* estimated. Total hospital demand variability is the sum of predicted and unpredicted demand. Demand uncertainty is defined as the unpredicted part of demand variability. Demand variability and particularly demand uncertainly has a major impact on hospital costs.

The 1960s and 1970s have seen an explosion in hospital cost analysis as part of a general interest in the estimation of cost functions in the early years of this period (Lave and Lave 1970). Martin Feldstein (1974) pointed out that the growth of research in hospital cost analysis can be explained by two factors: public policy and economic theory. The health sector is used as a data generation process to test economic models expanded to include nonstandard market characteristics, such as nonprofit objective functions, consumer uncertainty and third party financing. Apart from these two factors, a third factor can be added that relates to hospital managers' herculean effort to reduce costs and increase efficiency without reducing the quality of health care services. Planning in an environment of uncertain demand is a managers' nightmare when cost minimization is the

objective. Thus, empirical hospital cost analysis is a challenge for economic theory, policy makers and hospital managers.

The purpose of this paper is to review selected empirical studies of the impact of hospital demand variability and uncertainty on capacity and costs. This paper is organized in five sections, including this short introduction. Section two provides a concise presentation of the theory of hospital demand uncertainty and its impact on capacity and costs. Section three provides a selective review of the empirical literature of demand variability and uncertainty on hospital costs. In total, 12 studies are reviewed. In this review are included studies that directly assess the relationship between demand variability and hospital costs, and also studies that examine the topic of excess capacity on hospital costs and the cost of empty hospital beds. Section four discusses the implications of this literature review for future research in all three areas of interest: theory, policy and management. Section five summarizes the main issues and concludes.

2. Demand Variability, Demand Uncertainty and Hospital Cost: The Theory

2.1 The Simple Theory

The theory of firm facing an uncertain demand, the most important part of demand variability, was developed in the early 1970s by Baron (1971) and Leland (1972). In a more general form, Leland concluded that the presence of demand uncertainty alters the predictions of the theory of firm under certainty. Most importantly, Leland has shown that the hospitals' (firms') selection of prices and outputs will not be invariant to changes in fixed costs, i.e. hospital capacity. The traditional cost theory, which assumes that demand is certain, defines total hospital costs c as the sum of the product of input prices r and input quantities x. If n inputs were used by hospital managers, then the total hospital cost in period t is

$$c_t = r_{1t}x_{1t} + r_{2t}x_{2t} + \ldots + r_{nt}x_{nt} = \sum_{j=1}^{n} r_{jt}x_{jt}$$

Hospitals produce health care services y using the n inputs, and a technology described by the following hospital production function

$$y_t = f(\mathbf{x_t})$$

where $\mathbf{x_t}$ is the vector of the n inputs.

Hospital mangers minimize costs subject to a production function. This yields demand functions for the n inputs that depend on all input prices r, and the hospital output y. Thus, for the j_{th} input the demand is:

$$x_{jt} = \varphi_{jt}(r_{1t}, r_{2t}, \ldots, r_{jt}, \ldots, r_{nt}, y_t)$$

Substituting these demand functions into the cost function, we obtain hospital cost as a function of input and output

$$c_t = \psi(r_{1t}, r_{2t}, \ldots, r_{nt}, y_t)$$

2.2 Demand Variability and Uncertainty

The above approach does not take into consideration demand variability and uncertainty. Gaynor and Anderson (1995), using the work by Duncan (1990) and Friedman and Pauly (1981), developed a theoretical model that explicitly incorporates demand variability in the hospital sector. The starting point of this analysis is the decisions that hospital managers must make a decision on capacity and input uses. The latter includes staffing. When managers face an environment of demand variability, then they have to set what is known a "turnaway probability." This measures the risk that hospital managers are willing to take in case future demand will exceed capacity. In this case, new patients are turned away or some of the existing patients are discharged earlier. In the early empirical study by Joskow (1980), a Poisson distribution was assumed. However, as Joskow himself claims the normal distribution is a good approximation when the mean is relatively large, as is the case in hospitals. More recently, Baker *et al.* (2004) have concluded that the daily census probability has a characteristic normal shape for all hospitals they observed. Thus, hospital face demand z which is a random variable that can be approximated by the normal distribution

$$z \sim N(\mu, \sigma^2)$$

where μ is the mean and σ^2 is the variance of hospital demand.

In this model, hospital managers are assumed to know the parameters of the demand distribution. Given that demand is normal, the mean and variance gives full information on the structure of the demand variability. Managers decide on the turnaway probability α, which determines the level of hospital capacity given by

$$z_c = \mu + z_a * \sigma$$

where z_c is the maximum demand that can be satisfied with the existing hospital capacity and z_a is the one tail value of the z distribution of α probability. To illustrate, assume that the average demand for beds is 160, with a variance of 400 beds. If the distribution is normal and the turnaway probability is set equal to 5%, then the optimal capacity is given by

$$z_c = 160 + 1.28 * 20 = 186 \text{ beds}$$

Thus, to satisfy at least 95% of demand for hospital beds, given the demand variability, managers should build a capacity of 186 beds.

How does this variability affect cost? Following Gaynor and Anderson (1995) very closely, it is assumed that demand for a hospital's output is given by the following conditional distribution function

$$G(z_t \mid z_{t-k}), \quad k = 1, \ldots .t$$

The probability that demand exceeds capacity is given by

$$\Pr[z_t \geq y_t \mid z_{t-k}] = \Pr[z_t \geq f(\mathbf{x_t}) \mid z_{t-k}] = 1 - G(f(\mathbf{x_t}) \mid z_{t-k}) = \alpha$$

Hospitals minimize costs subject to the constraint that the demand can exceed capacity α percent of the time

$$\min_{\mathbf{x_t}} c_t = \sum_{j=1}^{n} r_{jt} x_{jt.}$$

$$\text{s.t. } 1 - G(f(\mathbf{x_t}) \mid z_{t-k}) = \alpha$$

Solving this constrained maximization problem, results to a variable cost function that depends on (a) hospital's output, (b) the target standby capacity $G^{-1}(1-\alpha \mid z_{t-k})$, which is a function of the "turnaway probability", (c) the prices of variable inputs and (d) the fixed factors of production k

$$c_t = c_t(y_t, G^{-1}(1-\alpha \mid z_{t-k}), \sum r_{jt}, k_t)$$

2.3 The Empirical Specification

The above theoretical cost function requires a very careful interpretation when it is used for empirical purposes. The most important variable is the output produced by the hospital given the variability and the uncertainty of demand. Capacity output is determined *ex ante* by the hospital given the turnaway probability. This enters directly the cost function by the vector of fixed inputs k. The realized (actual) output is the sum of the predicted and unpredicted demand for hospital services that cannot exceed capacity levels. Since unpredicted demand cannot be forecasted, it can only be estimated *ex post*. An *ex post* evaluation is given by the forecasting error of an econometric model. For example, Gaynor and Anderson (1995) use the mean and variance of the actual demand as a proxy for unpredicted demand. Hughes and McGuire (2003) use a simple autoregressive process to account for the demand uncertainty. The difference between realized and forecasted demand is then used as a measure of demand uncertainty.

Given the fixed factors of production k, Baker *et al.* (2004) have added another impact of demand variability on hospital costs. Hospitals use long-term contracts for the provision of variable inputs. To match realized demand that exceeds the one that can be satisfied with the inputs purchased in the long-term contracts market, hospitals must purchase these inputs in the spot market at a price higher than the long-term market. Thus, the interpretation of the theoretical cost function requires that a distinction should be made between inputs that are bought through long-term contracts and the same inputs that are bought from the spot market. For example, nursing staff can be permanent and on demand. The per hour wage rate of the permanent nursing staff is assumed to be lower than the per hour nursing service that is bought in the spot market. Thus, the vector of input prices in the above theoretical cost function must be interpreted as including long-term and spot input prices.

Demand uncertainty explains part of the observed excess capacity of hospitals. This excess capacity can be considered as insurance or more generally as service to the public should a citizen unpredictably require hospital

services. If this demand uncertainty is not taken into consideration, there is a misspecification error of the empirical cost function. However, the empirical cost functions include more variables as determinants of hospital cost. An empirical specification is given by

$$c_t = c_t(y_t, G^{-1}(1-\alpha \mid z_{t-k}), \Sigma\, r_{jt}, k_t, \theta_t)$$

where θ_t is a vector of other determinants of cost such as case mix, teaching status, urban/regional hospitals, public and for-profit hospitals, etc.

The economic theory does not provide a clear cut view on the functional form of the cost functions to be used in econometric estimations. The dominant functional form of cost function estimation is the translog specification, but others have been used as well such as quadratic or simple logarithmic transformations. Recently, empirical estimations use nonparametric estimation methods. These issues are discussed in length by Wilson and Carey (2004), Dranove (1998), Vita (1990) and Vitaliano (1987). These are general econometric problems that pertain to econometric methodology and are not specific to hospital cost function estimation. If one of the purposes of econometrics is to verify economic theories, then parametric approaches to estimation is the appropriate methodology.

3. Review of Empirical Studies

This section reviews empirical studies that in one way or another incorporate demand variability and demand uncertainty in determining hospital cost and the efficient allocation of resources within the hospital unit. In total, 12 studies are reviewed. Table 1 provides a comparative illustration of these studies in terms of whether they include demand variability or uncertainty, the period covered, the type of data, the number of hospitals, the country (area), the measurement of uncertainty and variability, the dependent variable and its determinants and the main conclusions reached. The publication dates extend over three decades from 1967 to the most recent one in 2008. Four studies were published in the 1980s, three in the 1990s and four after 2000.

An early review of the hospital cost literature was provided by Cowing *et al.* (1983). They dealt with all empirical studies of hospital cost that were available up to that period. There were only two empirical studies available for the treatment of demand variability and demand uncertainty. Recently, Smet (2002) reviewed selected studies of the empirical literature on hospital costs. His emphasis was on theoretical and methodological aspects of the hospital cost literature. Smet (2002, p. 902) included only the study by Gaynor and Anderson (1995) which treats uncertainty of demand. One important conclusion of this review is that "…hospitals generally over-invest in capital (capacity and equipment)." A possible explanation is that this apparent over-investment is an efficient response to uncertain demand.

Long and Feldstein (1967) examined techniques for estimating optimal hospital capacity levels given the peak load nature of demand. They restrict their analysis to only one hospital unit (i.e. obstetrics) and in the one area of Chicago, USA. They do so in order to keep the analysis simple from a mathematical point of view and because in this unit they observed the highest excess capacity. Their analysis includes two years, 1964 and 1965, and data were collected from 34 hospitals. The cost is regressed using a linear form on the number of beds, the number of patients and the length of stay. They concluded that improvements in handling demand variability in hospitals could reduce the costs of care. Another policy related conclusion of their study is that mergers to achieve efficiency (reduce excess capacity) may not be necessary if the hospital system permits transfer of patients. Their analysis shows the tremendous problems that the empirical economists face when there is a lack of appropriate data and they want to adhere to the theoretical fundamentals. Their econometric investigations do not include input prices and output quality because of lack of such data. The econometric findings may have multiple explanations. For example, they found a negative constant term which might be interpreted as (a) the functional form is not linear, (b) there were omitted variables and/or (c) there are diseconomies of scale or a combination of all three.

Joskow (1980) provides a simple model of capacity decisions. Excess capacity is costly and the model takes account of demand uncertainty, the internal organization of the hospital, non-price competitions among hospitals and hospital regulation. He assumes that demand (in-patients admissions) can be described as a Poisson process but in any case this is not necessary because for realistic parametric values of the Poisson distribution, it can be approximated by the normal density function. He defines the average reserve margin R for the hospital as

$$R = \text{BEDS-ADC} = k(\text{ADC})^{1/2}$$

where BEDS is the capacity, ADC is the average daily census that follows the Poisson distribution, k is the number of standard deviations from the mean (ADC) and the square root of ADC is the standard deviation of the

average daily census. Hospitals choose k and by doing so they reveal their preferences for the turnaway probability. Annual data from 346 hospitals in the USA in 1976 were used in a linear regression model. The impact of demand variability and uncertainty, they are measured by R and directly by k, on hospital costs depend on regulation and hospital concentration in one area. For example, Joskow (1980) found that a monopoly hospital will choose a k that is 0.5 standard deviation below the mean. This implies a considerable reduction in the cost of excess capacity. Thus, competition among hospitals may lead to excess bed supply and therefore higher hospital cost. On the other hand, a regulated hospital will choose a lower value of k which on average will reduce the supply of beds by 6% but the impact on hospital cost would be substantially less because the marginal cost saving is far below the average total cost per bed. This paper does not investigate directly the impact of demand uncertainty on hospital cost but the impact on the decision to supply more capacity. However, it links the analysis to hospital costs. It also talks about uncertainty when it actually investigates demand variability, i.e. the standard deviation of average daily census. In a follow up paper, Mulligan (1985) has showed that most of Joskow's simplifying assumptions were not necessary. If he has used the exact probabilities instead of the proxy, he would have found greater support for his empirical evidence and policy implications.

Friedman and Pauly (1981) examine cost functions when demand is stochastic and one of the firm's possible reactions is the deterioration of the quality of output. Their analysis of hospital costs starts with an observation that the short-run marginal cost of hospitals is below the average cost and the in the long-run they are about the same. This cannot be explained assuming certain demand. Friedman and Pauly (1981) developed a hospital cost function that is subject to demand uncertainty and long-term contracts in input purchases. Quarterly data from 800 hospitals in the USA over the 1973-1978 period were used to estimate a cost function using a number of determinants (Table 1). They found that costs depend on the ratio of expected to actual hospital output. In another study by the same authors, Friedman and Pauly (1983), monthly data were used from 870 hospitals in the USA. They found that unexpected (transitory) changes in demand have a positive impact on average cost per admission. The cost of excess bed capacity is positive but much smaller than the one thought by the regulatory authorities. A number of policy implications are drawn from the empirical evidence. First, policies to restrict excess capacity do not save costs. Second, the evidence supports proportional revenue allowances for volume change as part of hospitals' cost containment policies consisting of budget and price regulation.

Pauly and Wilson (1986) examined 176 hospitals from the Michigan area over a four year period (1979-1982). They used a linear model to estimate the average hospital cost as a function of the ration of forecast to actual discharges and other variables. They found that demand variability has a positive impact on average hospital cost. They also report results using a logarithmic specification. They found that the actual cost of excess capacity is lower than what was thought for the Michigan hospitals, particularly for small and medium sized hospitals. They also discussed the issue of closing down the operation of some hospitals and in such a way, increase the occupancy of the other hospitals in the area, cutting down the industry cost. However, the cost savings are estimated as very small. This result is similar to Joskow's and Mulligan's finding that hospital concentration may lead to a reduction in costs because the impact of demand variability and uncertainty on hospital cost is reduced.

Gaynor and Anderson (1995) reformulated the theory of hospital cost and production to include demand uncertainty. An extensive sample of 5000 hospitals over the years 1983-1987 is used to estimate short-run hospital cost functions. Annual data is used and their findings are substantial. They found that demand uncertainty affects input choices and therefore costs. They considered a number of econometric problems including the functional form of the cost specification. They employed a translog specification but they test a Cobb-Douglas specification that was rejected. Another important econometric issue that is fruitful field of future research was the exogeneity of important determinants of hospital costs: admissions, occupancy rate and outpatient visits and the case mix. Exogeneity was rejected for the first two, but they could not reject it for the other two. They used instrumental variables to account for the possible endogeneity of admissions and occupancy rate. They also found heteroskedasticity that they corrected using White's method. Since this econometric problem relates to economies of scale, nonparametric methods might be more appropriate when the sample includes hospitals with large differences in their means, i.e. bed capacity. Their results support the hypothesis that demand variability and uncertainty have a positive impact on average hospital cost. Forecasted admissions and the variance of forecasted admissions have a significant impact on cost. They interpreted that as being consistent with their theory of uncertain demand. The greater the uncertainty, the more hospitals hold excess capacity and therefore, the higher the cost. They also found that the cost of an empty bed is higher for hospitals with greater expected admissions. However, the difference in empty bed costs between hospitals facing high and low variability is not very large. As they acknowledged, their results are subject to the model limitation, such as the use of only one input price, lack of quality index and the estimation of a short-run cost function.

The next two studies of Table 1 use daily data to estimate the demand variability and uncertainty and its impact on hospital cost. Lynk (1995) used the standard deviation of daily census as a dependent variable regressing it on the average daily census and four hospital units. Their results only indirectly can be interpreted for the cost impact. They found that the reduction in demand volatility creates the necessary conditions to reduce costs through the consolidation of clinical services. The paper by Carey (1998) examined the excess bed capacity as a function of stochastic demand. She used daily data from over 2500 hospitals in the USA over the years between 1987 and 1992. A distinction is made between small (less than 100 beds) and large hospitals. This is critical because the assumption of demand uncertainty is supported by the evidence from the larger hospitals but it does not hold for smaller ones.

Hughes and McGuire (2003) considered the production responses to demand uncertainty, and its implications on hospital cost structures from 1993 to 1995 in a sample of 85 UK NHS hospitals. Demand uncertainty is measured as the forecast errors in monthly emergency admissions. This explanatory variable is included with other cost determinants in the total operating cost function. Other cost shifters included total inpatient elective and emergency admissions, number of staffed beds, allocated to the elective and emergency sector, the inverse of the occupancy rate, outpatient visits, day attendances, accident and emergency outpatient visits, a wage index, and a dummy variable for teaching hospitals. From regression analysis, they found a positive and significant relationship between uncertain demand and hospital cost: the higher the extent of the uncertainty, the higher the total variable cost will be. Also, it was calculated that the marginal costs of an emergency and elective admission were £1744 and £823, respectively, in 1995 prices. Thus, the cost arising from the inpatient treatment of an additional elective case is about 45% of the additional cost of an emergency case. The cost of uncertainty per emergency admission is estimated to be £77, or 5% of the total cost of an emergency admission.

Table 1

Baker *et al.* (2004) examined day-to-day variation in hospital utilization over a long time period from 1983 to 1995, and its effects on costs. They highlighted two ways that demand variability can influence costs: capacity and staffing decisions. In the former case, the capacity a hospital requires will be a function of demand variability, while in the latter case, input purchasing influences the hospital costs. The cost functions developed differ from that in the traditional cost theory by incorporating the average daily demand and the standard deviation of demand facing a hospital within the cost equation. To compute estimates of variation in hospital utilization, the daily 'census' for about 500 Californian hospitals for 12 years were used. Baker *et al.* (2004) found that only a negligible number of hospitals in their sample are almost full. Considerable day-to-day variation in hospital utilization and significant differences in utilization patterns among hospitals were found. The calculation of inpatient operating costs was based on annual financial data from the California Office of Statewide Health Planning and Development. A hybrid form of the trans-log cost function was used for estimating a short-run hospital cost function. This cost function was elaborated with measures of variance and controls for average census, beds, wages, and other cost shifters, such as case mix, the percentage of HMO discharges, a dummy variable for teaching hospital, and a dummy variable for ownership status. For controlling bed size, a proxy for capital stock and fixed costs, the number of available beds (but not necessarily staffed) in the hospital is used. The wage variable they use is the average wage for six categories of labor, weighted by the total hours worked by each. From their econometric analysis, they found that the variance coefficients are statistically significant, with small effects on total inpatient operating expenditures. To illustrate, a 1% increase in variance would lead to a 0.092% increase in inpatient operating costs.

Smet (2007) measured the performance of Belgian general care hospitals in the presence of stochastic demand for hospital services. To measure standby capacity held by hospitals to serve unexpected demand he used a waiting time indicator derived from queuing theory. He argued that such an indicator improves on the conventionally used inverse of the occupancy rate. The entire population of 187 Belgian general care hospitals for the year 1997 was used. Stochastic Frontier Analysis (SFA) was used to estimate a multi-product translog stochastic cost function. His hospital cost function had total operating cost as the dependent variable and the number of patient days for seven categories and the number of admissions as output measures. Additionally, other variables included in the cost model were the number of beds, the occupancy rate and the queuing indicator for standby capacity, region dummy variables, teaching dummy variables, and ownership dummy variables. He found that standby capacity is significant in explaining the level of hospital costs. Hospitals trying to reduce their waiting time indicator by providing sufficient extra standby capacity to cope with uncertain demand will incur higher costs than hospitals that provide this service to a lesser extent. The significance of the standby capacity indicator is an indication that uncertain demand should be accounted for. Finally, he stressed that the results of traditional multi-product cost function studies that hospitals generally tend to overinvest in capacity and

equipment could be explained by the fact that these empirical studies did not include the provision of standby capacity to serve unexpected demand for hospital care.

Lovell *et al.* (2008) extended the analysis of Rodriguez-Alvarez and Lovell (2004) by incorporating demand uncertainty into the technology to account for the hospital activity of providing standby capacity or insurance against the unexpected demand. By modeling a simple autoregressive (AR1) process where demand expectations are related to prior demand realizations and, also including bed size in the regression to control for hospital size, the authors measure demand uncertainty as the difference between actual demand and predicted demand (in absolute values). They applied their model in the sample of 67 Spanish general hospitals over the period 1987-1994. Their results showed that the roles of both stochastic demand and expense preference behaviour should be considered when studying hospital production and costs. Ignoring either of these effects, if they are present, will lead to biased parameter estimates and misleading inference, and if applied researchers have reason to suspect that these effects may be present that should test for them.

Summing up, the above brief review of the empirical studies three points can be made. First, authors use various specifications of costs, different data sets and make various assumptions concerning the impact of the various explanatory variables on hospital costs. Second, most studies use the standard explanatory variables of output and input prices but this is not a sufficient specification of the cost function. Additional explanatory variables are needed in order to improve the specification of the hospital cost function. Third, hospital costs functions are estimated to serve different purposes but if there is a common purpose this is to determine the level of hospital efficiency and how this can be improved by policy initiatives.

4. Implications for Future Research

There is no doubt that substantial theoretical and empirical work is necessary if we are to comprehend the impact of demand variability and uncertainty on hospital cost structure and how this is affected by important determinants of cost, such as input prices, hospital capacity and other characteristics of the hospital sector. Such research should have a sound theoretical economic background in order for the interpretation of the results to be useful to both hospital managers and policy makers. Economic theory can guide the empirical researcher in setting up an empirical model for both hospital cost and demand uncertainty. A better economic theory of the hospital is needed that will take into consideration a hospital's objective function, the control variables of the hospital decision making process and the hospital technology constraints. It is important to emphasize that this theory should take into consideration the entire nexus of the hospital market system, which includes both the nature of consumers of the hospital care services and the suppliers. Emphasis should be given on the exact nature of uncertainty. Thus, future research calls for better theoretical models of the hospital market.

The review of the empirical literature shows that the treatment of uncertainty is restricted to patients' demand, while other aspects of demand variability and uncertainty are not taken into consideration, such as the uncertainty of diagnosis and the uncertainty of the treatment of illness. Also, the uncertainty of the input supplies has been completely ignored and cannot be simply captured by differences in input prices between long-term and spot market purchases. Future research should take into consideration all of the uncertainties faced by hospitals related to the demand for health care services.

Relative to the treatment of demand uncertainty and the demand for hospital inputs is the treatment of excess capacity. Most empirical studies consider the reserve capacity or excess capacity of the hospital as being an indication of inefficient use of scarce resources and adding to the cost. There is a theoretical gap that future research should fill. Reserve capacity by hospitals to treat unexpected demand can theoretically be dealt with very easily using the economics of insurance. Hospitals can be considered as providing insurance and therefore, they keep an excess capacity to meet such a demand. The analysis is very similar to insurance provided by the fire services, the earthquake relief squads even the military. The theoretical challenge is to determine the optimal amount of such hospital insurance. Excess capacity can be thought as an efficient allocation of resources because hospitals want to meet the unexpected demand. On the other hand, excess capacity might indicate mismanagement. The empirical literature should account for these two cases. This is extremely important in a hospital unit that faces a soft budget constraint, as it is alleged as the case in public hospitals. However, if demand variability and uncertainty varies across time and hospitals, then what is efficient and inefficient allocation of resources is not easily identifiable.

The empirical methodology is another area for future research. Parametric and nonparametric methods should be used not because they provide better diagnostic tests, but because there are certain theoretical restrictions that one of the other methodologies cannot satisfy. For example, parametric methods can not account for economies of scale and it is only good as a local approximation. On the other hand, results from nonparametric methods

may be difficult to interpret since there are not based on a specific functional (theoretical) form of the hospital cost structure. Research on modeling techniques for the treatment of demand uncertainty and variability can improve our econometric results of the impact of demand on hospital cost. For example, using simple autoregressive processes may not be useful to a hospital manager that is taken into consideration all available information in deciding on input demands. A Bayesian analysis might be one way of incorporating this information revision process.

Many times the choice of the modeling techniques of the empirical research is determined by the availability of data. Most studies use aggregated data, monthly, quarterly and annually, loosing important information concerning the nature of demand variability and uncertainty. We need more studies using daily data because demand variations take place on a day by day basis. Relative to this area of research is the role of hospital managers and physicians. The data generation process may vary from hospital to hospital because managers and physicians may deal differently with demand variability and uncertainty. If demand uncertainty and variability is part of the cost, then cost minimization requires a minimization of demand variability and uncertainty. Efficient hospitals may show little impact of uncertainty and variability in costs, simply because in some hospitals the management has smoothed the variability, i.e. the physician decides if a particular case requires immediate hospitalization. In this case, the empirical evidence will show no impact of uncertainty and variability because this variable, from an exogenous one, has become a decision making variable. Future research might require the modeling of demand variability as a dependent variable and instead of one equation model we might need a simultaneous equations model. In one study reviewed in this paper, i.e. Lynk (1995), the standard deviation of daily census is the dependent variable. A useful area of future research is to employ a system of equations modeling the entire hospital market where the cost and demand variability, among other variables, are interdependent variables. In this case, the estimation techniques will be different and will relate to simultaneous econometric models.

Finally, future research should include more countries. All empirical studies reviewed in this paper that incorporate demand variability and uncertainty use data from the USA with one exception of the UK. We need more case studies in order to be able to generalize the conclusions concerning the impact of demand variability and uncertainty on hospital costs.

5. Conclusions

The purpose of this review was to present a small number of empirical studies that relate demand variability and uncertainty to hospital costs. In total, twelve studies were included in this survey. The most important conclusion that emerges from this literature is that there is a need for further research of this area. This study identified a number of research areas that future research can explore. First, the theory of hospital costs needs to take into consideration the impact of demand variability and uncertainty. Formal economic models should take this into consideration and derive hypotheses that can be verified empirically not only in two areas – the USA and the UK – but in other countries and under different contexts, i.e. different hospital management systems. The insurance role of hospital excess capacity should be treated endogenously in such an approach. Verification of this theory requires disaggregated data at the day level because otherwise important information is missing on the demand variability. Better econometric techniques and the application of simultaneous models can improve the efficiency of empirical results. Overall, this review has shown that the most recent studies are more sophisticated in dealing with demand variability. However, we need more studies to identify the mechanism and the magnitude of the demand variability impact on hospital costs.

Acknowledgements

References

Baker, L.C., Phibbs, C.S., Guarino, C., Supina, D. & Reynolds, J.L. (2004). Within-year variation in hospital utilization and its implications for hospital costs. *Journal of Health Economics*, 23(1), 191-211.

Baron, D. (1971). Demand uncertainty in imperfect competition. *International Economic Review,* 12(2), 196-208.

Carey, K. (1998). Stochastic demand for hospitals and optimising "excess" bed capacity. *Journal of Regulatory Economics,* 14(2), 165-187.

Cowing, T., Holtmann, A. & Powers, S. (1983). Hospital cost analysis: a survey and evaluation of recent studies. *Advances in Health Economics and Health Services Research,* 4, 257-303.

Dranove, D. (1998). Economies of scale in non-revenue producing cost centers: implications for hospital mergers. *Journal of Health Economics,* 17(1), 69-83.

Duncan, G.M. (1990). The effects of probabilistic demands on the structure of cost functions. *Journal of Risk and Uncertainty,* 3(3), 211-220.

Feldstein, M.S. (1974). Econometric studies of health economics. In M. Intrilligator & D. Kendrick (Ed.), *Frontiers of Quantitative Economics.* North Holland: Amsterdam.

Freidman, B. & Pauly, M. (1981). Cost functions for a service firm with variable quality and stochastic demand: the case of hospitals. *The Review of Economics and Statistics,* 63(4), 620-624.

Freidman, B. & Pauly, M. (1983). A new approach to hospital cost functions and some issues in revenue regulation. *Health Care Financing Review,* 4(3), 105–114.

Gaynor, M. & Anderson, G. (1995). Uncertain demand, the structure of hospital costs and the cost of an empty bed. *Journal of Health Economics,* 14(3), 291–317.

Hughes, D. & McGuire, A. (2003). Stochastic demand, production responses and hospital costs. *Journal of Health Economics,* 22(6), 999-1010.

Joskow, P.L. (1980). The effects of competition and regulation on hospital bed supply and the reservation quality of the hospital. *Bell Journal of Economics,* 11(2), 421-447.

Knox Lovell, C.A., Rodriguez-Alvarez, A. & Wall, A. (2008). The effects of stochastic demand and expense preference behaviour on public hospital costs and excess capacity. *Health Economics,* 18(2), 227-235.

Lave, J. & Lave, L. (1970). Hospital cost functions. *The American Economic Review,* 60, 379-395.

Leland, H. (1972). Theory of the firm under uncertainty. *The American Economic Review* 62(3), 278-291.

Long, M.F. & Feldstein, P.J. (1967). Economics of hospital systems: peak loads and regional coordination. *The American Economic Review,* 57(2), 119-129.

Lynk, W.J. (1995). The creation of economic efficiencies in hospital mergers. *Journal of Health Economics,* 14(5), 507-530.

Mulligan, J. (1985). The stochastic determinants of hospital-bed supply. *Journal of Health Economics,* 4(2), 177-185.

Pauly, M.V. & Wilson, P. (1986). Hospital output forecasts and the cost of empty hospital beds. *Health Services Research,* 21(3), 403-428.

Rodriguez-Alvarez, A. & Knox Lovell, C.A. (2004). Excess capacity and expense preference behaviour in National Health Systems: an application to the Spanish public hospitals. *Health Economics,* 13(2), 157-169.

Smet, M. (2002). Cost characteristics of hospitals. *Social Science and Medicine,* 55(6), 895-906.

Smet, M. (2007). Measuring performance in the presence of stochastic demand for hospital services: an analysis of Belgian general care hospitals. *Journal of Production Annals,* 27(1), 13-29.

Vita, M.G. (1990). Exploring hospital production relationships with flexible functional forms. *Journal of Health Economics,* 9(1), 1-21.

Vitaliano, D.F. (1987). On the estimation of hospital cost functions. *Journal of Health Economics,* 6(4), 305-318.

Wilson, P.W. & Carey, K. (2004). Nonparametric analysis of returns to scale in the US hospital industry. *Journal of Applied Econometrics* 19(4), 505-524.

Table 1. Empirical studies on hospital cost including measurements for demand uncertainty and/or variability

Author(s) (Year)	Period covered	Data	Number of Hospitals	Country	Dependent Variable	Independent Variables (in parenthesis those used to measure uncertainty or variability)	Main Conclusions
Long & Feldstein (1967)	1964-1965	Annual	34	USA	Total cost of obstetrical units	1. Number of beds 2. Number of patients 3. Length of stay	1. Demand variability affects costs 2. Mergers are not necessarily
Joskow (1980)	1976	Annual	346	USA	Reserve margins	1. (Standard deviations of average daily census) 2. Number of physicians 3. Regulatory variables 4. Competition measures	1. Regulation has a negative effect on bed supply, while competition has a positive effect
Friedman & Pauly (1981)	1973-1978	Quarterly	800	USA	Average cost per admission	1. (Forecasted to actual admissions. The forecasted admissions based on a time series univariate model on past admission values.) 2. Inverse of bed occupancy rate 3. Number of beds 4. Wage rate 5. The % of total days spend in ICU 6. The number of visits to the operative suite per admission 7. The number of outpatient clinic visits per admission 8. The number of radiology procedures per admission 9. The number of respiratory therapy per admission	1. Hospital costs are highly sensitive to unexpected demand. 2. Larger hospitals serve a more expensive case mix have higher occupancy rates, and higher average costs. 3. The marginal cost for an expected admission is 98% of its average cost, while for an unexpected admission the proportion is only 35%.
Friedman & Pauly (1983)	1973-1978	Monthly	870	USA	Average cost per admission	1. (Forecasted to actual admissions. The forecasted admissions based on a time series univariate model on past admission values.) 2. Inverse of bed occupancy rate 3. Number of beds 4. Wage rate 5. The % of total days spend in ICU 6. The number of visits to the operative suite per admission 7. The number of outpatient clinic visits/admission 8. The number of radiology procedures per admission 9. The number of respiratory therapy per admission	1. Unexpected changes in demand have a positive impact on average cost per admission

Table 1. (cont.)

Author(s) (Year)	Period covered	Data	Number of Hospitals	Country	Dependent Variable	Independent variables (in parenthesis those used to measure uncertainty or variability)	Main Conclusions
Pauly & Wilson (1986)	1979-1982	Quarterly	176	USA	Total inpatient cost per discharge	1. (Forecasted to actual discharges. The forecasted values are provided by the hospitals themselves as part of their budget packages submitted to payers.) 2. Input prices 3. Percent patient-days in special care units 4. Length of stay 5. Outpatient costs as a % of inpatient costs 6. Number of beds 7. Interns' and residents' salaries as a % of total cost	1. The actual average hospital cost increases when forecasted discharges are more than the actual.
Gaynor & Anderson (1995)	1983-1987	Annual	5000	USA	Total variable cost	1. (The mean and the standard error of forecasted admissions. Forecasted admissions are the econometric forecast of output based on a univariate time series model that uses the past values.) 2. Actual inpatient admissions and outpatient visits 3. Inverse of occupancy rate 4. Hospital beds 5. Wage rate 6. Medicare case mix index 7. Ownership and teaching dummy variables	1. Demand uncertainty affects input choices and therefore costs 2. The cost of empty bed is higher for hospitals with greater expected admissions.
Lynk (1995)	1991-1992	Daily	4	USA	The standard deviation of daily census	1. (Standard deviations of average daily census) 2. The average daily census	1. Hospital mergers decrease demand variability that reduce costs.
Carey (1998)	1987-1992	Daily	2568	USA	Total variable cost	1. (A regression model with the average daily census as the dependent variable and the lagged dependent variable, beds, case mix, and a dummy variable for large teaching hospitals as the independent variables was used to estimate demand uncertainty.) 2. Staffed beds 3. Number of discharges and outpatient visits 4. Wage rate 5. Average length of stay 6. Case mix index, and teaching, population density and ownership dummy variables	1. Demand uncertainty affects hospital cost of large hospitals (more than 100 beds), but this hypothesis does not hold for smaller hospitals (less than 100 beds).

Table 1. (cont.)

Author(s) (Year)	Period covered	Data	Number of Hospitals	Country	Dependent Variable	Independent variables (in parenthesis those used to measure uncertainty or variability)	Main Conclusions
Hughes & McGuire (2003)	1993-1995	Monthly	85	UK	Total variable cost	1. (Forecast errors in monthly emergency admissions. An autoregressive model based on past values is estimated.) 2. Elective and emergency admissions 3. Total number of beds allocated to emergency and elective sector 4. Inverse of occupancy rate 5. Accident and emergency outpatient visits, and day attendances 6. Wage index 7. Teaching dummy variable	1. The cost of an elective admission is approximately 45% of the cost of an emergency admission. Demand uncertainty imposes a direct cost equivalent to around 5% of the total cost of emergency admission.
Baker et al (2004)	1983-1995	Daily	500	USA	Total inpatient variable cost	1. (The mean and the variance of daily census) 2. Available beds 3. Mean wage 4. Case mix index 5. The % of discharges HMO 6. Teaching, ownership, year and hospital dummy variable	1. Considerable day-to-day variation in hospital utilization and significant differences in utilization patterns among hospitals. 2. The variance coefficients are statistically significant, with small effects on total inpatient operating cost. A 1% increase in variance lead to a 0.092% increase in cost.
Smet (2007)	1997	Annual	187	Belgium	Total variable cost	1. (Occupancy rate and a waiting time indicator) 2. Number of patients days, and admissions 3. Number of beds 4. Region, teaching and ownership dummy variables	1. The stand by capacity held hospitals in order to serve unexpected demand incur additional costs.
Lovell et al (2008)	1987-1994	Annual	67	Spain	Total variable cost	1. The insurance demand for inpatient services (and is approximated by the residual of the demand equation) 2. Number of discharges in general medicine, surgery, obstetrics, paediatrics, and intensive care 3. Number of outpatient visits and emergencies 4. Number of care graduates, technical personnel, other personnel and supplies 5. Number of beds 6. Number of students 7. Time dummy variables	1. Demand uncertainty affects hospital production decisions and increases costs. 2. Overcapitalization can be explained by hospitals providing insurance demand when faced with demand uncertainty. 3. Both stochastic demand and expense preference behaviour should be taken into account when analysing hospital costs and production.