



EUROPE IN THE SECOND CENTURY OF AUTO-MOBILITY

SERVICE NEEDS FORECASTING:

An Approach for the Automotive Industry Using Analogies with Medical ER Management Models

Miguel Ángel Pérez Salaverría

Jaguar – Land Rover of Spain & Portugal

ABSTRACT

During the past years, the industry has shifted position and moved towards “the luxury universe” whose customers are demanding, treating individuals as unique and valued customer for the business, offering vehicles produced with the state of the art technologies and implementing the highest finishing standards. Due to the competitive level in the market, car makers enable processes which equalizes customer services to E.R. management, being dealt with the maximum urgency that allows the comparison between both, car workshops and emergency rooms, where workshop bays or ramps will be equal to emergency boxes and skilled technicians are equivalent to the health care specialist, who will carry out tests and checks prior to afford any final operation, keeping the “patient” under control before it is back to normal utilization.

This paper establishes a valid model for the automotive industry to estimate customer service demand forecasting under variable demand conditions using analogies with patient demand models used for the medical ER.

KEYWORDS: Vehicle engineering; Body construction; Structural optimisation; Light weight body

INTRODUCTION

All car makers experience product-related requests from customers. These requests help align the product with customer's needs, shed light on how they use the product and generally improve the attractiveness of the product to its target market.

However, sometimes, these requests become disruptive. This can happen when providing a solution is done outside the product roadmap. Many product organizations do this because the "request" includes a threat that unless the feature is added, changed, fixed etc., the customer will not buy the product, will stop servicing it or sell it.

Although many car manufacturers increase their efforts to build strong service networks, a gap still remains between customer real demands and Authorized Service retailer's workload. Formerly, premium brands were being focused on reaching high service standards to match their customer expectations on service and maintenance, placing price on a secondary option, while volume car maker's acts basically on pricing and service times. According to premium brands point of view, a car entering the workshop is treated as a matter of urgency; it is like a patient entering an emergency room of a hospital and needs to be diagnosed with regards to the symptoms present in this moment, to offer the best solution for this particular case.

Car makers estimate their facilities capacity with simple methods with various restrictions, but typically arrivals occurs under uncertain conditions and variable demand that were not included in the calculations, producing work overloads, stocks backorders and customer dissatisfaction and complaints. Opposite to that, health services used different techniques to dimension hospital facilities according to the demographic distribution of the area to be serviced. Estimations are compared with a computer model simulation result and validated to create a model to be applied in future health services.

This article study the process to accommodate the existing models used for medical facilities to the service needs of a car manufacturer network.

The seminal references founded for the present document are based on the works of B. Liu, who, in 1976, established an input-output approach for regional hospital needs projection. Later, in 1982, R. D. Kamenetzky, L. J. Shuman and H. Wolfe studied a how to estimate needs and demands for prehospital care. Subsequently in 1989, D. M. Rhyne reviewed the applicability and a measure of forecasting systems in managing hospital services demand. In 1993, M. A. Badri and J. Hollingsworth published a simulation model for scheduling in the emergency room. Also, in 1996, Y. Gerchak, D. Gupta and M. studied a reservation planning under uncertain demand for emergency surgery.

Other authors afford the case from an operational research point of view, such as: A.

Bagust, M. Place and J. W. Posnett studied in 1999 a dynamic model to be used for accommodating emergency admissions applying stochastic simulation. Later in 2004, S. C. Brailsford, V. A. Lattimer, P. Tarnaras and J. C. S. Turnbull studied an emergency and on-demand health care model for large complex systems. Then in 2006, L. V. Green, S. Savin and B. Wang studied a model to manage patient service in a diagnostic medical facility.

Inclusion of more and more electronic devices interacting together in the car makes requires a better understanding of vehicle electrical architecture and has an impact on training needs, modern facilities with nice and clean workshops and, of course, a good management to ensure the required productivity and efficiency. Opposite to that, generally, low salaries still offered to the workshop technicians enabling a high personnel rotation. The former statement supposes any skilled technician will act as experienced doctor to diagnose a critical patient in an emergency box. The service receptionist will therefore assign jobs and times to the workshop according to pre-established priorities rivalling the medical ER. Customer requests can be a double-edged sword. On the one hand they can help point the way of where the market wants a company to go. On the other, requests can become disruptive and distracting. By understating the factors behind customer requests, the dynamics of the relationship and how these requests impact the process, companies can channel the "request energy" into positive channels leading to a better product that customers are excited about and willing to pay for.

1. THE GENERAL SERVICE MODEL

This paper explores experimental procedures used in ER management for comparing the capabilities of complex discrete event service systems. Instead of measuring system capability by analyzing or simulating the system with a constant rate of arriving work, system capability is measured as the maximum rate of work arrival for which the system has a steady state. Hence, we seek the arrival rate which causes the system to be at full capacity. This rate is arguably the best indication of the service system's capability.

The service systems considered all have the following features:

- centralized, controllable, processes which does not generate tasks at a rate A per unit time
- tasks are admitted upon generation and processed by the system
- completed task is ejected from the system
- the system has the capability to process as many tasks per unit time on average

Work-conserving queuing models do not allow:

- tasks to expire while in service
- tasks to create other tasks while in service
- tasks to be split or combined
- tasks which never finish service

Work-conserving queuing system models are common in both the practice and literature of applied probability. In a typical experiment, we generate input to the system at a constant rate; monitor the performance of the system either at fixed intervals or upon departure from the system, and employ well known methods of steady-state analysis to estimate the steady-state average of the performance measure.

A maxim of the analysis of service systems is that the system will have stationary long-run behaviour if and only if the number of arriving tasks are, on average, less than the number of tasks the system is capable of processing. If our overall system can work at a maximum of p tasks per unit time, we can input as many as p per unit time and the system will remain stationary. If A is our arrival rate for the system, we wish to manipulate A to expose p .

2. THE GEOMETRIC BROWNIAN MOTION (GBM) PROCESS

A Markov process is a particular type of stochastic process where only the present value of a variable is relevant for predicting the future. The past history of the variable and the way that the present has emerged from the past are irrelevant.

A Wiener process is a type of Markov stochastic process in which the mean change in the value of the variable is zero with the variance of change equal to one per unit time. The Wiener process was first applied in physics to describe the motion of a particle that is subject to a large number of small molecular shocks and was called Brownian motion (Hull, 2000). The mathematical description of the process was later developed by Wiener. To understand the equation, each of the components is considered separately.

Using this model, the service provider can optimize the parameters of the expansion policy according to numerical values of the model parameters observed in the industry in which the service provider operates. Relaxing the assumptions of the model suggests new directions in which this base model can be extended.

One of the most important assumptions made was that the demand process follows the GBM process, as it is also been used to represent future demand in capacity studies. Although this may be true for some industries, some bumpy demand processes may be

more closely represented by a probability distribution that incorporates sudden changes in demand values, for example a GBM process with jumps. Market saturation could be modelled by a process with nonhomogeneous drift. In the current model, only capacity expansions were considered.

The formulation of the service level constraint in this paper allows for expansion policies that either anticipates demand reaching the capacity position or react to demand having exceeded it. Its evaluation by using barrier option pricing tools is exact, and therefore the numerical results in this paper supersede those where timing and size decisions were made sequentially and evaluation of the service level constraint could err on the side of caution. We found that the optimal expansion parameters nearly always increased or decreased together. The delayed and infrequent expansion strategy that corresponds to large values of both parameters is optimal when greater shortages are permissible, lead times are short, economies of scale are significant, average demand growth is small, and/or demand volatility is low.

The opposite strategy, of small and frequent expansions that are initiated proactively, is optimal when the problem parameters reacts a more stringent service level, smaller economies of scale, and greater risk of shortage from the combination of long lead times and faster or more volatile demand growth.

Lastly, a deterministic lead time was considered for expansion. A probability distribution could be considered for the lead time to make it more realistic and the effect of stochastic lead time on the capacity expansion problem could be analyzed.

3. GENERATING DATA

There are two ways to generate data from a work-conserving system which will reveal the maximum processing rate in the system. They are:

- input tasks to the system at a rate known to be much higher than the system can handle
- fill the system, then input a new task every time that a task completes

In the former, the rate of outgoing jobs eventually converges to p . Instead of choosing a very high input rate and dealing with the problems of exploding buffer contents and a no recurrent system, we will simply close off the system and recalculate the tasks which finish. Hence, we take the second approach.

Hospitals reserve capacity to meet stochastic demand. This production response to demand uncertainty aids the specification of optimal capacity, which incorporates costly reserve capacity. Few estimates of hospital cost structures have taken account of this

aspect of hospital production and none have been applied, as undertaken below, to a heavily capacity constrained setting such as exists in the UK. Freidman and Pauly (1983), Gaynor and Anderson (1995), and Carey (1996), have all incorporated the impact of stochastic demand on hospital cost structures, while also recognising that hospitals control the output decisions, in response to such demand. In these studies the emphasis has been on estimating the cost of maintaining reserve capacity.

Running at full capacity also imposes a cost, however, in the form of production inflexibility, leading to patients being queued or turned away. There is therefore a trade-off between the cost of holding unused capacity in order to service stochastic demand, and operating at full capacity and turning patients away. This trade-off defines the optimal level of reserve capacity compatible with economically efficient utilisation. As Gaynor and Vogt (2000) note in any case, failure to take account of stochastic demand and the consequent production responses, leads to misspecification of hospital cost-output relations.

The determination of optimal capacity itself depends on an appropriate specification of output. One limitation of previous studies is that they have used an aggregate measure of hospital in-patient care, total admissions, to define in-patient output. The precise stochastic nature of demand will vary according to the type of case being serviced. A second limitation of previous studies is the reliance on annual or quarterly fluctuations in demand to model hospital responses to stochastic demand. It seems more realistic to model shorter-term fluctuations in demand to capture such responses. Use of aggregate measures, for both hospital output and demand fluctuation, will lead to a loss of information on the form and structure of the demand uncertainty.

In order to further explore the production responses to demand uncertainty, it is noted that hospitals distinguish between elective and emergency admissions.

Demand for emergency services is assumed randomly distributed with a known probability density function, while there is an assumed excess demand for elective treatments. This situation is readily observed in health care systems similar to that found in the UK where National Health Service (NHS) hospital capacity is limited through a budget constraint imposed on expenditure with total health care funding raised through general taxation. Patients are fully covered for all care within the NHS and treatment costs are not charged to the individual patient (see Cullis et al., 2000). NHS hospital referrals are designated to be emergency or elective cases with waiting lists used to explicitly ration the capacity allocated to elective treatments. Simultaneously each individual hospital retains some capacity to meet stochastic emergency demand, while also maintaining a waiting list for elective demand.

Within this system individual hospitals make a decision to allocate their fixed capacity based on their expectations of emergency demand turning into effective demand,

recognising that these expectations may not be realised ex post. In order to produce at any given level of output the hospital commits resources ex ante based on a forecast of emergency demand. Given seasonal fluctuations and the short-term nature of hospital planning such forecasts are based on within-year variations, even although budget allocations are tied to a yearly cycle.

4. FORECASTING WITH LIMITED DATA USING ARIMA MODELS

In statistics, an autoregressive integrated moving average (ARIMA) model is a generalisation of an autoregressive moving average or (ARMA) model.

These models are fitted to time series data either to better understand the data or to predict future points in the series. The model is generally referred to as an ARIMA(p,d,q) model where p, d, and q are integers greater than or equal to zero and refer to the order of the autoregressive, integrated, and moving average parts of the model respectively.

Given a time series of data X_t where t is an integer index and the X_t are real numbers, then an ARMA(p,q) model is given by

$$\left(1 - \sum_{i=1}^p \phi_i L^i\right) X_t = \left(1 + \sum_{i=1}^q \theta_i L^i\right) \varepsilon_t$$

where L is the lag operator, the ϕ_i are the parameters of the autoregressive part of the model, the θ_i are the parameters of the moving average part and the ε_t are error terms. The error terms are generally assumed to be independent, identically distributed variables sampled from a normal distribution with zero mean.

An ARIMA(p,d,q) process is obtained by integrating an ARMA(p,q) process. That is,

$$\left(1 - \sum_{i=1}^p \phi_i L^i\right) (1 - L)^d X_t = \left(1 + \sum_{i=1}^q \theta_i L^i\right) \varepsilon_t$$

where d is a positive integer that controls the level of differencing (or, if d = 0, this model is equivalent to an ARMA model). Conversely, applying term-by-term differencing d times to an ARMA(p,q) process gives an ARIMA(p,d,q) process. Note that it is only necessary to difference the AR side of the ARMA representation, because the MA component is always I(0).

It should be noted that not all choices of parameters produce well-behaved models. In particular, if the model is required to be stationary then conditions on these parameters

must be met.

Some well-known special cases arise naturally. For example, an ARIMA(0,1,0) model is given by:

$$X_t = X_{t-1} + \varepsilon_t$$

which is simply a random walk.

A number of variations on the ARIMA model are commonly used. For example, if multiple time series are used then the X_t can be thought of as vectors and a VARIMA model may be appropriate. Sometimes a seasonal effect is suspected in the model. For example, consider a model of daily road traffic volumes. Weekends clearly exhibit different behaviour from weekdays. In this case it is often considered better to use a SARIMA (seasonal ARIMA) model than to increase the order of the AR or MA parts of the model. If the time-series is suspected to exhibit long-range dependence then the d parameter may be replaced by certain non-integer values in a Fractional ARIMA (FARIMA also sometimes called ARFIMA) model.

The conceptual motivation for the empirical variable cost model estimated below follows that of Freidman and Pauly (1983) and Gaynor and Anderson (1995).

Following the latter a short-run cost model is estimated with attention focussed on how hospitals use existing fixed capacity to service unexpected demand. Variable hospital costs are specified as a function of the in-patient output, disaggregated into emergency and elective outputs, as well as other dimensions of output such as day case, accident and emergency and outpatient activity, and other characteristics of the hospital such as teaching status. An estimate of the level of fixed resource use, measuring the extent of excess capacity is incorporated through the inverse occupancy rate, which also controls for length of stay.

All these cost elements are conditioned on the hospital's estimate of unexpected demand as it relates to the probability of the hospital being full. This is controlled for through an estimate of unexpected emergency demand that enables empirical testing of whether or not uncertainty impacts hospital costs. It is hypothesised that if the coefficient on this variable is positive and significant, then demand uncertainty imposes a real cost on hospital production. It is this variable that differentiates the approach from the traditional cost function.

The small number of studies which have estimated such a variable have used different estimates of demand uncertainty as proxies for the standby capacity required to service unexpected demand. Gaynor and Anderson (1995) use the first two moments of the distribution of annual demand to proxy the relationship between unexpected demand and standby capacity. Of course the annual level of data smoothes within period fluctuations

while the focus on the described distribution emphasises the predictive content of the information used. Freidman and Pauly (1983) employ a measure of the ratio of expected to actual demand analysed on a quarterly basis.

Given that a ratio is estimated, the level of uncertainty is not captured. Indeed such measures of demand uncertainty reflect the expected fluctuations in demand, i.e. the ones the hospitals can predict. If hospitals do accurately predict the fluctuations then there is no reason to expect this to impact on costs. In the model estimated below demand uncertainty is based on a residual estimate of forecast monthly emergency demand.

The level of uncertainty faced by a hospital is thus defined as the difference between realised and forecast emergency demand. Such a measure captures the shocks imposed by stochastic excess demand while simultaneously avoiding possible multicollinearity between these demands.

Following Freidman and Pauly (1983), a simple autoregressive process was modelled assuming demand expectations are related to prior demand experience. Panel data were used to estimate the demand-forecast equation for emergency admissions, and the performance criteria rest on their ability to forecast, rather than explain behavioural relationships. An AR1 process was adopted specified as follows:

$$\text{Eq. (1)} \quad D_t = \alpha DV_t + \beta DVI + \rho(D_{t-1} - \alpha DV_{t-1})$$

where D_t is emergency demand in period t , DV_t represents a monthly dummy variable, DVI represents a hospital dummy variable, ρ represents the autocorrelation between periods, and α and β are constants. The variable representing unpredictable demand is specified as the differences between actual emergency demand and the forecast demand gained from Eq. (1).

The specified cost model relates total variable cost to the following explanatory variables:

$$\text{Eq. (2)} \quad TVC = f(ADM_{el}, ADM_{em}, RESem, VacSem, VacSel, INVOCC, OPATT, DAYATT, AEATT, WAGE, TD)$$

Where TVC is total variable costs, ADM_{el} and ADM_{em} are total in-patient elective and emergency admissions, respectively, representing the two major dimensions of output. $RESem$ is the variable that captures demand uncertainty, and represents the level of unexpected emergency arrivals, which is the difference between actual emergency demand and expected emergency demand (gained from the forecasting Eq. (1) above).

In the short-run, while the overall capacity is fixed, there is still a choice over the level of different outputs. Maintaining consistency with the theoretical specification, beds are separated into those allocated to the elective sector and those to the emergency sector. These are calculated on the basis that, under conditions of excess demand, occupancy

rate in the elective sector is assumed to be 100%, which is consistent with the existence of substantial hospital waiting lists for elective treatments as observed in the NHS. The level of staffed elective beds is therefore based on elective admissions and length of stay in that sector.

The remaining service availability is assumed to be used for urgent admissions, including an element of reserve capacity. This enables the staffed beds allocated to each sector and the level of reserve capacity in the emergency sector to be determined.

Eq. (2) includes the following independent output variables for ER – Hospital management where:

- **VacSem** is the total number of staffed beds allocated to the emergency
- **VacSel** is the total number of elective sectors
- **INVOCC** is the inverse of the occupancy rate, refers to the specific sector and relates to the fixed capacity servicing stochastic demand
- **OPATT** is the amount of outpatient visits
- **DAYATT** is the number of day attendances
- **AEATT** are the accident and emergency outpatient visits
- **WAGE** is the index for the NHS which represents a proxy for prices
- **TD** is a dummy variable for teaching hospitals

These variables can be easily accommodated to service applications:

- *VacSem* *Is the total number of work bays allocated to the Service Workshop*
- *VacSel* *Is the total number of elective sectors*
- *INVOCC* *Is the inverse of the occupancy rate, refers to the specific sector and relates to the fixed capacity servicing stochastic demand*
- *OPATT* *Total of vehicles entering the workshop*
- *DAYATT* *Number of day services*
- *AEATT* *Non scheduled vehicles entering the workshop*
- *WAGE* *Equal to 1; as every service will have the same internal cost*
- *TD* *variable for training technicians*

There is no theoretically accepted functional form for hospital cost functions consequently to determine an appropriate functional form a Box–Cox transformation applied to both dependent and explanatory variables was initially estimated. The results suggested a square root transformation on the dependent variable would fit the data with reasonable

Subsequent mapping of the two main output variables against this square root transformation of hospital variable costs are given in Fig. 1 and suggest a reasonable mapping, with some possible curvature indicated for the emergency cases variable.⁵ The final estimated specification is as follows:

$$\text{Eq. (3)} \quad \text{SNFCOST} = \alpha + \beta_1 \text{ADM}_{em} + \beta_2 \text{ADM}_{2em} + \beta_3 \text{ADM}_{el} + \beta_4 \text{RE}_{Sem} + \\ \beta_5 \text{Vac}_{Sem} + \beta_6 \text{Vac}_{Sel} + \beta_7 \text{INVOCC} + \beta_8 \text{OPATT} + \beta_9 \text{DAYATT} \\ + \beta_{10} \text{AEATT} + \beta_{11} \text{WAGE} + \beta_{12} \text{TD} + \varepsilon$$

where the variables are defined as in Eq. (1) with the addition of a quadratic term applied to the emergency admissions (ADM_{2em}) and the total variable costs transformed by 0.5 (SNFCOST).

The coefficient on unexpected demand is both positive and significant supporting the hypothesis that production responses to uncertain demand do impact on hospital costs: the higher the extent of the uncertainty, the higher these costs will be. The implication being that even after the hospital has chosen an optimal level of standby capacity shocks to the production process caused by unexpected emergency demand impose costs. The level of such costs are quantified below.

All other coefficients have expected a priori signs. The wage index is insignificant, which may not be surprising as the NHS is characterised by collective wage bargaining and thus may vary little between hospitals. The coefficients on emergency admissions, emergency admissions squared and occupancy rate do not attain conventional levels of statistical significance.

This could arise as a result of possible contamination from the inclusion of emergency bed levels, emergency occupancy rate and emergency admissions. It is clearly difficult to separate out the truly variable from the fixed cost elements associated with provision as Keeler and Ying (1996) note and if this is the case collinearity may be introduced. Indeed as Gaynor and Anderson (1995) discuss there may even be quasi-fixed elements within a hospital cost function reflecting rigidities in utilising existing capital structures.

As an alternative specification a transcendental logarithmic (translog) function was also estimated, but the results (which again can be obtained from the authors) were poor, with counterintuitive signs on the coefficients and insignificant t-statistics on almost all the independent variables.

The marginal costs of emergency and elective admissions are based on the variable cost element, which is taken from the estimated coefficient on the admission variables in the cost equation, and the quasi-fixed element taken from the beds variables. The quasi-fixed element is adjusted for length of stay in the emergency and elective sectors, respectively.

Aletras et al. (1997) review the literature with regards to economies of scale finding that

economies of scale are exploited at a relatively low level. As they point out, however, this conclusion is premised on the assumption that hospitals are operating on their efficiency frontier. Gaynor and Vogt (2000) note such conclusions are based on inconsistent estimates, and scale economies have to be related to demand uncertainty and production responses.

This is calculated in a manner similar to Gaynor and Anderson (1995) through the following formula:

$$\text{Eq. (4)} \quad S = (1 - (\delta \ln TVC / \delta \ln B_i)) / [\text{output cost elasticities}]$$

with the B_i representing elective and emergency beds and where the estimate is based on observed rather than optimal values.

5. CONCLUSIONS

This paper extends earlier work on production and cost responses to demand uncertainty. Therefore, the data used allows a more detailed specification of hospital output can be applied to the automotive service industry to forecast service needs.

The cost function also incorporates high informational content on demand uncertainty through the use of within-year fluctuations in demand and applies this to a sample of services.

The results are supportive of the earlier conclusions suggested by Gaynor and Anderson that hospitals do respond to demand uncertainty. Recently, Keeler and Ying (1996) interpreted the bias arising from demand uncertainty to be related to the regression fallacy problem. While this may be the case, adjustment at the individual hospital level is quite appropriate if interest is in the individual hospital responses to demand uncertainty.

In this application the various measures of marginal cost and scale economies seemed plausible and consistent with our conceptual arguments relating to production responses to demand uncertainty.

The results suggest that services do incur costs in holding reserve capacity to service stochastic demand. By separating out this stochastic demand from the excess elective demand it has been possible to quantify this cost. The rise in service needs is therefore of some concern, if only from a purely budgetary perspective. Moreover the marginal cost of treating a non scheduled admission includes an element of cost, directly attributable to the holding of reserve capacity to service this stochastic demand. The holding of reserve capacity in response to shocks experienced within the service system is also consistent with the finding of increasing returns to scale found.

These general findings are themselves consistent with the conclusions drawn by earlier

studies that an appropriate specification of service cost function will reflect the incorporation of production responses, in the form of holding reserve capacity, to unexpected demand.

If brand regulatory policies are to be guided by analysis of service costs such considerations are of paramount importance. The setting of labour fees and service levels depends on the accurate demand forecasting, cost of service and understanding of their influence.

In turn, fees should be set at a level that provides the appropriate incentives to workshops to hold reserve capacity where this is an efficient response to demand uncertainty.

Furthermore, apparent inefficiencies resulting from services operating within production possibility frontiers may be explained by the existence of uncertain demand, therefore, care should be taken in the interpretation of efficiency rankings without adequate adjustment for demand uncertainty and its impact on cost structures.

10. REFERENCES

1. Al, M. M. C., Claude Coray, Carol J Cuccaro, William K Green, & Et. (1972). The distribution system simulator. *Management Science* (pre-1986), 18(8), B425.
2. Alexander, M. (2003). 'The focus is on patients' clinical needs and distress'. *British medical journal*, 326(7402), 1332.
3. Alfredsson, P., & Verrijdt, J. (1999). Modelling emergency supply flexibility in a two-echelon inventory system. *Management Science*, 45(10), 1416.
4. Andrews, K. (2000). Factors that affect the demand for medical care services: A micro-macro econometric analysis. (Ph.D., Clemson University).
5. Arda, G, Barbaroso [gcaron] lu and Y. (2004). A two-stage stochastic programming framework for transportation planning in disaster response. *The Journal of the Operational Research Society*, 55(1), 43.
6. Badri, M. A., & Hollingsworth, J. (1993). A simulation model for scheduling in the emergency room. *International Journal of Operations & Production Management*, 13(3), 13.
7. Bagust, A., Place, M., & Posnett, J. W. (1999). Dynamics of bed use in accommodating emergency admissions: Stochastic simulation model. *British medical journal*, 319(7203), 155.
8. Ball, M. O., & Lin, F. L. (1993). A reliability model applied to emergency service vehicle location. *Operations research*, 41(1), 18.

9. Bazargan, M., Bazargan, S., & Baker, R. S. (1998). Emergency department utilization, hospital admissions, and physician visits among elderly African American persons. *The Gerontologist*, 38(1), 25.
10. Champion, R., Kinsman, L. D., Lee, G. A., & Masman, K. A. (2007). Forecasting emergency department presentations. *Australian Health Review*, 31(1), 83.
11. Dawson, D., Jacobs, R., Martin, S., & Smith, P. (2006). The impact of patient choice and waiting time on the demand for health care: Results from the London patient choice project. *Applied Economics*, 38(12), 5.
12. Duclos, L. K. (1993). Hospital inventory management for emergency demand. *International Journal of Purchasing and Materials Management*, 29(4), 30.
13. Gerchak, Y., Gupta, D., & Henig, M. (1996). Reservation planning for elective surgery under uncertain demand for emergency surgery. *Management Science*, 42(3), 321.
14. Green, L. V., Savin, S., & Wang, B. (2006). Managing patient service in a diagnostic medical facility. *Operations research*, 54(1), 11.
15. Jones, S. A., Joy, M. P., & Pearson, J. (2002). Forecasting demand of emergency care. *Health care management science*, 5(4), 297.
16. Kamenetzky, R. D., Shuman, L. J., & Wolfe, H. (1982). Estimating need and demand for prehospital care. *Operations research*, 30(6), 1148.
17. Khadem, R., & Schultski, A. (1983). Planning and forecasting using a corporate model. *Managerial Planning*, 31(4), 37.
18. Liu, B. (1976). Regional hospital needs projection: An input-output approach. *Socio-Economic Planning Sciences*, 10(1), 37-42.
19. Melnick, G. A., Nawathe, A. C., Bamezai, A., & Green, L. (2004). Emergency department capacity and access in California, 1990-2001: An economic analysis. *Health affairs*, , W136.
20. Minner, S., Silver, E. A., & Robb, D. J. (2003). An improved heuristic for deciding on emergency trans-shipments. *European Journal of Operational Research*, 148(2), 384.
21. New skills for capacity planners.(1994). *Capacity Management Review*, 22(5), 9.
22. Perkins, W. J. (1996). An analysis of the demand for emergency medical service in a medium sized city. (M.P.H., New York Medical College).
23. Powell, S. (1999). Using linear programming to simulate service engineers. *The Journal of the Operational Research Society*, 50(12), 1252-1255.

24. Rhyne, D. M. (1989). Forecasting systems in managing hospital services demand: A review of applicability and a measure of utility. *Socio-Economic Planning Sciences*, 23(3), 115-123.
25. Turnbull, S C Brailsford and V A Lattimer and P Tarnaras and J C. (2004). Emergency and on-demand health care: Modelling a large complex system. *The Journal of the Operational Research Society*, 55(1), 34.
26. Urbanos-Garrido, R. M. (2001). Explaining inequality in the use of public health care services: Evidence from Spain. *Health care management science*, 4(2), 143.
27. Utley, M., Gallivan, S., Treasure, T., & Valencia, O. (2003). Analytical methods for calculating the capacity required to operate an effective booked admissions policy for elective inpatient services. *Health care management science*, 6(2), 97.
28. Wasserman, G. S. (1992). An application of dynamic linear models for predicting warranty claims. *Computers & Industrial Engineering*, 22(1), 37-47.
29. Weerahandi, S., Hisiger, R. S., & Chien, V. (1994). A framework for forecasting demand for new services and their cross effects on existing services. *Information Economics and Policy*, 6(2), 143-162.
30. White, S. L. D. (1995). Algorithm-directed deferred paediatric care and reported parent/caretaker response: An exploratory study. (Ph.D., The George Washington University)