THE "DEVELOPMENT-DRIVES-DEMAND" EQUATION: AN ECONOMETRIC APPROACH FOR QUANTIFYING THE EFFECTS OF LATENT DEMAND WITHIN LODGING MARKETS

D. Clayton Peister
Senior Financial Analyst
Dclaytonp@yahoo.com

ABSTRACT

Estimating lodging demand within a given market has traditionally been a highly subjective activity. Despite the importance of an accurate forecast, there has been little research into developing a substantive methodology. This paper forwards an econometric process for estimating three elements of latent demand, namely the room-night demand related to development, the relevant timeframe, and a determination of whether development, in fact, causes demand. The conclusions gleaned from this process fosters informed decisions related to the effect of development upon incumbent properties' operating results, market-wide Revenue per Available Room, and prevailing yield management assumptions.

Key Words: latent demand, lodging market research, feasibility studies, econometric analysis

INTRODUCTION

A principal component of the hotel feasibility process is the assumption that the creation of hotel rooms drives demand for those very accommodations (Rushmore, 1997). This assumption is termed in the hotel industry as latent demand. Latent demand is an *'if you build it, they will come'* assumption; ergo- supply causes demand. When developers speak of latent demand they are referring to the demand for hotel rooms that exists, but is not currently being satisfied (Rushmore, 1997). In that sense, latent demand is akin to excess demand. Latent demand is an umbrella term that encompasses demand that is not being accommodated (unaccommodated demand) and demand that is enticed to the hotel/market (induced demand). Unaccommodated demand is the pent-up demand that exists, but most seek alternative accommodations or dates because the existing rooms are unavailable or unattractive (Rushmore, 1990). Induced demand describes the room nights lured to the market by a particular hotel's or brand's entry.

Traditionally, consultants estimate latent demand by reviewing how many fill nights or denied reservations a market has experienced. Consultants also evaluate latent demand by evaluating the cyclic behavior of a market and the potential draw from new demand drivers (Wheaton & Rossoff, 1997). Despite the importance of latent demand in the feasibility equation there is an underdeveloped process for studying the relevant parameters related to the effect (Sanders, 1999a) Consequently, quantifying latent demand in terms of room nights sold remains less a science than an art.

THE SCOPE AND OUTLINE OF THE CURRENT PAPER

The purpose of this paper is to propose a simple and flexible framework for studying the effect of latent demand within lodging markets. The following is divided into two broad segments. The first segment describes the composition and applications of the proposed tool, while the latter segment describes the dataset and illustrate an application of the proposed statistical process (for the sake of brevity, the statistical process which is the focus of the present paper is referred to as simply the 'process' or 'tool' henceforth). Whenever possible, both segments follow a logical sequence that parallels the flow of the analysis process. The composition of this process entails three sub-components. The first one examines the effect created from supply-side additions in terms of magnitude. The second sub-component analyzes the time-frame related to this effect; and the last one estimates if causation exists between demand and occupancy. The examples described here cover only the effect supply has on demand.

LATENT DEMAND ESTIMATION AT PRESENT AND THE NEED FOR A CONSISTENT APPROACH

When developing an estimate of room demand consultants often use an approach known as the *build-up approach based upon existing demand* (Rushmore, 1990). This approach first establishes a market baseline, i.e. the sum of accommodated rooms, for the competitive-set hotels over a specified period. The consultant then adds a percentage to the baseline figure to recognize the influence of latent demand. Once the consultant prepares this adjusted demand figure, she or he then creates a fair-market penetration estimate of this baseline for the proposed property over a set time-period. This penetration estimate is the basis for the occupancy forecast portion of the proforma projection. Part simple math and part intuition, this approach is highly subjective. The build-up approach is not only subjective but ignores basic economic theory regarding supply and demand (as described in a section to follow).

There is a dire need for a consistent and universally accepted approach to judge the effect of a market's latent demand. Without a consensus on how to estimate latent demand, there is little consistency in the lodging feasibility industry. A high profile example of two major consultancies arriving at different conclusions regarding a market occurred in August of 2000. One firm stated the subject city was overbuilt, while the other found that sufficient demand existed to warrant room expansion. These two firms squared off publicly regarding the validity of their opinions (Poole, 2000). Such situations result from a consultant tailoring findings to the client's needs and ignoring broader market indicators. The willingness of consultants to 'pad' their findings brings an air of illegitimacy to the field.

A series of white papers produced by the University of Texas at Austin develops a critical portrait of the hospitality feasibility industry. The papers cite apparent inconsistencies in convention center feasibility studies where 'invariably' consultants recommend expansion and construction despite obvious indicators that development is not warranted (Sanders, 1999b). The study cites an apparent eagerness to recommend development and a hesitation to advocate caution by many prominent industry advisors. This issue is becoming increasingly important with the scandals surrounding some of the major accounting firms and the separation of their advisory and assurance practices (Sarbanes & Oxley, 2002).

CONCEPTUAL AND EMPIRICAL DEVELOPMENT OF LATENT DEMAND

Little scholarly research has been devoted to studying the metrics of hotel latent demand. Initial publications outlining the concept and implications of latent demand were published in the early 1980s to mid 1990s. In these publications the authors postulate that latent demand is composed of induced demand and unaccommodated demand (Rushmore, 1978). Further, early practitioners estimated that latent demand should compose 0% to 30% of all accommodated demand for a new unit (Rushmore, 1990). Several recent publications have made significant headway into describing the room demand environment in a fuller and more descriptive fashion. A recent article applied a method long utilized in the office and residential leasing industry to estimate excess demand: a metric known as Natural Occupancy Rates (NOR) (DeRoos, 1999). The article estimated NOR by price-point for several major metropolitan areas. The difference between a market's NOR and average occupancy rate is termed as the *occupancy gap*. Of the cities reviewed, long-run occupancy gaps ranged from 6% to almost 7%.

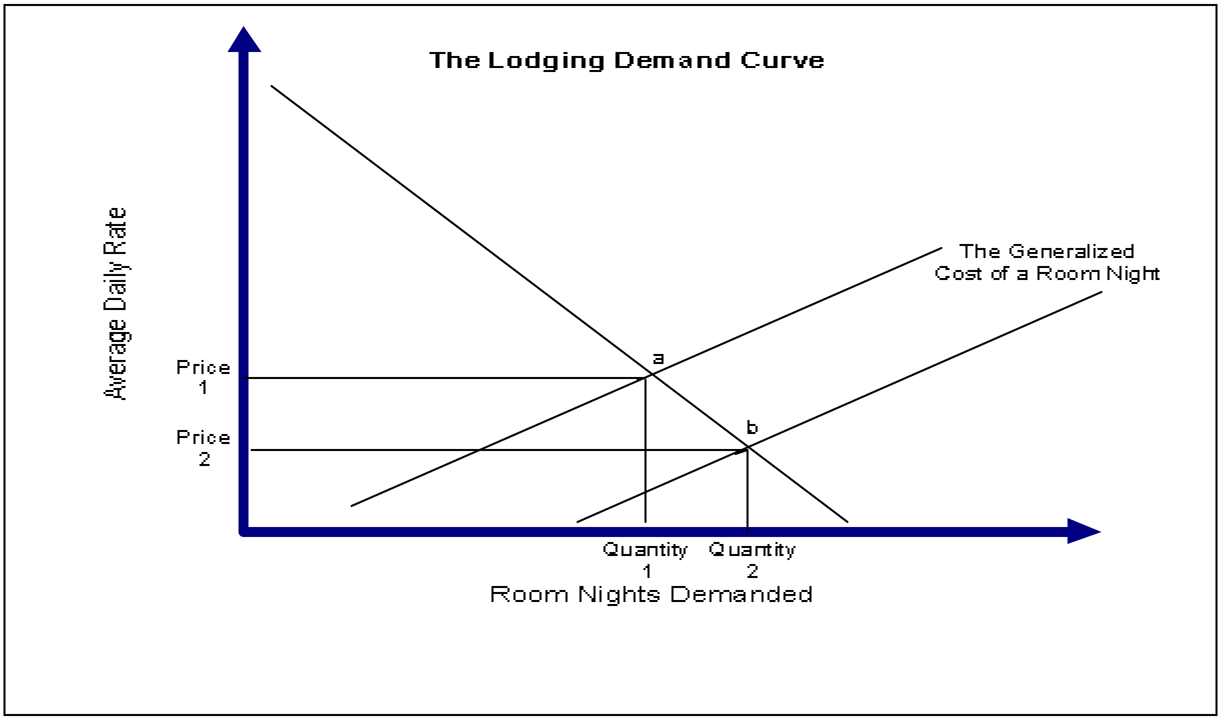
An oft-cited paper found that over the past 25 years, domestic hotel demand has closely followed the movement of the greater US economy (Wheaton & Rossoff, 1997). By contrast, hotel investment has a secular or long-term pattern, which is independent of the macroeconomic environment. In addition, this paper showed that room development has an immediate effect of decreasing market-wide occupancies. The effect of room development on Average Daily Rates (ADR) is negative and trails the impact on occupancy by a given period.

Outside of the industry, the concepts of latent and induced demand have received more attention. Within the field of urban planning econometric type analysis has been used to determine if road development eases congestion or induces more road usage. In investigating this hypothesis, urban planners have utilized multivariate regression models and structural path modeling to draw defensible and substantive conclusions (Franses, 2003; Fulton et al., 2000).

PROCESS DESCRIPTION AND RESEARCH QUESTIONS

The economic basis of the lodging supply and demand relationship is simple: a change in the (generalized) cost of staying in a hotel room will change the demand for those accommodations. Of course, the most dominant component of the generalized cost for accommodations is room tariff. Other costs are notable as well, however. A reduction in the travel time between attractions and the property reduces the cost of travel, albeit, in a less acute fashion. A purely classical perspective of the hotel room supply and demand relationship is improper, as there is not a perfect relationship between room price and room demand. Indeed the service nature of the industry requires analysts to consider the qualitative draw as well. The quality of the amenities, the level of the services and brand associations are all points guests consider when making decisions (Hinkin & Tracy, 2003). The process described here draws conclusions in the quantitative realm even though qualitative aspects receive some consideration as well. The process incorporates qualitative elements (such as service) in an implicit fashion only. For instance, economic theory predicts that Average Daily Rate should decrease with increases in supply (see Figure 1 for a graphical description of this relationship) (Corgel, 2004). If this assumption does not hold and ADR increases with increases in supply then there should be other mechanisms in play. In this example, newer hotels command higher rates and thus supply additions may move market-wide ADR higher.

Figure 1: The Lodging Market Demand Curve



Econometric techniques are the most appropriate tool-set for studying the latent demand environment associated with the lodging industry. Econometrics is a broad term, which grew out of the quantitative study of the economy. In a general sense, econometrics is the analysis of actual data to yield tangible conclusions regarding an event or trend (Franses, 2003). More specifically, econometrics is simply an applied field of statistics.

The process described here establishes an approach for understanding the market dynamics that drive market-wide occupancy (in the form of demand over supply). The applications for this process are varied and only a few are covered here. Analyzing a market with this process yields insights into the effect that proposed projects will have on the market indices. The process yields conclusions regarding the effect on incumbent properties from development and/or the financial feasibility of proposed hotels. This process is also useful in calibrating revenue management assumptions and proforma forecasts or budgets in the face of an unknown and changing environment. For instance, new development will likely affect the booking rate for existing properties. Moreover, development is not in the best interest of the market if market-wide revenue per available room (RevPAR) will drop when new units enter the market. Associations and city offices are keenly interested in this relationship, as are investors and asset managers.

As introduced above, the process is segmented into three components. The first step is to estimate the magnitude of the effect using a statistical model, which is similar to an ordinary least squares (OLS) model. The second component is a distributed lag analysis – used for estimating timeframe – and the final segment of this process is a system for identifying causation (albeit in a specific sense of the word). The descriptions provided below are an overview of these three components.

ESTIMATING THE EFFECT OF LATENT DEMAND

In order to approximate the size of the effect created when new rooms enter the market, the process uses a linear model. The model is similar to standard OLS; however, several key components are specified. As with time-series data, there is a need to differentiate among years. A general linear model provides a system for such differentiation as the general linear model does not have the rigid parametric constraints of OLS. Utilizing this model allows us to specify a categorical variable, which enables estimates by category. The data are also transformed into a natural log base. This transformation not only reduces unequal error variances (known as hetroskedasity) among the variables, but also allows one to interpret the coefficients as elasticities. All the estimates specify the inclusion of an intercept.

The model regresses demand against a series of explanatory variables. For the example below the continuous predictors are personal income, total employment, Market ADR (MADR) and supply. The included categorical variable is years, decomposed into quarters. Please note that the explanatory variables included here are just some of the many possible factors of room demand that may be appropriate in this model. As this paper does not focus on how to estimate demand elasticities, (except in an implicit fashion required for the latter portions of the process) the paper does not elaborate beyond these basic variables. Several studies forward a substantive approach for elasticity estimation within the field of lodging and the log-log approach; for interested readers, the following resources are recommended (Canina & Carvell, 2003; Franses, 2003; Wheaton & Rossoff, 1997). This process specifies an additive difference of the log model (i.e. year t minus t-1) (Noland, 1999). In essence, this model runs a regression on the growth/contraction over the examined years. The advantage of the differenced model is a reduction in the effect of correlation amongst the variables, a condition termed as multi-collinearity. The differenced model often picks up some of the randomness inherent in the base model. Therefore, results may vary slightly from the base model. The difference/growth model also is the basis of the causation testing.

After transforming and analyzing the data, the result is an overview of the effective magnitude that the explanatory variables have on the demand. Since the data are transformed into a natural log base, the coefficients are interpreted as point elasticities. Further, issues related to correlation among the regression errors (known as auto-correlation) can be investigated with the Durbin-Watson statistic. If auto-correlation exists, further transformation may be required. By the end of the first component, one should have a clear portrait of the relevant model elasticity and validity. Note that the base and first-difference model does not include a provision for a lag structure. Therefore, the coefficients and correlations should be viewed as the effect over the long-term period.

APPROXIMATING THE TIMEFRAME RELATED TO LATENT DEMAND

A conventional premise in hotel feasibility analysis is that a new hotel takes two to four years to achieve full fair market share penetration (Rushmore, 1978). During this ramp-up period, the hotel is developing a sales clientele, establishing repeat patronage and fine-tuning market position. The property is therefore not at full potential until the end of this period. By the end of the ramp-up period, owners expect the hotel to be at or above (assuming financial feasibility) fair-market share. By this wisdom, during the ramp-up period total demand is not at maximum levels. Thus, the <u>full</u> effect of supply-side additions on demand will trail supply changes (although a truncated effect should be immediate). ADR by contrast should respond much more quickly and negatively to supply additions than does demand. Recognizing that new hotels often buy market share by offering discounted rates in the first year of operation, the expectation is that the effect from new hotel rooms on ADR would occur simultaneously and decrease the market-wide ADR (Corgel, 2004).

Therefore, an understanding of market dynamics is not complete without a portrait of the lag duration resulting from supply changes. In order to estimate this timeline analysts can utilize a statistical technique know as distributed lag analysis. Distributed lag analysis is similar in composition to OLS. However, this analysis is an iterative process, where the dependent variable is regressed against a predictive variable repeatedly, with each iteration having a different lag duration.

Essentially, this model is looking for the strongest correlation between the movement in dependent variables and the movement in independent variables. The period with the strongest statistical significance is determined to be the best estimate of lag duration. If demand is a dependent (endogenous) variable and supply is an independent (exogenous) variable, then this relationship is best described as a linear relationship. A proper expression of this relationship can be seen below:

$$Y_t = \Sigma\ b_i * x_{t-i}$$

In the above equation, the value of the dependent variable at time t is a linear function of x measured at times t, t-1, t-2, etc (Statsoft Corporation, 2002). As with any OLS model statistical significance is judged by the highest *t*-values. For the example included in the next section, the analysis was completed with the aid of a statistical package (Statistica 6.0). However, a distributed lag analysis can easily be completed manually or with the aid of macro (see VBA for Excel- looping procedure and the *linest* function).

In an analysis with no constraints on the lag, the effect of the lag is sudden and undistributed. As such, an unconstrained model is awkward and does not fit many datasets. There is, therefore, a need for a method of distributing the effect of the lags appropriately over a period. Two approaches to structuring the distribution of the lag weights are common in econometrics. The most established approach is the Koyck procedure (Wallace & Silver, 1998). This procedure distributes the lagged weights in an exponentially declining fashion. As a result, the lagged weights peak and then diminish quickly. Therefore, this approach is not well suited to quarterly or monthly time-series with many observations (Wallace & Silver, 1998). The Almon approach is an alternative to the Koyck procedure and is attractive because the analyst controls the behavior of the lags (Almon, 1965). This approach truncates the effect of the lag at a definitive point. The concept is to control the shape of the weight's distribution by estimating the weight's values with a polynomial equation. If the analyst expects the effect to rise smoothly, peak and then decline smoothly a lower order polynomial is appropriate.

After running the distributed lag analysis, the analyst should have an understanding of the lag structure intrinsic to the supply/demand relationship. The expectation is that supply will have some effect on demand in the first period; however, the bulk of the effect should trail the supply changes by some period. The distributed lag analysis described here analyzes this relationship and draws substantive conclusions regarding the prevailing structure. The conclusions gathered here will assist managers and owners in making decisions regarding when supply changes will affect market or property returns.

DETERMINING IF DEVELOPMENT CAUSES DEMAND

Is supply really driving demand, or are supply changes just a response to demand pressures? The assumption that supply is causing demand is a central premise in most feasibility study's occupancy forecast. Establishing the clear relationship required for this estimate is a subjective and highly speculative task. While the concept of causality is straightforward, an investigation into causal relationships can yield only ad hoc conclusions. If room development were a cause of room demand then a logical expectation is that changes in supply will precede changes in demand. This time precedent relationship is obvious and readily tested for using econometric techniques. Remembering that simple correlation is not a true indicator of causation, there is a need for a more definitive approach for determining cause and effect.

In order to establish a causal relationship between supply and demand, analysts often utilize two similar statistical approaches: *Granger causality* and *Two-Stage Least Squared Regression (2SLS).* The purpose of these 'tests' is to establish if supply is an exogenous variable in the supply/demand relationship – a situation know as supply exogeneity. In addition, 2SLS helps correct for false conclusions resulting from simultaneity bias. Simultaneity bias is an error that can occur because a set of variables move in unison over a time period, but are in fact not related.

To draw cause and effect conclusions, econometricians often utilize the *Granger Test for Causality.* In economics if, in the presence of a lagged dependent variable, a lagged independent variable significantly helps explain the dependent variable at a given year (t_x) then the independent variable *Granger causes* the dependent variable. Causality in the philosophical sense is not analogous to Granger causality. The distinction is important, as Granger causation is a test for temporal sequence and not for the intrinsic structure of the dataset.

In time-series analysis, two obvious influences can invalidate an analysis: cointegration and non-stationarity. A variable should be stationary, i.e. have a consistent mean and variance, to be valid. Seasonal influences often obscure non-stationary variables, which make trends difficult to interpret. To remove seasonal influences and induce stationarity analysts can treat the variable with first-order differencing (first-order differencing is subtracting the second period from a previous period) (Noland, 1999).

Cointegration is the second influence that can invalidate a regression model and by extension a test for causality. Cointegration between variables exists when the regression residuals correlate over time. The contention among statisticians is that cointegration invalidates a causality test because 'common relationships' are present between the variables (Vahid & Engle, 1993). Therefore, conclusions regarding causal relationships are potentially false, as the model confuses common movements for cause and effect.

A good example of how cointegration can cause false correlation is occupancy levels between two hotels within the same chain, but in separate cities. If an analyst were to test that demand at hotel X in city X causes changes in occupancy for hotel Y in city Y causal relationship might well exist. However, spillover between the properties could not exist because of geographic distance. Franchise programs, such as discounting and marketing could cause the occupancy at each of the properties to move in similar fashions. When datasets are non-stationary and cointegrated, conclusions gleaned from regression models using these data may be spurious. As with non-stationary variables, cointegration is minimized by differencing the observations.

Testing for Granger causality is straightforward. The test involves two steps. First, a constrained model is estimated. Then an unconstrained model is estimated and compared with the unconstrained model. The unconstrained model regresses lagged demand and lagged supply as independent variables against an un-lagged demand variable (Y). The constrained model regresses an un-lagged demand variable (Y) against a lagged demand variable. The sum-of-squares-errors (SSE) of the regression runs are the basis of the test statistic (w). The test statistic is a comparison to determine if lagged supply increases the accuracy of the model. The test statistic equation is provided below:

$$w = \frac{(SSEr - SSEu)/s}{SSEu/(T-K)}$$

Where the denotation for the restricted sum-of-squares error is SSE_r, while, the unrestricted sum-of-squares error is denoted as SSE_u. The degrees of freedom in the denominator are the number of observations (T) less the number of regressors (K). The degrees of freedom in the numerator are the number of coefficients restricted to zero (s). The resulting test statistic (w) can then be compared against a criteria F-value (F_{crit}) for the chosen level of significance. The F_{crit} value is easily determined using Excel's FINV function. Using the FINV function, with s as the numerator and $T - K$ as the denominator at the chosen level of significance (normally a = .05 for a 95% confidence interval) the appropriate F_{crit} value is calculated. If the $w > F_{crit}$ then supply Granger causes demand. A *p*-value can also be specified using Excel's FDIST function. Again *s* and $T - K$ are the degrees of freedom, but the w is specified as X. If $p <$ level of significance then supply Granger causes demand. Again, causality is not truly a test of cause and effect. The causality described by this tool is a test of time precedence, which is (of course) essential for *real* causality to exist.

The most consistent means of testing if supply truly influences demand is to estimate a two-stage-least-squared (2SLS) model. A two-stage-least-squared model regresses demand as the dependent variable against an estimated independent variable. The estimated independent variable is a composite of supply and an instrumental variable. The instrumental variable correlates with supply (x) but not demand (y). A common approach in econometrics is to use lagged x as an instrumental variable in the equation. The model cited in this paper is staged in the following manner:

1. Stage one: estimate x-hat by regressing x on z
2. Stage two: estimate y-hat by regressing x-hat on y

Exogeneity can then be assessed by comparing the coefficients of the first-differences model to the 2SLS coefficients. If the 2SLS coefficients are markedly different from the first-difference model, then supply is not an exogenous predictor of demand and conclusions regarding causality are impossible.

By the conclusion of the third component, a clear portrait of Granger causality and exogeneity should be obvious. This analysis demonstrates if supply Granger causes demand, in the sense that changes in the number of rooms must precede changes in demand. In addition, it shows whether supply is a truly exogenous variable in a supply/demand regression. By the end of this component, analysts will be able to answer the question of whether latent demand exists in the analyzed market. If latent demand exists then the results of the first two components hold and room creation/deletion should have the expected results. If Granger causality is absent and supply is an endogenous variable, then determination of latent demand's presence is not possible.

AN ILLUSTRATION OF THE PROCESS UTILIZING SIMULATED DATA

To evaluate how well the process estimates a market's trends a simulated dataset that mirrors the supply and demand relationship is required. Using a simulated dataset (as opposed to a real dataset) is intuitively appealing for several reasons. A simulated dataset lets one control the intrinsic structure of the data. Therefore, the coefficients of the model can be checked against the structure known to be present. Simulated data also lets one check for consistency and robustness. Creating a dataset with simulated random variables allows the model to be run multiple times and checked for consistency. As such, the example included in the following section utilizes a simulated dataset.

Proper simulation techniques require two elements. First, simulations require a method to set the value and probability of all possible outcomes. Second, a simulated experiment is an iterative process requiring the experiment to be run multiple times. Thus, simulations must repetitiously model the variable and save the outcome for future comparisons in a stable and consistent process. The essence of any simulation is a model that reflects the possible variation between the set parameters but centers around the mean (Mooney, 1997). Therefore, the variable's mode must be near the mean and the least occurring observations have values that differ from the mean greatly (relatively speaking) (Mooney, 1997). This description is identical to the traits of a probability distribution.

Modeling a simulated variable as a probability distribution is attractive as well. Thinking of ADR in the high/low terms required for a uniform distribution is logical (Sturman, 2003). Likewise, using a triangular distribution to model marginal-room-cost estimates is reasonable as well (Ganchev, 2000). The reservation-booking rate is similar in composition to a Poisson distribution (Talluri & Van Ryzin, 2000). Therefore, the nature of the variable must drive the modeling of the simulated variable itself. This approach also aids in the interpretation of the simulation results.

Excel simulates variables using two approaches. The first approach is to use the standard functions available within Excel. The random number generator lets users model a variable as a discrete, Poisson, normal, uniform, binomial, Bernoulli and patterned distribution. Users must specify the number of simulations and random numbers. This tool is powerful, but is not 'live.' Most analysts prefer to simulate a dataset using the live functions. Excel includes several functions for modeling elements of a probability distribution. The most popular of these functions is FINV function used for modeling a normal distribution. FINV becomes stochastic if the rand() notation is specified in the probability field. In addition to the functions included standard in Excel, analysts can create a simulation using an aftermarket add-in. These packages are more user friendly, but carry a substantial upfront cost.

THE STRUCTURE OF THE ANALYZED SIMULATED DATA

This section details the specific structure of the random variables, which compose the example analysis in the following section. The simulated dataset has six variables: date, demand, supply, ADR, personal income and total area employment. Each of the variables has 64 observations, which compose the years 1988 to 2003 in quarters.

Dummy Variable: The behavior of the simulated variables (except date) corresponds to a random dummy variable. This dummy variable is coded one if random variable A is greater than random variable B, otherwise, the dummy variable is zero. In this instance, the dummy variable is used to control the corresponding variable's values.

Supply: The simulation splits the supply variable into two buckets: high and low. Observations that correspond to a dummy variable zero belong to the high bucket. Conversely, observations that correspond to a dummy variable one belong to the low bucket. Observations take on a random value between 63,000 and 65,000 in the low bucket. In the high bucket, observations take on a random variable between 64,000 and 70,000 rooms. The Rand() function used for this variable produces a uniform random variable; the resulting dataset is thus uniformly distributed. The goal of this variable is to segment supply into high and low years.

Demand: As the focus of this paper is the investigation of latent demand, the example dataset has a latent demand structure. To model latent demand the demand variable is specified so that changes in supply will induce changes in demand. In addition, changes in demand follow changes in supply with an eight-quarter lag. To create the demand variable, the first eight observations are random between values (Excel function randbetween(,)), the logic being that this period is a reaction to previous supply changes not specified in the model. After the eighth period the observation is modeled as a logical statement. If period one corresponds to a supply observation coded as one, then the observation become a random uniform variable of 64-66%, (with a mean of 65%) of period one supply. If the corresponding supply variable is coded zero, then the observation becomes a uniform random variable of 65-74% (with a mean of 69.5%) of period one supply. For the simulation, the dataset is in essence modeling occupancy rates (supply over demand). It is important when analyzing a factual dataset to model demand in room nights, as opposed to occupancy. Several reasons exist for analyzing room sales versus occupancy. First, the causal models respond better to larger increment changes. Second, occupancy is relative to supply, whereas demand is a stand-alone figure. Since analysts are interested in demand as the number of room nights sold, occupancy is not appropriate.

ADR: According to economic theory, the cost of a hotel room night is the largest detriment to demand. As such, the expectation is that an increase in ADR will correspond to a decrease in demand. Moreover, this relationship is expected to react in a near simultaneous fashion relative to demand (see previous section describing the cyclical nature of the hotel market). As with demand, ADR is modeled using a uniform random distribution, within a logical statement format. If demand corresponds to a supply value of zero, then ADR becomes a uniform random variable between $61 and $66. If the corresponding supply variable is coded one then ADR becomes a uniform random variable between $65 and $74. Thus, ADR follows demand explicitly and supply implicitly.

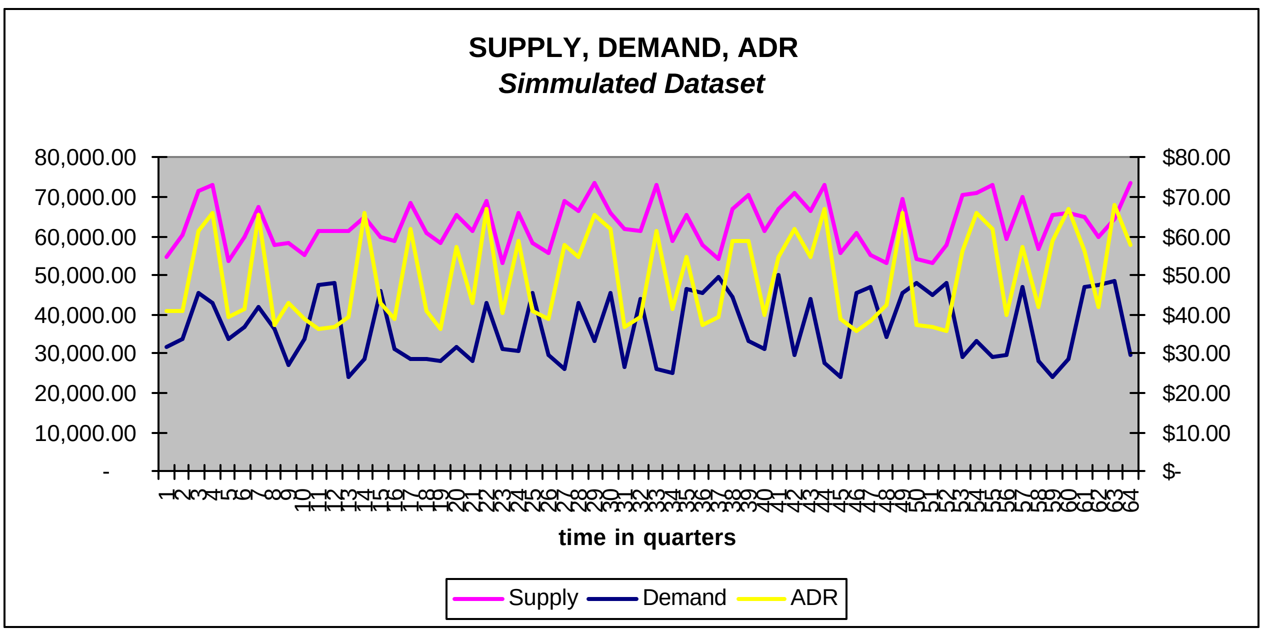
Personal Income and Total Employment: Also included in the base and first-difference models are the variables personal income (PI) and total area employment. The structures of both of these variables are the same. As with supply, PI and total employment are uniformly distributed simulations, with a high and low bucket. Bucket grouping in these variables differs from demand in that if the corresponding demand observation is greater than the average demand observations, then a high bucket grouping (and vice versa) is assigned. There is no lag structure integrated into these variables.

Summary of the Simulation Model: The structure outlined above is a major simplification of the market environment. In this model, demand responds only once and drastically to supply changes. As introduced in the section describing distributed lag analysis the effect of changes in supply is expected to rise smoothly, peak, and diminish. This model does not reflect this distributed relationship. The validity of the simulation rests in the ability to establish that the process works consistently and appropriately. Therefore, the exaggerated relationships represented here serve merely as a proxy to demonstrate the ability of the process. A graphic illustrating the movement of the simulated data over time is included in Figure 2 below.

Specific hypotheses can now be stated as:

1. Hypothesis One - Demand should increase with increases in supply, PI and total employment and decrease with ADR.
2. Hypothesis Two - Demand should lag supply in an eight-quarter pattern.
3. Hypothesis Three - Supply should Granger cause demand and is an exogenous variable.

Figure 2. Simulated Dataset



ANALYSIS PROCESS APPLIED TO THE SIMULATED DATA

This section summarizes the analyzed simulated dataset. As with the previous sections, this segment will follow the sequence of the process specifications. The base and first-difference model is overviewed first, after which the distributed lag model is discussed. Finally, there is a description of the conclusions regarding the causal nature of the supply and demand relationship.

The Magnitude of Latent Demand

Analyzing the dataset with the general linear model produces the expected elasticity directions. Holding all else equal a ten percent increase in supply relates to a four percent increase in demand. Likewise, a ten percent increase in ADR decreases demand by three percentage points. Ceteris paribus, a ten percent increase in personal income relates to a one percent increase in demand and a ten percent increase in total employment relates to a two percent increase in demand.

Differencing the dataset and analyzing the variables with the same approach as the base model produced consistent coefficients. The relevant coefficients and model statistics from the first-difference regression are included in Table 1 below. All the directions are consistent with the initial hypotheses and the base model. The size of the effect for the personal income and total employment variables are slightly smaller than in the base model. This pattern is consistent with the transformation process, as differencing tends to exaggerate the variable's stochastic elements (Fulton et al, 2000). In order to check for the influence of auto-correlation, the Durbin-Watson statistic was utilized. Auto-correlation likely exists if the Durbin-Watson statistic is less than 2.0 (Wallace & Silver, 1998). The computed Durbin-Watson statistic for the first-differences model is 3.054. Therefore, auto-correlation is not an issue.

Table 1. The Coefficient Output from the First-Differences Model

	Level of Effect/ Year	Demand Parameter	Demand Std Error	Demand t	Demand p	(95%) confidence level	95% confidence level	Demand Beta	Demand St Error
Intercept		0.000152	0.041943	0.003634	0.997117	-0.084433	0.084738		
Supply		0.458942	0.652123	0.703765	0.485373	-0.85619	1.774073	0.215489	0.306194
ADR		-0.267144	0.274243	-0.974115	0.335448	-0.820207	0.285919	-0.29333	0.301122
Pi		0.125291	0.068816	1.820655	0.075623	-0.013491	0.264072	0.329238	0.180835
EE		0.223672	0.340663	0.65658	0.514949	-0.463339	0.910684	0.118956	0.181175
Date	1988	-0.044827	0.187846	-0.238638	0.81252	-0.423655	0.334001	-0.05043	0.211333
Date	1989	0.011067	0.162027	0.068307	0.945858	-0.31569	0.337825	0.013326	0.195092
Date	1990	-0.065822	0.161893	-0.40658	0.686333	-0.39231	0.260665	-0.07926	0.194931
Date	1991	0.008604	0.163142	0.052741	0.958182	-0.320403	0.337611	0.01036	0.196435
Date	1992	0.084603	0.164768	0.513466	0.610254	-0.247684	0.41689	0.101868	0.198393
Date	1993	0.022194	0.162773	0.136351	0.89218	-0.306069	0.350458	0.026724	0.195991
Date	1994	0.000991	0.161871	0.006122	0.995144	-0.325453	0.327435	0.001193	0.194904
Date	1995	-0.017375	0.162045	-0.107226	0.915109	-0.34417	0.309419	-0.02092	0.195114
Date	1996	-0.085044	0.162186	-0.52436	0.602721	-0.412124	0.242036	-0.1024	0.195284
Date	1997	0.106331	0.161737	0.657433	0.514406	-0.219843	0.432505	0.128031	0.194743
Date	1998	-0.097319	0.161249	-0.603531	0.549326	-0.422509	0.227871	-0.11718	0.194156
Date	1999	0.124988	0.162955	0.767014	0.447264	-0.203641	0.453618	0.150495	0.196209
Date	2000	-0.130586	0.161624	-0.807962	0.42356	-0.456532	0.19536	-0.15724	0.194607
Date	2001	0.105203	0.160713	0.654603	0.51621	-0.218905	0.429311	0.126672	0.19351
Date	2002	-0.015937	0.162646	-0.097987	0.922398	-0.343943	0.312069	-0.01919	0.195837

The first component illustrated that an increase in supply relates to an increase in demand. Demand does not move in the same proportion as supply, however. Supply outstrips demand by almost 25% in the simulated dataset. For owners and investors in this market supply changes are an important determinant of demand. In addition to supply, ADR is also a major influence on demand. The results indicate that as ADR rises demand falls. The thought being that ADR often dips as marginal properties are forced to fight for guests by discounting. To draw a meaningful conclusion from this analysis, RevPAR should be reviewed. Holding personal income, employment and ADR the same, a ten percent increase in supply requires a five percent increase in ADR to hold RevPAR at current levels. Raising ADR by 5% is unlikely in most markets, thus new rooms are not in the best interest of the incumbent properties. Table 2 provided below illustrates this concept in detail.

Table 2. The Effect on Revenue per Available Room

	Room Nights			
	Supply	Demand	ADR	RevPAR
Base	73,623	29,495	$ 58.00	$ 23.24
10% supply increase	80,985	30,823	$ 61.05	$ 23.24

Conclusions Regarding the Temporal Structure of Latent Demand in the Simulated Data

The goal of the distributed lag analysis is to develop a portrait of the effective timeframe resulting from supply additions. In general, the Almon approach provides a more consistent platform for the analysis of time-series data (review introduction above). However, the simulated data has an undistributed lag structure, which makes an unconstrained analysis preferable. As expected, the eighth quarter has the largest *t*-value, t= 6.94, which is considerably larger than the other observations. A summary output of this model has been included below in Table 3. Therefore, the lag duration that best models the effect supply has on demand is a two-year lag. As expected, ADR moves in tandem with supply in the same period as does personal income and employment (see Table 4).

Table 3. Distributed Lag Analysis – Demand Regressed Against Supply

Lag	Regression Coeff.	Standard Error	t(52)	P
0	0.206864	0.186476	1.109328	0.275556
1	0.356077	0.251743	1.414450	0.166887
2	-0.057719	0.272877	-0.211520	0.833823
3	-0.014812	0.272449	-0.054366	0.956981
4	-0.050748	0.289428	-0.175340	0.861917
5	0.280996	0.319703	0.878927	0.385993
6	0.178448	0.326623	0.546342	0.588621
7	-0.128742	0.315729	-0.407762	0.686164
8	1.825645	0.311388	5.86292*	0.000002
9	0.078184	0.320744	0.243758	0.808974
10	0.172660	0.323352	0.533969	0.597053
11	0.165757	0.316330	0.523999	0.603890
12	0.005690	0.305555	0.018623	0.985257
13	0.001380	0.300509	0.004591	0.996366
14	0.204880	0.263944	0.776226	0.443315

15	-0.028546	0.194077	-0.147085	0.883988
R=.95				
R^2=.904				

* Largest t-value

Table 4. Distributed Lag Analysis - ADR Regressed Against Supply

Lag	Regression Coeff.	Standard Error	T(52)	P
0	1.975244	0.198694	9.94116*	0.000000
1	0.063453	0.239619	0.264809	0.792203
2	-0.010490	0.258083	-0.040647	0.967733
3	0.482757	0.249389	1.935758	0.058343
4	0.051826	0.233889	0.221582	0.825507
5	-0.062995	0.187075	-0.336738	0.737670
R=.8713				
R^2=.7591				

* Largest t-value

Conclusions Regarding Causation and Exogeneity from the Simulated Data

The findings here support the hypothesis that supply causes demand in the Granger sense. As expected, an eight-quarter lag is statistically significant. A Granger Test for Causality yields a test F statistic of 517.734, which easily surpasses the critical hurdle value (F_{crit}) of 3.179 and with 63 observations the resulting P-level is 0.00. The statement can therefore be made that supply Granger causes demand at above the 99% significance level. A summary of this test has been included below in Table 5. Of course, this test is more for time-precedence than true causality. To evaluate if supply is an exogenous variable in the supply/demand relationship a two-stage-least-square model (2SLS) is run.

Table 5. Output from the Granger Test for Causality

Restricted sum of squares (no lag X)	5.074	
Unrestricted (w/ lag X)	0.455	
Number of obs (T)	53	
Number of regressors - unrestricted model (K)	2	
Number of coeffecients restricted to zero (s)	1	
SSR (r) - SSR (U)	4.619	Numerator
SSR (r) - SSR (U)/S	4.619	
T-K	51	Denominator
SSR (U)/ T-K	0.00892157	
Test Stat	517.734	
F-Criteria	3.179	
F-Dist	0.00%	
Alpha	5%	
Lagged variable improves analysis		

2SLS is a more powerful tool for determining if supply is a reaction to increased demand or if supply is creating demand. The findings from the 2SLS model support the third hypothesis that supply 'causes' demand. A linear-regression of supply and supply lagged against demand as the dependent variable yields the following coefficients: $ß_1$ = -.00433 and $ß_{t-8}$ = 2.186. A 2SLS regression with the same variables produced nearly identical coefficients ($ß_1$ = -.004814 and $ß_{t-8}$ = 2.185). Table 6 below details this output. Since these two estimates produce similar results, supply is exogenous and supply causes demand.

Table 6. Two-Stage Least Squared Model

	B	SE Beta	Beta	T Sig	T
Supply	-0.004814	0.092979	-0.002134	-0.052	0.9589
Supply Lagged	2.18538	0.0944	0.954124	23.15	0
R= .9553					
R^2=.91254					

CONCLUSIONS AND PROCESS APPLICATIONS

The assumption that room development causes incremental room night demand is a fundamental premise of lodging market research (Rushmore, 1997). This dormant room night demand is termed as latent demand. Accurately quantifying the amount of latent demand within a lodging market has long been an elusive activity. Although the concept of latent demand is well established, there exists a poorly defined system for evaluating this demand within a given lodging market. Consequently, lodging feasibility studies are notoriously inaccurate and often contentious. This study described a process for quantifying the magnitude, timeframe and causation related to latent demand. The ability of the process to estimate a latent demand structure was then verified with a simulated dataset. The latter portion of the paper described the process parameters from analyzing the simulated dataset with the process.

The process depicted in the preceding paragraphs is a substantive approach to estimating the effect of latent demand. Analysts can use this process to determine if development is a response to owners anticipating demand or if demand is 'caused' by development. In addition, this process develops a portrait of the magnitude and timeframe associated with supply additions. The magnitude of the effect associated with room development is estimated by means of a general linear model. This model lets the analyst specify a categorical variable that picks up some of the noise associated with different years. This paper also forwards a simple framework for estimating the appropriate lag duration associated with supply changes. Last, the process outlined above helps define the causal nature of the supply/demand relationship.

This tool will assist in drawing defensible conclusions regarding market or project viability. By estimating the effect, causation and timeline associated with supply changes analysts can make appropriate decisions and reasonable forecasts. For instance, the elasticity conclusions gleaned from this process will help define the robustness of the market. The resulting project proforma is a more appropriate reflection of the market, which leads to sound, well rationed decisions. Further, asset managers and revenue managers can use this process to adjust their operating assumptions to market dynamics. For instance, the booking curve utilized in the revenue management process might change with the addition of new hotels (Harris, 1999). This tool can be utilized to understand when (if at all) the booking rate will change and what the magnitude of the effect will be. This tool is also useful for planning boards to decide if new additions will create undue competition or foster needed demand. Planning boards can determine what effects a new hotel's entry will be on market-wide RevPAR for the incumbent hotels and the proposed entry. Markets where the market-wide RevPAR is expected to decrease are poor candidates for expansion.

REFERENCES

Almon, S. (1965). The distributed lag between capital appropriations and expenditures. *Econometrica*, *33*, 178-196.

Canina, L., & Carvell, S. (2003). Lodging demand for urban hotels in major metropolitan markets. *CHR Reports, 3(3)*, 5-24.

Corgel, J. (2004). *Predictive powers of hotel cycles.* Retrieved September 15, 2004 from http://www.hospitalitynet.org/news/154000320/4020379.html.

DeRoos, J. (1999). Natural occupancy rates and development gaps. *Cornell Hotel and Restaurant Administration Quarterly, 40(2)*, 14-23.

Franses, P. (2003). *A concise introduction to econometrics: An intuitive guide (1st ed.).* New York, NY: Cambridge University Press.

Fulton, L., Noland, R., Meszler, D. & Thomas, J. (2000). *A statistical analysis of induced travel effects in the U.S. Mid-Atlantic region*. Presentation at the 79th Annual Meeting of the Transportation Research Board, Washington, D.C.

Ganchev, O. (2000). Applying value drivers to hotel valuation. *Cornell Hotel and Restaurant Administration Quarterly, 41(5)*, 78-89.

Harris, P. (1999). *Profit planning.* Woburn, MA: Butterworth-Heinemann.

Hinkin, T. & Tracy, B. (2003). Continued relevance of "Factors driving meeting effectiveness". *Cornell Hotel and Restaurant Administration Quarterly, 44*(5/6), 27-30.

Mooney, C. Z. (1997). *Monte Carlo Simulations (Quantitative applications in the social sciences).* Thousand Oaks, CA: Sage Publications Inc.

Noland, R. (1999). Relationships between highway capacity and induced travel vehicle traffic. Presentation at the 78th Annual Meeting of the Transportation Research Board, Washington, D.C.

Poole, S. (2000, July 13). Officials skeptical of study on hotels. *Atlanta Journal-Constitution*, G1

Rushmore, S. (1978). *The valuation of hotels and motels.* Chicago, IL: American Institute of Real Estate Appraisers of the National Association of Realtors.

Rushmore, S. (1990). *The computerized income approach to hotel and motel market studies and valuations (rev. ed.).* Chicago, IL: Appraisal Institute.

Rushmore, S. (1997). *Hotel investments handbook: A guide for lenders and owners.* New York, NY: Warren, Gorham & Lamont.

Sanders, H. (1999a). *If we build it, will they come? And other questions about the proposed Boston Convention Center*. Boston, MA: The Pioneer Institute.

Sanders, H. (1999b). *Flawed forecast: A study of convention center feasibility studies.* Boston, MA: The Pioneer Institute.

Sarbanes, P. & Oxley, M. (7-30-2002). Sarbanes-Oxley Act of 2002. Ref Type: Bill/Resolution.

Statsoft Corporation. (2002). *Statistica 6.0 Contents and Advisor* [Computer software].

Sturman, M. (2003). Computer simulation in hospitality teaching, practice, and research. *Cornell Hotel and Restaurant Administration Quarterly*, *44*(2) 85-93

Talluri, K., & Van Ryzin, G. (2000, September 1). A discrete choice model of yield management. Unpublished work

Vahid, F., & Engle, R. (1993). Common trends and common cycles. *Journal of Applied Econometrics, 8(4),* 341-360.

Wallace, D. & Silver, L. (1998). *Econometrics: an introduction.* New York, NY: Addison-Wesley

Wheaton, W. & Rossoff, L. (1997). The cyclic behavior of the U. S. lodging market. *Hotel Real Estate Economics, 26(1),* 67-82

Submitted March 3rd, 2004
First revision received November 7, 2004
Second revision received June 13, 2005
Third revision received July 6, 2005
Accepted August 22, 2005