DEMAND FORECASTING IN REGIONAL AIRPORTS: DYNAMIC TOBIT MODELS WITH GARCH ERRORS

Matthew G. Karlaftis, Ph.D.
Department of Transportation Planning and Engineering
National Technical University of Athens
Tel: +302107721280, e-mail: mgk@central.ntua.gr

ABSTRACT
In this paper we discuss the general issue of forecasting highly seasonal demand in regional airports, where peak flows approach airport capacity. For this, we propose a modeling combination, dynamic Tobit models with GARCH errors/disturbances, that is able to capture many of the shortcomings of most traditional models. Models are calibrated using monthly passenger and flight data for a 20 year period for the airport of Corfu in Greece, where traffic over the summer approaches airport capacity and seasonal fluctuations in demand are very intense. Results show that: i. Not explicitly accounting for seasonal variations in demand or for traffic approaching capacity may significantly bias model parameter estimates and affect demand predictions; and, ii. Improved demand model specifications are an invaluable tool in obtaining more accurate demand estimates.

1. INTRODUCTION
Forecasting is at the heart of the planning and design process at airports. Airport terminals, runways, freight storage facilities, parking lots, and even the roadway network to and from an airport are all based on the forecasts for the airport. Forecasts of passenger volumes are translated to space requirements for the terminal building facilities, while forecasts of aircraft movements are translated to the runway, taxiway, and apron needs, as well as to the need for air traffic control systems.

When planning and designing airport terminal there is a significant number of traffic characteristics that must be forecasted: total number of passengers for the design period, domestic, commuter and international passenger ratios at peak hours, seasonal variations in demand, volumes of transfer and/or transit passengers for each type of traffic, number of passengers, bags and well wishers, distribution of dwelling times of passengers, and sometimes origin and destination of flights for immigration, customs, and health control purposes (Tösic, 1992). These forecasts are used to determine the space requirement for new terminals or the expansion of existing facilities.

It is evident that the forecasting process can be the most critical factor in the development of the airport (Howard, 1974). Mistakes made in this phase of the process can be very costly and damaging for local economies. Underestimating demand can lead to increased congestion, delay, and lack of storage facilities, as it happened in Venezuela in 1974. The discovery of oil resulted in a dramatic and unforeseen increase of the freight volumes handled by the Caracas airport. The planned storage facilities were insufficient to handle this increased demand, and so the cargo was stored in areas where it was either destroyed or stolen. Overestimating demand could also create significant problems. Forecasts of passenger demand for the Newark airport were so high that the newly constructed airport remained empty for a number of years (de Neufville, 1976). Errors in the forecasting process can lead to long delays or to empty terminals: both these cases can cause significant damage to the economy of an airport’s hinterland that depends on the successful operation of that airport.
There are three main approaches for forecasting civil aviation traffic: market surveys, trend projections and econometric relationships (models). Particularly the former, econometric methods, with regression as the main tool, have enjoyed years of extensive use internationally; however, there are cases where traditional regression time-series or ARIMA analyses yield questionable forecasts, mainly as a result of their inherent assumptions. Such a case is in forecasting heavily seasonal passenger traffic in regional airports, where intense fluctuations (volatility) exists in daily and monthly passenger traffic, and where both regression and classical time-series analyses cannot successfully capture the nature of demand.

In this paper we propose a modeling combination, dynamic Tobit models with GARCH errors/disturbances, that is able to capture many of the shortcomings of traditional models. A tobit model assumes a dependent variable with censoring (where, for example, passenger traffic in certain months approaches airport capacity). The GARCH extension for the Tobit model disturbances allows us to model the volatility in seasonal passenger traffic without jeopardizing model robustness. Models are calibrated using monthly passenger and flight data for a 20 year period for the airport of Corfu in Greece, where traffic over the summer approaches airport capacity and seasonal fluctuations in demand are very intense.

2. BACKGROUND

The goal of the review is to identify the work done for airport-specific forecasting, which is the focus of this paper, rather than for corridor forecasting. Corridor forecasting is concerned with the demand for air-travel in a given travel route (corridor), say the Rome (Italy) - Athens (Greece) corridor. While this type of forecasting is certainly of great interest to airline companies, it is not as useful to airport planners who are mainly concerned with the aggregate passenger forecasts for a given airport (Whitford, 1990). Knowledge of the aggregate passenger forecasts will provide planners with the necessary guidelines for the planning and design of the entire airport facilities.

The key issue in the development of successful airport forecasts is the ability to identify those factors that influence the change in the volume of passenger traffic at an airport. There are two major approaches in the development of forecasting models: Simple Time Series (STS) and Causal Modeling. STS methods are the most widely used methods for predicting charter air-travel demand, and assume that “history repeats itself,” in that the underlying stochastic structure of the data does not change with time. The common characteristic, and weakness, of all time-series methods is that they ignore the determinants of demand such as fares, income, GNP, and do not attempt to explain the causes of change in demand. Causal modeling attempts to determine, from given data, causality or least explanation when analyzing the relationship between certain data; that is, the relationship between two or more variables is used to predict one variable (the dependent) from the others (independent).

One of the first published causal modeling efforts concerned with airport-specific forecasting was by Jacobson (1970). His linear regression based model predicted the trips generated at an airport (dependent variable) based only on the average airfare per mile for all routes in the United States, and the total income per capita for the airport catchment area. The model was
calibrated with eighteen years of data from the airports of Virginia. A rather similar regression model was developed by Haney (1975) to analyze airport specific demand for the St. Louis, Missouri, airport. The model is a typical airport forecasting model in that the explanatory variables are socio-economic in nature and refer to the metropolitan area served by the airport.

A different type of forecasting model was developed by Thomet and Sultan (1979) for the Riyadh International Airport in Saudi Arabia. The model used imports and exports of crude oil and petroleum products as the sole independent variable, to account for the fact that most travel to and from Saudi Arabia was related to the oil emporium.

Another interesting example of placing a special effort in the direction of identifying key growth factors unique to an area is the case study of Mexico city’s airport traffic by Zuñiga et al. (1978). Approximately 20% of the population and wealth of Mexico are concentrated in the capital, and the fast growth rates of these characteristics certainly influence the future airport traffic. The models that were developed for domestic traffic included as variables the population of Mexico City and the cost of transportation in constant pesos. Mellman et al. (1980) developed a model for Boston’s Logan airport. Unlike many other airports where transit passengers are a significant portion of the airport’s traffic, Logan’s traffic is mostly Boston based. Explanatory variables in this case related air-travel to the socio-economic characteristics of the Boston region.

Using a different approach, several models were built using STS modeling for air-travel demand forecasting. STS modeling has been used extensively in the literature in those cases where causality could not be established, or where economic information on the airport’s catchment area was not readily available (ICAO, 1991). As an example, passenger forecasts for the Cornavin Airport in Geneva, Switzerland, have been obtained by STS methods (ICAO, 1985); STS methods were also used extensively in the forecasts for the fourth airport in the Chicago region (al Chalabi and Mumayiz 1991, al Chalabi 1993, Whitford 1990). Moreover, forecasts of air-travel growth for the Airbus industries are often obtained using STS methods (Lenormand, 1989).

Finally, Karlaftis et al. (1996) examined the predictive ability and forecasting accuracy of air-travel demand models. In particular, they developed an analytical framework for developing econometric models, and used post-fact analysis to test the accuracy of the models. Statistical data describing air-travel patterns for two major international airports, the Miami and Frankfurt International Airports, were used to demonstrate the effectiveness of the proposed models. In addition, the effect of external factors such as the deregulation of the air-transport industry was examined. The results suggested that simple models with few independent variables perform as well as more complicated and costly models, and that external factors have a pronounced effect on air-travel demand.

3. DATA AND METHODOLOGY

3.1. The Data
The Greek airport system comprises of 35 airports that service commercial flights. From these airports approximately one third are international; that is, they are open to direct flights from airports outside of Greece. The airports that were mostly affected by the increase in direct
charter flights were the major airports located in popular resort areas. These airports include the airports of Zakynthos and Kerkyra (Corfu) islands (both located on the Ionian Sea off the west coast of mainland Greece), the airports of Kos and Rhodes (both located on the Aegean Sea off the east coast of mainland Greece), and the airport of Chania (in the island of Crete, at the south end of Greece). The airports off the west (east) coast serve as destinations to passengers visiting the “west” (“east”) islands, because tourists can either remain on those islands or take small boats to visit the neighboring islands (a type of intermodal hub-and-spoke system). Similarly, passengers to the airport of Chania are visitors of Crete, the largest Greek island. Figures 1 and 2 depict the evolution of annual and monthly data for a variety of traffic categories (domestic flight and passenger arrivals, international passenger arrivals, and so on). What is of particular interest is the notion (Figure 2), of very intense seasonal characteristics for all the measures examined, coupled with similar peak values on passenger traffic. Seasonality suggests that international passenger arrivals is almost zero in the winter months, while it approaches airport capacity in the summer months.

Figure 1: Annual Data on Domestic and International Arrivals

Figure 2: Monthly Data on Domestic and International Arrivals
The origin of the tourists visiting Greece is important in attempting to identify the key factors affecting the dramatic increase in island airport traffic. From Figure 3 (OECD, 1994), it is evident that approximately 90% of the tourists originate in Europe. Therefore, explanatory variables were selected that account for socio-economic characteristics for Europe. Interestingly, the Greek economy is closely tied to the revenues generated by the tourist industry; the ratio of private consumption in foreign currency to total private consumption, a strong indicator of the importance of tourism to the National Economy (Figure 2), clearly indicates this importance.

Figure 3: Origin of Tourists Visiting Greece

Figure 4: Ratio of Private Consumption in Foreign Currency to Total Private Consumption
The sample period for the monthly analysis presented is the period 1989-2006. The dependent variables in the models were: i. Domestic flight arrivals; ii. Domestic passenger arrivals; iii. International flight arrivals; iv. International passenger arrivals. Data were provided by the Greek Civil Aeronautics Board; data are collected on a monthly basis for all Greek airports. The independent variables that were considered for inclusion in the models were the Gross Domestic Product (GNP) per capita for Europe, the Personal Consumption Expenditures (PCE) per capita for Europe, the Disposable Income (DI) per capita for Europe (all measured in constant 2006 €), Regional Product for Corfu, GDP for Greece, Number of Tourist arrivals on the island, population, a time trend (taking the value of 1 for 1989, 2 for 1990 and so on), and monthly dummy variables (to capture monthly variability in the dependent variables).

The above mentioned explanatory variables were selected because they are ideally considered as some of the “best” descriptors of the ability (and desire) for tourists to travel. The development of air-travel is also affected by a large number of external economic factors, government policies and regulations (such as travel advisories and strikes of the air-traffic controllers), physical phenomena (Greece has faced a large decrease in the number of international tourists the years when earthquakes have occurred), technological developments and the operating economics of the air-transportation industry. Unfortunately, these occurrences cannot be forecasted and, as a result, the models have to be built on measures that have not only been diligently observed in the past, but also for which forecasts exist (Kanafani 1981). In preparation and application of the forecasts, these uncertainties have to be recognized; nevertheless, a forecaster’s task is to use the data sources readily available, and apply methods that minimize the range of uncertainty.

3.2. The Methodology

The initial thought and approach in modeling demand for the Corfu airport was to develop Simple Time Series (STS) models (Karlaftis et al. 1997). But, while STS modeling has been extensively used in charter passenger travel, questions still remain as to the factors that affect its demand, as well as its elasticity with respect to those factors. Further, and besides modeling external influencing factors, the demand characteristics for regional Greek airports, i.e. intense seasonal variation in demand and peak demand approaching airport capacity, necessitated the development of a new approach. In this respect, we propose a modeling combination, dynamic Tobit models with GARCH errors/disturbances, that is able to capture many of the shortcomings of traditional models. A tobit model assumes a dependent variable with censoring (where, for example, passenger traffic in certain months approaches airport capacity); the GARCH extension allows us to model the volatility in seasonal passenger traffic without jeopardizing model robustness.

Limited dependent variables were introduced in econometrics in Tobin (1958) and Amemiya (1973). Let

\[ y_i^* = x_i \beta + \epsilon_i \]  

(1)

be the underlying latent equation before censoring. The \( x_i \) is a vector of exogenous variables. The observed variable is
where the \( c_i \)'s are known constants. To allow serial correlation in \( \epsilon_i \) and dynamic structure in the model, one can consider the general formulation

\[
y_i^* = \psi(\bar{y}_{t-1}, \bar{y}_t, \bar{f}_t) + \epsilon_i,
\]

where \( \bar{y}_t \) denotes a vector of \( y \) consisting of its observed current and lagged values from time 1 to \( t \), \( \bar{y}_t^* \) is the vector consisting of all latent-dependent variables of \( y^* \) preceding and up to period \( t \), and \( \bar{f}_t \) is defined accordingly. For simplicity, unknown parameters in \( \psi \) of the model are suppressed. As one knows whether \( y_t^* \) has been censored or not at each \( t \), it is convenient to define a dichotomous indicator \( I_i \) such that

\[
I_i = \begin{cases} 
1 & \text{if } y_t^* > c_i, \\
0 & \text{if } y_t^* \leq c_i.
\end{cases}
\]

If \( I_i = 0 \), \( y_t^* \) has been censored and its value is not observed. When \( I_i = 1 \), \( y_t^* \) is observed. A model with serially correlated disturbances such as

\[
y_t^* = x_t \beta + u_t, \quad u_t = \rho u_{t-1} + \epsilon_t
\]

is a member of the general case in Eq. (3), because it implies that

\[
y_t^* = \rho y_{t-1}^* + (x_t - \rho x_{t-1}) \beta + \epsilon_t
\]

after a quasi-difference transformation. Other relevant and interesting examples are censoring models with ARCH or GARCH disturbances (Engle, 1982; Bollerslev, 1986). In a Tobit ARCH model with \( y_t^* = x_t \beta + \epsilon_t \), the conditional variance \( \sigma_t^2 \) of \( \epsilon_t \) conditional on past information depends on past disturbances \( \epsilon_{t-1}, \epsilon_{t-2}, \) etc. For an ARCH(p) process,

\[
\sigma_t^2 = \alpha_0 + \sum_{j=1}^{p} \alpha_j \epsilon_{t-j}^2
\]

In a GARCH(p,q) process, additional lagged variances introduced as

\[
\sigma_t^2 = \alpha_0 + \sum_{j=1}^{p} \alpha_j \epsilon_{t-j}^2 + \sum_{k=1}^{q} \alpha_{p+k} \sigma_{t-k}^2
\]

Dynamic structures can be incorporated into an ARCH or GARCH model. The ARCH in the mean (ARCH-M) model introduced by Engle et al. (1987) is specified as

\[
y_t^* = x_t \beta + \delta \sigma_t + \epsilon_t
\]

in addition to Eq. (5). A survey of ARCH models is given in Bollerslev et al. (1995). Without censoring, popular ARCH and GARCH models are estimated by the ML or generalized methods of moments (GMM). Additional complication is introduced with censoring. Under censoring, the calculation of underlying residuals for censored observations are infeasible, as exact values of these residuals are not observed. To avoid this difficulty, Calzolari and Fiorentini (1996) modified the ARCH process by replacing the lagged \( \epsilon_{t-j}^2 \)
formulation in Eq. (6) by its conditional expected value $E_{t-1}(\epsilon_t^2)$, where $E_{t-1}$ means the conditional expectation conditional on past information up to $t-1$.

4. MODEL ESTIMATION

In Tables 1 and 2 we provide the models that provide the best fit to the data after extensive testing (for details on model development and testing see Washington et al. 2003). Coefficients for all variables, including GDP and tourist arrivals, have the expected signs and are significantly different from 0 at the 90% in most cases. We note here that, in all cases, both dependent and independent variables were taken in their log forms; this, as the literature clearly shows (Washington et al. 2003), leads to models were coefficient estimates are elasticities. In most cases, all dependent variables are inelastic with respect to explanatory variables such as GDP and arrivals. For example, during the period of the analysis, a 10% increase in Per Capita GDP for Greece increases domestic passenger traffic at the airport by 4.8%. We also note that, in all models, we included both an annual time trend (to capture overall movement in the series), and monthly dummy variables to capture monthly variations in overall traffic levels.

Despite the above corrections, from Tables 1 and 2, we note that neither monthly variation nor traffic levels approaching capacity where fully captured. This is evident by the statistical significance (almost all at the 95% significance level), of all Tobit-GARCH parameters for both domestic and international traffic variables.

<table>
<thead>
<tr>
<th>Table 1: Domestic Flight and Passenger Arrivals (Monthly Data)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Independent Variables$^a$</strong></td>
</tr>
<tr>
<td>Constant</td>
</tr>
<tr>
<td>National GDP</td>
</tr>
<tr>
<td>European GDP$^b$</td>
</tr>
<tr>
<td>Tourist Arrivals</td>
</tr>
<tr>
<td>Time Trend</td>
</tr>
<tr>
<td><strong>Monthly Dummies$^c$</strong></td>
</tr>
<tr>
<td>Unconditional Variance$^d$</td>
</tr>
<tr>
<td>Lagged Variance</td>
</tr>
<tr>
<td>Lagged Squared Disturbance</td>
</tr>
</tbody>
</table>

a All variables in log form (double log model; all coefficients are elasticities)
b Weighted average of the GDP of countries of origin for tourists
c All dummies are statistically significant
d GARCH parameters from Eq. (6)

To further examine the differences in modeling airport demand with the two different models used (OLS versus Tobit-GARCH), we examine the models highlighted in Table 2 (monthly...
data on international flight arrivals). As is obvious, differences in coefficient estimates for the two models (and hence of the elasticity estimates for international flights with respect to the parameters examined) are vastly different. For example, if analysts had used OLS with monthly corrections (the most frequently used approach for modeling airport demand), there would have been severe underestimation of both parameter significance (i.e. what the important parameters for modeling airport demand are) and of the magnitude of elasticity estimates with respect to a number of parameters. Interestingly, while in the case of international flights modeled via OLS, the only statistically significant parameter is the yearly trend, indicating an annual 0.5% increase in international flights, in the Tobit-GARCH estimation approach the results are different. First, the annual trend is decreasing (by 0.2% annually); second, European GDP is statistically significant, suggesting that an increase in European GDP increases Corfu Airport International flights (a 10% increase in per Capita GDP for Europe increases Corfu Airport International flights by 10.2%); and, third, increases in monthly tourist arrivals to Corfu also ‘help’ in increasing international flights (an expected result; a 10% increase in tourist arrivals increase Corfu Airport International flights by 0.17%).

Table 2: International Flight and Passenger Arrivals (Monthly Data)

<table>
<thead>
<tr>
<th>Independent Variables(^d)</th>
<th>OLS Flights</th>
<th>OLS Passenger Arrivals</th>
<th>Tobit - GARCH Flights</th>
<th>Tobit - GARCH Passenger Arrivals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-0.18</td>
<td>-0.32</td>
<td>-10.86</td>
<td>5.79</td>
</tr>
<tr>
<td>National GDP</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>European GDP(^b)</td>
<td>0.13</td>
<td>0.22</td>
<td>1.24</td>
<td>0.24</td>
</tr>
<tr>
<td>Tourist Arrivals</td>
<td>0.056</td>
<td>1.06</td>
<td>0.017</td>
<td>0.08</td>
</tr>
<tr>
<td>Time Trend</td>
<td>0.044</td>
<td>2.91</td>
<td>-0.022</td>
<td>0.04</td>
</tr>
<tr>
<td>Monthly Dummies(^c)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Unconditional Variance\(^d\) | 0.00025     | 3.93                   | 0.01                   | 4.01                            |
| Lagged Variance              | 0.105       | 2.63                   | 0.006                  | 1.79                            |
| Lagged Squared Disturbance   | 2.939       | 4.27                   | 2.92                   | 4.35                            |

\(^a\) All variables in log form (double log model; all coefficients are elasticities)
\(^b\) Weighted average of the GDP of countries of origin for tourists
\(^c\) All dummies are statistically significant
\(^d\) GARCH parameters from Eq. (6)

5. IMPLEMENTATION ISSUES

When planning and designing airport terminals there is a significant number of traffic characteristics that must be forecasted: total number of passengers for the design period, domestic, commuter, and international passenger ratios at peak hours, seasonal variations in demand, volumes of transfer and/or transit passengers for each type of traffic, and sometimes origin and destination of flights for immigration, customs, and health control purposes. The aggregate passenger forecasts, such as those presented in this paper, are used to determine the space requirements for new terminals or the expansion of existing facilities.
Forecasting is also very important for the economic development of airline fleets which have grown significantly in the last few decades. Orders for commercial transport aircraft have reached an all time high, with delivery rates now backed-up into the next century for some aircraft types. Future fleet development has an important role in the design of terminals since the aircraft size is critical to geometric characteristics, such as the main deck elevation and the lateral clearance at gates due to the aircraft wing-span.

The basis for the development of new and for the expansion of already existing terminal facilities in the Greek airport system is based on these forecasts. The data necessary for these forecasts is readily available through the Greek Civil Aeronautics Board and our analysis has attempted to establish causal relationships between airport demand and a set of explanatory variables. These explanatory variables are statistically adequate; that is, the data is collected by the Greek government in a consistent manner at equally spaced time intervals. It is also important to note that the analysis of time-series data should be done cautiously to account for the presence of: i. error term serial correlation, ii. intense variability in monthly evolution of the variables, and iii. cases were demand approaches capacity. If any of these three issues is systematically ignored, conclusions could be reached that parameter estimates are more precise than they actually are (loss of parameter efficiency) and, as shown in this paper, biased parameter estimates could also be obtained (Washington et al. 2003). It is evident that until actual values of the future air-travel demand are known, a model accuracy cannot be precisely ascertained. Nevertheless, it is important to ensure the statistical validity of the forecasting models that will be used in practice, so that a rational basis for the selection of the appropriate model exists.

6. CONCLUDING REMARKS

The goal of this paper was to examine and analyze the demand for, and the ability to, forecast passenger traffic and flights at regional airports. In this paper we proposed a modeling combination, dynamic Tobit models with GARCH errors/disturbances, that is able to capture many of the shortcomings of traditional airport demand forecasting models. A tobit model assumes a dependent variable with censoring (where, for example, passenger traffic in certain months approaches airport capacity). The GARCH extension for the Tobit model disturbances allows us to model the volatility in seasonal passenger traffic without jeopardizing model robustness. Models were calibrated using monthly passenger and flight data for a 20 year period for the airport of Corfu in Greece, where traffic over the summer approaches airport capacity and seasonal fluctuations in demand are very intense.

Given the nature of the data base and the rather uncertain and variable conditions that exists at regional airports, the results are surprisingly good. The model is statistically significant and exhibits good fit to the historical trends. Also, and very important from a statistical perspective, there is clear evidence that not explicitly accounting for both demand approaching capacity and for intense monthly variation, affects the efficiency of parameter estimates and leads to biased – and erroneous – elasticity measures. These errors could lead to significant underestimation of future demand at airports.

Further work on charter passenger traffic is certainly warranted, given the generally diverse and volatile nature of travel at regional airports with high tourist related travel. Forecasts are
of course uncertain, yet the establishment of causal relationships gives a more rational and scientific basis for model development, and makes airport planners more secure in planning and designing for airports based on the available forecasts. Moreover, careful work should be done in the direction of international chartered traffic to identify the factors that affect this type of demand; this will become increasingly important as passenger volumes approach airports capacities and low cost carriers become an even more dominating force in international tourist travel.

BIBLIOGRAPHIC REFERENCES

al Chalabi, S. 1993, “Aviation Demand Forecasting - Traditional and New Approaches (The application of ground transport methodologies to enplanement forecasting).” Submitted to the Transportation Research Board.


International Civil Aviation Organization 1985, Airport Economics, Doc. 8991-AT/722/2, Montreal, Canada.


