Border Region Bridge and Air Transport Predictability

Fullerton, Thomas M., Jr. and Mukhopadhyay, Somnath

University of Texas at El Paso, University of Texas at El Paso

11 April 2013

Border Region Bridge and Air Transport Predictability

Journal of Business & Economics*, 2013, Volume 4, Number 11, Pages 1089-1104*

Thomas M. Fullerton, Jr.[1], Somnath Mukhopadhyay[2]

(1. Department of Economics & Finance, University of Texas at El Paso, El Paso, TX 79968-0543, USA;

2. Department of Information & Decision Sciences, University of Texas at El Paso, El Paso, TX 79968-6282, USA)

Abstract: Border region transportation forecast analysis is fraught with difficulty. In the case of El Paso, Texas and Ciudad Juarez, Chihuahua, Mexico, dual national business cycles and currency market fluctuations further complicate modeling efforts. Incomplete data samples and asymmetric data reporting conventions further confound forecasting exercises. Under these conditions, a natural alternative to structural econometric models to consider is neural network analysis. Neural network forecasts of air transportation and international bridge activity are developed using a multi-layered perceptron approach. Those out-of sample simulations are then compared to previously published forecasts produced with a system of simultaneous econometric equations. Empirical results indicate that the econometric approach is generally more accurate. In several cases, the two sets of forecasts are found to contain complementary information.

Key Words: regional transport demand; neural networks; econometric forecasting

JEL Codes: R15, R41

1. Introduction

Traditional econometric forecasts of surface and air transportation traffic flows in border regions have been found to contain errors. One potential alternative to structural econometric models is provided by neural networks. Neural network forecasts have proven helpful in a variety of different settings, but have not been extensively tested using data for border metropolitan economies. This study carries out such an exercise using data for the Borderplex economy of El Paso, Texas and Ciudad Juarez, Mexico.

Border region transportation analyses are complicated by numerous factors. Some of the more prominent examples include dual national business cycles, cross jurisdictional metropolitan business cycles, incomplete data coverage, data asymmetries, and currency market fluctuations. Regional demographic and labor market conditions represent further obstacles to predictive accuracy in the case of El Paso, Texas and Ciudad Juarez, Chihuahua, Mexico (West, 2003). Given the preceding, it is not surprising that border region econometric transportation forecasts have historically posted mixed accuracy records, relative to random walk benchmarks, in at least some markets (Fullerton, 2004).

Under these circumstances, one potential alternative approach is provided by neural network analysis (Vlahogianni, Golias, & Karlaftis, 2004). Neural networks have been applied to a number of different economic and transport forecasting problems and frequently are found to provide accurate predictions relative to other methodologies. In particular, they are well suited to handling situations in which data generating processes are unknown and may involve nonlinearities (Jagric & Strasek, 2005).

Thomas M. Fullerton, Jr., Ph.D., Professor & Trade in Americas Chair, University of Texas at El Paso; research areas: border economics, urban economics, applied econometrics. E-mail: tomf@utep.edu.

Somnath Mukhopadhyay, Ph.D., Associate Professor, University of Texas at El Paso; research areas: geometry of classifiers. E-mail: smukhopadhyay@utep.edu.

This study examines the applicability of neural networks to forecasting international port-of-entry bridge traffic as well as major airport transportation flows in the El Paso and Ciudad Juarez Borderplex regional economy that straddles the boundary separating the United States from Mexico. Section 2 provides an overview of related studies. Section 3 discusses data and methodology. Section 4 summarizes empirical results. Concluding remarks follow.

2. Related Studies

Traffic forecasting is an area of substantial academic and practitioner research interest. That is due to the critical role of transportation in the global economy as well as the resource limitations which constrain infrastructure development. At present, a variety of time series and econometric are generally utilized these efforts, with no single method consistently outperforming the others (Taylor, 2010; Fildes, Wei & Ismail, 2011; Tsekeris & Tsekeris, 2011). Several studies have also employed neural network approaches in attempting to improve traffic planning and forecasting effectiveness (Dunne & Ghosh, 2012; Wei & Chen, 2012).

Recognition of the importance of cross-border traffic and transport planning has intensified in recent years as a consequence of reduced trade barriers, greater demands for imported products, and increased cargo vehicle traffic volumes (McCray, 1998; Fullerton & Tinajero, 2002). Much of the literature on this topic deals with the infrastructure capacity constraints, personnel staffing shortages, and other problems associated with the boundary between the United States and Mexico (Saintgermain, 1995; Nozick, 1996). In addition to cargo vehicles, substantial attention has also been paid to light vehicle and pedestrian flows through the international ports of entry (Fullerton, 2000).

Accompanying the growth in merchandise trade has been an increased focus on delivery times and transportation expenditures as a percentage of the cost of doing business (Stank & Crum, 1997). Not surprisingly, these efforts led to substantial concern over how to address periodic congestion problems and infrastructure development issues (Harrison, Sanchez-Ruiz, & Lee, 1998; Bradbury, 2002). Security delays and port staffing practices frequently serve to raise transportation costs and proposals in favor of “seamless borders” form a substantial plank within this research landscape (Figliozzi, Harrison, & McCray, 2001; Ashur, Weismann, Perez, & Weismann, 2001).

Concerns over management and administrative practices are well placed. Following the 11 September 2001 terrorist attacks, North American border transportation costs grew and substantial economic displacements occurred in response to new security measures, regulatory, and administrative changes (Taylor, Robideaux, & Jackson, 2004; Fullerton, 2007). In spite of temporary declines associated with business cycle downturns, merchandise trade continues to expand, leaving traffic and transportation issues squarely on the agenda regarding border development policies (Villa, 2006). The ongoing presumption is that, in spite of trade growth, infrastructure bottlenecks and institutional barriers at the border cause international commerce to underperform its natural level by large amounts (McCallum, 1995; An & Puttitanun, 2009).

Similar to domestic infrastructure, planning efforts regarding maintenance and new facility construction of border ports of entry inevitably run into questions over funding (Levinson, 2005). Three of the four international bridges linking El Paso and Ciudad Juarez charge tolls to pedestrians, light vehicles, and cargo vehicles. The price elasticities for these various traffic flows are similar in magnitude to what has been documented for other types of bridges and highways (De Leon, Fullerton, & Kelley, 2009). Pricing and funding issues are likely to play prominent roles in future discussions involving port of entry bottlenecks.

Neural network analysis has not previously been utilized to model and forecast Borderplex transportation data. It has, however, been applied in several similar contexts to those considered below (Smith & Demetsky, 1997; Lam, Ng, Seabrooke, & Hui, 2004; Celikoglu & Akad, 2005). Among those studies, Lam, Ng, Seabrooke, and Hui (2004) report empirical results in favor of neural network cargo forecasts using data for the port of Hong Kong. Celikoglu & Akad (2005) also obtain relatively accurate neural network out-of-sample simulations of public transport volumes for Istanbul. The Northern Virginia highway traffic flow predictions for the neural network exercise of Smith and Demetsky (1997) are not as successful.

These and other research endeavors make it easy to see why transportation forecasting research continues to receive substantial academic and practitioner attention (Flyvbjerg, Holm, & Buhl, 2005; Flyvbjerg, Holm, & Buhl, 2006; Bain, 2009). This study examines whether neural network analysis can help forecast traffic flows across the international bridges between El Paso and Ciudad Juarez. Prior research indicates that econometric approaches toward this objective meet with mixed success in terms of predictive accuracy (Fullerton, 2004; De Leon, Fullerton, & Kelley, 2009). Neural network analysis may provide one means for improving forecast accuracies.

3. Data and Methodology

Transportation flows analyzed are the same as in Fullerton (2004). They include eight categories of air transportation data from El Paso International Airport. They also include eight categories of northbound traffic flows across the international bridges from Ciudad Juarez into El Paso. These data are forecast every year using the structural econometric system of simultaneous equations described in Fullerton (2001). Historical data employed in the Borderplex modeling system can be accessed via the web site at the University of Texas at El Paso (http://academics.utep.edu/Default.aspx?tabid=14417). As might be expected for a regional economy in which unemployment is relatively high and demographic data are subject to large periodic revisions, the track record of these structural forecasts is mixed (Charney & Taylor, 1984; West, 2003).

Table 1 summarizes the variable names, time span, and units of measure for each of the transportation series that are modeled and forecasted (Fullerton, 2004). Air passenger data are only available from 1979 forward. International air passenger data are only available from 1979 to 2006, the year in which international commercial passenger flights to El Paso were suspended. Air freight and air mail data are available from 1974 to 2007. Beginning in 2007, reporting conventions at El Paso International Airport changed and these series are now reported together rather separately. All of the international bridge time series are available from 1974 through 2011.

Table 2 provides an overview of the explanatory variables employed in the specifications for the various transportation variables listed in Table 1. Similar to other regions of the United States, transportation flows to and from El Paso were disrupted by infrastructure administrative practices that changed subsequent to 11 September 2011 (Fullerton, 2007). Accordingly, many of the Borderplex transportation equations now include a post-9/11 dummy variable either to allow for intercept adjustments or interacted with other explanatory variables. All of the econometric equations are dynamic in nature and contain one-period autoregressive lags of the dependent variables and/or ARMAX autoregressive and moving average serial correlation correction coefficients (Pagan, 1974).

Table 1. Transportation Variables Forecasted and Units of Measure

Variable Names	Definitions	Units
APDD	Domestic Air Passenger Arrivals, 1979-2011	Thousands
APDI	International Air Passenger Arrivals, 1979-2006	Thousands
APED	Domestic Air Passenger Departures, 1979-2011	Thousands
APEI	International Air Passenger Departures, 1979-2006	Thousands
AFDT	In-Bound Air Cargo, Total, 1974-2007	1,000 Tons
AFET	Out-Bound Air Cargo, Total, 1974-2007	1,000 Tons
AMD	In-Bound Air Mail, 1974-2007	1,000 Tons
AME	Out-Bound Air Mail, 1974-2007	1,000 Tons
BAC	Bridge of the Americas Northbound Cars, 1974-2011	Millions
BAT	Bridge of the Americas Northbound Trucks, 1974-2011	Millions
BAW	Bridge of Americas Northbound Pedestrians, 1974-2011	Millions
BPC	Paso del Norte Bridge Northbound Cars, 1974-2011	Millions
BPW	Paso del Norte Northbound Pedestrians, 1974-2011	Millions
BYC	Ysleta Bridge Northbound Cars, 1974-2011	Millions
BYT	Ysleta Bridge Northbound Trucks, 1974-2011	Millions
BYW	Ysleta Bridge Northbound pedestrians, 1974-2011	Millions

Table 2. Transportation Variables and Regressors

Equation	Independent Variables
APDD	USA Real Consumer Transportation Expenditures scaled by ratio of El Pas Population relative to USA Population Dummy Variable set to 1 for post-9/11 Period, set to 0 for 1974-1990 Dummy Variable for post-9/11 Period interacted with USA Real Transport. Exp.
APED	El Paso Wage & Salary Disbursements deflated by El Paso Air Travel Price Index
AFDT	El Paso Labor Income deflated by USA Transportation Price Index Dummy Variable for post-9/11 Period interacted with El Paso Real Labor Income
AFET	USA Real Gross Domestic Product Total Inbound Air Freight Cargo to El Paso International Airport Dummy Variable set to 1 for post-9/11 Period, set to 0 for 1974-1990
AMD	El Paso Real Gross Metropolitan Product Dummy Variable for post-9/11 Period interacted with El Paso Real GMP First Class Mail Price deflated by USA GDP Implicit Price Deflator
AME	El Paso Real Gross Metropolitan Product Dummy Variable for post-9/11 Period interacted with El Paso Real GMP First Class Mail Price deflated by USA GDP Implicit Price Deflator Dummy Variable set to 1 for post-9/11 Period, set to 0 for 1974-1990
BAC	Mexico Real Exchange Rate Index, Pesos per Dollar, Inflation Adjusted One Period Lag of El Paso Population plus Ciudad Juarez Population Dummy Variable set to 1 for post-9/11 Period, set to 0 for 1974-1990
BAT	Mexico Real Exchange Rate Index, Pesos per Dollar, Inflation Adjusted Dummy Variable for post-9/11 Period interacted with USA Real GDP Dummy Variable set to 1 for post-9/11 Period, set to 0 for 1974-1990
BAW	Mexico Real Exchange Rate Index, Pesos per Dollar, Inflation Adjusted Dummy Variable for post-9/11 Period interacted with USA Real GDP Dummy Variable set to 1 for post-9/11 Period, set to 0 for 1974-1990
BPC	One Period Lag of El Paso Population plus Ciudad Juarez Population Dummy Variable set to 1 for post-9/11 Period, set to 0 for 1974-1990 Dummy Variable for post-9/11 Period interacted with Regional Population Lag Mexico Real Exchange Rate Index, Pesos per Dollar, Inflation Adjusted
BPW	Dummy Variable set to 1 for post-9/11 Period, set to 0 for 1974-1990 Dummy Variable for post-9/11 Period interacted with Ciudad Juarez Population
BYC	One Period Lag of Ciudad Juarez Population
BYT	USA Real Gross Domestic Product Mexico Real Exchange Rate Index, Pesos per Dollar, Inflation Adjusted Dummy Variable set to 1 for post-9/11 Period, set to 0 for 1974-1990 Dummy Variable for post-9/11 Period interacted with Mexico Real Exch. Rate
BYW	Mexico Real Exchange Rate Index, Pesos per Dollar, Inflation Adjusted Dummy Variable for post-9/11 Period interacted with Mexico Real Exch. Rate Dummy Variable for post-9/11 Period interacted with USA Real GDP Dummy Variable set to 1 for post-9/11 Period, set to 0 for 1974 - 1990

Table 3 reports summary statistics for the variables listed above. As can be seen, there is substantial variability in the sample. All of the bridge series have 38 annual observations. As noted above, historical data for the air passenger data only extend back to 1979. Commercial international flights to El Paso were phased out in 2006 and have yet to resume. Air freight and air mail data are no longer reported separately. For those reasons, there are fewer historical observations for the air transport series included in the sample.

Table 3. Historical Data Summary Statistics

Variable	Mean	Standard Deviation	Maximum	Minimum	Total Observations
APDD	1,452.9	276.3	1,830.8	913.0	33
APDI	14.168	11.140	46.054	0.106	28
APED	1,483.4	288.4	1,862.6	920.3	33
APEI	10.809	7.898	34.891	0.137	28
AFDT	26.494	17.654	55.600	5.002	35
AFET	21.042	12.624	41.697	5.467	35
AMD	2.535	0.949	4.337	0.739	35
AME	1.490	0.659	2.331	0.046	35
BAC	6.691	1.331	8.802	3.268	38
BAT	0.259	0.131	0.492	0.053	38
BAW	0.660	0.200	1.208	0.403	38
BPC	4.223	0.920	6.039	2.011	38
BPW	5.291	1.188	7.738	3.466	38
BYC	2.389	0.845	3.871	1.166	38
BYT	0.173	0.159	0.386	0.002	38
BYW	0.418	0.353	1.256	0.027	38

Figures 2-6 illustrate the behavior of several of the variables in the sample over time. From those graphs it is easy to see that data in the sample exhibit substantial variability. That is, in part, because the series respond to business cycle developments in both countries as well as currency market fluctuations.

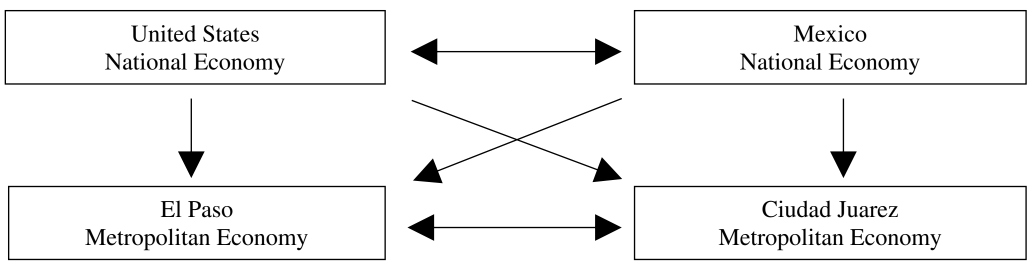

Figure 1 Borderplex Model Design

The structural econometric forecasts are published annually by the University of Texas at El Paso. Parameter re-estimation is carried out for all of the model equations every year once data bank updating is completed. The forecasts are generated for three-year time periods. At the point at which the neural network utilized for this study was developed, that made 27 previously published, ex-ante, transportation forecasts per variable available for analysis during the sample period under consideration. The structural econometric model used to generate those forecasts has been discussed in several prior studies (Fullerton, 2001; 2004). Figure 1 provides an overview of the basic modeling strategy underlying it. Economic conditions in El Paso and Ciudad Juarez are affected by regional business cycles and economic trends, as well as by their national counterparts in Mexico and the United States. The neural network developed for this study has not previously been analyzed. A summary of that model follows.

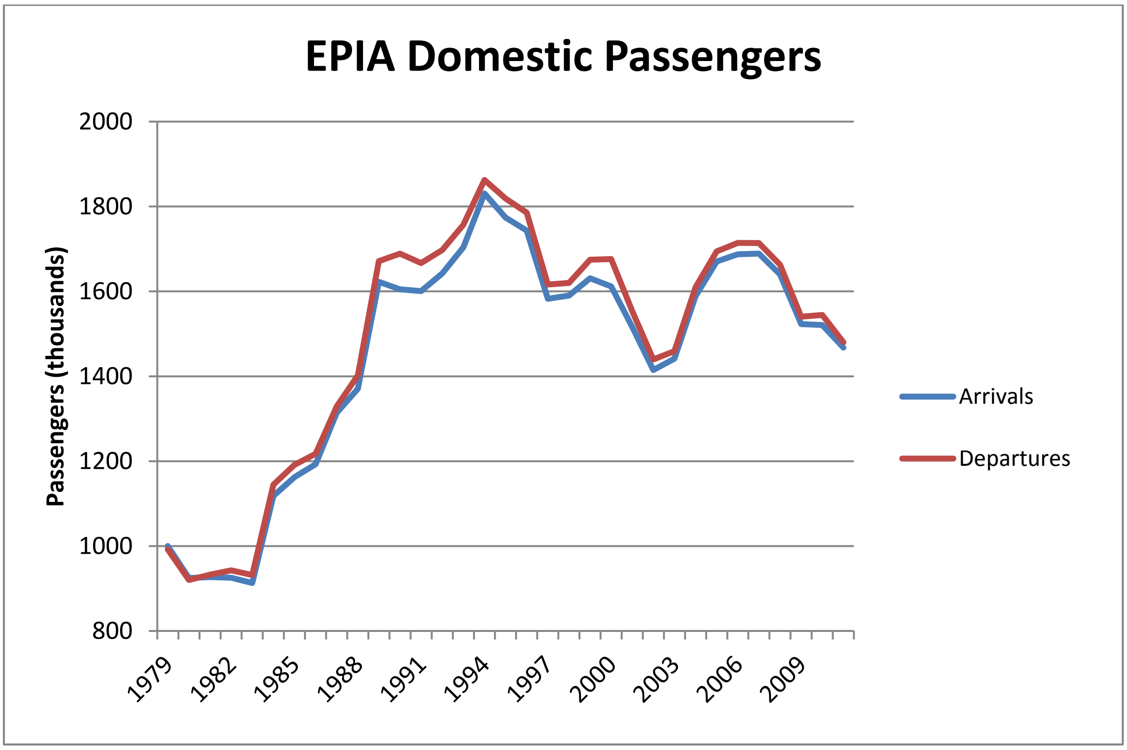

Figure 2 **El Paso International Airport Domestic Flight Passenger Arrivals And Deparutres In Thousands**

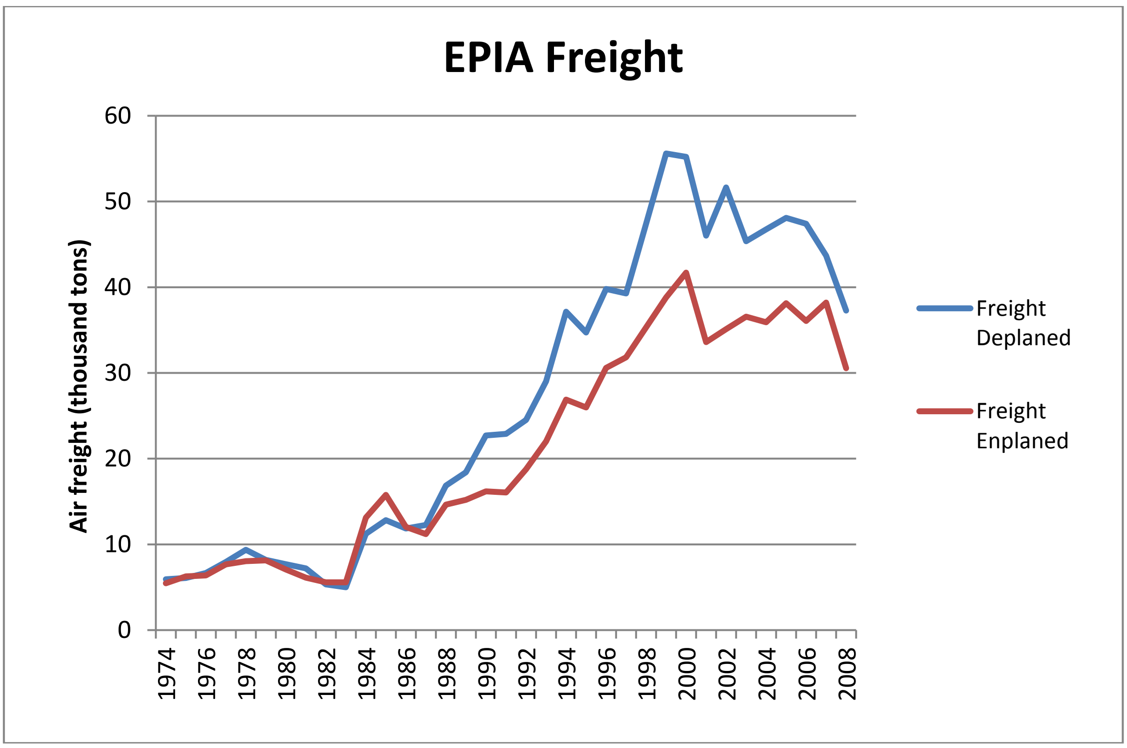

Figure 3. **El Paso International Airport Incoming and Outgoing Freight Volumes in Thousand Tons**

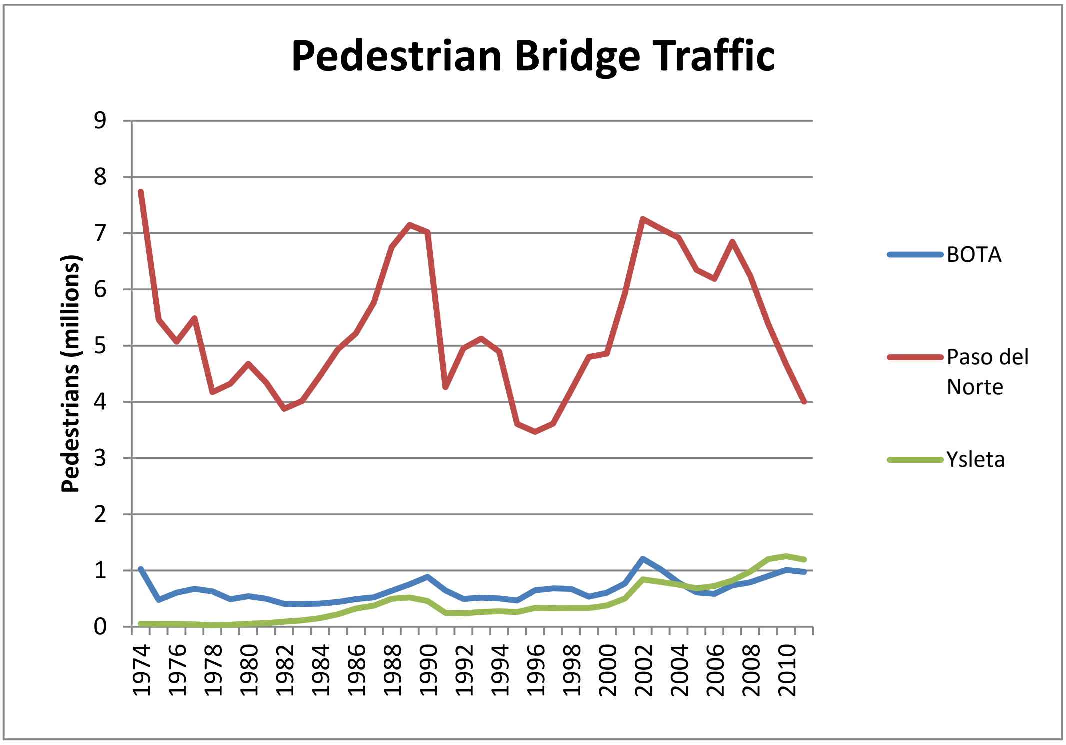

Figure 4. El Paso International Bridge Ports of Entry Pedestrian Flows in Millions of Persons

BOTA – Bridge of the Americas, near central El Paso. Paso del Norte Bridge, near downtown El Paso. Ysleta Bridge, near east El Paso.

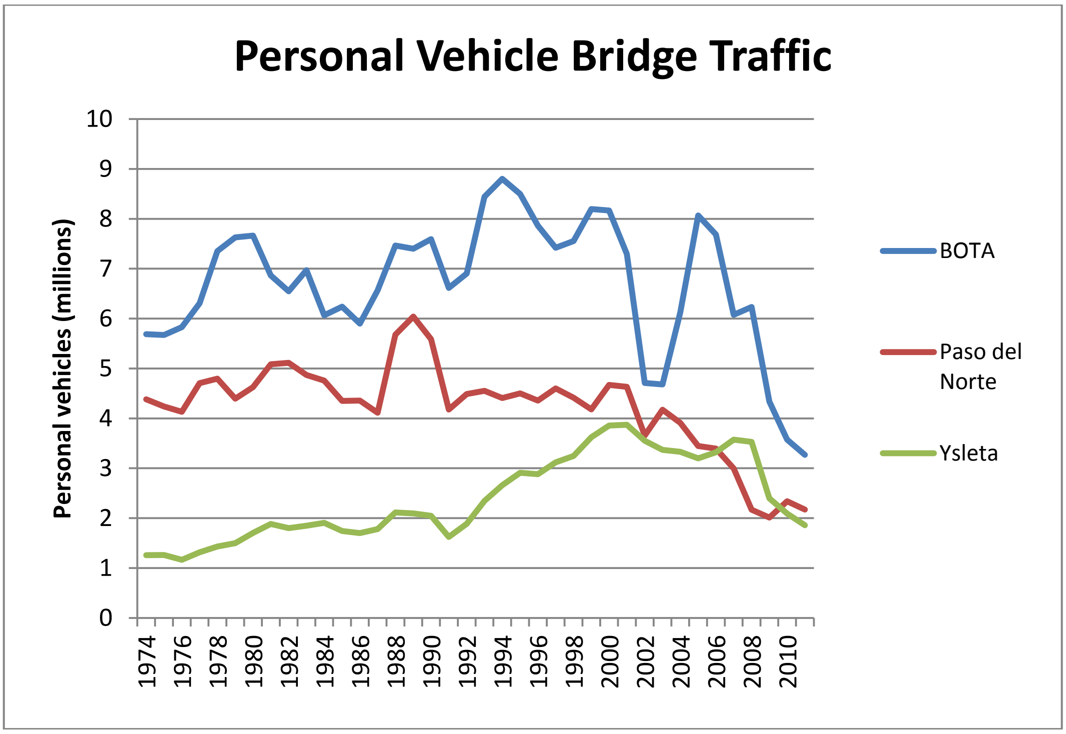

Figure 5. El Paso International Bridge Ports of Entry Personal Vehicle Flows in Millions of Cars

BOTA – Bridge of the Americas, near central El Paso. Paso del Norte Bridge, near downtown El Paso. Ysleta Bridge, near east El Paso.

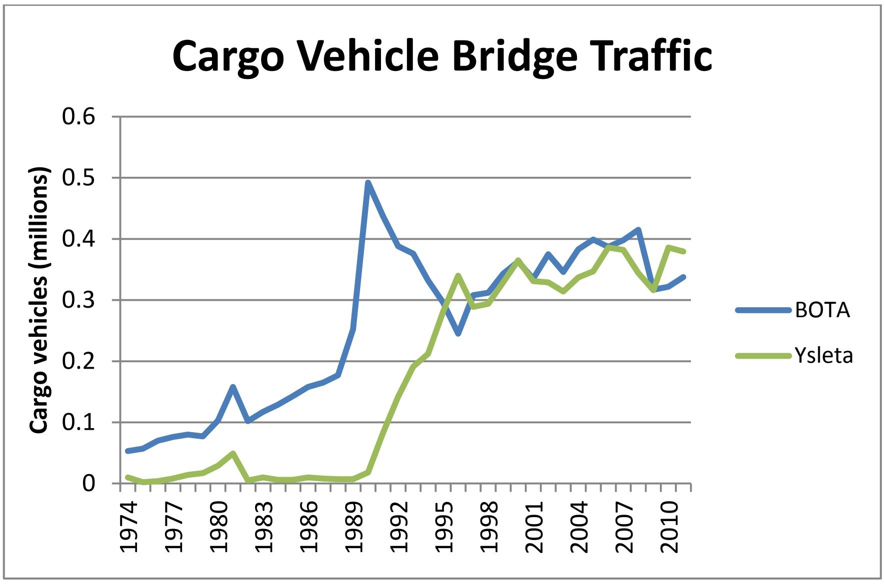

Figure 6. El Paso International Bridge Ports of Entry Cargo Vehicle Flows in Millions of Trucks

BOTA – Bridge of the Americas, near central El Paso. Ysleta Bridge, near east El Paso. Cargo vehicles cannot cross the Paso del Norte Bridge near downtown El Paso.

4. Multi-Layered Perceptron (MLP) NN Model

MLP offers two major advantages over traditional non-pattern-seeking mathematical models. First, MLP is flexible in looking for nonlinear patterns in data. Second, MLP does not require *prior* knowledge of relationships or distributional assumptions about the data. Among other things, MLP forecasting models have been shown to do well in simulating time series data that are subject to business cycle fluctuations (Heravi, Osborn & Birchenhall, 2004). Transportation data, of course, are relatively sensitive to prevailing economic conditions.

Mukhopadhyay (2006) describes the generic topology of an MLP comprising a layer of input nodes, one or more layers of hidden nodes, and a layer of output nodes. First hidden layer nodes connect with input layer nodes. Output layer nodes connect with the last layer of hidden nodes. Connection strengths, called *weights*, are connection values.

The main feature of connectionist NN models reflects the learning mechanism in the brain where knowledge is in the connections of neurons rather than in the neurons themselves. The learning is assumed to be in modifying the connection strengths. Communication among neurons involves either excitatory or inhibitory messages. Mathematical learning algorithms attempt to mimic that. Greater values for the weights between two nodes represent more meaningful relationships between the nodes (excitation). Lower values reflect lesser association between the nodes. For example, if all the estimated weights of an input variable are insignificant in the first layer, the input variable is not significant in the model.

The output of each node in an MLP, called *activation value*, is a function of its inputs from the previous layer and the corresponding weights. The function is called *activation function*. The activation value of an input layer node is the value of the input variable. The activation value of the output layer unit is the estimated value of the dependent variable (target). A training algorithm learns the mathematical relationship between input variables and the target by assigning proper weights to all network connections.

4.1 Training Algorithm

A back-propagation (BP) algorithm based on Rumelhart, Hinton, and Williams (1988) is utilized. The BP training algorithm estimates a target value from input variable values by initially assigning a set of arbitrary weights. The algorithm then compares actual target value with the estimated value. The difference between the actual value and the estimated value is called *error signal*. The training process changes all weights in proportion to the error signal. The constant of proportionality is called the *learning rate*. The method produces no error signal if there is no difference between the actual and the estimated value.

The training method starts changing weights from the top layer connections. The process of updating weights propagates back through the network from top layer to the first layer connections. The larger the learning rate, the larger the corresponding weight change. The process of updating weights repeats over all sample points to complete a full *iteration*. The method computes an error sum of squares value over all sample points upon completing each iteration. Training stops when the error sum of squares value is less than a predefined criterion.

The nonlinear regression equation form of one hidden layered MLP is:

$$y_{t+h} = \beta_{\phi,h} + \sum_{j=1}^{n} \beta_{j,h} \ \mathrm{f}\ (\mathrm{I}_t\ ,\ w_{h,j}) + \mathrm{E}_\mathrm{p} \tag{1}$$

where, h is the length of forecast period. I_t is input vector of current time period value of y_t. $w_{h,j}$ is the network weight vector corresponding to forecast horizon h and the *jth* hidden node. E_p is the measure of the error on input/output pattern p. The logistic form of the activation function *f* at each node is employed:

$$\mathrm{f}\ (\ \mathrm{I}_t\ ,\ w_{h,j}) = (1+\mathrm{e}^{-\mathrm{z}}\ )^{-1} \tag{2}$$

where,

$$\mathrm{z} = w_{h,j,\phi} + w_{h,j} * t \tag{3}$$

and n is the number of hidden nodes. Logistic activation functions (Equations 2 and 3) introduce nonlinearity to the model. Activation functions must be differentiable for usage in the BP training algorithm. A differentiable sigmoid function (Equations 2 and 3) is used to compute activation values for hidden and output layer nodes.

4.2 MLP Network Architecture and Parameter Values

The MLP architecture selection guidelines proposed by Xiang, Ding, and Lee (2005) are the ones followed. That study suggests utilizing a simple, three-layered MLP network. In it, the number of hidden units should match the minimum number of line segments (hyper planes in high dimensional cases) required to approximate the target function for a minimal architecture. Functions learned by a minimal net over calibration sample points work well on new samples. Three network layers are used - one input layer for input variables (time t and a bias unit), one hidden unit layer with three units, and one output layer of one unit. The network connects all hidden nodes with all input nodes. The output node connects to all hidden nodes. Optimal network selection is based on model performance using training file samples.

With high learning rate values, faster learning can be achieved. However, the learning process can jump back and forth along the error surface for learning rates that are too high. This process of high error sum of squares error fluctuation during calibration is called *oscillation*. One way to increase the learning rate without leading to oscillation is to include a *momentum* factor in the weight change formula. This determines the effect of past weight changes on the current direction of movement in the weight space. The addition of momentum factor effectively filters out high frequency variations of the error surface in weight space. By trying out small learning rates and no momentum factor can achieve similar results. However, the learning time will increase for small learning rates. A value of 0.1 is used for the learning rate and 0.9 is employed for the momentum factor as recommended by previous research (Rumelhart, Hinton, & Williams, 1988).

A three period lag is utilized for the neural network analyses. All of the variables employed in the structural econometric modeling system are also used for the MLP neural networks. The process was repeated sequentially by adding one additional year of actual historical variable values prior to developing each successive set of three-year forecasts for the different transportation series being analyzed. That mimics how the data become available every year for econometric model parameter estimation and forecast generation.

5. Empirical Results

Table 4 summarizes the predictive accuracy results for the air transport variables included in the sample. The metrics utilized to gauge forecast accuracy are the root mean squared error (RMSE), the Theil inequality coefficient (also known as the U-statistic), and the second moment decomposition of the U-statistic. The advantage of the U-statistic over RMSE is its scale-free characteristic that avoids unbounded, from above, accuracy metrics. The second moment decomposition of the U-statistic indicates whether forecast errors are due to systematic out-of-sample simulation flaws or, ideally, due to unforeseeable random events (Pindyck & Rubinfeld, 1998; Theil, 1961). As pointed out by Theil (1961), the optimal distribution of the second moment inequality proportions is $U^M = 0$, $U^S = 0$, and $U^C = 1$.

Table 4. Air Series Predictive Accuracy

Series	RMSE	U-Statistic	U-Bias	U-Variance	U-Covariance
El Paso International Airport Domestic Passenger Arrivals					
APDD[1]	**123.7009**	**0.0385**	0.3125	0.0445	0.6430
APDD[2]	132.4120	0.0411	0.3973	0.1925	0.4102
El Paso International Airport International Passenger Arrivals, 1998-2006					
APDI[1]	6.4606	0.2657	0.6705	0.0640	0.2654
APDI[2]	**4.7764**	**0.2158**	0.4774	0.0225	0.5001
El Paso International Airport Domestic Passenger Departures					
APED[1]	**133.9325**	**0.0411**	0.3110	0.0213	0.6678
APED[2]	149.4551	0.0454	0.4994	0.1680	0.3327
El Paso International Airport International Passenger Departures, 1998-2006					
APEI[1]	**5.5429**	**0.2464**	0.6326	0.02297	0.3444
APEI[2]	5.6168	0.2606	0.3543	0.0225	0.6232
El Paso International Airport In-Bound Freight, 1998-2008					
AFDT[1]	8.1135	0.0819	0.2000	0.0057	0.7943
AFDT[2]	**7.4492**	**0.0764**	0.0888	0.0070	0.9042
El Paso International Airport In-Bound Freight, 1998-2008					
AFET[1]	**4.8859**	**0.0648**	0.2592	0.0005	0.7403
AFET[2]	5.5981	0.0732	0.3712	0.0001	0.7693
El Paso International Airport In-Bound Air Mail, 1998-2008					
AMD[1]	**0.9614**	**0.1815**	0.2381	0.0002	0.7618
AMD[2]	1.1568	0.2210	0.1043	0.0026	0.8931
El Paso International Airport Out-Bound Air Mail, 1998-2008					
AME[1]	**0.6874**	**0.2727**	0.2248	0.0009	0.7742
AME[2]	0.7489	0.2931	0.2664	0.0000	0.7336

Note: Sample Simulation Period: 1998 – 2011, unless otherwise noted. **Boldface** type indicates greatest predictive accuracy.

1. Previously published Borderplex structural model forecasts. 2. Multi-Layered Perceptron (MLP) Neural Network forecasts.

Equation (4) shows how the RMSEs are computed. In (4), Y^s is the out-of-sample simulation or forecast value for variable Y, Y^a is the actual historical value for Y, and T is the total number of forecasts for Y.

$$\mathrm{RMSE} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(Y_t^s - Y_t^a)^2} \tag{4}$$

Equation (5) provides the details for calculating the U-statistics. The denominator in (5)

$$U = \frac{\sqrt{\frac{1}{T}\sum_{t=1}^{T}(Y_t^s - Y_t^a)^2}}{\sqrt{\sum_{t=1}^{T}(Y_t^s)^2} + \sqrt{\sum_{t=1}^{T}(Y_t^a)^2}} \tag{5}$$

ensures that the inequality coefficients to vary between 0 and 1. When U = 0, $Y_t^s = Y_t^a$ for all *t*, a perfect forecast has been obtained. At the other extreme, if U = 1, the predictive performance of the model is as bad as it can get (Pindyck & Rubinfeld, 1998). Because it covers a finite range, the inequality statistic is easier to interpret than other accuracy gauges such as the mean absolute percentage error and also avoids the risk of division by zero during severe lulls in economic activity or abnormal disruptions in commerce.

Equation (6) summarizes the formulae for the second moment inequality proportions. In Equation (6), "ρ" is the correlation coefficient for the forecast data and the actual data. Also in Equation (6), "σ_s" is the standard deviation of the prediction values and "σ_a" is the standard deviation of the actual values of the variable being forecast. U^M, U^S, and U^C represent bias, variance, and covariance proportions, respectively, of the second moment of the prediction errors (Theil, 1961). The bias proportion measures the extent to which the average values of the simulated and actual series deviate from each other. It thus provides an indication of systematic error. Optimally, the bias proportion will approach zero. The variance proportion indicates the ability of the model to replicate the degree of variability in the variable of interest. Again, as simulation performance improves, the variance proportion approaches zero. The covariance proportion measures unsystematic error. As noted above, when simulation accuracy improves, the covariance proportion approaches one.

$$U^M = \frac{(\overline{Y}_t^s - \overline{Y}_t^a)^2}{(1/T)\sum_{t=1}^{T}(Y_t^s - Y_t^a)^2}, \quad U^S = \frac{(\sigma_s - \sigma_a)^2}{(1/T)\sum_{t=1}^{T}(Y_t^s - Y_t^a)^2}, \text{ and}$$
$$U^C = \frac{2(1-\rho)\sigma_s\sigma_a}{(1/T)\sum_{t=1}^{T}(Y_t^s - Y_t^a)^2} \tag{6}$$

In Table 4, the overall superiority of the econometric forecasts is underscored by lower RMSE and U-statistics for six of the eight air transportation series. Those variables include domestic passenger arrivals, domestic passenger departures, international passenger departures, out-bound freight, in-bound air mail, and out-bound air mail. The only variables for which the MLP neural network forecasts prove more accurate are international passenger arrivals and in-bound air freight. In all eight cases, replicating series variability is achieved by both methods. For all but the international flight passenger flows, prediction bias is not problematic and the greatest source of forecast error, according to the results obtained, is generally due to random factors. With respect to the international passenger data, the econometric forecasts were consistently biased upward and overly optimistic during the last four years during which those services were offered at El Paso International Airport.

Table 5 reports the forecast accuracy for the international bridge variables. Again, the econometric forecasts achieve greater accuracy than those of the MLP neural network approach. For six of the eight variables, the RMSE and U-statistics are lower for the structural econometric forecasts. Those variables include all northbound pedestrian traffic and all light vehicle traffic across all three international bridges. In contrast, both cargo vehicle traffic series across two of the three bridges are forecast more accurately by the MLP neural network models. Bias does not represent an obstacle for either set of forecasts. In fact, non-random simulation errors account for less than 50 percent of the inaccuracies that are tabulated for any of the eight variables by either method.

Table 5. Bridge Series Predictive Accuracy

Series	RMSE	U-Statistic	U-Bias	U-Variance	U-Covariance
Bridge of the Americas Northbound Light Vehicle Traffic					
BAC[1]	**1.6682**	**0.1270**	0.1538	0.0041	0.8420
BAC[2]	1.8130	0.1385	0.1854	0.5359	0.2787
Bridge of the Americas Northbound Cargo Vehicle Traffic					
BAT[1]	0.0668	0.0876	0.2410	0.1056	0.6533
BAT[2]	**0.0343**	**0.0479**	0.0676	0.6276	0.3048
Bridge of the Americas Northbound Pedestrian Traffic					
BAW[1]	**0.2248**	**0.1368**	0.0150	0.0030	0.6792
BAW[2]	0.2477	0.1582	0.1292	0.2097	0.6611
Paso del Norte Bridge Northbound Light Vehicle Traffic					
BPC[1]	**0.5416**	**0.0744**	0.3108	0.0100	0.6792
BPC[2]	0.6411	0.0876	0.1428	0.2139	0.6433
Paso del Norte Bridge Northbound Pedestrian Traffic					
BPW[1]	1**.0333**	**0.0878**	0.0273	0.0110	0.9617
BPW[2]	1.3278	0.1163	0.1087	0.1738	0.7175
Ysleta-Zaragoza Bridge Northbound Light Vehicle Traffic					
BYC[1]	**0.5824**	**0.0876**	0.0929	0.0839	0.8231
BYC[2]	0.7059	0.1056	0.1049	0.1289	0.7662
Ysleta-Zaragoza Bridge Northbound Cargo Vehicle Traffic					
BYT[1]	0.0399	0.0561	0.0838	0.0089	0.9073
BYT[2]	**0.0386**	**0.0554**	0.0112	0.1054	0.8833
Ysleta-Zaragoza Bridge Northbound Pedestrian Traffic					
BYW[1]	**0.2042**	**0.1252**	0.3554	0.0569	0.5877
BYW[2]	0.2641	0.1599	0.3667	0.0875	0.5457

Note: Sample Simulation Period: 1998 – 2011

Boldface type indicates greatest predictive accuracy.

1. Previously published Borderplex structural model forecasts.

2. Multi-Layered Perceptron (MLP) Neural Network forecasts.

The bridge results contain an interesting dichotomy. Traffic related to commercial activities is most accurately predicted by the structural econometric model. Cargo traffic, a function of industrial economic activities, is most accurately forecast by their respective MLP neural network models.

Although the bulk of the evidence reported in Tables 4 and 5 indicates that the MLP neural net forecasts are relatively less accurate than the econometric model forecasts, it does not rule out the possibility that the two sets of simulations might complement each other. The latter possibility is certainly plausible due to the distinct methodological steps involved. To formally test this proposition, a series of regression equations are estimated following a general approach previously utilized in various different forecasting applications and contexts (Cooper & Nelson, 1975; Granger & Ramanathan, 1984).

In order to calculate the combination coefficients, each composite weights equation models the actual historical values of the individual transportation series as functions of each forecast. In cases where the information contents of the two sets of forecasts are complementary, estimated coefficients for both regressor prediction series will have significant t-statistics associated with them. The regression equation used to test for complementarity is shown in Equation (7). In Equation (7), Y^a_t represents the actual value of Y in period t, Y^e_t stands for the structural econometric forecast of Y in period t, Y^n_t represents the neural network forecast for Y in period t, and U_t is a random disturbance term for period t. Outcomes from that test, using the forecast data from 1998 through 2007, are summarized in Table 6.

$$Y^a_t = C_0 + C_1 Y^e_t + C_2 Y^n_t + U_t \tag{7}$$

Table 6. Neural Network and Econometric Forecast Complement t-Test Outcomes

Series	Reject Null Hypothesis that Forecast Information is not Complementary	Fail to Reject Null Hypothesis that Forecast Information is not Complementary
APDD		Fail to Reject
APDI		Fail to Reject
APED		Fail to Reject
APEI		Fail to Reject
AFDT		Fail to Reject
AFET		Fail to Reject
AMD		Fail to Reject
AME		Fail to Reject
BAC		Fail to Reject
BAT		Fail to Reject
BAW		Fail to Reject
BPC	Reject*	
BPW	Reject**	
BYC		Fail to Reject
BYT	Reject**	
BYW	Reject**	

Note: Sample Simulation Period: 1998-2011.

* Complementary information test passed at 5-percent significance level.

** Complementary information test passed at 1-percent significance level.

In 4 of the 16 cases, the complementary forecast information hypothesis is supported. The 4 variables for which the forecast information provided by each set of predictions is complementary in nature are all northbound bridge traffic categories. They include pedestrian bridge crossers at the downtown Paso del Norte port of entry and the eastside Ysleta port of entry. They also include light vehicles across the downtown Paso del Norte Bridge, as well as cargo vehicles across the eastside Ysleta Bridge. In all four cases, support for the complementarity hypothesis occurs at the 5-percent or 1-percent levels.

No evidence in favor of the complementary information hypothesis is provided by any of the other sets of forecasts. None of the airport variables would obtain improved out-of-sample simulation accuracy by combining the neural network and econometric forecasts. The same outcome is also tallied for all three northbound traffic series forecasts for the Bridge of the Americas near central El Paso. That is also the case for light vehicles coming across the Ysleta Bridge in east El Paso.

Because they can handle data nonlinearities and do not require imposing any distributional assumptions, neural network modeling offers several potential advantages over other quantitative procedures for analyzing traffic flows. In the case of the transportation variables associated with the Borderplex economy of El Paso, Texas and Ciudad Juarez, Mexico, forecasts from an MLP neural network methodology sometimes do achieve better accuracy than previously published econometric forecasts. In most cases, however, the structural econometric method tallies more accurate out-of-sample simulations. For the air and bridge transport variables in this region, the evidence is generally in line with the outcomes obtained by Smith and Demetsky (1997) and differs from the results that favor neural network forecasts reported by Lam, Ng, Seabrooke, and Hui (2004) and Celikoglu and Akad (2005).

6. Conclusion

Prior research has documented that traditional econometric forecasts of both surface and air transportation traffic in border regions are difficult to carry out accurately. One alternative to structural econometric models is provided by neural networks. Neural network forecasts have proven helpful in a variety of different settings, but have not been extensively tested using data for border metropolitan economies. This study carries out such an exercise using transportation data for the Borderplex economy of El Paso, Texas and Ciudad Juarez, Mexico.

The sample period for which the forecasts are considered is 1998-2011. The data frequency is annual. For six of the air transport variables analyzed, the structural econometric forecasts are more accurate than those of the multi-layered perceptron neural network method utilized. For the international bridge traffic series, the econometric forecasts are also relatively more accurate for six of the eight series analyzed. In twelve of the sixteen cases, no evidence of forecast complementarity is uncovered.

Empirical results obtained in this study indicate neural network approaches may not prove more accurate than traditional econometric modeling frameworks. Whether these results are unique to the Borderplex sample used in this effort is not known. Other border regions between the United States and Mexico that offer potentially interesting transportation series to analyze include Brownsville–Matamoros, McAllen–Reynosa, Laredo–Nuevo Laredo, and San Diego–Tijuana. Similar studies for the United States border with Canada might also yield interesting results. Evidence reported above also indicates that, in a few cases, neural network transport forecasts may contain information that complements that embodied in structural econometric out-of-sample simulations.

Acknowledgements

Financial support for this research was provided by El Paso Water Utilities, Hunt Companies, Texas Department of Transportation, JPMorgan Chase Bank of El Paso, UTEP Center for the Study of Western Hemispheric Trade, and a UTEP College of Business Administration Faculty Research Grant. Econometric research assistance was provided by Teodulo Soto, Carlos Morales, Francisco Pallares, Alan Jimenez, and Ashley Molson. Helpful comments and suggestions were provided by Karen Fierro, Leo Gemoets, Enedina Licerio, Angel Molina, Emmanuel Villalobos, Peg Young, and three anonymous referees.

References:

An G. and Puttitanun T. (2009). "Revisiting McCallum's Border puzzle", *Economic Development Quarterly*, Vol. 23, pp. 167-170.

Ashur S., Weismann J., Perez S. and Weismann A. J. (2001). "Traffic simulation at international ports of entry—El Paso-Mexico case study", *Transportation Research Record*, Vol. 1763, pp. 48-56.

Bain R. (2009). "Error and optimism bias in toll road traffic forecasts", *Transportation*, Vol. 36, pp. 469-482.

Bradbury S. L. (2002). "Planning transportation corridors in Post-NAFTA North America", *Journal of the American Planning Association*, Vol. 68, pp. 137-150.

Celikoglu H. B. and Akad M. (2005). "Estimation of public transport trips by feed forward back propogation artificial neural networks: A case study for Istanbul", in: *Proceedings of the 8th Online World Conference on Soft Computing Methodologies & Applications*, pp. 27-36.

Charney A. H. and Taylor C. A. (1984). "Decomposition of Ex–Ante state model forecasting errors", *Journal of Regional Science*, Vol. 24, pp. 229-247.

Cooper J. P. and Nelson C. R. (1975). "The Ex-Ante prediction performance of the St. Louis and FRB-MIT-Penn econometric models and some results on composite predictors", *Journal of Money, Credit, and Banking*, Vol. 7, pp. 1-32.

De Leon M., Fullerton T. M. Jr. and Kelley B. W. (2009). "Tolls, exchange rates, and borderplex international bridge traffic", *International Journal of Transport Economics*, Vol. 36, pp. 223-259.

Dunne S. and Ghosh B. (2012). "Regime-based short-term multivariate traffic condition forecasting algorithm", *Journal of Transportation Engineering*, Vol. 138, pp. 455-466.

Figliozzi M. A., Harrison R. and Mccray J. P. (2001). "Estimating Texas-Mexico North American free trade agreement truck volumes", *Transportation Research Record*, Vol. 1763, pp. 42-47.

Fildes R., Wei Y. and Ismail S. (2011). "Evaluating the forecasting performance of econometric models of air passenger traffic flows using multiple error measures", *International Journal of Forecasting*, Vol. 27, pp. 902-922.

Flyvbjerg B., Holm M. K. S. and Buhl S. L. (2005). "How (in)accurate are demand forecasts in public works projects? The case of transportation", *Journal of the American Planning Association*, Vol. 71, pp. 131-146.

Flyvbjerg B., Holm M. K. S. and Buhl S. L. (2006). "Inaccuracy in traffic forecasts", *Transport Reviews*, Vol. 26, pp. 1-24.

Fullerton T. M. Jr. (2000). "Currency movements and international border crossings", *International Journal of Public Administration*, Vol. 23, pp. 1113-1123.

Fullerton T. M. Jr. (2001). "Specification of a borderplex econometric forecasting model", *International Regional Science Review*, Vol. 24, pp. 245-260.

Fullerton T. M. Jr. (2004). "Borderplex Bridge and air econometric forecast accuracy", *Journal of Transportation and Statistics*, Vol. 7, pp. 7-21.

Fullerton T. M. Jr. (2007). "Empirical evidence regarding 9/11 impacts on the borderplex economy", *Regional & Sectoral Economic Studies*, Vol. 7, July-December, pp. 51-64.

Fullerton T. M. Jr. and R. Tinajero (2002). "Cross border cargo vehicle flows", *International Journal of Transport Economics*, Vol. 29, pp. 201-213.

Granger C. W. J. and Ramanathan R. (1984). "Improved methods of combining forecasts", *Journal of Forecasting*, Vol. 3, pp. 197-204.

Harrison R., Sanchez-Ruiz L. A. and Lee C. E. (1998). "Truck traffic crossing Texas-Mexico border", *Transportation Research Record*, Vol. 1643, pp. 136-142.

Heravi S., Osborn D. R. and Birchenhall C. R. (2004). "Linear versus neural network forecasts for European industrial production series", *International Journal of Forecasting*, Vol. 20, pp. 435-446.

Jagric T. and Strasek S. (2005). "A nonlinear extension of the NBER model for short-run forecasting of business cycles", *South African Journal of Economics*, Vol. 73, pp. 435-448.

Lam W. H. K., Ng P. L. P., Seabrooke W. and Hui E .C. M. (2004). "Forecasts and reliability analysis of Port Cargo throughput in Hong Kong", *Journal of Urban Planning and Development*, Vol. 130, pp. 133-144.

Levinson D. (2005). "Paying for the fixed costs of roads", *Journal of Transport Economics and Policy*, Vol. 39, pp. 279-294.

McCallum J. (1995). "National borders matter—Canada-US regional trade patterns", *American Economic Review*, Vol. 85, pp. 615-623.

McCray J. P. (1998). "North American free trade truck highway corridors—US-Mexican truck rivers of trade", *Transportation Research Record*, Vol. 1613, pp. 71-78.

Mukhopadhyay S. (2006). "Predicting global diffusion of the internet: An alternative to diffusion models", *Communications of the Association for Information Systems* (January).

Nozick L. K. (1996). "Trade between the United States and Mexico", *Transportation Quarterly*, Vol. 50, pp. 111-133.

Pagan A. R. (1974). "A generalised approach to the treatment of autocorrelation", *Australian Economic Papers*, Vol. 13, pp. 267-280.

Pindyck R. S. and Rubinfeld D. L. (1998). *Econometric Models and Economic Forecasts* (4th ed.), Irwin McGraw-Hill: New York, NY.

Rumelhart D. E., Hinton G. E. and Williams R. J. (1988). "Learning internal representations by error propagation", in: *Parallel Distributed Processing Explorations in the Microstructure of Cognition*, Volume 1, MIT Press: Cambridge, MA.

Saintgermain M. A. (1995). "Problems and opportunities for cooperation among public managers on the U.S.-Mexico border", *American Review of Public Administration*, Vol. 25, pp. 93-117.

Smith B. L. and Demetsky M. J. (1997). "Traffic flow forecasting: Comparison of modeling approaches", *Journal of Transportation Engineering*, Vol. 123, pp. 261-266.

Stank T. P. and Crum. M. R. (1997). "Just-in-time management and transportation service performance in a cross-border setting", *Transportation Journal*, Vol. 36, pp. 31-42.

Taylor J. W. (2010). "Exponentially weighted methods for forecasting intraday time series with multiple seasonal cycles", *International Journal of Forecasting*, Vol. 26, pp. 627-646.

Taylor J. C., Robideaux J. R. and Jackson G. C. (2004). "US-Canada transportation and logistics: Border impacts and costs, causes, and possible solutions", *Transportation Journal*, Vol. 43, pp. 5-21.

Theil H. (1961). *Economic Forecasts and Policy* (2nd ed.), North Holland Publishing Company: Amsterdam, NE.

Tsekeris T. and C. Tsekeris (2011). "Demand forecasting in transport: Overview and modeling advances", *Ekonomska Istrazivanja – Economic Research*, Vol. 24, pp. 82-94.

Villa J. C. (2006). "Status of the US-Mexico commercial border crossing process—Analysis of recent studies and research", *Transportation Research Record*, Vol. 1966, pp. 10-15.

Vlahogianni E. I., Golias J. C. and Karlaftis M. G. (2004). "Short-term traffic forecasting: Overview of objectives and methods", *Transport Reviews*, Vol. 24, pp. 533-557.

Wei Y. and Chen M. C. (2012). "Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks", *Transportation Research Part C – Emerging Technologies*, Vol. 21, pp. 148-162.

West C. T. (2003). "Structural Regional Factors that Determine Absolute and Relative Accuracy of U.S. Regional Labor Market Forecasts", *Journal of Agricultural & Applied Economics*, Vol. 35, Supplement, pp. 121-135.

Xiang C., Ding S. Q. and Lee T. H. (2005). "Geometrical interpretation and architecture selection of MLP", *IEEE Transactions on Neural Networks*, Vol. 16, pp. 84-96.