

ECONOMETRIC VERSUS NEURAL NETWORK TRANSPORTATION FORECASTS

***Thomas M. Fullerton, Jr.* , Somnath Mukhopadhyay** and
Adam G. Walke******

Abstract: This analysis compares the performance of an econometric model, a neural network, and random walk benchmarks in forecasting international bridge and airport traffic in El Paso, Texas, a major gateway for trade between the United States and Mexico. For 7 of the 16 variables analyzed, the econometric forecasts outperform the alternatives as judged by forecast error summary statistics. The neural network forecasts are more accurate than the competing models in four cases and the random walk is the best alternative in five cases. To provide a more comprehensive assessment of predictive accuracy, the forecast errors are decomposed into proportions of inequality. Systematic forecast error is problematic for several of the econometric and neural network forecasts. Finally, a statistical test of directional accuracy is applied. The results of that test largely corroborate the relative accuracy of the econometric model.

Keywords: Transportation Forecasts, Econometric Models, Neural Networks, Directional Accuracy

1. INTRODUCTION

Decision-makers use traffic forecasts in a variety of contexts. Traffic forecasts facilitate operational and budgetary planning activities for airport and tollway administration. Estimates of traffic flows are also critical inputs to the cost-benefit analyses and environmental impact assessments that are frequently employed for transportation infrastructure planning (Nicolaisen and Driscoll, 2014). Overestimation of demand can lead to infrastructure overbuilding and revenue shortfalls while underestimation can result in underinvestment in and excessive traffic congestion. Traffic forecasts sometimes suffer from large inaccuracies (Antoniou, Psarianos, and Brilon, 2011). Because predictive inaccuracies impose real costs for public and private institutions, much effort has been devoted to improving transportation modeling and forecasting.

In some time-series data, biologically motivated pattern-seeking algorithms perform better than traditional econometric models for prediction. As a means toward potentially improving transportation modeling and forecasting forecasts from econometric models can be compared with those from pattern-seeking methods such as neural networks that attempt to mimic multi-

* Department of Economics & Finance, University of Texas at El Paso, El Paso, TX 79968, E-mail tomf@utep.edu

** Department of Marketing & Management, University of Texas at El Paso, El Paso, TX 79968, E-mail smukhopadhyay@utep.edu

*** Department of Economics & Finance, University of Texas at El Paso, El Paso, TX 79968, E-mail agwalke@utep.edu

layered perceptron (MLP) brain-like learning. MLP offers two major advantages over traditional non-pattern-seeking mathematical models (Fullerton and Mukhopadhyay, 2013). First, MLP is flexible and allows for nonlinear patterns in data. Second, MLP does not require *prior* knowledge of relationships or distributional assumptions about the data. However, researchers sometimes over-calibrate MLP models, a step that can result in poor predictive performance. In this study, care is taken so that the MLP models are calibrated in a manner that helps ensure accurate MLP forecasts. Details are described below.

This study evaluates the accuracy of econometric and artificial neural network forecasts for 16 categories of international bridge and airport traffic in El Paso, Texas. El Paso is located on an important trade corridor linking export processing plants in neighboring Ciudad Juárez, Mexico with twin plants in the United States. Both airport and bridge traffic in El Paso are closely linked with international manufacturing activity in the region. The vibrant cross-border trade in merchandise results in large volumes of cargo truck traffic across the international bridges (Figliozi, Harrison, and McCray, 2001). Many cross-border shoppers and commuters also traverse the international bridges on a regular basis (Fullerton, Jimenez, and Walke, 2015). Finally, much of the cargo and passenger traffic that passes through the El Paso International Airport is connected in some way with the border region trade and industrial complex. Given the importance of cross-border supply chain linkages, traffic forecasts in this region have notable implications for international trade and commercial flows.

The next section briefly summarizes the literature on the comparative predictive accuracy of econometric and neural network forecasts. The following section describes the models used for this study and the variables that are forecasted. That is followed by predictive accuracy results and a concluding section.

2. LITERATURE REVIEW

Some prior studies compare neural network forecasts of economic variables against econometric model predictions. In an analysis of consumer expenditure forecasts, Church and Curram (1996) find that neural networks are competitive with traditional econometric models, but do not outperform those models. The choice of relevant explanatory variables is found to be critical to the performance of both methodologies. Hann and Steurer (1996) report that a neural network outperforms a linear regression model of the exchange rate between the United States dollar and the German deutschemark when that variable is measured at a weekly frequency. When monthly data are used instead of weekly data, the two forecasting approaches yield similar results in terms of accuracy. It is noteworthy that the monthly-frequency regression models include more economic predictor variables than the weekly-frequency alternatives due to constraints on the availability of high-frequency economic data.

Within the literature on air and bridge traffic forecasting, several studies assess the simulation performance of econometric models in comparison with alternative methodologies. Fullerton (2004) evaluates *ex-ante* econometric forecasts of airport and international bridge traffic in El Paso. The econometric model exhibits a mixed forecast accuracy performance compared to random walk benchmark projections. In a study comparing various econometric and benchmark forecasts, Fildes, Wei and Ismail (2011) find that an econometric model,

specifically an autoregressive distributed lag model, generates comparatively accurate forecasts of air passenger traffic between the United Kingdom and five other countries.

Other studies apply artificial neural networks to transportation forecasting (see, for example, Faraway and Chatfield, 1998; Wei and Chen, 2012; Chen *et al.*, 2012). Xiao *et al.* (2015) develop a hybrid forecasting approach in which the trend component of airport passenger traffic is predicted using a generalized regression neural network and seasonal oscillations are forecast using radial basis function networks. The model performs relatively well in comparison to several alternatives, as judged by forecast error summary statistics. Vlahogianni and Karlaftis (2013) find that neural network predictions of vehicle speeds on a tolled motorway generally outperform predictions generated using classical statistical approaches. Overall, the results of these studies suggest that econometric models and neural networks are both potentially useful for predicting air and roadway traffic.

This study represents one of the few attempts to directly compare econometric and neural network forecasts of transportation demand. Another distinguishing feature of the analysis is the use of directional accuracy metrics in addition to forecast error summary statistics to evaluate the quality of forecasts. Directional accuracy refers to the capacity to correctly predict increases or decreases in the variable of interest. Different accuracy measures sometimes yield divergent conclusions regarding the forecasting track records of competing models. Using multiple forecast accuracy measures provides a more balanced assessment of overall predictive performance (Fildes, Wei, and Ismail, 2011). Previous research finds that neural network forecasts are sometimes better than linear models at predicting directional changes, even when the latter have smaller average errors (Heravi, Osborn, and Birchenhall, 2004). The evaluation framework described in the next section includes measures of directional accuracy as well as metrics based on the size of forecast errors.

3. DATA AND METHODOLOGY

The 16 variables analyzed in this study measure different components of air and bridge traffic. The eight international bridge traffic variables include pedestrian and light vehicle bridge crossings at the three ports of entry connecting El Paso with Ciudad Juárez, along with cargo vehicle crossings at two of the bridges. Cargo vehicles do not, for the most part, use the downtown international bridges. The eight air traffic variables include the number of passengers on domestic and international flights, the volume of freight transported by plane, and the volume of air mail shipments. The air traffic data are collected from El Paso International Airport and are disaggregated into arrivals and departures.

Table 1A provides information on the sample period and units of measure for the forecasted variables. The sample periods shown in Table 1 are for historical data rather than forecasts. Six of the series were discontinued during the timeframe covered in this analysis. First, the airport suspended international flights in 2006 so only data on domestic flights are available after that point. Second, separate datasets on air freight and air mail were discontinued after 2007. The forecast sample used for this effort begins in 1998 and ends in the last year for which historical data are available. Table 1A also shows summary statistics for the full historical sample period. Annual average traffic across all three bridges during the full historical period

(1974-2015), amounted to nearly 12.9 million light vehicle crossings, 6.4 million pedestrian crossings, and 461,000 cargo truck crossings. Domestic passenger flows from 1979 to 2015 averaged around 1.4 million departures and a nearly equal number of arrivals each year. Finally, air freight and mail deplanements exceeded 28,000 tons per year during the 1974-2007 sample period, on average, while enplanements surpassed 22,000 tons.

Table 1A
Variable definitions and descriptive statistics

<i>Variable name</i>	<i>Sample period</i>	<i>Units</i>	<i>Mean</i>	<i>Standard deviation</i>
<i>Bridge of the Americas</i>				
Light Vehicles	1974-2015	Millions	6.400	1.558
Cargo Vehicles	1974-2015	Millions	0.268	0.130
Pedestrians	1974-2015	Millions	0.683	0.203
<i>Paso del Norte Bridge</i>				
Light Vehicles	1974-2015	Millions	4.056	1.021
Pedestrians	1974-2015	Millions	5.210	1.159
<i>Ysleta Bridge</i>				
Light Vehicles	1974-2015	Millions	2.439	0.831
Cargo Vehicles	1974-2015	Millions	0.193	0.165
Pedestrians	1974-2015	Millions	0.478	0.385
<i>El Paso International Airport</i>				
Domestic Passenger Arrivals	1979-2015	Thousands	1,446.584	261.280
International Passenger Arrivals	1979-2006	Thousands	14.164	11.140
Domestic Passenger Departures	1979-2015	Thousands	1,475.036	273.204
International Passenger Departures	1979-2006	Thousands	10.809	7.898
In-Bound Air Freight Shipments	1974-2007	1,000 Tons	26.178	17.819
Out-Bound Air Freight Shipments	1974-2007	1,000 Tons	20.762	12.703
In-Bound Air Mail Shipments	1974-2007	1,000 Tons	2.587	0.912
Out-Bound Air Mail Shipments	1974-2007	1,000 Tons	1.532	0.619

Table 1B reports additional descriptive statistics for the sample data. With the exception of three variables (international passenger arrivals, international passenger departures, out-bound air shipments), the distributions of the data are fairly symmetric. With the exception of the same three variables (international passenger arrivals, international passenger departures, out-bound air shipments), the data are also platykurtic. The three series that depart from those patterns have all been discontinued. Commercial international passenger flights are no longer scheduled in or out of El Paso International Airport. Air mail data are now reported as part of a broader airmail and freight category.

Each variable is forecast using a structural econometric model and an artificial neural network. The regional econometric model uses a variety of indicators to explain movements in bridge and air traffic as described in Fullerton (2004). Explanatory variables that appear in one

Table 1b
Variable definitions and descriptive statistics

<i>Variable name</i>	<i>Maximum</i>	<i>Minimum</i>	<i>Skewness</i>	<i>Kurtosis</i>
<i>Bridge of the Americas</i>				
Light Vehicles	8.802	3.268	-0.638	2.442
Cargo Vehicles	0.497	0.053	-0.270	1.812
Pedestrians	1.208	0.403	0.623	2.555
<i>Paso del Norte Bridge</i>				
Light Vehicles	6.039	2.011	-0.616	2.779
Pedestrians	7.738	3.466	0.564	2.232
<i>Ysleta Bridge</i>				
Light Vehicles	3.871	1.166	0.216	1.620
Cargo Vehicles	0.438	0.002	-0.037	1.192
Pedestrians	1.256	0.027	0.642	2.133
<i>El Paso International Airport</i>				
Domestic Passenger Arrivals	1,830.811	913.023	-0.857	2.742
International Passenger Arrivals	46.054	0.106	1.316	4.602
Domestic Passenger Departures	1,862.582	920.268	-0.833	2.680
International Passenger Departures	34.891	0.137	1.129	4.790
In-Bound Air Freight Shipments	55.600	5.002	0.267	1.417
Out-Bound Air Freight Shipments	41.697	5.467	0.253	1.416
In-Bound Air Mail Shipments	4.337	0.739	0.147	2.186
Out-Bound Air Mail Shipments	2.331	0.051	-1.205	3.569

or more of the international bridge traffic equations include United States gross domestic output, local population levels on both sides of the border, and a real peso-to-dollar exchange rate index. One of the innovative features of the model is the incorporation of cross-border variables, which reflects the interdependence between the economies of El Paso, Texas, and neighboring Ciudad Juárez, Mexico. Air traffic explanatory variables include wages, output, price indices, and aggregate transportation expenditures. Binary variables for the period following the 11 September 2001 terrorist attacks are included in several of the equations due to traffic flow disruptions that occurred in the aftermath of those events (Walke and Fullerton, 2014).

The econometric model has been used to generate published forecasts since 1998 (Fullerton, 2001). The parameters of the model are re-estimated each year and a new set of published *ex-ante* forecasts is then produced. Each set of published forecasts covers a three year timeframe. The 16 indicators listed in Table 1 are among the variables forecasted. Pooling all of the forecasts together results in a sample size of 51 forecast observations for 10 of the variables. Due to discontinuation of international flights to and from El Paso in 2006, the in-bound and out-bound international passenger series only have 24 forecast observations. Also, because the air freight and mail variables were merged for record-keeping purposes in 2008, there are only 30 forecast observations each for those variables.

To ensure the best forecasts from the MLP model, the artificial intelligence (AI) pattern-seeking software -Statistical Analytical Systems (SAS) is employed. The SAS software generates

three randomly created data sets - calibration, validation, and test. The software finalizes the calibration models when it achieves the best forecast accuracy on the validation model without going to test file.

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 are connected with input layer nodes. Output layer nodes are connected with the last hidden layer nodes. Connection strengths, called *weights*, are connection values. The output of each node in an MLP, called an *activation value*, is a function of its inputs from the previous layer and the corresponding weights. That function is called an *activation function*. An activation value of an input layer node is the value of the input variable. An 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.

The back-propagation (BP) training algorithm (Rumelhart et al., 1986) is used to develop the MLP. The BP training algorithm estimates an output value from input variable values by assigning a set of initial weights (connection strengths). The algorithm compares actual output values with estimated values. The difference between an actual value and an estimated value is called *error*. The training algorithm changes all weights in proportion to error magnitudes. 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. If there is a difference, the training method starts changing weights from the top layer connections. The process of updating weights propagates back through the network from the 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 in the calibration data to complete a full *iteration*. The method computes a sum of squared errors value over all sample points upon completion of each iteration. Training stops when the sum of squared errors value is less than a predefined low value.

The nonlinear regression equation form (Fullerton and Mukhopadhyay, 2013) of one hidden layered MLP is:

$$\hat{y}_{t+h} = \hat{\beta}_{\phi,h} + \sum_{j=1}^n \hat{\beta}_{j,h} f(I_t, \hat{w}_{h,j}) \quad (1)$$

where, h is forecast horizon. I_t is input vector of current time period value of $\hat{Y}_t$. $\hat{w}_{h,j}$ is the network weight vector corresponding to forecast horizon h and j th hidden node. A logistic form of the activation function f is used at each node:

$$f(I_t, \hat{w}_{h,j}) = (1 + e^{-z})^{-1} \quad (2)$$

where,

$$z = \hat{w}_{h,j,\phi} + \hat{w}_{h,j} * t \quad (3)$$

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

MLP Network Architecture and Parameter Values

The guidelines followed are those proposed by a prior study on MLP architecture selection that has been shown to be useful (Xiang *et al.*, 2005). The study suggests starting with a simple, three-layered MLP network with only hidden layer. The number of hidden units should match the minimum number of line segments (or hyper-sigmoid surfaces in high dimensional cases) required to approximate the target function for a minimal architecture. Functions learned using a minimal net over calibration sample points work well on new samples. Three network layers are employed: one input layer for input variables (time t and a bias unit), one hidden unit layer, 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. With high learning rate values, faster learning can be achieved. However, the learning process can jump back and forth on the error surface for high learning rates. This process of high error sum of squares 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. Similar to Rumelhart *et al.* (1986), 0.1 is used for the learning rate and 0.9 for the momentum factor.

When *ex-ante* econometric forecasts are generated each year, recent historical values of the explanatory variables must often be estimated due to time lags in the publication of economic and demographic data series. To avoid biasing the results in favor of *ex-post* simulations estimated with revised and complete historical data, a three-year time lag is used to develop the neural network forecasts. For each set of previously published econometric forecasts, a corresponding set of neural network forecasts is developed. Finally, random walk benchmark projections are also developed for the same time periods. The three sets of forecasts are compared against each other to assess relative predictive performance.

Several approaches are employed to compare the accuracy of the three sets of forecasts for each variable. The U -statistic measures the average size of forecast errors. It is obtained by scaling the root mean squared error such that all possible values of the statistic lie within the interval between zero and one. Lower values of the U -statistic indicate greater forecast accuracy. The statistic is shown in Equation (4), where F denotes the forecast of a time series, A denotes the value of the same series that is actually observed, and T denotes the total number of time periods forecasted.

$$U = \sqrt{\frac{1}{T} \sum_{t=1}^T (F_t - A_t)^2} / \left(\sqrt{\frac{1}{T} \sum_{t=1}^T (F_t)^2} + \sqrt{\frac{1}{T} \sum_{t=1}^T (A_t)^2} \right) \quad (4)$$

In addition to the average size of forecast errors, it is often useful to have information regarding the composition of those errors. The mean squared forecast error can be decomposed into three parts, known as the proportions of inequality. The first of these is the bias proportion,

denoted *U-bias*, which is calculated as shown in Equation (5). The bias proportion measures the deviation of the mean forecast value from the mean of the actual observed series. The second component is the variance proportion, denoted *U-var*. As shown in Equation (6), the variance proportion measures the deviation of the forecast standard deviation, σ_F , from the actual standard deviation, σ_A . The third component of the mean squared forecast error, which measures random or unsystematic error, is the covariance proportion, denoted *U-cov*. This is shown in Equation (7), where ρ is the correlation coefficient for F and A . If there is some nonzero forecast error, then the ideal distribution of the error is *U-bias* = *U-var* = 0 and *U-cov* = 1 (Pindyck and Rubinfeld, 1998).

$$U - bias = (\bar{F} - \bar{A})^2 / \left(\frac{1}{T} \right) \sum_{t=1}^T (F_t - A_t)^2 \quad (5)$$

$$U - var. = (\sigma_F - \sigma_A)^2 / \left(\frac{1}{T} \right) \sum_{t=1}^T (F_t - A_t)^2 \quad (6)$$

$$U - cov. = 2(1 - \rho) \sigma_F \sigma_A / \left(\frac{1}{T} \right) \sum_{t=1}^T (F_t - A_t)^2 \quad (7)$$

In addition to the size and composition of forecast errors, another important dimension of overall forecasting performance is directional accuracy. In a study of airport traffic forecasting, Xiao *et al.* (2015) measure directional accuracy as the percentage of directional changes (increases versus decreases in the variable of interest) that are correctly predicted. In this study, directional changes are defined with respect to the value of each variable in the previous year, which is equal to the random walk forecast. For the econometric and neural network forecasts, a statistical test of directional predictive accuracy is carried out (Pesaran and Timmermann, 1992; 1994). The null hypothesis of the test is that predicted directional changes are independent of actual directional changes. If the null hypothesis can be rejected using a one-sided normal test, that provides evidence in favor of the hypothesis that the forecasts contribute useful information for predicting the direction of change (Granger and Pesaran, 2000).

4. EMPIRICAL RESULTS

Tables 2 and 3 report summary statistics for the econometric (EC), neural network (NN), and random walk (RW) forecast errors. For each variable, the three sets of forecasts are ranked from one to three in terms of predictive accuracy, as measured by the *U*-statistic. The rankings are shown in the third column of each table. The econometric forecasts rank first for seven out of 16 variables, while the neural network projections are more accurate than the alternatives in four cases and the random walks provide the best alternatives in five cases.

Among the bridge traffic variables (Table 2), the econometric forecasts outperform the alternatives in five cases, the neural network models rank first in two cases and the random walk is superior in only one case. The neural network performs best in forecasting the two

cargo vehicle traffic series, while the econometric model outperforms the neural network in forecasting pedestrian and light vehicle traffic. These results are similar to those documented in Fullerton and Mukhopadhyay (2013). The random walk forecast errors have higher covariance proportions than either of the alternatives in six out of eight traffic categories. The two exceptions to this pattern are pedestrian and cargo truck traffic at the Ysleta Bridge. In both of the latter categories, the econometric forecasts have the highest covariance proportion of the three alternatives. High covariance proportions indicate that the forecast error is dominated by random components rather than systematic deviations between the forecasts and the actual historical series.

Table 2
Bridge traffic forecast error metrics

<i>Variable</i>	<i>Method</i>	<i>Rank</i>	<i>U-Stat.</i>	<i>U-Bias</i>	<i>U-Var.</i>	<i>U-Cov.</i>
Bridge of the Americas	EC	1	0.1240	0.1435	0.0005	0.8561
Light Vehicles	NN	3	0.1527	0.3066	0.0004	0.6929
	RW	2	0.1450	0.0734	0.0003	0.9263
Bridge of the Americas	EC	3	0.0924	0.1156	0.0067	0.8777
Cargo Vehicles	NN	1	0.0681	0.0172	0.7884	0.1944
	RW	2	0.0788	0.0210	0.0643	0.9147
Bridge of the Americas	EC	1	0.1200	0.0270	0.0024	0.9706
Pedestrians	NN	2	0.1354	0.0524	0.0173	0.9303
	RW	3	0.1465	0.0057	0.0003	0.9940
Paso del Norte Bridge	EC	1	0.0727	0.1707	0.0098	0.8195
Light Vehicles	NN	3	0.1226	0.0096	0.4699	0.5205
	RW	2	0.0835	0.1504	0.0223	0.8272
Paso del Norte Bridge	EC	1	0.0829	0.0299	0.0056	0.9645
Pedestrians	NN	3	0.1351	0.0205	0.0006	0.9789
	RW	2	0.0968	0.0085	0.0117	0.9797
Ysleta Bridge	EC	2	0.1057	0.0024	0.0018	0.9958
Light Vehicles	NN	3	0.1185	0.3421	0.0455	0.6123
	RW	1	0.1033	0.0001	0.0000	0.9999
Ysleta Bridge Cargo	EC	2	0.0757	0.0027	0.0904	0.9070
Vehicles	NN	1	0.0694	0.1911	0.4331	0.3758
	RW	3	0.0852	0.1784	0.1077	0.7139
Ysleta Bridge Pedestrians	EC	1	0.1043	0.1272	0.0119	0.8608
	NN	3	0.1558	0.1024	0.1470	0.7506
	RW	2	0.1240	0.1642	0.0198	0.8160

Note: EC, NN, and RW stand for the econometric model, neural network, and random walk, respectively. Accuracy rankings are based on *U*-statistics. *U-Bias*, *U-Var.*, and *U-Cov.* represent the bias, variance, and covariance proportions of inequality.

Forecast error analyses for the eight categories of air traffic are summarized in Table 3. The random walk forecasts are relatively accurate in four of the eight cases while the econometric and neural network forecasts are more accurate in two cases each. The econometric model outperforms the other methods in forecasting air mail shipments while the neural network is more accurate for international air passenger arrivals and in-bound air freight shipments. In the other four passenger and freight categories, the random walk performs best, followed by

the econometric model, and then the neural network. The random walk forecasts also have higher covariance proportions than either of the alternatives in five of eight cases. For the three remaining categories, the neural network forecasts have the highest covariance proportions. Those categories are in-bound air mail shipments and international passenger arrivals and departures.

Table 3
Air traffic forecast error metrics

<i>Variable</i>	<i>Method</i>	<i>Rank</i>	<i>U-Stat.</i>	<i>U-Bias</i>	<i>U-Var.</i>	<i>U-Cov.</i>
Domestic Air Passenger Arrivals	EC	2	0.0380	0.3188	0.0490	0.6323
	NN	3	0.0491	0.5290	0.2325	0.2385
	RW	1	0.0363	0.0543	0.0126	0.9330
International Air Passenger Arrivals	EC	3	0.2657	0.6705	0.0640	0.2654
	NN	1	0.2158	0.4774	0.0225	0.5001
	RW	2	0.2191	0.7236	0.0000	0.2764
Domestic Air Passenger Departures	EC	2	0.0384	0.8710	0.0305	0.0985
	NN	3	0.0529	0.5799	0.2163	0.2037
	RW	1	0.0376	0.0540	0.0066	0.9394
International Air Passenger Departures	EC	2	0.2464	0.6326	0.0230	0.3444
	NN	3	0.2606	0.3543	0.0225	0.6232
	RW	1	0.2172	0.6821	0.0003	0.3176
In-Bound Air Freight Shipments	EC	3	0.0819	0.2000	0.0057	0.7943
	NN	1	0.0764	0.0888	0.0070	0.9042
	RW	2	0.0814	0.0000	0.0017	0.9983
Out-Bound Air Freight Shipments	EC	2	0.0648	0.2592	0.0005	0.7403
	NN	3	0.0732	0.3713	0.0017	0.6271
	RW	1	0.0625	0.0008	0.0012	0.9980
In-Bound Air Mail Shipments	EC	1	0.1815	0.2342	0.0002	0.7656
	NN	3	0.2210	0.1043	0.0026	0.8931
	RW	2	0.1994	0.2910	0.0018	0.7073
Out-Bound Air Mail Shipments	EC	1	0.2727	0.2248	0.0009	0.7742
	NN	2	0.2931	0.2664	0.0000	0.7336
	RW	3	0.3032	0.2172	0.0000	0.7828

Note: EC, NN, and RW stand for the econometric model, neural network, and random walk, respectively. Accuracy rankings are based on *U*-statistics. *U-Bias*, *U-Var.*, and *U-Cov.* represent the bias, variance, and covariance proportions of inequality.

The statistics in Tables 2 and 3 measure the average size and composition of forecast errors. To provide a more comprehensive assessment of forecasting performance, a statistical test of directional accuracy is also employed (Pesaran and Timmermann, 1992; 1994). Table 4 presents the test statistics for the null hypothesis that the indicated set of forecasts does not provide useful information for predicting the direction of change in a given variable. Random walk forecasts are not considered because those forecasts assume no change, either positive or negative, in the variables. For the econometric forecasts, it is possible to reject the null hypothesis at the 5 percent significance level in 13 out of 16 cases. Using the same significance criterion, it is only possible to reject the null hypothesis for three out of 16 sets of neural network forecasts. If a 10 percent significance criterion is used instead, then the null hypothesis can be rejected for six out of the 16 sets of neural network forecasts.

Table 4
Pesaran-Timmermann (PT) directional accuracy test results

<i>Variable</i>	<i>EC PT Statistic</i>	<i>NN PT Statistic</i>
Bridge of the Americas Light Vehicles	3.8072**	0.4689
Bridge of the Americas Cargo Vehicles	1.9007**	1.5143*
Bridge of the Americas Pedestrians	4.2065**	2.0864**
Paso del Norte Bridge Light Vehicles	3.9424**	-0.0357
Paso del Norte Bridge Pedestrians	5.8011**	0.3139
Ysleta Bridge Light Vehicles	2.4005**	-2.5825
Ysleta Bridge Cargo Vehicles	2.3608**	3.0052**
Ysleta Bridge Pedestrians	4.8596**	1.5271*
Domestic Air Passenger Arrivals	2.4602**	1.1649
International Air Passenger Arrivals	0.8083	-0.3944
Domestic Air Passenger Departures	3.1038**	1.1621
International Air Passenger Departures	1.4651*	0.3640
In-Bound Air Freight Shipments	2.6875**	2.4358**
Out-Bound Air Freight Shipments	2.6762**	1.2670
In-Bound Air Mail Shipments	0.3955	0.6857
Out-Bound Air Mail Shipments	3.6395**	1.3234*

* Surpasses the 10-percent critical value for a one-tailed normal test.

** Surpasses the 5-percent critical value for a one-tailed normal test.

Notes: The null hypothesis is that the forecasts do not provide useful information for predicting directional changes. Negative PT statistics indicate that the forecasts systematically move in the opposite direction of the actual values.

The directional accuracy results are largely consistent with the *U*-statistics in the sense that the econometric model generally outperforms the neural network. This contrasts with the finding in Heravi, Osborn, and Birchenhall (2004) that neural network forecasts of European industrial production variables appear in a more favorable light when evaluated by measures of directional accuracy as compared to measures of the average size of forecast errors. For the sample examined in this analysis, both types of evaluation metrics suggest better predictive performance by the econometric model.

5. CONCLUSION

Forecasts of metropolitan area bridge and air traffic are important aids in planning, policymaking, and investment decisions. This analysis assesses the performance of previously published econometric forecasts against predictions generated by a neural network model and random walk benchmarks. Theil *U*-statistics indicate that the neural network forecasts improve upon their econometric counterparts in four out of 16 cases. However, the econometric forecasts outperform the alternatives in seven cases. These results indicate that, while no single method is consistently superior to the others, the econometric forecasting model performs reasonably well for at least a substantial subset of the time series examined.

While metrics such as the *U*-statistic that gauge the average size of forecast errors represent useful summary measures of forecasting performance, other statistics can offer a more holistic

assessment of predictive accuracy. The proportions of inequality provide information on the composition of forecast errors. Across 11 of the 16 traffic categories examined, the random walk forecasts are associated with the highest covariance proportions, suggesting that they are less prone to systematic deviations from the mean and variance of the actual time series. However, the econometric forecasts have higher covariance proportions than any of the alternatives for two of the traffic categories. The neural network forecasts have the highest covariance proportion in three cases. The analysis of forecast error decompositions indicate somewhat dissatisfying outcomes for the relatively sophisticated forecasting models.

Finally, planners are also often concerned with correctly anticipating the direction of change in traffic variables. Statistical tests of directional accuracy may prove useful in this regard. The null hypothesis that econometric forecasts fail to provide useful information on directional accuracy is rejected for a majority of the traffic categories. However, for the neural network forecasts, the null hypothesis can only be rejected in a minority of cases. The directional accuracy results lend further support to the regional econometric model. Overall, the latter model appears to be fairly reliable, although both neural networks and random walks are sometimes more accurate.

Acknowledgements

Financial support for this research was provided by El Paso Water, City of El Paso Office of Management & Budget, National Science Foundation Grant DRL-1740695, the UTEP Center for the Study of Western Hemispheric Trade, and the Hunt Institute for Global Competitiveness at UTEP. Econometric research assistance was provided by Ernesto Duarte and Omar Solis.

References

- Antoniou, C., Psarianos, B. and Brilon., W. (2011), Induced Traffic Prediction Inaccuracies as a Source of Traffic Forecasting Failure. *Transportation Letters*, Vol. 3 (4): 253-264.
- Chen, S.-C., Kuo, S.-Y., Chang, K.-W., and Wang, Y.-T. (2012), Improving the Forecasting Accuracy of Air Passenger and Air Cargo Demand: The Application of Back-Propagation Neural Networks. *Transportation Planning and Technology*, Vol. 35 (3): 373-392.
- Church, K.B. and Curram., S.P. (1996), Forecasting Consumers' Expenditure: A Comparison between Econometric and Neural Network Models. *International Journal of Forecasting*, Vol. 12 (2): 255-267.
- Faraway, J. and Chatfield, C. (1998), Time Series Forecasting with Neural Networks: A Comparative Study Using the Airline Data. *Journal of the Royal Statistical Society Series C (Applied Statistics)*, Vol. 47 (2): 231-250.
- Figliozi, M.A., Harrison, R. and McCray, J.P. (2001), Estimating Texas-Mexico North American Free Trade Agreement Truck Volumes. *Transportation Research Record*, 1763: 42-47.
- Fildes, R., Y. 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 (3): 902-922.
- Fullerton, T.M., Jr. (2001), Specification of a Borderplex Econometric Forecasting Model. *International Regional Science Review*, Vol. 24 (2): 245-260.
- Fullerton, T.M., Jr. (2004), "Borderplex Bridge and Air Econometric Forecast Accuracy." *Journal of Transportation & Statistics* 7 (1): 7-21.

- Fullerton, T.M., Jr. and Mukhopadhyay, S. (2013), Border Region Bridge and Air Transport Predictability. *Journal of Business & Economics*, Vol. 4 (11): 1089-1104.
- Fullerton, T.M., Jr., Jimenez, A.A. and, Walke, A.G. (2015), An Econometric Analysis of Retail Gasoline Prices in a Border Metropolitan Economy. *North American Journal of Economics & Finance*, Vol. 34 (SI): 450-461.
- Granger, C.W.J. and Pesaran, M.H. (2000), Economic and Statistical Measures of Forecast Accuracy. *Journal of Forecasting*, Vol. 19 (7): 537-560.
- Hann, T.H. and Steurer, E. (1996), Much Ado about Nothing? Exchange Rate Forecasting: Neural Networks vs. Linear Models Using Monthly and Weekly Data. *Neurocomputing*, Vol. 10 (4): 323-339.
- 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 (3): 435-446.
- Mukhopadhyay, S. (2006), Predicting Global Diffusion of the Internet: An Alternative to Diffusion Models. *Communications of the Association for Information Systems*, Vol. 17: 106-122.
- Nicolaisen, M.S. and Driscoll, P.A. (2014), *Ex-Post* Evaluations of Demand Forecast Accuracy: A Literature Review. *Transport Reviews*, Vol. 34 (4): 540-557.
- Pesaran, M.H. and Timmermann, A.G. (1992), A Simple Nonparametric Test of Predictive Performance. *Journal of Business & Economic Statistics*, Vol. 10 (4): 461-465.
- Pesaran, M.H. and Timmermann, A.G. (1994), A Generalization of the Non-Parametric Henriksson-Merton Test of Market Timing. *Economics Letters*, Vol. 44 (1-2): 1-7.
- Pindyck, R.S. and Rubinfeld, D.L. (1998), *Econometric Models and Economic Forecasts*, 4th edition, Boston, MA: Irwin McGraw-Hill.
- Rumelhart, D. E., Hinton G. E. and, Williams, R. J. (1986), Learning Internal Representations by Error Propagation. Chapter 8 in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vol. 1, D.E. Rumelhart and J.L. McClelland, Editors, Cambridge, MA: MIT Press.
- Vlahogianni, E.I., and M.G. Karlaftis. (2013), "Testing and Comparing Neural Network and Statistical Approaches for Predicting Transportation Time Series." *Transportation Research Record* 2399: 9-22.
- Walke, A.G., and Fullerton, T.M. Jr. (2014), Freight Transportation Costs and the Thickening of the US-Mexico Border. *Applied Economics*, Vol. 46 (11): 1248-1258.
- 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*, Vol. 21 (1): 148-162.
- Xiang, C., Ding, S.Q. and Lee, T.H. (2005), Geometrical Interpretation and Architecture Selection of MLP, *IEEE Transactions on Neural Networks*, Vol. 16 (1): 84-96.
- Xiao, Y., Liu, J.J., Xiao, J., Hu, Y., Bu, H., and Wang, S. (2015), Application of Multiscale Analysis-Based Intelligent Ensemble Modelling on Airport Traffic Forecast. *Transportation Letters*, Vol. 7 (2): 73-79.