Volume 7 Number 1, 2004
ISSN 1094-8848

SPECIAL ISSUE ON FORECASTING

BUREAU OF TRANSPORTATION STATISTICS   UNITED STATES DEPARTMENT OF TRANSPORTATION

JOURNAL OF TRANSPORTATION AND STATISTICS

JOURNAL OF TRANSPORTATION AND STATISTICS

JOURNAL OF TRANSPORTATION AND STATISTICS

Volume 7 Number 1, 2004
ISSN 1094-8848

BUREAU OF TRANSPORTATION STATISTICS UNITED STATES DEPARTMENT OF TRANSPORTATION

Bureau of Transportation Statistics

Our mission: To lead in developing transportation data and information of high quality and to advance their effective use in both public and private transportation decisionmaking.

Our vision for the future: Data and information of high quality supporting every significant transportation policy decision, thus advancing the quality of life and the economic well-being of all Americans.

The *Journal of Transportation and Statistics* releases three numbered issues a year and is published by the
Bureau of Transportation Statistics
U.S. Department of Transportation
Room 7412
400 7th Street SW
Washington, DC 20590
USA
journal@bts.gov

Subscription information
To receive a complimentary subscription:

mail	Product Orders Bureau of Transportation Statistics U.S. Department of Transportation Room 7412 400 7th Street SW Washington, DC 20590 USA
phone	202.366.DATA
fax	202.366.3197
internet	www.bts.gov

Information Service

email	answers@bts.gov
phone	800.853.1351

Cover and text design	Susan JZ Hoffmeyer
Cover photo	Close-up of a jet engine www.acclaimimages.com

The Secretary of Transportation has determined that the publication of this periodical is necessary in the transaction of the public business required by law of this Department.

JOURNAL OF TRANSPORTATION AND STATISTICS

Volume 7 Number 1
2004

Contents

Letter from the Editor-in-Chief

Dear JTS Readers,

As guest editors Keith Ord and Peg Young point out, forecasting is at the heart of policymaking. To make policy, one must forecast what will happen if current policies continue, and one must then forecast what will happen if policies change. The papers in this issue give policymakers—and those who provide analysis for them—a variety of approaches to forecasting outcomes under various policy scenarios.

The structural forecasting models offer the most complex analysis of the interaction of the various factors in the evolution of a policy outcome, and are helpful in analyzing what happens when several policy variables are changing simultaneously. Auto-regressive models, like ARIMA, sometimes perform better as pure forecasting models, in the sense that they often can produce exceptionally high predictive accuracy and are particularly well-suited to modeling the influence of interventions of varying degrees of duration.

In some cases, the models are useful not so much for forecasting the future, but for "forecasting" the present—or even the past. The Liu and Vilain paper shows how a forecasting model can be employed to estimate data at a more disaggregated geographical level than the reported data permit. In the world of transportation data, where the water glass of data often seems more empty than full, this application of forecasting may be extremely valuable.

The application of forecasting to policymaking will be of even greater interest to me as I leave the Bureau of Transportation Statistics (BTS) to take a new position as Chief Economist in DOT's Office of Policy. The Department is investigating a variety of policy proposals—truck-only lanes, congestion pricing, and increased private sector financing of transportation infrastructure, for example—and forecasting techniques will be essential in determining the likely effects of these proposals.

Editing a journal is a wonderful intellectual experience, and I leave the editorship with considerable regret. I will continue to participate as one of the journal's family of reviewers and readers and will continue to make use of the journal's papers in my work. Just this month, for example, I used the estimates of price-elasticities of demand for tolled highways that Anna Matas and José-Luis Raymond presented in their recent article in volume 6 numbers 2/3. Congestion pricing is a topic of considerable interest in the policymaking circles of the Department of Transportation.

I am delighted that I leave the *Journal of Transportation and Statistics* in the capable hands of Peg Young as Editor-in-Chief and Marsha Fenn as Managing

Editor. Peg has worked closely with me as Associate Editor, and we have educated each other in economics and statistics in a mutually edifying partnership. Marsha Fenn, of course, has been with *JTS* since the beginning, and the high quality of the journal is primarily due to her hard work, painstaking thoroughness, and high standards. Peg and Marsha will be assisted by David Chien and Caesar Singh as Associate Editors, by Jennifer Brady as the Data Review Editor, and by Alpha Glass as the Editorial Assistant.

I want to thank the members of the JTS Editorial Board, our reviewers, our authors, and our readers, along with the staff, for making my work on the journal such an enjoyable experience. I look forward to joining our readership and benefiting from the work of our new editorial staff.

JOHN V. WELLS
Editor-in-Chief

INTRODUCTION TO THE SPECIAL ISSUE ON FORECASTING

The key to effective decisionmaking, in transportation as elsewhere, is to understand the workings of the system in order to make accurate assessments of future developments. In this special issue of JTS on transportation forecasting, we have tried to select papers that both examine transportation issues of interest and provide examples of state-of-the-art approaches to forecasting.

SUMMARY OF TOPICS IN THIS ISSUE

The six papers in this special issue are grouped into three areas: economic modeling, surface transportation, and air transportation. The forecasting methods range from single equation time series procedures to detailed econometric models, and the forecast horizons include the short term (typically a few months) through to the long term (five years or more).

Economic Studies

Time series methods typically emphasize effective forecasting, whereas econometric models allow the decisionmaker to gain a deeper understanding of the structure of a system. Thus, the two approaches are complementary and, of course, partially overlapping. In the first paper, Fullerton develops an econometric forecasting system to study two cross-border metropolitan areas: El Paso, Texas, and Ciudad Juárez, Mexico. Within this system, the author models two blocks of transportation equations: northbound surface traffic across the bridge at Ciudad Juárez; and passenger, cargo, and mail flows at the El Paso airport. The model is then used to forecast surface and air traffic in the region.

Liu and Vilain use input-output analysis to estimate commodity inflows in the United States. Using data from the 1993 Commodity Flow Survey (CFS), the authors demonstrate a method for estimating freight inflows on a smaller, substate, regional basis and base the estimates on the industrial structure of the region. Because the CFS only disaggregates data to the state level, Liu and Vilain test the accuracy of their method by estimating flows at the state level and then comparing their results to the actual state results.

Surface Transportation

Surface transportation systems often require rather detailed forecasts of passenger flows over relatively short time horizons. García-Ferrer et al. develop forecast models for monthly bus and Metro ticket demand in Madrid, Spain. Incorporating changing seasonality, calendar effects, and several interventions, the authors compare forecast results from a dynamic transfer-function model and a variant of an unobserved component model.

In an article that studies safety concerns, Raeside studies trends in highway casualties in Great Britain for both travelers and pedestrians. From these time series models, he provides forecasts through 2010 of the casualty rate per kilometer and then compares these forecasts to government-set targets to assess if they are achievable.

Air Transportation

Estimating the impact of September 11, 2001, on U.S. air travel, Ord and Young provide a method for quickly estimating three aspects of the intervention for monthly time series measures of air travel. By creating three separate components of the event—an additive outlier, a level shift, and a temporary decay—the authors illustrate how the combination of the three interventions can be used to adjust a time series only a few months after a major event occurs.

The paper by Bhadra and Texter examines changes in the structure of airline networks in the United States. The growth of low-cost airlines and the increased use of regional jets have provided the impetus for the industry to reconsider the value of the traditional hub-and-spoke system and the authors examine the impact of these changes on airline networks.

TERMS AND ISSUES UNIQUE TO FORECASTING

If the reader is a novice to the forecasting literature, he or she may be confronted with vocabulary or ideas that are new. We hope the following brief synopsis of some of the terms used in this issue will be helpful.[1]

Calendar Effects

In addition to measures of seasonality of the data, some models incorporate additional interventions that reflect consequences of the calendar. Terms such as trading days and holiday effects indicate interventions that reflect the differing number of business days in a particular month (trading days) and the differing position of some holidays in the month (e.g., Easter and Thanksgiving do not fall on the same date every year). Some seasonality procedures may not be able to handle the changes in these effects, so dummy variables may be introduced to reflect their impact.

Hold-Out Samples

In order to measure a model's ability to forecast unknown future values, a set of data points from the end of the series is sometimes withheld during model estimation. The withheld data points, called hold-out observations, can then be compared with forecasted values of this period to evaluate the accuracy of the forecast.

[1] An excellent glossary of forecast terms can be found in Makridakis et al. (1998) or on the Principles of Forecasting website maintained by Professor Scott Armstrong of the University of Pennsylvania, available at http://morris.wharton.upenn.edu/forecast/dictionary, as of November 2004.

Ex-Ante and Ex-Post Forecasts

In ex-ante forecasting, a hold-out sample of both the explanatory and dependent variables is created and removed. Forecasts are generated for the explanatory variables and then used to forecast the dependent variables. The result is a true forecast. Ex-post forecasts typically use the actual values of the explanatory variables. In addition, models producing ex-post forecasts may not use any hold-out sample at all, resulting in all the data being included in the model estimation.

Information Criteria

Model fit measures like MSE (mean square error) may not be very informative when trying to compare models that have different numbers of parameters. In order to compare models, measures such as Akaike's Information Criterion (AIC) and Bayesian Information Criterion (BIC) may be used. The general structure of such measures is: "goodness of fit measure" + "penalty function" and represents a tradeoff between fit and model complexity. For example, the AIC may be written as

$$AIC = n \log_e MSE + 2k$$

The coefficient k denotes the number of parameters fitted. These measures allow quality-of-fit comparisons across models with differing numbers of variables.[2]

Measures of Fit and Accuracy

In addition to the familiar figures of MSE and R^2 (the coefficient of determination or proportion of variance explained), forecast procedures employ statistics that measure different aspects of the quality of the model. Since model fit is not an adequate way to assess forecasting performance, in these articles forecast performance may be assessed either by using information criteria or by using measures based on the hold-out sample. In addition to MSE, authors also use the mean absolute deviation (MAE) and the mean absolute percent error (MAPE).[3]

Theil's U

If the researcher has developed only one model, he or she could still compare the results against the simplest of the forecast methods—termed the "naïve" model—which usually consists of a forecast repeating the most recent value of the variable (e.g., the best forecast of a stock price today is the price of that stock yesterday). The model underlying this naïve forecast is the random walk, which can be specified as

$$y_t = y_{t-1} + \varepsilon_t, \text{ where } \varepsilon_t \sim \text{i.i.d. N}(0, \sigma^2).$$

[2] For an interesting series of forecast competitions, we suggest the reader pursue the literature on the M-competition (e.g., Makridakis et al. 1982, 1993; and Makridakis and Hibon 2000).

[3] Armstrong (2001), Harvey (1997), and Kennedy (1998) are just a small subset of articles dealing with the model fit versus forecast accuracy debate.

That is, each value in the time series is the previous value plus some noise. We may then compare a selected model to the random walk. Behind this notion is the belief that if a forecasting model cannot do better than a naïve forecast, then the model is not doing an adequate job. Theil's U is a statistic that uses the random walk as a benchmark for comparing the quality of forecast models.

UPCOMING FORECASTING ARTICLES

As expected, we received more articles than could be incorporated into one issue. We want to call your attention to three articles, in particular, that we expect to publish in the near future. We originally thought this issue would contain both general transportation forecasting research as well as special articles describing current forecasting models used by the Department of Transportation (DOT). But space required that we delay publishing the two articles dealing with the DOT models until a later issue. A paper by David Chien of the Bureau of Transportation Statistics will present an evaluation of some of the models for greenhouse gas emissions. Roger Schaufele will summarize the models used by the Federal Aviation Administration to forecast large U.S. air carrier domestic revenue passenger-miles, domestic passenger enplanements, and domestic revenues.

The third article we plan to publish is by Miriam Scaglione (Institute of Economy and Tourism, Switzerland) and Andrew Mungall (Lausanne Institute for Hospitality Research, Switzerland), who study interventions with respect to international air travel. They analyzed the impact on Swiss air traffic of Swissair's decision to concentrate all long-haul flights through Zurich and its subsequent filing for bankruptcy. They also look at how air traffic in Switzerland was affected by the terrorist attacks in the United States.

The special issue has generated considerable interest in transportation and forecasting in such groups as the Transportation Research Board and the International Institute of Forecasters. We expect this interest to result in a number of forecasting papers in future issues of JTS.

Keith Ord
Guest Editor
The McDonough School of Business
Georgetown University

Peg Young
Guest Editor
Bureau of Transportation Statistics
U.S. Department of Transportation

ACKNOWLEDGMENTS

We wish to thank all the diligent work by the numerous referees as well as by the JTS editorial staff. While we cannot list the referees by name at this time, we can at least comment that their insight and careful review made each and every article a better piece of research, as also noted by every author. Names of the JTS staff we can mention and publicly thank: Jack Wells, Marsha Fenn, Alpha Glass, Jennifer Brady, Dorinda Edmondson, Martha Courtney, and Lorisa Smith. With all their work and advice, we feel confident that transportation forecasting will be appreciated by a wider audience.

REFERENCES

Armstrong, J.S., ed. 2001. *Principles of Forecasting: A Handbook for Researchers and Practitioners*. Boston, MA: Kluwer Academic Press.

Harvey, A. 1997. Trends, Cycles and Autoregressions. *The Economic Journal* 107:192–201, January.

Kennedy, P. 1998. *A Guide to Econometrics,* 4th ed. Cambridge, MA: The MIT Press.

Makridakis, S., A. Anderson, R. Carbone, R. Fildes, M. Hibon, R. Lewandowski, J. Newton, E. Parzen, and R. Winkler. 1982. The Accuracy of Extrapolation (Time Series) Methods: Results of a Forecasting Competition. *Journal of Forecasting* 1:111–153.

Makridakis, S., C. Chatfield, M. Hibon, M. Lawrence, T. Mills, J.K. Ord, and L. Simmons. 1993. The M-2 Competition: A Real-Life Judgmentally Based Forecasting Study with Discussion. *International Journal of Forecasting* 9:5–22.

Makridakis, S. and M. Hibon. 2000. The M3-Competition: Results, Conclusions and Implications. *International Journal of Forecasting* 16(4):451–476.

Makridakis, S., S.C. Wheelwright, and R.J. Hyndman. 1998. *Forecasting Methods and Applications,* 3rd ed. New York, NY: John Wiley & Sons., Inc.

Borderplex Bridge and Air Econometric Forecast Accuracy

THOMAS M. FULLERTON, JR.

Department of Economics & Finance
University of Texas at El Paso,
El Paso, TX 79968-0543
Email: tomf@utep.edu

ABSTRACT

El Paso, Texas, and Ciudad Juárez, Mexico, jointly comprise a large cross-border metropolitan economy. El Paso is an important port-of-entry for international cargo, as well as a key transit point for regional trade flows in the southwestern United States. Reflective of those traits, the borderplex econometric forecasting system includes two blocks of transportation equations. One subsystem models northbound surface traffic across the international bridges from Ciudad Juárez. The other deals with passenger, cargo, and mail flows at El Paso International Airport. To gauge model reliability, an analysis of borderplex transportation variable forecast accuracy relative to a random walk benchmark is completed. Empirical evidence is mixed with respect to model precision for the 1998 to 2003 sample period for which data are currently available.

INTRODUCTION

Given the historical importance of regional and international trade flows through El Paso, Texas, transportation variables have formed part of the borderplex econometric model from its inception in 1997. Currently comprising 218 individual equations, two sets of transportation equations are

KEYWORDS: Econometric forecasts, air transportation, border economics. JEL Category R15: Regional Econometrics

included in the borderplex model (Fullerton 2001). One block of transportation equations is for northbound traffic categories on the international bridges connecting El Paso with Ciudad Juárez, Mexico. Another subsystem models passenger, cargo, and mail flows at El Paso International Airport.

Fullerton and Tinajero (2003) and others generated short-term cyclical forecasts of borderplex business and economic conditions using the model from 1998 forward. The three-year out-of-sample forecast period for transportation variables is included in these publications. For some other variables, the effective simulation period is longer due to lags in data collection and dissemination. When missing values for the last historical period of any series occur, a model simulation provides estimates of the missing observations. Data release delays occasionally contribute to that circumstance. To date, formal prediction accuracy assessments have not been conducted for the transportation variables included in the border region system of simultaneous equations.

This paper first examines the accuracy of the borderplex transportation variable forecasts published for 1998 through 2003. Predictive accuracy is assessed relative to a random walk benchmark. Subsequent sections of the paper include discussions of regional econometric forecasting research, borderplex model attributes, and an empirical analysis of transportation forecast accuracy between 1998 and 2003. A summary and suggestions for future research are provided in the conclusion.

REGIONAL AND BORDER ECONOMETRIC FORECASTING RESEARCH

Structural model forecasting analysis for regional and national economies can be traced back to 1936 (Dhane and Barten 1989). Overall design flexibility makes it a widely used tool for corporate planning efforts and public policy analysis. Structural models have been applied to a wide variety of regional and metropolitan economies in the United States, Europe, and Asia (Klein 1969; Bolton 1985; Kim 1995; Hunt and Snell 1997). Since 1997, one such model has been utilized to simulate economic and business conditions in the cross-border regional economy comprised by El Paso, Texas; Las Cruces, New Mexico; Ciudad Juárez, Mexico; and Ciudad Chihuahua, Mexico (Fullerton 2001).

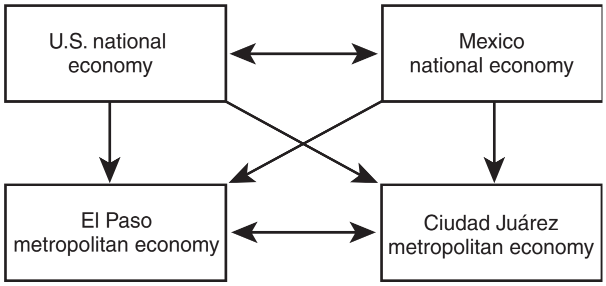
Several authors have suggested that out-of-sample forecasting accuracy and simulation analyses can be important tools for assessing econometric model reliability (Leamer 1983; West 1995; Granger 1996; McCloskey and Ziliak 1996). Out-of-sample forecasts are model simulations that go beyond the end of the sample period for which historical data are available. A growing number of studies have used these forecasts for the regional class of structural equation models (West and Fullerton 1996; Fullerton and West 1998; Fullerton et al. 2000; Lenze 2000; Fullerton et al. 2001; West 2003). Those studies indicate that regional forecasts for many variables, such as employment, income, and population, are relatively accurate. However, the track records for regional housing activity are less successful.

Infrastructure planning has long required systematic forecasting efforts for transportation systems. Numerous methodologies have been examined over the years (Schneider 1975; Beenstock and Vergottis 1989; Matthews 1995; Caves 1997; Dennis 2002). However, to date, relatively few regional transportation forecasting programs have been examined for historical accuracy. This gap in the literature is partially addressed in this paper by examining the accuracy of borderplex air and bridge traffic forecasts published between 1998 and 2003.

Figure 1 depicts the basic strategy deployed in the borderplex system of simultaneous equations. In addition to being affected by national economic trends in the United States, El Paso feels the effects of the national business cycle fluctuations in Mexico, as well as regional business cycles in Ciudad Juárez. The reverse circumstance also holds true for Ciudad Juárez. Consequently, individual equation specifications in the model may contain national macroeconomic, domestic regional, international macroeconomic, and/or cross-border metropolitan variables (Fullerton 2001).

Recent economic history along the Mexican border in Texas succinctly reflects the attributes shown in figure 1. The "Tequila Effect" peso devaluation of December 1994 precipitated a severe loss in Mexican consumer purchasing power that led to a

FIGURE 1 Borderplex Model Design



decline in international bridge crossings and a reduction in retail sales in El Paso (Fullerton 1998; Fullerton 2000). Macroeconomic shocks on the north side of the border also affect the local economy. The 2001 U.S. recession hurt manufacturing plants throughout the borderplex. In response to that, plus a changed inspection environment subsequent to September 11, 2001, cargo vehicle traffic from Mexico sagged (Fullerton and Tinajero 2002).

The borderplex model is used for a variety of purposes, with the most important being regional business trend monitoring and econometric forecasting analysis. While there are a small number of commercially available forecasts for El Paso County (Leppold 2002; Shankar 2003), those efforts generally ignore cross-border business conditions and omit transportation flows. The model is also used in a variety of public policy analysis exercises, such as the provision of simulation data utilized in testimony provided to Texas State Senate committees regarding local income trends and North American Free Trade Agreement (NAFTA) adjustment efforts. Local governmental units and public utilities have engaged a series of annual contracts with the University of Texas at El Paso Border Region Modeling Project for special simulation exercises designed to assist infrastructure planning efforts. Access to comprehensive forecasts involving both sides of the border using a common set of exogenous variable assumptions are critical to those endeavors.

To accomplish the model's goals, sectoral coverage is necessarily broad. Twelve separate equation blocks are utilized: demographics, employment, personal income, labor earnings, disaggregated retail sales, residential real estate, nonresidential construction, maquiladora activity, northbound international border crossings, air transportation, water consumption, and regional bankruptcy trends. The structure of the model contains numerous direct and indirect feedback loops connecting the various equation blocks (Fullerton 2001). Because annual data are used, lag structures are fairly short in all of the different blocks. They are confirmed every year via goodness-of-fit tests.

The 218 equations in the current version of the border forecasting system offer at least partial detail for each of the 12 blocks of endogenous equations.[1] The 218 equations contain 40 identities and 178 stochastic equations. Over the years, some equation specifications have remained unchanged, while multiple variations have been tested for others. Specification updates occur due to new data acquisitions, alternative possibilities identified in the literature, and/or as a consequence of previous relationships not performing well following the annual data bank updates and parameter re-estimation exercises. Heightened security inspection efforts and post-9/11 travel disruptions also caused equation modifications in both 2002 and 2003. Of the 178 fitted results ultimately selected every year, most exhibit good statistical traits, but nearly all contain at least partial design and/or empirical flaws.

Of the 178 regression equations, 51 required serial correlation correction techniques. Three categories of data generating processes can be seen in the affected residual series: 18 autoregressive, 28 moving average, and 5 mixed autoregressive-moving average sets of parameters. Given the variety of autocorrelation processes involved, parameter estimation was accomplished using a nonlinear ARMAX procedure (Pagan 1974). That more than one-fourth of all the border model stochastic specifications required serial correlation correction in part reflects widespread data constraints that have long affected regional econometric modeling efforts. Unavailable data series occasionally prevent some systematic variation in dependent variables from being handled as satisfactorily as they can be in macroeconometric models (Fullerton and West 1998). As with national econometric models, persistence effects probably also contribute to the preva-

[1] Statistical output for the econometric equations currently comprising the borderplex model are available from the author.

TABLE 1 Borderplex Model Variables

Series	Description
Endogenous variables	
AIRP	U.S. Air Travel Price Index[1] (1982–1984 = 100)
ELAFDT	El Paso International Airport Inbound Freight Shipments (1,000 tons)
ELAFET	El Paso International Airport Outbound Freight Shipments (1,000 tons)
ELAMD	El Paso International Airport Inbound U.S. Mail (1,000 tons)
ELAME	El Paso International Airport Outbound U.S. Mail (1,000 tons)
ELBAC	Bridge of the Americas Northbound Light Vehicle Traffic (millions)
ELBAT	Bridge of the Americas Northbound Cargo Vehicle Traffic (millions)
ELBAW	Bridge of the Americas Northbound Pedestrian Traffic (millions)
ELBPC	Paso del Norte Northbound Light Vehicle Bridge Traffic (millions)
ELBPW	Paso del Norte Northbound Pedestrian Bridge Traffic (millions)
ELBTC	Total Northbound Light Vehicles (millions)
ELBTT	Total Northbound Cargo Vehicles (millions)
ELBTW	Total Northbound Pedestrians (millions)
ELBYC	Ysleta-Zaragoza Northbound Light Vehicle Bridge Traffic (millions)
ELBYT	Ysleta-Zaragoza Northbound Cargo Vehicle Bridge Traffic (millions)
ELBYW	Ysleta-Zaragoza Northbound Pedestrian Bridge Traffic (millions)
ELAPDD	El Paso International Airport Domestic Passenger Arrivals (thousands)
ELAPDI	El Paso International Airport International Passenger Arrivals (thousands)
ELAPDT	El Paso International Airport Total Passenger Arrivals (thousands)
ELAPED	El Paso International Airport Domestic Passenger Departures (thousands)
ELAPEI	El Paso International Airport International Passenger Departures (thousands)
ELAPET	El Paso International Airport Total Passenger Departures (thousands)
MAILP	U.S. Air Mail Price Index (1996 = 100)

Series	Description
Exogenous variables	
CESTRNPI	U.S. Consumption Expenditures, Intercity Travel (billion $)
CJMQM	Ciudad Juárez Maquiladora Employment (thousands)
CJPOP	Ciudad Juárez July 1 Population (thousands)
DV911	International Bridge Security Dummy Var. = 1 for 2001, 2002, …
ELBDC	Dedicated Commuter Lane Northbound Light Vehicles (millions)
ELGMP96	El Paso Gross Metropolitan Product (billion 1996 $)
ELPPOP	El Paso July 1 Population (thousands)
ELYWSD	El Paso Wage and Salary Disbursements (million $)
ELYLP	El Paso Labor and Proprietor Earnings (million $)
GDP96	U.S. Gross Domestic Product (GDP) (billion 1996 $)
MXREX	Mexico, Real Exchange Rate Index, Peso/$ (1997 = 100)
PDCCESTRNPI	U.S. Intercity Chained Travel Deflator (1996 = 100)
PDCGDP	U.S. GDP Chained Price Deflator (1996 = 100)
Equation statistics	
SUM SQ	Error Sum of Squares
STD ERR	Standard Error of Regression
LHS MEAN	Left Hand Side Dependent Variable Sample Mean
R SQ	R-Squared Coefficient of Determination
R BAR SQ	Adjusted R-Squared Coefficient of Determination
F	F Statistic for Joint Slope Coefficient Equality to Zero Hypothesis
DW	Durbin Watson Serial Correlation Statistic for logs 1 (DW(1)) and 2 (DW(2))
H	Durbin H Lagged Dependent Variable Serial Correlation Statistic

[1] This index is the Consumer Price Index Component for Personal Travel. U.S. Department of Commerce, U.S. Census Bureau, *Statistical Abstract of the United States* (Washington, DC: Various issues).

lence autocorrelation in borderplex equation residuals (Fair 1984; Campbell and Mankiw 1987).

Table 1 lists and describes the variables included in the transportation blocks of the border model and empirical summaries for all of the air and bridge equation parameter estimates. Table 2, which reports Border Region Modeling Project historical data used to estimate those blocks of equations, can

TABLE 2 Borderplex Bridge and Air Equation Estimation Results

EQUATIONS 1–2 are Air Transportation Identities
EQUATIONS 3–12 are Air Transportation Stochastic Equations

Equation 1 Air Passengers, Total Departures (thousands)

ELAPET = ELAPED + ELAPEI

Equation 2 Air Passengers, Total Arrivals (thousands)

ELAPDT = ELAPDD + ELAPDI

Equation 3 U.S. Air Travel Price Index, 1982–1984 = 100

AIRP = f(AIRP.1, PDCGDP)

Ordinary Least Squares, annual data for 39 periods from 1964–2002

airp = 0.84754 * airp[–1] + 44.5066 * pdcgdp – 8.24792
(17.0174) (3.72458) (2.51809)

Sum Sq	946.932	Std Err	5.1287	LHS Mean	105.877
R Sq	0.9955	R Bar Sq	0.9952	F 2, 36	3947.49
DW(1)	1.8062	DW(2)	2.5556	H	0.5858

Equation 4 Air Passengers, Domestic Departures (thousands)

ELAPED = f(ELAPED.1, ELYWSD/AIRP, DV 911)

Nonlinear Least Squares, annual data for 22 periods from 1981–2002

Elaped = 0.82612 * elaped[–1] + 21.0483 * elywsd/airp – 150.688 * dv911 – 324.617
(7.04844) (1.03279) (1.98167) (0.60654)

Sum Sq	128582	Std Err	86.9694	LHS Mean	1482.80
R Sq	0.9322	R Bar Sq	0.9163	F 4, 17	58.4537
DW(1)	2.2001	DW(2)	2.0029	H	–0.6211

AR_0 = 0.27511 * AR_1
(1.06312)

Equation 5 Air Passengers, International Departures (thousands)

ELAPEI = f(ELAPEI.1, MXREX, PDCCESTRNPI/PDCGDP)

Ordinary Least Squares, annual data for 23 periods from 1980–2002

elapei = 0.49491 * elapei [–1] – 0.19892 * mxrex –6.81160 * dv911 – 13.6134 * pdccestrnpi/pdcgdp + 41.6922
(3.23979) (2.54669) (1.48202) (1.15825) (2.61212)

Sum Sq	445.716	Std Err	4.9761	LHS Mean	11.9377
R Sq	0.6691	R Bar Sq	0.5956	F 4, 18	9.1001
DW(1)	1.2223	DW(2)	1.8125	H	0.5157

Equation 6 Air Passengers, Domestic Arrivals (thousands)

ELAPDD = f(ELAPDD.1, CESTRNPI/PDCCESTRNPI, DV911)

Ordinary Least Squares, annual data for 23 periods from 1980–2002

elapdd = 0.91955 * elapdd[–1] + 0.88772 * cestrnpi/pdccestrnpi – 121.591 * dv911 + 115.683
(9.43567) (0.16531) (1.57280) (1.13256)

Sum Sq	158515	Std Err	91.3394	LHS Mean	1422.51
R Sq	0.9206	R Bar Sq	0.9080	F 3, 19	73.3817
DW(1)	1.7249	DW(2)	1.6678	H	0.4568

Equation 7 Air Passengers, International Arrivals (thousands)

ELAPDI = f(ELAPDI.1, MXREX, PDCCESTRNPI/PDCGDP)

Ordinary Least Squares, annual data for 23 periods from 1980–2002

elapdi = 0.77877 * elapdi[–1] – 0.11833 * mxrex + 0.21759 * pdccestrnpi/pdcgdp + 14.4671
(6.00170) (1.89065) (0.01736) (1.09570)

Sum Sq	493.062	Std Err	5.0942	LHS Mean	15.3853
R Sq	0.8035	R Bar Sq	0.7725	F 3, 19	25.8982
DW(1)	1.6222	DW(2)	1.7017	H	–0.4845

(continued on next page)

TABLE 2 Borderplex Bridge and Air Equation Estimation Results (Continued)

Equation 8 Air Freight, Outbound (thousand tons)

ELAFET = f(ELAFET.1, GDP96, DV911)

Ordinary Least Squares, annual data for 28 periods from 1975–2002

elafet = 0.34113 * elafet[–1] + 0.00576 * gdp96 – 6.9705 * dv911 – 25.2374

(2.02867) (4.38056) (3.70933) (4.19841)

Sum Sq	89.6066	Std Err	2.1717	LHS Mean	20.8295
R Sq	0.9687	R Bar Sq	0.9637	F 3, 19	195.943
DW(1)	1.4562	DW(2)	1.8892	H	2.0793

Equation 9 Air Freight, Inbound (thousand tons)

ELAFDT = f(ELAFDT.1, ELYLP/PDCCESTRNPI, ELGMP96, DV911)

Ordinary Least Squares, annual data for 23 periods from 1980–2002

elafdt = 0.41934 * elafdt[–1] + 0.00034 * elylp/pdccestrnpi + 5.25436 * elgmp96 – 5.90542 * dv911 – 43.0747

(4.22055) (0.30584) (3.24465) (3.66921) (4.5494)

Sum Sq	62.8499	Std Err	1.9073	LHS Mean	26.7198
R Sq	0.9900	R Bar Sq	0.9871	F 5, 17	337.956
DW(1)	1.8939	DW(2)	2.3568	H	0.1479
MA_0 =	– 0.93588 * MA_1				
	(3.38261)				

Equation 10 Air Mail Price Index (1996 = 100)

MAILP = f(MAILP.1, PDCGDP)

Ordinary Least Squares, annual data for 39 periods from 1964–2002

mailp = 0.06430 * mailp[–1] + 32.1832 * pdcgdp – 3.01038

(0.43204) (6.34350) (4.71268)

Sum Sq	21.9588	Std Err	0.7810	LHS Mean	19.2099
R Sq	0.9946	R Bar Sq	0.9943	F 2, 36	3299.64
DW(1)	1.9237	DW(2)	2.1208	H	0.4443

Equation 11 Air Mail, Outbound (thousand tons)

ELAME = f(ELAME.1, ELGMP96, MAILP/PDCGDP, DV911)

Ordinary Least Squares, annual data for 28 periods from 1975–2002

elame = 0.65473 * elame[–1] + 0.04247 * elgmp96 – 0.01800 * mailp/pdcgdp – 1.14069 * dv911 + 0.75765

(3.94033) (1.44456) (0.54124) (6.30882) (0.92028)

Sum Sq	1.0553	Std Err	0.2142	LHS Mean	1.7473
R Sq	0.7345	R Bar Sq	0.6883	F 4, 23	15.9039
DW(1)	1.9430	DW(2)	2.1283	H	0.1106

Equation 12 Airmail, Inbound (thousand tons)

ELAMD = f(ELAMD.1, ELGMP96, MAILP/PDCGDP, DV911)

Ordinary Least Squares, annual data for 28 periods from 1975–2002

elamd = 0.83160 * elamd[–1] + 0.06354 * elgmp96 + 0.01042 * mailp/pdcgdp – 1.44147 * dv911 – 0.40496

(6.21740) (0.95369) (0.14702) (3.70439) (0.22777)

Sum Sq	4.8725	Std Err	0.4603	LHS Mean	2.8202
R Sq	0.7252	R Bar Sq	0.6774	F 4, 23	15.1713
DW(1)	1.6837	DW(2)	1.8787	H	1.0486

EQUATIONS 13–15 are International Bridge Identities
EQUATIONS 16–23 are International Bridge Stochastic Equations

Equation 13 International Bridges, Total Cargo Vehicles (millions)

ELBTT = ELBAT + ELBYT

Equation 14 International Bridges, Total Light Vehicles (millions)

ELBTC = ELBAC + ELBDC + ELBYC + ELBPC

TABLE 2 Borderplex Bridge and Air Equation Estimation Results (Continued)

Equation 15 International Bridges, Total Pedestrians (millions)

ELBTW = ELBAW + ELBYW + ELBPW

Equation 16 Bridge of the Americas, Cargo Vehicles (millions)

ELBAT = f(ELBAT.1, CJMQM, DV911)

Ordinary Least Squares, annual data for 26 periods from 1977–2002

elbat = 0.72275 * elbat[–1] + 0.00042 * cjmqm – 0.01454 * dv911 + 0.02716

(5.27650) (1.53320) (0.31027) (1.06493)

Sum Sq	0.0733	Std Err	0.0577	LHS Mean	0.2438
R Sq	0.8138	R Bar Sq	0.7885	F 3, 22	32.0614
DW(1)	1.7214	DW(2)	2.0725	H	0.9385

Equation 17 Bridge of the Americas, Light Vehicles (millions)

ELBAC = f(MXREX, ELPPOP.1 + CJPOP.1, DV911)

Ordinary Least Squares, annual data for 26 periods from 1977–2002

elbac = –0.01704 * mxrex + 0.00165 * elppop[–1] + cjpop[–1] – 2.44885 * dv911 + 6.78609

(2.12403) (3.23975) (4.05759) (6.36514)

Sum Sq	10.9309	Std Err	0.7049	LHS Mean	7.1934
R Sq	0.4886	R Bar Sq	0.4189	F 3, 22	7.0068
DW(1)	1.6757	DW(2)	1.9535		

Equation 18 Bridge of the Americas, Pedestrians (millions)

ELBAW = f(ELBAW.1, CJMQM, DV911)

Ordinary Least Squares, annual data for 26 periods from 1977–2002

elbaw = 0.36724 * elbaw[-1] + 0.00047 * cjmqm + 0.33277 * dv911 + 0.30128

(2.33034) (1.32460) (3.45088) (2.99741)

Sum Sq	0.3336	Std Err	0.1179	LHS Mean	0.5942
R Sq	0.5671	R Bar Sq	0.5130	F 3, 24	10.4817
DW(1)	1.5787	DW(2)	1.6586	H	0.9508

Equation 19 Paso del Norte Bridge, Light Vehicles (millions)

ELBPC = f(ELBPC.1, ELPPOP.1+CJPOP.1, MXREX, DV911)

Ordinary Least Squares, annual data for 33 periods from 1970–2002

elbpc = 0.50442 * elbpc[–1] + 0.00001 * elppop[–1] + cjpop[–1] + 0.00141 * mxrex – 0.48639 * dv911 + 2.14509

(3.00439) (0.03346) (0.30880) (1.33118) (2.55494)

Sum Sq	5.2607	Std Err	0.4335	LHS Mean	4.5782
R Sq	0.2961	R Bar Sq	0.1955	F 4, 28	2.9441
DW(1)	1.7126	DW(2)	2.4887	H	2.6796

Equation 20 Paso del Norte Bridge, Pedestrians (millions)

ELBPW = f(ELBPW.1, CJMQM, DV911)

Ordinary Least Squares, annual data for 33 periods from 1970–2002

elbpw = 0.59209 * elbpw[–1] + 0.00113 * cjmqm + 1.35238 * dv911 + 1.79973

(4.29580) (0.47648) (2.15409) (2.25472)

Sum Sq	13.9889	Std Err	0.7635	LHS Mean	4.9874
R Sq	0.5408	R Bar Sq	0.4834	F 3, 24	9.4228
DW(1)	1.5463	DW(2)	1.9588	H	1.4861

Equation 21 Ysleta-Zaragoza Bridge, Cargo Vehicles (millions)

ELBYT = f(ELBYT.1, CJMQM, MXREX, DV911)

Ordinary Least Squares, annual data for 33 periods from 1970–2002

elbyt = 0.89245 * elbyt[–1] + 0.00033 * cjmqm – 0.00046 * mxrex – 0.04697 * dv911 + 0.03436

(10.9453) (2.39713) (1.60623) (2.17479) (1.33619)

Sum Sq	0.0180	Std Err	0.0253	LHS Mean	0.1041
R Sq	0.9691	R Bar Sq	0.9646	F 4, 28	219.208
DW(1)	1.6956	DW(2)	2.4523	H	0.9429

(continued on next page)

TABLE 2 Borderplex Bridge and Air Equation Estimation Results (Continued)

Equation 22 Ysleta-Zaragoza Bridge, Light Vehicles (millions)

ELBYC = f(ELBYC.1, CJPOP.1, DV911)

Nonlinear Least Squares, annual data for 33 periods from 1970–2002

elbyc = 0.73020 * elbyc[–1] + 0.00109 * cjppop[–1] – 0.27012 * dv911 + 0.17834

(4.67118) (2.31758) (1.59233) (1.39335)

Sum Sq	0.7507	Std Err	0.1769	LHS Mean	2.2453
R Sq	0.9585	R Bar Sq	0.9533	F 3, 24	184.796
DW(1)	1.5615	DW(2)	2.0065	H	1.9514
MA_0 =	0.20841 * MA_1				
	(0.75853)				

Equation 23 Ysleta-Zaragoza Bridge, Pedestrians (millions)

ELBYW = f(ELBYW.1, MXREX, CJMQM, DV911)

Ordinary Least Squares, annual data for 33 periods from 1970–2002

elbyw = 0.82936 * elbyw[–1] + 0.00143 * mxrex + 0.00020 * cjmqm + 0.25096 * dv911 – 0.11255

(7.42629) (2.26824) (0.83996) (5.25357) (1.92377)

Sum Sq	0.0931	Std Err	0.0577	LHS Mean	0.2314
R Sq	0.9226	R Bar Sq	0.9116	F 4, 28	83.4640
DW(1)	2.0700	DW(2)	1.9894	H	–1.0059

be accessed via the University of Texas at El Paso College of Business Administration website (www.utep.edu). The statistical diagnostics for these two groups of equations worsened notably once 2001 observations were included in the historical sample set in 2002. That circumstance continued in 2003 even after 2002 data became available. Most notably, the 18 stochastic equations include 22 separate slope coefficients that fail to satisfy the 5 percent significance criterion. This development is most likely temporary and is expected to fade eventually as the aftermath of the post-September 11, 2001, air travel and border disruptions dissipates.

Multi-equation regional econometric forecasting systems usually omit air transportation activities. The borderplex model partially overcomes this customary gap with a 12-equation subsystem encompassing air passenger, freight, and mail flows through the El Paso International Airport. Domestic passenger arrivals and departures are modeled as functions of metropolitan real wage and salary disbursements and a real price variable for air travel (Howry 1969). International passenger traffic flows are dependent on the inflation-adjusted value of the peso and the relative price index for air transportation (González and Moral 1995). Equations 1 through 12 in table 2 show that a combination of national and border region variables are used to model both freight and airmail shipments and deliveries.

International bridge traffic from Mexico is modeled with a block of 11 equations, 8 of which are stochastic. Coverage in this portion of the model is confined to northbound border commuting across the three bridges within the El Paso city limits and excludes other regional crossings data as a consequence of time series information constraints. Merchandise trade statistics for El Paso extend back only to 1993, precluding the estimation of trade flow equations that might otherwise be of interest to policy analysts and corporate planners. Three categories of traffic flows are included in the current version of the border model: pedestrians, personal automobiles, and cargo vehicles. More than 9 million pedestrians, 11 million light vehicles, and 700,000 cargo vehicles crossed the border using these arteries in 2001 (Fullerton and Tinajero 2003). Not surprisingly, a mixture of national and international exogenous variables, plus border region endogenous data, is used in the specifications shown in equations 13 through 23 in table 2 (Sawyer and Sprinkle 1986; Cobb et al. 1989; Fullerton 2001).

In addition to the transportation endogenous variables that are analyzed for historical predictive accuracy, air travel and air mail price indices are also included in the borderplex model. Similar to other equations in the transportation blocks of the model, their respective empirical traits continued to be acceptable in both 2002 and 2003. Curiously, however, the estimated parameter for the autoregressive lag of the mail price index is not statistically significant. Given the nature of postal service price increases, the partial adjustment specification is probably correct. Prior to 2003, it obtained better estimation results for the lagged dependent variable regression coefficient.

HISTORICAL ACCURACY ASSESSMENT: 1998–2003

The preceding section provides descriptive insights with respect to the overall structure of the borderplex model. It does not shed any light on its general forecast reliability for the transportation equations. To examine this question, a straightforward accuracy assessment was devised along the regional modeling guidelines proposed by West (1995). Historically, extrapolations from univariate ARIMA equations are regarded as the most reliable benchmarks against which structural model performance should be measured (Granger 1996). Because annual data are used in the border model, small sample sizes preclude estimating ARIMA equations.

To circumvent that obstacle, random walk projections were used to provide the backdrop against which a comparison can be made with the previously published structural model forecasts. Figures 2 and 3 illustrate the variable growth rates observed for many borderplex transportation series. Given that variability, the random walks utilized only the last historical observation available for each variable. While apparently simple, this type of benchmark has proven surprisingly effective in other regional forecasting contexts where sawtooth growth patterns occur (Fullerton et al. 2001). The 1998 to 2003 outlook publications allow assembling the original data for each dependent variable, thus avoiding the common problem of inadvertently handicapping the structural simulations when revised data must be used to generate the random walk forecasts (West and Fullerton 1996).

Using borderplex model data for 1998 to 2003, three-year forecasts are shown for selected transportation variables (see table 3). The forecasts are ex ante dynamic simulations and do not employ historical data for the right-hand-side variables. National

FIGURE 2 Borderplex Air Activity: 1979–2003

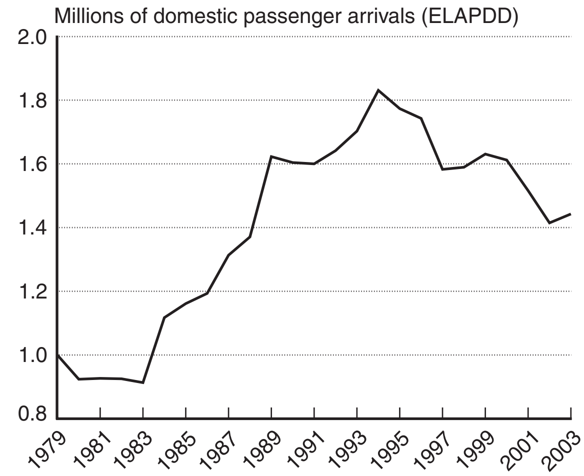
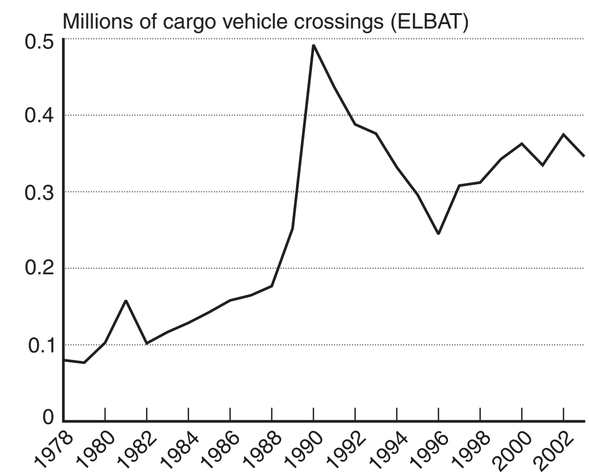


FIGURE 3 Borderplex Bridge Activity: 1978–2003



consultant service subscriptions provide forecast data for U.S. and Mexico macroeconomic variables used as exogenous regressors (Alemán 2003; Behravesh et al. 2003; Zandi 2003). For the 1998 to 2003 sample period, this allows 15 observations to be assembled for each of the air and bridge dependent variables of interest.

The previously published forecasts for each transportation variable are compared with random walk benchmarks. As shown for four representative variables in table 3, both sets of three-year forecasts are listed in order of publication. Accordingly, prediction data for 1998, 1999, and 2000 are followed by similar numbers for 1999, 2000, and 2001, next, and so forth. For the last two sets of previously published forecasts, only two and one historical data points are currently available for accuracy comparisons. Accuracy measures applied to the data include

TABLE 3 Borderplex Transportation Historical and Forecast Data

Year	Actual data	Structural model forecast	Random walk
Domestic Passenger Air Arrivals (thousands)			
1998	1,590.138	1,627.3	1,600.5
1999	1,631.010	1,677.7	1,600.5
2000	1,611.738	1,729.6	1,600.5
1999	1,631.010	1,630.9	1,607.1
2000	1,611.738	1,654.9	1,607.1
2001	1,516.602	1,679.8	1,607.1
2000	1,611.738	1,664.2	1,631.0
2001	1,516.602	1,701.5	1,631.0
2002	1,414.823	1,743.5	1,631.0
2001	1,516.602	1,573.2	1,611.7
2002	1,414.823	1,560.4	1,611.7
2003	1,443.058	1,581.0	1,611.7
2002	1,414.823	1,451.1	1,451.1
2003	1,443.058	1,476.0	1,451.1
2003	1,443.058	1,413.2	1,414.8
Inbound Air Freight (thousand tons)			
1998	47.396	41.475	39.273
1999	55.600	44.802	39.273
2000	55.204	48.565	39.273
1999	55.600	44.237	40.317
2000	55.204	48.800	40.317
2001	46.013	53.467	40.317
2000	55.204	58.673	56.131
2001	46.013	63.098	56.131
2002	51.637	67.427	56.131
2001	46.013	48.556	54.958
2002	51.637	47.683	54.958
2003	45.366	50.941	54.958
2002	51.637	52.066	46.013
2003	45.366	54.573	46.013
2003	45.366	46.854	51.637
Bridge of the Americas Northbound Light Vehicle Traffic (millions)			
1998	7.553	7.870	7.421
1999	8.196	8.048	7.421
2000	8.168	8.229	7.421
1999	8.196	8.090	7.942
2000	8.168	8.209	7.942
2001	7.295	8.327	7.942
2000	8.168	8.276	8.196
2001	7.295	8.394	8.196
2002	4.708	8.534	8.196
2001	7.295	7.723	8.168
2002	4.708	7.800	8.168
2003	4.680	7.889	8.168
2002	4.708	4.655	7.295
2003	4.680	4.771	7.295
2003	4.680	4.820	3.658
Ysleta-Zaragoza Bridge Northbound Cargo Vehicles (millions)			
1998	0.294	0.327	0.289
1999	0.329	0.368	0.289
2000	0.365	0.411	0.289
1999	0.329	0.345	0.314
2000	0.365	0.374	0.314
2001	0.331	0.403	0.314
2000	0.365	0.356	0.329
2001	0.331	0.379	0.329
2002	0.329	0.400	0.329
2001	0.331	0.367	0.365
2002	0.329	0.369	0.365
2003	0.311	0.372	0.365
2002	0.329	0.319	0.331
2003	0.311	0.326	0.331
2003	0.311	0.319	0.329

root mean square error statistics (RMSEs) and Theil inequality coefficients, also known as U-statistics (Pindyck and Rubinfeld 1998). U-statistics can take values between 0 and 1. A value of 0 indicates a perfect fit. For the covariance proportions of the prediction error second moments, the optimal values are 0, 0, and 1. See the appendix for specific information on the calculation of those measures.

Table 4 summarizes predictive accuracy results for air passenger traffic, air freight, and airmail flows in and out of El Paso International Airport. Passenger traffic variables analyzed for out-of-sample forecast precision include inbound passengers from domestic flights, inbound passengers from international flights, outbound passengers to domestic destinations, and outbound passengers to international destinations. The other stochastic equation series examined include inbound freight, outbound freight, inbound mail, and outbound mail.

TABLE 4 Air Series Predictive Accuracy
Sample period: 1998–2003

Series	RMSE	U-statistic	U-bias	U-variance	U-covariance
El Paso International Airport Domestic Passenger Arrivals					
ELAPDD[1]	125.9	0.004	0.51	0.01	0.48
ELAPDD[2]	99.7	0.003	0.33	0.02	0.65
El Paso International Airport International Passenger Arrivals					
ELAPDI[1]	6.947	0.276	0.74	0.02	0.24
ELAPDI[2]	4.867	0.203	0.75	0.03	0.22
El Paso International Airport Domestic Passenger Departures					
ELAPED[1]	140.9	0.004	0.38	0.01	0.62
ELAPED[2]	118.1	0.004	0.26	0.11	0.63
El Paso International Airport International Passenger Departures					
ELAPEI[1]	5.796	0.255	0.69	0.00	0.31
ELAPEI[2]	4.657	0.202	0.68	0.03	0.29
El Paso International Airport Inbound Freight					
ELAFDT[1]	8.622	0.055	0.02	0.10	0.88
ELAFDT[2]	9.838	0.063	0.07	0.09	0.84
El Paso International Airport Outbound Freight					
ELAFET[1]	4.889	0.046	0.05	0.01	0.94
ELAFET[2]	5.399	0.052	0.03	0.02	0.95
El Paso International Airport Inbound U.S. Mail					
ELAMD[1]	1.231	0.507	0.23	0.04	0.74
ELAMD[2]	1.361	0.552	0.28	0.15	0.57
El Paso International Airport Outbound U.S. Mail					
ELAME[1]	0.934	0.957	0.25	0.06	0.69
ELAME[2]	1.035	0.978	0.24	0.08	0.68

[1] Previously published borderplex structural model forecast.
[2] Random walk forecast calculated as last available historical observation.

In table 4, the first row for each variable contains the structural model predictive summary statistics and the second row reports the same estimates for the random walk extrapolations. With the exception of the airmail series, U-statistics close to 0 are obtained for both sets of airport activity forecasts. Examination of the second moment prediction error proportions reveals that the passenger variable structural model forecasts tend to be biased, but the same problem afflicts the random walk counterparts for those series. Similar to the regional housing starts results obtained for Florida (Fullerton and West 1998; Fullerton et al. 2000), the borderplex model passenger forecasts obtain higher U-statistic values than do their respective random walk alternatives.

In contrast, the structural model air cargo and airmail variables also analyzed in table 4 obtain U-coefficients that are lower than those of their respective random walk counterparts. Those outcomes are more in line with regional results previously obtained for variables such as employment, population, or personal income (West and Fullerton 1996; Lenze 2000). Interestingly, the distributions of the inequality coefficient second moment proportions (U-bias, U-variance, and U-covariance) are much closer to the optimum 0, 0, 1 distribution for the nonpassenger variables. However, those improve-

TABLE 5 **Bridge Series Predictive Accuracy**
Sample period: 1998–2003

Series	RMSE	U-statistic	U-bias	U-variance	U-covariance
Bridge of the Americas Northbound Light Vehicle Traffic					
ELBAC[1]	1.573	0.181	0.31	0.02	0.67
ELBAC[2]	1.900	0.241	0.27	0.06	0.67
Bridge of the Americas Northbound Cargo Vehicle Traffic					
ELBAT[1]	0.089	0.831	0.45	0.31	0.24
ELBAT[2]	0.026	0.269	0.08	0.01	0.91
Bridge of the Americas Northbound Pedestrian Traffic					
ELBAW[1]	0.273	0.760	0.01	0.04	0.95
ELBAW[2]	0.327	0.883	0.12	0.05	0.83
Paso del Norte Bridge Northbound Light Vehicle Traffic					
ELBPC[1]	0.430	0.091	0.26	0.01	0.73
ELBPC[2]	0.506	0.110	0.15	0.03	0.82
Paso del Norte Bridge Northbound Pedestrian Traffic					
ELBPW[1]	1.237	0.228	0.62	0.00	0.38
ELBPW[2]	1.469	0.272	0.74	0.00	0.26
Ysleta-Zaragoza Bridge Northbound Light Vehicle Traffic					
ELBYC[1]	0.339	0.110	0.00	0.02	0.98
ELBYC[2]	0.432	0.141	0.08	0.05	0.87
Ysleta-Zaragoza Bridge Northbound Cargo Vehicle Traffic					
ELBYT[1]	0.041	0.451	0.61	0.05	0.34
ELBYT[2]	0.035	0.406	0.02	0.02	0.96
Ysleta-Zaragoza Bridge Northbound Pedestrian Traffic					
ELBYW[1]	0.209	0.929	0.27	0.02	0.71
ELBYW[2]	0.248	0.995	0.46	0.12	0.42

[1] Previously published borderplex structural model forecast.
[2] Random walk forecast calculated as last available historical observation.

ments are observed for the structural model and the random walk data.

Table 5 reports the accuracy estimates for the international bridge data included as part of the borderplex modeling system. For the eight series modeled, results point to superior accuracy by the random walk benchmarks in only two cases. Both of those cases, however, involve cargo vehicle traffic, an increasingly important traffic category as trade liberalization occurs under the auspices of NAFTA (Orrenius et al. 2001; Fullerton and Tinajero 2002). Table 5 indicates that bias is a problem for both sets of cargo vehicle structural forecasts. Additional testing is obviously warranted for the specifications associated with both of the bridge cargo econometric specifications.

Results shown in table 5 for pedestrian and personal vehicle traffic flows from Ciudad Juárez to El Paso all point to relatively better simulation precision by the econometric model. Those outcomes are encouraging, because both categories influence retail sales performance in El Paso in noticeable ways and represent key indicators for the regional economy (Fullerton 2001). Personal vehicles are also important in terms of emissions impacts on the environment (Roderick 1993; Funk et al. 2001). As with earlier documented regional employment and income results (West and Fullerton 1996; Fullerton et al. 2004), outcomes shown in table 5 indicate that borderplex model forecasts of automobile and pedestrian categories of northbound bridge traffic are accurate relative to random walk benchmarks.

These results are encouraging, because simulations from the model are being used in transportation planning exercises conducted by the El Paso Metropolitan Planning Organization. The presence of bias in two of the pedestrian and two of the automobile sets of forecasts indicates, however, that even these equations experience simulation flaws.

Due to the small numbers of similar studies for regional transportation forecasting efforts and for other border economies, it is hard to assess whether the outcomes shown in tables 4 and 5 are unique to the borderplex economy. Given their relatively high U-coefficients, caution should be exercised with respect to using the out-of-sample air passenger and bridge cargo forecasts published using the borderplex model. At a minimum, subscribers and other users should use the latest available historical observations as "sanity checks" for those extrapolations (Fullerton and West 1998). Although the random walk approach using the latest historical observations has been presented here as a competitive benchmark, practice has shown that the information content of random walk forecasts frequently complements that contained in structural model counterparts (Granger 1996). Over time, it will become possible to assess whether structural model simulation reliability improves for these variable categories.

CONCLUSION

Transportation variables have formed integral components of the borderplex econometric forecasting effort from its inception in 1997. Included among the 218 equations in the border model are 2 blocks of transportation equations. The latter cover international bridge crossings from Ciudad Juárez as well as air traffic activity at El Paso International Airport. To examine out-of-sample forecast reliability, extrapolation accuracy is examined for those variables between 1998 and 2003.

Results indicate that the air freight, airmail, bridge auto, and bridge pedestrian series forecasts are somewhat more accurate than random walk benchmarks over the course of the sample period. Outcomes for the air passenger and bridge cargo simulations are less encouraging. In each of those cases, the random walk benchmarks obtain lower root mean square error statistics and Theil inequality coefficients. Care should be exercised when assessing the usefulness of forecasts for those variables. Future forecasts for those variables should probably be compared with the last available historical observations. That step can potentially help ensure that the model simulations do not stray too far what might be reasonably expected during multi-step prediction periods.

Border region econometric forecasting analysis is still a relatively new endeavor. As additional outlooks are published, greater numbers of observations will eventually permit more formal testing to be engaged. The sample used here is also geographically limited in scope. Replication for other border areas such as San Diego-Tijuana and Laredo-Nuevo Laredo would be helpful. Should similar efforts be carried out for other international metropolitan economies, evidence obtained for the borderplex indicates that transportation forecasting accuracy can be achieved in some cases. Because accuracy relative to random walk benchmarks is not achieved for all of the variables examined, evidence from other regions will help document whether that is a problem specific to the borderplex or one that is general in nature.

ACKNOWLEDGMENTS

Partial funding support for this research was provided by El Paso Electric Company, El Paso Metropolitan Planning Organization, Wells Fargo Bank of El Paso, National Science Foundation Grant SES-0332001, and the University of Texas at El Paso College of Business Administration. Helpful comments were provided by Peg Young, Keith Ord, Roberto Tinajero, Marsha Fenn, and three anonymous referees. Econometric research assistance was provided by Armando Aguilar and Brian Kelley.

REFERENCES

Alemán, E. 2003. Síntesis. *Global Insight Perspectivas Económicas de México*. Marzo:1.1–1.10.

Beenstock, M. and A. Vergottis. 1989. An Econometric Model of the World Market for Dry Cargo Freight and Shipping. *Applied Economics* 21:339–356.

Behravesh, N., A. Hodge, and C. Latta. 2003. So Far, So Good. *Global Insight U.S. Economic Outlook* April:1–14.

Bolton, R. 1985. Regional Econometric Models. *Journal of Regional Science* 25:130–165.

Campbell, J.Y. and N.G. Mankiw. 1987. Permanent and Transitory Components in Macroeconomic Fluctuations. *American Economic Review Papers and Proceedings* 77:111–117.

Caves, R.E. 1997. European Airline Networks and Their Implications for Airport Planning. *Transport Reviews* 17:121–144.

Cobb, S.L., D.J. Molina, and K. Sokulsky. 1989. The Impact of Maquiladoras on Commuter Flows on the Texas-Mexico Border. *Journal of Borderlands Studies* 4:71–88.

Dennis, N.P.S. 2002. Long-Term Traffic Forecasts and Flight Schedule Patterns for a Medium-Sized European Airport. *Journal of Air Transport Management* 8:313–324.

Dhane, G. and A.P. Barten. 1989. Where It All Began: The 1936 Tinbergen Model Revisited. *Economic Modelling* 6:203–219.

Fair, R.C. 1984. *Specification, Estimation, and Analysis of Macroeconometric Models*. Cambridge, MA: Harvard University Press.

Fullerton, T.M., Jr. 1998. Cross Border Business Cycle Impacts on Commercial Electricity Demand. *Frontera Norte* 10:53–66.

______. 2000. Currency Movements and International Border Crossings. *International Journal of Public Administration* 23:1113–1123.

______. 2001. Specification of a Borderplex Econometric Forecasting Model. *International Regional Science Review* 24:245–260.

Fullerton, T.M., Jr., M.M. Laaksonen, and C.T. West. 2001. Regional Multi-Family Housing Start Forecast Accuracy. *International Journal of Forecasting* 17:171–180.

Fullerton, T.M., Jr., J. Luevano, and C.T. West. 2000. Accuracy of Regional Single-Family Housing Start Forecasts. *Journal of Housing Research* 11:109–120.

Fullerton, T.M., Jr., and R. Tinajero. 2002. Cross Border Cargo Vehicle Flows. *International Journal of Transport Economics* 29:201–213.

______. 2003. Borderplex Economic Outlook: 2003–2005. *Business Report SR03-2*. El Paso, TX: University of Texas at El Paso, Border Region Modeling Project.

Fullerton, T.M., Jr., R. Tinajero, and L. Waldman. 2004. Regional Econometric Income Forecast Accuracy. *Journal of Forecasting* (forthcoming).

Fullerton, T.M., Jr., and C.T. West. 1998. Regional Econometric Housing Start Forecast Accuracy in Florida. *Review of Regional Studies* 28:15–42.

Funk, T.H., L.R. Chinkin, P.T. Roberts, M. Saeger, S. Mulligan, V.H.P. Figueroa, and J. Yarbrough. 2001. Compilation and Evaluation of a Paso del Norte Emission Inventory. *Science of the Total Environment* 276:135–151.

González, P. and P. Moral. 1995. An Analysis of the International Tourism Demand in Spain. *International Journal of Forecasting* 11:233–251.

Granger, C.W.J. 1996. Can We Improve the Perceived Quality of Economic Forecasts? *Journal of Applied Econometrics* 11:455–473.

Howry, E.P. 1969. On the Choice of Forecasting Models for Air Travel. *Journal of Regional Science* 9:215–224.

Hunt, L.C. and M.C. Snell. 1997. Comparative Properties of Local Econometric Models in the UK. *Regional Studies* 31:891–901.

Kim, C.H. 1995. Regional Forecasting Analysis for the Kyunggi Province. *Korean Economic Journal* 11:147–162.

Klein, L.R. 1969. The Specification of Regional Econometric Models. *Papers of the Regional Science Association* 23:105–115.

Leamer, E.E. 1983. Let's Take the Con Out of Econometrics. *American Economic Review* 73:31–43.

Lenze, D.G. 2000. Forecast Accuracy and Efficiency: An Evaluation of Ex Ante Substate Long-Term Forecasts. *International Regional Science Review* 23:201–226.

Leppold, K. 2002. El Paso, TX. *Global Insight U.S. Markets Metro Economies—South*. Fall:67–72.

Matthews, L. 1995. Forecasting Peak Passenger Flows at Airports. *Transportation* 22:55–72.

McCloskey, D.N. and S.T. Ziliak. 1996. The Standard Error of Regressions. *Journal of Economic Literature* 34:97–114.

Orrenius, P.M., K. Phillips, and B. Blackburn. 2001. Beating Border Bottlenecks in U.S.-Mexico Trade. *Federal Reserve Bank of Dallas Southwest Economy*. September/October: 1–8.

Pagan, A.R. 1974. A Generalised Approach to the Treatment of Autocorrelation. *Australian Economic Papers* 13:260–280.

Pindyck, R.S. and D. L. Rubinfeld. 1998. *Econometric Models and Economic Forecasts*. Boston, MA: Irwin/McGraw-Hill.

Roderick, L.M. 1993. A Computer Simulation of the Impact of the Cordova Bridge Traffic Delays on the Environment. *Journal of Environmental Science & Health, Part A* 28:1927–1946.

Sawyer, W.C. and R.L. Sprinkle. 1986. The Effects of Mexico's Devaluations and Tariff Changes on U.S. Exports. *Social Science Journal* 23:55–62.

Schneider, J.B. 1975. Self-Fulfillment of Long-Range Transportation Forecasts. *Traffic Quarterly* 29:555–569.

Shankar, R. 2003. El Paso. *Economy.com Précis: METRO*. March:1–3.

Theil, H. 1961. *Economic Forecasts and Policy,* 2nd edition. Amsterdam, Netherlands: North-Holland.

West, C.T. 1995. Regional Economic Forecasting: Keeping the Crystal Ball Rolling. *International Regional Science Review* 18:195–200.

______.2003. Structural Regional Factors that Determine Absolute and Relative Accuracy of U.S. Regional Labor Market Forecasts. *Journal of Agricultural & Applied Economics* 35 (Supplement):121–135.

West, C.T. and T.M. Fullerton, Jr. 1996. An Assessment of the Historical Accuracy of Regional Economic Forecasts. *Journal of Forecasting* 15:19–36.

Zandi, M.M. 2003. Reasons for Optimism. *Economy.com Regional Finance Review* 14:11–16. April.

APPENDIX

Equation (A1) shows how the RMSEs are computed. In (A1), Y^s is the forecast value for variable Y, Y^a is

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(Y_t^s - Y_t^a)^2} \tag{A1}$$

the actual historical value for Y, and T is the total number of forecasts for Y.

Equation (A2) provides the details for calculating the U-statistics. The

$$U = \frac{\sqrt{\frac{1}{T}\sum_{t=1}^{T}(Y_t^s - Y_t^a)^2}}{\sqrt{\sum_{t=1}^{T}(Y_t^s)^2} + \sqrt{\sum_{t=1}^{T}(Y_t^a)^2}} \tag{A2}$$

denominator in (A2) causes inequality coefficients to vary between 0 and 1. When $U = 0$, $Y_t^s = Y_t^a$ for all t and a perfect fit is obtained. At the other extreme, if $U = 1$, the predictive performance of the model cannot be any worse (Pindyck and Rubinfeld 1998).

Equation (A3) illustrates the formulae for the second moment inequality proportions. U^M, U^S, and U^C represent bias, variance, and covariance proportions,

$$U^M = \frac{\left(\overline{Y}_t^s - \overline{Y}_t^a\right)^2}{(1/T)\sum_{t=1}^{T}(Y_t^s - Y_t^a)^2},$$

$$U^S = \frac{(\sigma_s - \sigma_a)^2}{(1/T)\sum_{t=1}^{T}(Y_t^s - Y_t^a)^2},$$

and

$$U^C = \frac{2(1-\rho)\sigma_s\,\sigma_a}{(1/T)\sum_{t=1}^{T}(Y_t^s - Y_t^a)^2} \tag{A3}$$

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 0. The covariance proportion measures unsystematic error. As simulation accuracy improves, the covariance proportion approaches 1. As noted by Theil (1961), the optimal distribution of the second moment inequality proportions is $U^M = 0$, $U^S = 0$, and $U^C = 1$.

Estimating Commodity Inflows to a Substate Region Using Input-Output Data: Commodity Flow Survey Accuracy Tests

LOUIE NAN LIU[1]
PIERRE VILAIN[2, *]

[1] College of Management
Zhejiang University
Hangzhou 310027
People's Republic of China

[2] The Louis Berger Group
199 Water Street
New York, NY 10038

ABSTRACT

This paper describes a methodology to estimate current U.S. commodity inflows to a substate region using a supply-side, commodity-by-industry, input-output model and commodity flow data for U.S. states. Because the 1993 Commodity Flow Survey does not capture data below the state level, the estimation of commodity flows to a particular substate region of the United States has always proven difficult. By combining state-level commodity flow data with the supply-side, commodity-by-industry, input-output model, an estimate of commodity flows to smaller regions can be carried out entirely based on the regional industrial structure. Since the actual substate flows are typically unobserved, the accuracy of the methodology is unknown. However, by applying the same methodology to larger regions, with actual states used as the forecast region, the estimates can be compared with actual flows while maintaining an acceptable level of accuracy.

INTRODUCTION

A typical problem faced by transportation planners is being able to anticipate the need for expanded or new transportation infrastructure, facilities, or services. Estimates of freight flows between regions can

Email addresses:
* Corresponding author—pvilain@louisberger.com
L.N. Liu—nliu@mail.hz.zj.cn

KEYWORDS: Estimating commodity flows, freight planning, input-output model applications.

provide much-needed information for decisionmaking. In the United States, estimates of freight flows exist between individual states,[1] but little data exist for flows between areas below the state level, which we refer to as substate regions. Estimates of freight flows between substate regions could be generated based on costly direct surveys or by using secondary sources of data to infer patterns based on characteristics of the areas in question (Holguín-Veras 2000; Ortúzar and Willumsen 1994). The approach described here falls into the second category, deriving estimated freight flows from secondary data on the region's industrial structure. As these data are often readily and cheaply available (e.g., in the United States down to the county level), this approach is both simple and cost-effective.

In general, the process of estimating freight *outflows* from substate regions is fairly simple. Relatively accurate estimates can be produced based on data on the region's industrial structure and its state-to-state trade. By mechanically assigning freight commodity exports to the producing industries, estimates can be made of the share of a substate region's state exports based on the presence of these industries.

However, the estimation of the second category, freight *inflows*, is considerably more complicated. While we can roughly assign the production of commodities to certain industries, the consumption of commodities by various industries requires far more detailed knowledge of their input use. Fortunately, this type of information is readily available in input-output models, and some simple manipulations of standard input-output data yield a tool that can then be used to assign state-level commodity inflows to any substate region.

In the following sections we outline a methodology to estimate commodity inflows to smaller regions, which was initially described in Vilain et al. (1999). The methodology was devised specifically to regionalize inflows to substate regions in the United States, but it has also been used in other countries. In general, the methodology can be used in any country or region, given the availability of the requisite data on input-output accounts described below.

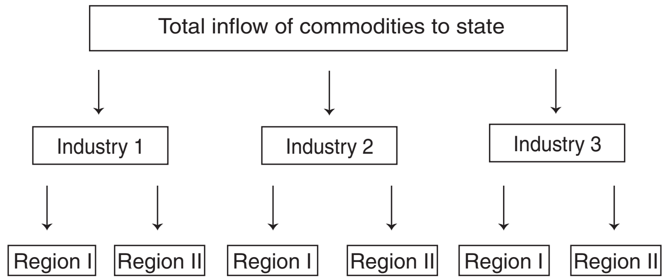
[1] The U.S. Department of Commerce's 1993 Commodity Flow Survey contains state-level data on commodity flows.

Having proposed a methodology to estimate freight inflows that is simple to use, it is of interest to examine the accuracy of the technique. In this paper, we carry out a series of simulations that we then test for their predictive accuracy. The key to being able to determine the accuracy of simulations is to carry them out for states as if they were smaller substate regions. Since states are regions for which commodity flow data does exist, we can then compare the predicted inflows with the actual observed inflows. Our results show that, excluding inflows of mining, petroleum, or coal products, the methodology leads to relatively accurate forecasts. Total inflows of all commodities to a state are typically predicted within 10% error, but the accuracy of forecasts for individual commodities is far more variable. Despite the mixed results, we argue that the methodology described here is valid, yielding predictions of commodity inflows that have an acceptable level of accuracy. The relative accuracy of the methodology must also be considered keeping in mind that, in the absence of expensive origin-destination surveys, there are really no alternatives that yield reliable estimates of commodity inflows.

SUPPLY-SIDE INPUT-OUTPUT MODEL AND COMMODITY FORECASTING

The gravity model is a widely used technique for estimating commodity inflows to a region. In this approach, observed freight flows between areas are encouraged by demand factors (e.g., concentrations of population) and accessibility, while transportation costs between regions act to inhibit such flows. Gravity models have been applied extensively to the analysis of passenger trip generation, and examples exist of their application to freight demand modeling (Ortúzar and Willumsen 1994). In terms of our problem of predicting commodity inflows to a substate area, this model could be estimated at the state level and the estimated parameters used to predict inflows to a substate region. However, data requirements to calibrate such a model (notably transportation costs) are significant. This same conclusion applies to other, closely related models based on discrete choice analysis, including disaggregated freight generation models.

FIGURE 1 Commodity Inflow



We propose an alternative approach here, one that bases estimates of actual commodity inflows to a substate region entirely on the region's industrial structure. The details of the industrial structure are themselves obtained from regional input-output data. The procedure can be carried out fairly easily, relying entirely on published national input-output data, existing state-level commodity flow data from the 1993 Commodity Flow Survey (USDOC 1993), and regional data on employment or earnings by industry.

The procedure involves two steps. First, using regional input-output data (USDOC 1997), we define the proportion of commodities used by various industries in a region of interest. Then we apply these proportions to existing state-level commodity inflow data from the 1993 Commodity Flow Survey to share down the state-level flows to the region. One significant advantage of the methodology is that it takes into account the possibility that the input needs of a regional industry are met, in whole or in part, by regional suppliers. By accounting for existing patterns of regional inter-industry freight flows, the accuracy of estimated regional freight inflows is presumably increased greatly.[2]

The procedure can be represented schematically. In essence, data on commodity inflows to a region (e.g., a state) are divided into the various industries (including households) that are the likely users of these commodities as inputs. Once the inflows have been divided among the various inflow-consuming industries, they are then disaggregated to the appropriate substate regions based on their industrial structure. Let us suppose there are three industries and two substate regions, called I and II. This would then produce an assignment of commodity inflows that would follow the pattern shown in figure 1.

To describe the supply-side, commodity-by-industry model, consider a set of accounts that details the sales of each commodity to the various industries that use it as inputs in production as well as sales of that commodity to final demand. (Details on the input-output accounts we describe are contained in the literature; see, e.g., Miller and Blair 1985.) For each of the commodities consumed in the economy we can write the following equation:

$$q_i = u_{i1} + u_{i2} + \dots + u_{in} + e_i \tag{1}$$

Equation (1) defines an identity, namely that the total production of commodity i is equal to the sales of that commodity to each of the n industries in the economy (e.g., u_{i2} is sales of commodity i to Industry 2) and commodity sales to final demand, e_i. In input-output accounts, final demand is consumption by households and governments as well as investment expenditures and the difference between imports and exports.

If there are m commodities being produced and consumed in the economy, we can represent all sales of commodities to industries as a matrix of dimensions m x n:

$$U = [u_{ij}] \tag{2}$$

$\mathbf{U}$ is composed of commodity sales to industries, with each u_{ij} representing the amount of commodity i (expressed in monetary units) used by industry j as an input in its production. In other words, each of the m rows of $\mathbf{U}$ details the total industrial destina-

[2] This aspect of the methodology contrasts with the approach suggested by Memmott (1983). While also based on input-output models, his suggested procedure for estimating regional freight flows does not account for the possibility of freight inflows being supplied regionally. As a result, the applicability of the approach for accurately estimating inflows from outside the region is limited.

tions of each of the m commodities represented in the accounts.

We then transform $\boldsymbol{U}$ into a matrix $\boldsymbol{\beta}$ whose elements are those in $\boldsymbol{U}$ divided by their row sum. Formally, this is defined as

$$b_{ij} = \frac{u_{ij}}{q_i} \tag{3}$$

where b_{ij} is equal to the share of commodity i sold to industry j. In matrix terms, the derivation of $\boldsymbol{\beta}$ is obtained with the following simple operation:

$$\beta = (\hat{Q})^{-1} U \tag{4}$$

where $\boldsymbol{Q}$ is an m by 1 vector of all commodity gross outputs as individually defined in equation (1), and ^ indicates a diagonalized matrix.

By dividing row elements by their total production for industrial or final demand uses, we obtain the commodity-by-industry equivalent of the "supply-side" input-output model (Augustinovics 1970). With each b_{ij} element of $\boldsymbol{\beta}$ representing the share of commodity i sold to industry j, the information in $\boldsymbol{\beta}$ will allow us to disaggregate state-level commodity inflows to the appropriate industries that use the commodities as inputs. Several further steps are required to do so.

The matrix $\boldsymbol{\beta}$, which is a matrix representing national data, must be regionalized to the state level. In order to share commodity inflows to the regional level, a procedure based on *location quotients* is used.[3] We define a simple state-level location quotient as the relative representation of a national commodity-producing industry in a particular state s:

$$l_i^s = \frac{Earnings_i^s / Earnings^s}{Earnings_i^N / Earnings^N} \tag{5}$$

$Earnings_i$ is earnings in the industry that produce commodity i in either state s (indicated by a superscript s) or in the nation (superscript N). $Earnings^s$ is total regional earnings and $Earnings^N$ is total national earnings. Note that the location quotient could also be based on employment data.

[3] *Location quotients* are widely used as a method of regionalizing national data, in particular input-output data. The measure indicates the relative concentration of an industry in a region, where values for equation (5) that are larger than 1 indicate a greater than average concentration and values less than 1 the opposite.

Location quotients are calculated for each of the n industries producing the m commodities in $\boldsymbol{\beta}$. These are then used to regionalize the elements of $\boldsymbol{\beta}$. Generating a vector of n state-level location quotients for state s, $\boldsymbol{L_s}$, the following multiplication is carried out:

$$\beta_s = \beta \hat{L}_s \tag{6}$$

where ^ again indicates the diagonalized matrix formed from the vector $\boldsymbol{L_s}$. Each element of $\boldsymbol{\beta_s}$ adjusts the national values of $\boldsymbol{\beta}$ downward if the state contains a presence of the industry that is less than the national average. Specifically, each element of $\boldsymbol{\beta_s}$ is equal to

$$b_{ij\,s} = b_{ij}\, l_j^s \tag{7}$$

A final adjustment is then carried out. The row sums of $\boldsymbol{\beta_s}$ (as opposed to each $b_{ij\,s}$ element of $\boldsymbol{\beta_s}$) should then be adjusted to equal 1. The reason for this is simple. Because the matrix will be used to apportion freight flows to different industries, we are interested in the relative values of the elements of $\boldsymbol{\beta_s}$ rather than their absolute values. To ensure that row sums equal 1, we carry out a balancing procedure:

$$c_{ij\,s} = b_{ij\,s} \frac{1}{\sum_j b_{ij\,s}} \tag{8}$$

This balancing procedure now ensures that the row sums of a new matrix, $\boldsymbol{C_s}$, sum to 1. This procedure is necessary in order to ensure that all commodity inflows to state s can be assigned an end user. It essentially reflects the following assumption: if an industry, say industry j, is not present in state s, the inflows of any commodity that it uses as an input are simply assumed to be used by other industries that are both present in the state and use the commodity as an input.

This same procedure is also carried out if industry j is present in the state but its presence is below the national average; whatever inputs are not used by industry j are simply allocated to all the other industries that use the commodity and are present in state s.

Each $c_{ij\,s}$ element of matrix $\boldsymbol{C_s}$ now can be said to approximate the *proportion* of commodity i that is shipped to state s that will be used by industry j. In other words, $\boldsymbol{C_s}$ directs the commodities entering state s to the industries that can be expected to use the commodities as inputs. Mathematically, the operation involves a simple post-multiplication of the state-level commodity inflows by $\boldsymbol{C_s}$, resulting in a disaggregation of these inflows into the industries that use them as inputs. If we define the vector ϕ_s that contains the inflows of the m commodity to state s, we perform the following matrix multiplication:

$$\rho_s = (\hat{\phi}_s) C_s \tag{9}$$

Again, ^ indicates the vector ϕ_s is converted to a diagonalized matrix. The operation produces the matrix $\boldsymbol{\rho_s}$ of dimension m x n, which apportions freight inflows among the state industries that will use them as inputs. Specifically, each $\rho_{ij\ s}$ element of matrix $\boldsymbol{\rho_s}$ details the *amount* of commodity i flowing to industry j in state s.

To further regionalize these flows to the substate level, another procedure needs to be carried out. In a manner similar to the previous regionalization, we calculate a matrix of regional earnings shares, $\boldsymbol{L_{region}}$, which measures the relative representation of each industry in the substate region. Multiplying $\boldsymbol{\rho_s}$ by a matrix produced from diagonalizing the vector $\boldsymbol{L_{region}}$ produces the matrix $\boldsymbol{\rho_{reg}}$.

$$\begin{aligned} \rho_{reg} &= \rho_s \hat{L}_{reg} \\ &= (\phi_s C_s) \hat{L}_{reg} \end{aligned} \tag{10}$$

Each $\rho_{ij\ reg}$ element of the matrix $\boldsymbol{\rho_{reg}}$ gives an approximation of the amount of a commodity shipped to state s that is used by a regional industry.[4] The state-level commodity inflows are, thus, directed to a substate region, depending on the location of industries using the commodities as inputs. Any row sum of $\boldsymbol{\rho_{reg}}$ gives an estimate of the total amount of a given commodity that is shipped to the region. The resulting vector of estimated regional inflows is denoted as ϕ_{reg} and the total inflow of any given commodity as $\phi_{i\ reg}$.

$$\phi_{i\ reg} = \sum_j \rho_{ij\ reg} \tag{11}$$

[4] Note that because L_{reg} contains simple regional shares of an industry, no balancing procedure is required. Our procedure differs from previous regionalizations in that we have already apportioned commodities to industries in state s and only need to share the flows between regions in the state based on the presence of the industry.

An important assumption embodied in the use of $\boldsymbol{\rho_{reg}}$ is that each regional industry that uses a given commodity as an input will use it in the same proportions as the industry nationally. In other words, it is assumed that local industries use commodity inputs in relation to the relative proportions in $\boldsymbol{\beta}$, a standard assumption when regionalizing national input-output flows with location quotients.

Another assumption implicit in the methodology is that all firms purchase locally produced commodity inputs in the same proportions. For example, if commodity i is produced in the state and satisfies 10% of local state needs, it is assumed that all firms that use commodity i will purchase 10% of their input needs locally. This assumption can presumably create bias in estimates of regional inflows. To the degree that the local production of i is concentrated in certain substate regions, some local industries might purchase more than 10% of their needs from the local state suppliers. Finally, in addition to assuming that all firms purchase locally produced inputs in the same proportions, the methodology further assumes that industries purchase their extra-regional inputs from any given region in the same proportion.

ACCURACY TESTS

Having described a relatively simple methodology to estimate freight inflows to a substate region, we want to determine its accuracy. As mentioned previously, the approach suggested here is intended for estimates of freight flows to substate regions where, by definition, little or no data exist to permit validation of the estimates. This would imply that validating the results of the methodology would require actual survey data on freight inflows to the region. The lack of such surveys for small regions is precisely what motivated the elaboration of the supply-side, commodity-by-industry methodology.[5]

An alternative approach to determining the accuracy of the methodology is possible, however. This involves treating states as if they were substate regions and creating larger regions comprised of a series of individual states. Then, the total freight

[5] An exception is the for-fee Transearch® freight data provided by Reebie Associates for the United States (available at www.reebie.com, as of Sept. 27, 2004). Future research could carry out accuracy tests based on this data.

inflows to these several states can be used as if they were inflows to an individual state. In so doing, one must be careful to remove the freight flows between the various states that make up the larger region. The result is data detailing all inflows of commodities to the larger region from outside this region.

It should be pointed out that the Commodity Flow Survey is comprehensive in that all modes are covered. For the 1993 Commodity Flow Survey, the U.S. Census Bureau used a sample of 200,000 establishments in manufacturing, mining, wholesale, and retail.[6] Each establishment was asked to report shipments for two-week periods in each of the four calendar quarters identifying domestic origin and destination, commodity type, weight, value, and mode of transport. The Commodity Flow Survey does exclude certain commodities, notably crude petroleum. Also, while imports and exports are included, commodities shipped from a foreign location through the United States to another destination are excluded.

In carrying out our tests, we selected four large regions in the United States that each contain a number of states. The regions are as follows:

- **Northeast Region:** Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, and Vermont;
- **Middle Atlantic Region:** Delaware, Maryland, New Jersey, New York, and Pennsylvania;
- **Great Lakes Region:** Illinois, Indiana, Michigan, Ohio, and Wisconsin; and
- **West Coast Region:** California, Nevada, Washington, Oregon, Alaska, and Hawaii.

We applied the procedure to seven states: Massachusetts, New York, Pennsylvania, Ohio, Illinois, California, and Washington.

For each of these four regions, we estimated a measure equivalent to $\boldsymbol{\rho}_s$ for the entire group of component states, as defined in equation (9). For the purposes of this analysis, the seven states we analyzed were treated as if they were substate regions. For each of these states, both $\boldsymbol{\rho}_{reg}$ and ϕ_{reg} were calculated according to the definitions of equations (10) and (11), as if the detailed state-level data in the Commodity Flow Survey did not exist.

How do these estimates compare with actual freight inflows to the states? Details of the estimates and the actual observed inflows for each of the seven states are reported in appendix tables A1 to A7, along with the percentage error of the forecasts. In general, the methodology performs well for *total* forecasts of different commodities, but the forecast of specific commodities is variable. Total commodity inflows to a state are forecast within 10% accuracy for all states except California and Ohio. For example, in the case of Massachusetts, the forecast error for total commodity inflows is 9.6% below the actual observed inflow. This figure, however, obscures the fact that while some commodities are forecast with less than 5% error, others are forecast with as much as 56% error (e.g., transportation equipment). This is due to the fact that a simple summation of the percentage error of individual commodities will see negative and positive forecast errors canceling each other.

Two commodities, mineral products and petroleum and coal products, tended to predict very poorly (in the case of New York, e.g., the forecast was off by over 800%) and were not included in tables A2 through A7. This can be partly explained by the different patterns of energy consumption in various regions of the United States. In particular, the use of such energy sources as oil, coal, hydroelectric, and nuclear power can vary across regions regardless of industries.[7] Because of the consistently large error in predicting these commodities, we do

[6] We rely on the older 1993 Commodity Flow Survey rather than the more recent 1997 Commodity Flow Survey, because the sample size was twice as large in the earlier survey, which we believe increases its reliability.

[7] The bias of assuming national patterns of energy use or production to regions has been discussed by other authors, in particular Miller and Blair (1985) wrote: "Electricity produced in Eastern Washington by water power (Coulee Dam) represents quite a different mix of inputs from electricity that is produced from coal in the greater Philadelphia area or by means of nuclear power elsewhere." They allude to a problem inherent in using national input-output data regionalized on the basis of nonsurvey techniques. This issue also affects the procedure we have suggested for estimating commodity flows. Because the methodology presented here relies on national input-output data, it will tend to assume that energy sources reflect the national "average."

not include them in our discussion and note that our methodology is inappropriate to forecast them.

In general, simply averaging the percentage error of individual commodities will be a poor measure of overall accuracy that will tend to overstate the accuracy of the methodology. Because individual commodities will generate both negative and positive forecast error, these will tend to cancel each other in a simple averaging over the sample. To account for the presence of both negative and positive forecast error, we relied on weighted average error (WAE) and mean absolute error (MAE). The definitions of the measures for the *m* commodities are:

$$WAE = \sum_{i=1}^{m} \frac{|Estimated_i - Observed_i|}{Observed_i} RelativeWeight_i$$

$$MAE = \frac{1}{m}\sum_{i=1}^{m} \frac{|Estimated_i - Observed_i|}{Observed_i}$$

Tables A1 through A7 report these measures. The WAE ranged from 16.8% to 29.1%, depending on the state. The MAE ranged from 15.6% to 71%, with the latter a relatively extreme result for the state of Washington and uncharacteristic of the sample.

The tables include a measure of the relative distance the commodity is being shipped (*distance ratio*) derived from data in the Commodity Flow Survey. The measure relates the average shipping distance for a commodity to a state relative to the average shipping distance for that commodity nationally. In other words, if commodity *i* when shipped to Massachusetts travels an average of 500 miles and the national average for the commodity is 250 miles, the distance ratio will be equal to 2.

One reason for measuring the distance ratio is that geography may well play a role in the export activity of firms. Specifically, as mentioned earlier, we assumed by necessity that firms all purchase locally produced commodity inputs in the same proportions. But geography may encourage different patterns of local versus nonlocal sales: if the transportation costs to the next largest concentrations of potential purchasers are great, firms may be particularly oriented to their local market. If the costs to potential nonlocal purchasers are not high, firms may be shipping outside the immediate region to a greater degree.

A cursory glance at the results in tables A6 and A7 suggests that states with larger distance ratios, in this case California and Washington, tend to have greater commodity forecast errors. In order to test for the effect of distance on shipments, and potentially on the accuracy of our method, we included distance ratio in a simple regression that measured the explanatory power of this variable on the accuracy of forecasts. In essence, we wanted to find out if a large deviation of the distance ratio from one is associated with an increased forecast error.[8]

Similarly, we were also interested in the effect of commodity volume on forecast accuracy. If the actual volume of a specific commodity shipped to a particular state is low, will this lower the forecast accuracy? We tested for both effects in a multivariate regression where observed tonnage shipped and distance ratio by commodity were regressed on the forecast error of the particular commodity. Table 1 presents the results of this regression. While the coefficients show the expected signs, both are only significant at the 15% confidence level. We interpreted this result to mean that there is no significant inherent bias in the forecast method due to the distance of shipments or the actual tonnage of commodities shipped.[9]

CONCLUSIONS

The procedure described above offers a relatively easy tool to estimate substate commodity inflows, one that can be used by transportation planners for relatively accurate "back of the envelope" predictions of aggregate commodity inflows to smaller regions. Further, the procedure has the important

[8] We thank an anonymous referee for suggesting a test for the effect of distance on forecast accuracy. Note that the Commodity Flow Survey does not allow us to calculate distance ratios for all commodities forecast, as distance estimates are not always reported for all commodities shipped to all states.

[9] It should be pointed out that the relatively aggregated commodity classifications dealt with here result in very few commodity shipments of small tonnage. Forecast error may be significantly related to commodity tonnage below certain thresholds. We also regressed the forecast error for a commodity on a variable that represents the importance of that commodity in total inflows to the state. The results, in table 2, are again of the expected sign, with the relative importance of a commodity as a percentage of shipments reducing error.

TABLE 1 Regression Results for the Effect of Distance and Commodity Tonnage on Forecast Accuracy

Dependent variable: commodity forecast error

Independent variable	Coefficient	*t*-statistic
Constant	0.24	2.4
Distance ratio for commodity	0.10	1.4
Observed inflows of commodity to state	–1.65	–1.4

Adjusted R^2: 0.03

Note: Dependent variable measured in absolute value terms; observed inflows as reported in appendix tables A1–A7.

TABLE 2 Regression Results for the Effect of Distance and the Relative Importance of a Commodity on Forecast Accuracy

Dependent variable: commodity forecast error

Independent variable	Coefficient	*t*-statistic
Constant	0.23	2.4
Distance ratio for commodity	0.11	1.5
Inflows as % of total inflows to state	–0.86	–1.6

Adjusted R^2: 0.04

Note: Dependent variable measured in absolute value terms; inflows as a percentage of total inflows derived from observed inflows by commodity and total state inflows.

advantage of using the appropriate *observed* state-level commodity inflows as a starting point to estimate substate flows, something that cannot be claimed by econometric or gravity models that generalize inflow patterns observed in one region to another region. Though somewhat laborious, the calculations are relatively simple, using data that are widely available and low cost, at least in the United States and European Union countries.

While estimates of total freight inflows were in some cases surprisingly accurate, the estimate errors of individual commodities were often significantly greater. In particular, commodities, such as energy inputs, whose use could vary significantly across regions in the United States, were predicted very poorly. Excluding these commodities, the MAE for all commodities to all states is 31% while the corresponding WAE is 21%, arguably acceptable imprecision for the suggested uses of the approach.

As discussed, the method entails two crucial assumptions. First, all firms at the state and regional level are assumed to display the same input use as their counterparts nationally, a necessary assumption in nonsurvey regional input-output modeling. This assumption appears to be a significant flaw in the estimate of inflows of energy inputs, as mentioned. Second, all firms in a regional industry are assumed to purchase locally produced commodity inputs in the same proportions. This could introduce bias, particularly in the case of large states where firms located near a local supplier of a given commodity could consume significantly more inputs produced locally than those located farther away from the supplier.

Our method cannot differentiate these differences among firms in the same industry. This could in turn lead to overestimates of inflows of a given commodity to regions with an important local producer of that commodity. Conversely, it could also lead to underestimates of inflows to the other regions. As opposed to gravity models, for example, our approach does not incorporate distance as a potential influence on trade flows. The econometric analyses reported in tables 1 and 2 indicate distance may affect our model's accuracy, although the significance of this bias appears modest in our accuracy tests.

Despite this imperfection, it is argued that the method of estimating substate inflows using input-output data is sound. We further argue that, in the absence of detailed and costly surveys, our approach estimates the most elusive component of regional trade, commodity inflows, with acceptable levels of accuracy.

REFERENCES

Augustinovics, M. 1970. Methods of International and Intertemporal Comparisons of Structure. *Contributions to Input-Output Analysis.* Edited by A.P. Carter and A. Brody. Amsterdam, The Netherlands: North Holland.

Holguín-Veras, J. 2000. A Framework for an Integrative Freight Market Simulation, paper presented at the IEEE/ITSC Conference, Dearborn, MI. October.

Memmott, F.W. 1983. *Application of Statewide Freight Demand Forecasting Techniques,* National Cooperative Highway Research Program Report 260. Washington, DC: Transportation Research Board and National Research Council.

Miller, R.E. and P. D. Blair. 1985. *Input-Output Analysis: Foundations and Extensions*. Englewood Cliffs, NJ: Prentice-Hall.

Ortúzar, J.D. and L.G. Willumsen. 1994. *Modelling Transport*. Chichester, England: John Wiley & Sons.

U.S. Department of Commerce (USDOC), Census Bureau. 1993. *1993 Commodity Flow Survey*, CD-CFS-93-1. Washington, DC. December.

_____. Bureau of Economic Analysis. 1997. *Input-Output Accounts of the US Economy, 1992 Benchmark*. Washington, DC. November.

Vilain, P., L.N. Liu, and D. Aimen. 1999. Estimation of Commodity Inflows to a Substate Region: An Input-Output Based Approach. *Transportation Research Record 1653*. Washington, DC: Transportation Research Board and National Research Council.

APPENDIX

TABLE A1 Estimates of Annual Domestic Commodity Inflows to Massachusetts
Excluding energy[1]

Commodity	Observed inflows (000s of metric tons)	Estimated inflows (000s of metric tons)	Percentage error	Distance ratio[2]
Apparel or other finished textile products	238	204	–16.8	0.8
Chemicals or allied products	3,332	2,519	–32.3	1.1
Clay, concrete, glass, or stone products	1,512	1,777	14.9	1.4
Electrical machinery, equipment, or supplies	276	269	–2.6	1.3
Fabricated metal products	644	585	–10.1	0.7
Farm products	259	228	–13.5	1.6
Food or kindred products	6,931	6,682	–3.7	1.3
Furniture or fixtures	244	219	–11.5	1.3
Instruments, photographic goods, optical goods, watches, or clocks	33	26	–26.5	1.1
Leather or leather products	41	36	–13.2	N/A
Lumber or wood products, excluding furniture	1,340	1,013	–32.3	4.5
Machinery, excluding electrical	221	195	–13.6	N/A
Miscellaneous products of manufacturing	333	269	–23.9	1.1
Primary metal products	1,051	1,351	22.2	1.0
Pulp, paper, or allied products	2,713	2,395	–13.2	1.3
Rubber or miscellaneous plastics products	557	483	–15.2	1.5
Textile mill products	391	337	–15.9	N/A
Transportation equipment	903	578	–56.1	1.1
Waste or scrap materials	32	36	10.7	N/A
Total, all commodities	**21,052**	**19,203**	**–9.6**	**1.1**

Weighted average error: 16.8%

Mean absolute error: 15.6%

[1] Energy commodities include mining, petroleum, and coal products.
[2] Distance ratio is the average shipment distance for the commodity to the state divided by the national average.

Key: N/A = not applicable.

Note: Observed inflows are obtained from the 1993 Commodity Flow Survey. Estimates are derived from the inflows to the six-state Northeastern Region (Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, and Vermont).

TABLE A2 Estimates of Annual Domestic Commodity Inflows to New York
Excluding energy[1]

Commodity	Observed inflows (000s of metric tons)	Estimated inflows (000s of metric tons)	Percentage error	Distance ratio[2]
Apparel or other finished textile products	429	442	2.9	0.8
Chemicals or allied products	6,508	7,328	11.2	0.8
Clay, concrete, glass, or stone products	3,079	2,769	–11.2	0.7
Electrical machinery, equipment, or supplies	1,067	1,037	–2.9	1.0
Fabricated metal products	2,264	1,901	–19.1	1.3
Farm products	1,222	3,880	68.5	1.8
Food or kindred products	14,785	15,421	4.1	0.5
Forest and fishing products	113	122	7.7	0.8
Furniture or fixtures	570	654	12.9	0.8
Instruments, photographic goods, optical goods, watches, or clocks	202	180	–11.7	N/A
Leather or leather products	43	40	–9.5	N/A
Lumber or wood products, excluding furniture	7,962	4,103	–94.1	3.3
Machinery, excluding electrical	848	830	–2.2	N/A
Miscellaneous products of manufacturing	528	446	–18.4	0.8
Primary metal products	4,358	4,167	–4.6	N/A
Pulp, paper, or allied products	7,241	6,969	–3.9	0.6
Rubber or miscellaneous plastics products	1,381	1,309	–5.5	1.1
Textile mill products	1,161	971	–19.6	N/A
Tobacco products, excluding insecticides	17	27	35.1	0.6
Transportation equipment	2,872	2,437	–17.9	0.8
Waste or scrap materials	408	1,053	61.2	0.9
Total, all commodities	**57,060**	**56,086**	**–1.7**	**0.8**

Weighted average error: 21.5%

Mean absolute error: 30.1%

[1] Energy commodities include mining, petroleum, and coal products.
[2] Distance ratio is the average shipment distance for the commodity to the state divided by the national average.

Key: N/A = not applicable.

Note: Observed inflows are obtained from the 1993 Commodity Flow Survey. Estimates are derived from the inflows to the five-state Middle Atlantic Region (Delaware, Maryland, New Jersey, New York, and Pennsylvania).

TABLE A3 Estimates of Annual Domestic Commodity Inflows to Pennsylvania
Excluding energy[1]

Commodity	Observed inflows (000s of metric tons)	Estimated inflows (000s of metric tons)	Percentage error	Distance ratio[2]
Apparel or other finished textile products	254	232	–9.9	0.8
Chemicals or allied products	7,125	5,819	–22.4	0.9
Clay, concrete, glass, or stone products	3,046	2,579	–18.1	0.7
Electrical machinery, equipment, or supplies	675	604	–11.7	N/A
Fabricated metal products	1,511	1,541	1.9	0.8
Farm products	1,523	3,788	59.8	0.7
Food or kindred products	12,683	10,131	–25.2	0.8
Forest and fishing products	93	62	–50.1	0.6
Furniture or fixtures	407	380	–7.3	N/A
Instruments, photographic goods, optical goods, watches, or clocks	92	98	6.1	N/A
Leather or leather products	28	23	–22.9	0.8
Lumber or wood products, excluding furniture	2,618	4,073	35.7	0.7
Machinery, excluding electrical	677	558	–21.4	N/A
Miscellaneous products of manufacturing	281	245	–14.7	N/A
Primary metal products	7,860	6,405	–22.7	0.7
Pulp, paper, or allied products	5,724	4,412	–29.7	0.6
Rubber or miscellaneous plastics products	973	920	–5.7	1.0
Textile mill products	553	522	–5.8	0.8
Tobacco products, excluding insecticides	14	15	6.1	0.6
Transportation equipment	1,305	1,432	8.9	N/A
Waste or scrap materials	1,873	703	–166.3	N/A
Total, all commodities	**49,315**	**44,542**	**–10.7**	**0.7**

Weighted average error: 29.1%

Mean absolute error: 24.1%

1 Energy commodities include mining, petroleum, and coal products.
2 Distance ratio is the average shipment distance for the commodity to the state divided by the national average.

Key: N/A = not applicable.

Note: Observed inflows are obtained from the 1993 Commodity Flow Survey. Estimates are derived from the inflows to the five-state Middle Atlantic Region (Delaware, Maryland, New Jersey, New York, and Pennsylvania).

TABLE A4 Estimates of Annual Domestic Commodity Inflows to Ohio
Excluding energy[1]

Commodity	Observed inflows (000s of metric tons)	Estimated inflows (000s of metric tons)	Percentage error	Distance ratio[2]
Apparel or other finished textile products	300	285	–5.2	0.8
Chemicals or allied products	12,893	9,612	–25.5	0.9
Clay, concrete, glass, or stone products	4,594	3,908	–14.9	0.9
Electrical machinery, equipment, or supplies	707	697	–1.5	0.9
Fabricated metal products	1,483	1,192	–19.6	0.7
Farm products	2,306	2,585	12.1	0.5
Food or kindred products	20,585	13,676	–33.6	0.7
Forest and fishing products	23	12	–50.0	0.9
Furniture or fixtures	440	361	–18.1	0.7
Instruments, photographic goods, optical goods, watches, or clocks	161	88	–45.2	0.9
Leather or leather products	23	34	45.3	0.8
Lumber or wood products, excluding furniture	3,655	2,919	–20.1	0.8
Machinery, excluding electrical	918	648	–29.4	0.9
Miscellaneous products of manufacturing	323	326	1.0	0.9
Ordnance or accessories	15	10	–32.6	N/A
Primary metal products	6,013	3,812	–36.6	0.9
Pulp, paper, or allied products	7,506	5,872	–21.8	0.8
Rubber or miscellaneous plastics products	1,138	1,022	–10.2	0.8
Textile mill products	323	280	–13.3	0.8
Tobacco products, excluding insecticides	43	14	–66.8	0.6
Transportation equipment	1,265	1,408	11.3	1.0
Waste or scrap materials	1,117	882	–21.0	1.6
Total, all commodities	**65,831**	**49,641**	**–24.6**	**0.7**

Weighted average error: 25.9%

Mean absolute error: 24.3%

[1] Energy commodities include mining, petroleum, and coal products.
[2] Distance ratio is the average shipment distance for the commodity to the state divided by the national average.

Key: N/A = not applicable.

Note: Observed inflows are obtained from the 1993 Commodity Flow Survey. Estimates are derived from the inflows to the five-state Great Lakes Region (Illinois, Indiana, Michigan, Ohio, and Wisconsin).

TABLE A5 Estimates of Annual Domestic Commodity Inflows to Illinois
Excluding energy[1]

Commodity	Observed inflows	Estimated inflows	Percentage error	Distance ratio[2]
	(000s of metric tons)			
Apparel or other finished textile products	276	257	–6.7	0.7
Chemicals or allied products	10,705	9,858	–7.9	0.9
Clay, concrete, glass, or stone products	4,372	4,014	–8.2	0.6
Electrical machinery, equipment, or supplies	677	598	–11.7	0.6
Fabricated metal products	1,269	1,272	0.2	0.7
Farm products	646	1,709	164.5	0.9
Food or kindred products	7,956	10,725	34.8	1.0
Forest and fishing products	5	8	61.0	0.7
Furniture or fixtures	338	324	–4.4	0.7
Instruments, photographic goods, optical goods, watches, or clocks	55	75	38.0	0.7
Leather or leather products	43	28	–35.0	0.7
Lumber or wood products, excluding furniture	3,309	2,910	–12.1	1.0
Machinery, excluding electrical	610	669	9.7	0.6
Miscellaneous products of manufacturing	278	261	–6.0	0.6
Ordnance or accessories	15	8	–42.7	N/A
Primary metal products	6,710	5,591	–16.7	0.7
Pulp, paper, or allied products	4,117	5,008	21.6	0.3
Rubber or miscellaneous plastics products	1,190	1,054	–11.5	0.8
Textile mill products	336	277	–17.6	0.7
Tobacco products, excluding insecticides	6	12	105.0	0.6
Transportation equipment	1,460	1,597	9.4	0.8
Waste or scrap materials	1,115	811	–27.3	0.7
Total, all commodities	**45,488**	**47,066**	**3.5**	**0.7**

Weighted average error: 18.3%

Mean absolute error: 29.6%

[1] Energy commodities include mining, petroleum, and coal products.
[2] Distance ratio is the average shipment distance for the commodity to the state divided by the national average.

Key: N/A = not applicable.

Note: Observed inflows are obtained from the 1993 Commodity Flow Survey. Estimates are derived from the inflows to the five-state Great Lakes Region (Illinois, Indiana, Michigan, Ohio, and Wisconsin).

TABLE A6 Estimates of Annual Domestic Commodity Inflows to California
Excluding energy[1]

Commodity	Observed inflows (000s of metric tons)	Estimated inflows (000s of metric tons)	Percentage error	Distance ratio[2]
Apparel or other finished textile products	506	622	23.0	1.6
Chemicals or allied products	8,252	10,022	21.5	1.4
Clay, concrete, glass, or stone products	1,945	3,304	69.9	1.0
Electrical machinery, equipment, or supplies	1,440	1,516	5.3	1.5
Fabricated metal products	1,605	1,880	17.1	1.2
Farm products	6,372	7,212	13.2	0.9
Food or kindred products	18,437	21,967	19.1	1.9
Forest and fishing products	8	8	–1.6	N/A
Furniture or fixtures	602	679	12.8	1.4
Instruments, photographic goods, optical goods, watches, or clocks	370	336	–9.2	1.7
Leather or leather products	104	98	–5.8	1.6
Lumber or wood products, excluding furniture	2,116	3,165	49.6	1.6
Machinery, excluding electrical	994	1,190	19.8	1.5
Miscellaneous products or manufacturing	535	574	7.3	1.6
Ordnance or accessories	6	5	–12.4	N/A
Primary metal products	4,638	4,878	5.2	1.7
Pulp, paper, or allied products	5,087	5,299	4.2	1.2
Rubber or miscellaneous plastics products	1,529	1,688	10.4	1.2
Textile mill products	896	975	8.8	1.6
Tobacco products, excluding insecticides	22	19	–12.4	3.1
Transportation equipment	2,544	2,639	3.7	1.7
Waste or scrap materials	122	347	184.9	0.7
Total, all commodities	**58,130**	**68,425**	**17.7**	**1.6**

Weighted average error: 17.9%

Mean absolute error: 23.5%

[1] Energy commodities include mining, petroleum, and coal products.
[2] Distance ratio is the average shipment distance for the commodity to the state divided by the national average.

Key: N/A = not applicable.

Note: Observed inflows are obtained from the 1993 Commodity Flow Survey. Estimates are derived from the inflows to the six-state West Coast Region (California, Nevada, Oregon, Washington, Alaska, and Hawaii).

TABLE A7 Estimates of Annual Domestic Commodity Inflows to Washington
Excluding energy[1]

Commodity	Observed inflows (000s of metric tons)	Estimated inflows (000s of metric tons)	Percentage error	Distance ratio[2]
Apparel or other finished textile products	64	98	53.7	1.5
Chemicals or allied products	1,146	1,612	40.6	1.4
Clay, concrete, glass, or stone products	991	687	–30.8	0.8
Electrical machinery, equipment, or supplies	160	227	41.9	2.1
Fabricated metal products	319	417	30.5	1.8
Farm products	12,771	11,207	–12.3	1.6
Food or kindred products	3,871	3,778	–2.4	1.9
Forest and fishing products	1	2	27.3	N/A
Furniture or fixtures	125	123	–1.3	1.4
Instruments, photographic goods, optical goods, watches, or clocks	10	56	480.7	1.8
Leather or leather products	7	15	113.0	1.5
Lumber or wood products, excluding furniture	1,940	1,112	–42.7	0.6
Machinery, excluding electrical	239	229	–4.1	2.1
Miscellaneous products of manufacturing	52	92	78.4	1.3
Primary metal products	586	920	57.0	1.8
Pulp, paper, or allied products	519	1,010	94.8	2.5
Rubber or miscellaneous plastics products	163	317	94.4	1.4
Textile mill products	67	101	50.5	1.5
Transportation equipment	325	624	91.8	1.8
Waste or scrap materials	214	61	–71.5	0.7
Total, all commodities	**23,569**	**22,686**	**–3.7**	**1.5**
Weighted average error: 21.3%				
Mean absolute error: 71.0%				

[1] Energy commodities include mining, petroleum, and coal products.
[2] Distance ratio is the average shipment distance for the commodity to the state divided by the national average.

Key: N/A = not applicable.

Note: Observed inflows are obtained from the 1993 Commodity Flow Survey. Estimates are derived from the inflows to the two-state West Coast Region (California and Washington).

Monthly Forecasts of Integrated Public Transport Systems: The Case of the Madrid Metropolitan Area

ANTONIO GARCÍA-FERRER[1]
ARÁNZAZU DE JUAN[1]
PILAR PONCELA[1,*]
MARCOS BUJOSA[2]

[1] Departamento de Análisis Económico: Economía Cuantitativa
Universidad Autónoma de Madrid
28049 Cantoblanco, Madrid, Spain

[2] Departamento de Fundamentos del Análisis Económio II
Facultad de Ciencias Económicas
Universidad Complutense de Madrid
Campus de Somosaguas
28223 Madrid, Spain

ABSTRACT

The Madrid public transportation system has been integrated since 1986 and includes bus, Metro, and suburban trains. This paper addresses the problem of forecasting the demand for a large number of bus and Metro tickets in the Madrid metropolitan area using monthly data from 1987 to 2002. The database is subject to several calendar effects, outliers, changing levels of service, and changing seasonality effects that further complicate the analysis and the models' forecasts. The transport agency needs estimates of all effects, as well as a forecast, of the pattern of monthly revenues and usage of the transport network. We use both traditional dynamic transfer-function causal models as well as new variants of unobserved component models estimated by least squares using automatic identification and linear techniques in the optimization on the frequency domain (BGF algorithm). Both types of models provide some interesting forecasting comparisons (using several forecast horizons that include turning points) where the pooling forecast is also used. Forecast accuracy is assessed using traditional root mean squared error and mean absolute error measures as well as variants of the Diebold-Mariano test.

Email addresses:
[*] Corresponding author—pilar.poncela@uam.es
A. García-Ferrer—antonio.garcia@uam.es
A. De Juan—aranzazu.dejuan@uam.es
M. Bujosa—marcos.bujosa@ccee.ucm.es

KEYWORDS: Public transport forecasts, forecast accuracy, pooling forecasts.

INTRODUCTION

Worldwide, traffic management is one of the main problems that large cities face. Given the costs associated with private car usage, most metropolitan transport agencies try to coordinate services, networks, and fares so as to offer consumers a higher capacity and higher quality service, with the aim of promoting public transport use and shifting demand away from private cars.

When planning transportation facilities, it is necessary to forecast how much they will be used. Also, to price them appropriately and determine the best operating policies, it is important to understand how users respond to changes in prices and service characteristics. Because the issue of price elasticities has received considerable attention in the literature (see, e.g., García-Ferrer et al. 2003; Matas and Raymond 2003, for the case of Spain), in this paper we will concentrate on alternative models to produce efficient predictions on public transportation using recent monthly data for the Madrid (Spain) region.

Public transportation in the Madrid metropolitan area encompasses four basic modes: the subway system (Metro), the municipal bus company (EMT), the RENFE suburban train service, and interurban buses. Because of data restrictions and service characteristics, the last two modes are not included in our analytical framework.[1] *Consorcio de Transportes de Madrid* (CTM), which was created in 1986, manages the entire public transportation system. CTM coordinates the efforts of public and private institutions related to public transport. Examples of similar policies in other European countries are well documented in Pucher and Kurth (1996).

As happens in many European cities, public transport fares in the Madrid region are based on both single-mode and multiple-mode tickets, the principal one being CTM's travel card. The travel card is a multimodal pass and coupon that can be used without limit during its valid time period, being directed mainly toward regular and prepaid passengers. Fares vary by travel zone. In particular, the introduction of less expensive season tickets can provide a powerful incentive to shift transport modes and is not without theoretical support (see, e.g., Carbajo 1988; FitzRoy and Smith 1999).

At the end of 2001, 175 municipalities representing practically the entire population of the Madrid region belonged to CTM. Despite decreasing population in recent years, the number of passengers using these services has grown from 951 million in 1986 to 1,549 million in 2001. The entire public transportation system had a 4.2% increase in demand during 2001. Suburban train services experienced a 9.5% increase in demand and suburban bus use rose 6.5%. This was not only a reflection of the recent investment and coordination efforts, but also the result of a growing residential suburbanization process in the Madrid region, rail and bus service improvements, and economic incentives that travel cards provide for long-distance travelers.

Although all modes have experienced steady growth during the past few years, the Metro system showed significant increases, recording a rise in demand of 9.6% in 1999, 9.3% in 2000, and 3.7% in 2001. As we discuss later, this growth has been fueled primarily by a large increase in Metro services in terms of the number of lines, stations, and rolling stock. In particular, the Metro route length grew 43% during 1998 and 1999, and the recently inaugurated 2003 *Metro-Sur Circle Line* added 36% more capacity that is expected to affect demand in the near future.

In this paper, we focus on monthly forecasts of the number of tickets/cards sold. Monthly forecasts are desirable for several reasons. First, such forecasts would allow CTM to obtain a pattern of future monthly revenues (just multiplying prices by the number of tickets/cards sold), which are lower, for instance, in summer months, and thus plan how much cash it will need in addition to monthly revenues. This target information can be gained with 1- to 12-steps-ahead forecasts in an on-time exercise. Second, we will be able to forecast the degree of usage of the transport system. For example, the frequency and size of the trains are lower in August. Again, for this purpose, monthly forecasts are needed. Third, by using causal models with monthly data, we are able to evaluate the effect of the number of working days within a month and the effect

[1] At the end of 2001, the Metro/bus share represented 72% of total passengers.

TABLE 1 Definitions of the Main Variables

Name	Definition	Sample Period	Sample Size	Estimation Period	Estimation Size	Forecasting Period	Forecasting Size
Number of tickets/cards sold							
SMT	Single-trip Metro	1987:1–2002:12	192	1987:1–1999:12	156	2000:1–2002:12	36
10MT	10-trip Metro	1987:1–2002:12	192	1987:1–1999:12	156	2000:1–2002:12	36
SBT	Single-trip bus	1988:1–2002:12	180	1988:1–1999:12	144	2000:1–2002:12	36
10BT	10-trip bus	1988:1–2002:12	180	1988:1–1999:12	144	2000:1–2002:12	36
TC	Monthly travel card	1987:1–2002:12	192	1991:1–1999:12	108	2000.1–2002:12	36
JTC	Junior monthly travel card	1988:2–2002:12	179	1991:1–1999:12	108	2000:1–2002:12	36
Rate of growth of deflated prices							
ΔPSMT	SMT	1987:1–2002:12	192	1987:1–1999:12	156	2000:1–2002:12	36
ΔP10MT	10MT	1987:1–2002:12	192	1987:1–1999:12	156	2000:1–2002:12	36
ΔPSBT	SBT	1988:1–2002:12	180	1988:1–1999:12	144	2000:1–2002:12	36
ΔP10BT	10BT	1988:1–2002:12	180	1988:1–1999:12	144	2000:1–2002:12	36
ΔPTC	TC	1987:1–2002:12	192	1991:1–1999:12	108	2000:1–2002:12	36
ΔPJTC	JTC	1988:2–2002:12	179	1991:1–1999:12	108	2000:1–2002:12	36

Notes: All the variables measure the number of tickets/cards sold. TC and JTC are monthly passes that can be used without limit. The price variables are measured in rates of growth of deflated prices. The deflator used was the Consumer Price Index published monthly by the Instituto Nacional de Estadística of Spain.

of Easter, as well as any other exogenous effects. Overall, short- and medium-term forecasts are our main target, which will be evaluated through 1-, 6-, 12-, and 24-steps-ahead forecasts.

The paper is organized as follows: the next section presents the characteristics of the database and tests for the presence of many outliers that create considerable estimation instability and complicate the posterior forecasting exercise. We next describe the theoretical framework and present the estimation results, after which we analyze the predictive performance of alternative models based on several forecasting horizons and different predictive accuracy criteria. The issue of optimal forecast combinations is also addressed. Finally, the last section provides our conclusions.

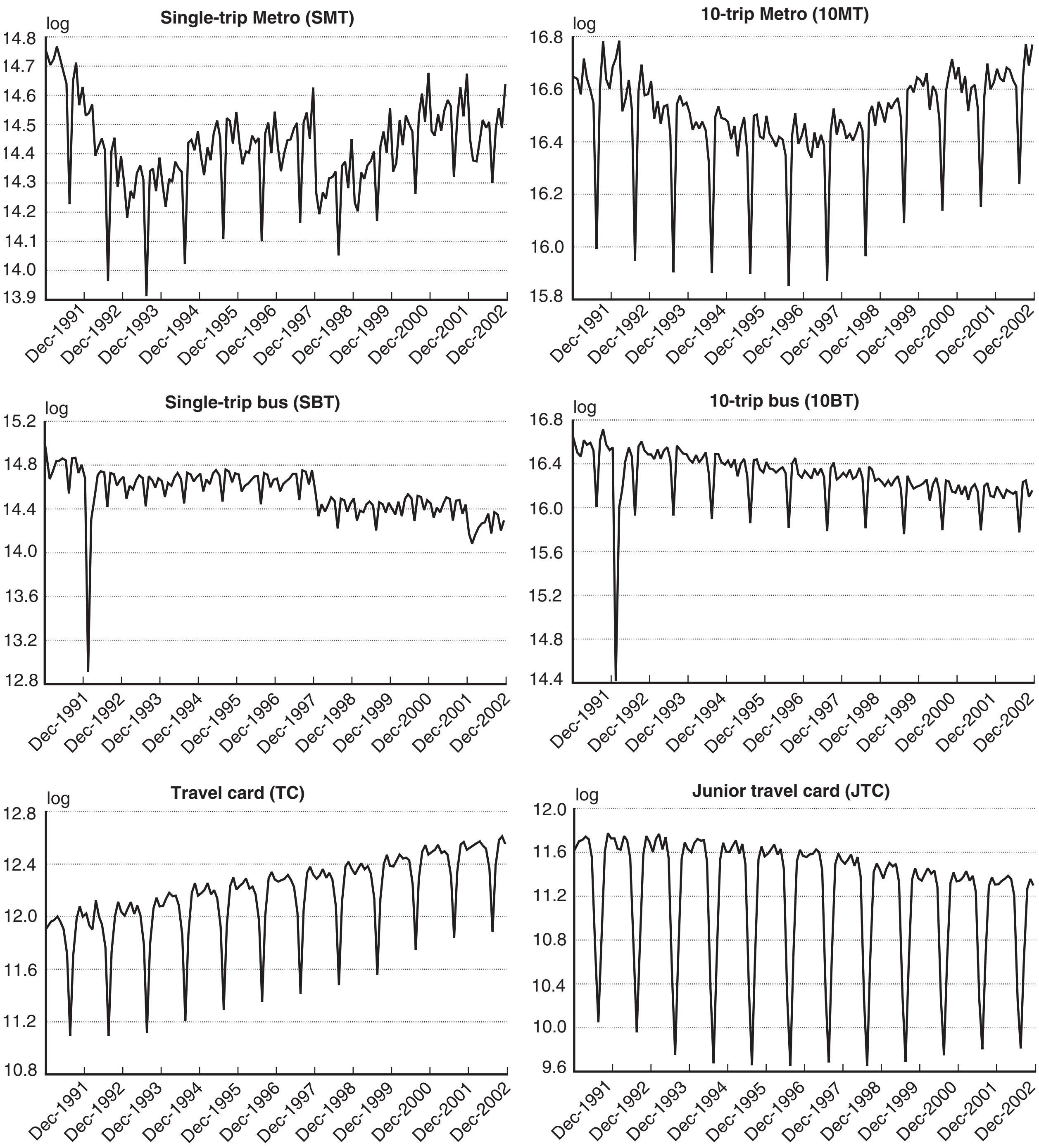
THE DATABASE

Our database includes CTM monthly data from January 1987 to December 2002 for the main public transport variables:

single-trip Metro tickets (SMT),
10-trip Metro tickets (10MT),
single trip bus tickets (SBT),
10-trip bus tickets (10BT),
regular travel card (TC), and
junior travel card (JTC).

The data are the number of tickets/cards sold in each category. Therefore, while the number of single tickets sold equals the number of trips, we do not know how many trips a holder of a travel card actually took. Definitions of the variables are given in table 1, and their plots (in logs) are shown in figure 1, where both nonstationarity as well as strong seasonality are clearly evident. The aberrant observations in both SBT and 10BT in the early months of 1992 correspond to a severe general strike on the bus network that took place during that period. Table 1 also includes the definition of the price variables that will be used later in the causal models. The descriptive statistics of the variables are given in table 2.

FIGURE 1 Logs Showing the Main Variables: January 1991–December 2002



The introduction of travel cards in 1987 caused a drastic change in the trends of use for the remaining tickets. This process has been the logical consequence of a price policy that penalizes the users of single- and 10-trip tickets through large price increases, while holding the price of travel cards constant from their introduction in 1987 until 1992. During this period, for instance, single-ticket prices increased 130% and, although this price trend has been reversed recently, regular users found multimodal travel cards much more advantageous when purchasing tickets. As a result, the demand shares have shifted dramatically during the last 15 years. Single-ticket use fell to just over 4% for both

TABLE 2 Descriptive Statistics of the Logs of the Main Variables

Name	Mean	Median	Maximum/ minimum	Standard deviation	Skewness	Kurtosis	Jarque-Bera
Number of tickets/cards sold							
LSMT	14.85	14.47	16.59/13.91	0.76	1.11	2.64	35.20
L10MT	16.38	16.49	16.81/14.78	0.33	–2.02	7.37	247.65
LSBT	14.80	14.67	15.90/12.91	0.46	0.62	4.44	23.67
L10BT	16.36	16.41	16.81/14.41	0.27	–2.69	18.89	7,219.84
LTC	12.10	12.16	12.54/11.09	0.30	–1.47	5.36	71.45
LJTC	11.24	11.52	11.77/9.65	0.61	–1.43	3.74	43.82
Rate of growth of deflated prices							
ΔPSMT	0.0032	–0.0028	0.3815/–0.0160	0.0395	7.41	62.47	27,386.79
ΔP10MT	0.0000	–0.0027	0.1120/–0.0160	0.0159	4.86	28.17	5,310.73
ΔPSBT	0.0032	–0.0028	0.3815/–0.0160	0.0395	7.41	62.47	27,386.79
ΔP10BT	0.0018	–0.0027	0.1850/–0.0160	0.0225	5.23	34.22	7,903.69
ΔPTC	0.0009	–0.0026	0.0953/–0.0123	0.0152	4.14	23.50	2,490.59
ΔPJTC	0.0006	–0.0026	0.0760/–0.0123	0.0133	4.02	19.63	1,706.63

Metro and EMT, while the market share of 10-ride tickets dropped significantly among bus users and remained almost unchanged in the case of Metro.[2]

Finally, the market share for travel cards in 2002 was over 60%, which was about the figure that CTM had in mind when the integrated system was established in 1986. In this regard, recent empirical studies (Garcia-Ferrer et al. 2003) have shown that, with the exception of travel cards, the remaining tickets show significant negative own-price elasticities and moderate estimated cross-elasticities, indicating that there is room for alternative pricing policies aimed at having positive effects on demand while minimizing the negative effects on revenues.

METHODOLOGIES AND ESTIMATION RESULTS

The plots of the main variables in figure 1 indicate that the statistical characteristics of such series changed considerably over the sample interval, so that the series can be considered nonstationary in a statistical sense. All series exhibited clear upward or downward trends, together with pronounced annual periodicity. This trend behavior is a classic example of a local mean value changing markedly over time. The nature of the seasonality varied over the six series but, in general, signs of changing seasonality patterns in the trend were evident.

These various kinds of nonstationarity are indicative of changes in the underlying statistical properties of the data. Therefore, we decided to use two alternative statistical approaches capable of characterizing the nonstationary features in an acceptable manner. One is the well-known Dynamic Transfer Function Causal Model obtained using the Intervention ARIMA (IARIMA) model developed by Box and Tiao (1975) as a starting model. The second alternative is the Dynamic Harmonic Regression (DHR) model developed by Young et al. (1999). For the latter, however, both the identification and estimation stages were carried out using the Bujosa-García-Ferrer (BGF) algorithm implemented in Bujosa et al. (2002), which is summarized in the appendix. Although the previous approaches are, basically, univariate alternatives, they also allow for the possibility of including exogenous inputs associated with intervention effects (e.g., strikes, working days, Easter effects) as well as price and service changes in the system.

Dynamic Transfer Function Causal Model

When using annual or quarterly data, demand equations can be based on causal models where the demand for transport services can be assumed to

[2] This behavior can be explained by the fact that subway transfers are free while bus transfers are not, thus penalizing the user.

depend on the attributes of each mode (monetary costs and quality of service), the competing modes of transport, and certain socioeconomic and demographic variables (e.g., income, employment, and population). Using this approach means care must be taken when interpreting estimation results due to a large number of potential econometric problems. Issues of dynamic specification of the model, small sample size, large number of regressors (small number of degrees of freedom), and extreme multicollinearity should be seriously considered before interpreting the estimates as current elasticities. Multicollinearity in this context is a problem, especially when models are specified in levels. It is well known that multicollinearity may affect estimated standard errors and signs of the coefficients, providing misleading results. Far from being a good solution, the often-used alternative of eliminating statistically insignificant regressors may be even worse in terms of interpreting results. Let us emphasize, however, that this is not a model problem (whether our model is causal or not) but a data problem that can equally affect other modeling alternatives, as we will show later.

When using monthly data, as we did in this paper, the use of causal models is restricted by data availability. Lacking monthly income, employment, and even population data for the Madrid region, our causal model was based only on fare and service quality changes plus the corresponding deterministic intervention effects and stochastic seasonality variation. Nevertheless, when short- and medium-term forecasting is the objective, we can assume that the first set of factors are reasonably incorporated in the stochastic long-term trends of the monthly data. In the case of independent submarkets for each type of ticket, our dynamic causal model can be written as:

$$\begin{aligned}\phi(L)\Phi(L^s)\nabla^d\nabla^D y_t &= \sum_{i=1}^{k} v_i(L)\nabla^d\nabla^D x_{it} \\ &\quad + \sum_{j=1}^{m} \partial_j(L)\nabla^d\nabla^D z_{jt} \\ &\quad + \theta(L)\Theta(L^s)a_t \qquad (1)\end{aligned}$$

where x_{it} are the exogenous or intervention variables, $v_i(L)$ includes the dynamic model for the i-th intervention variable, $\theta(L)$ and $\Theta(L^s)$ are the regular and seasonal moving average operators, and $\phi(L)$ and $\Phi(L^s)$ are the regular and seasonal autoregressive operators. The process a_t is assumed to be white noise, and ∇^D and ∇^d allow for seasonal and nonseasonal differencing. The usual stationarity and invertibility conditions apply for the *AR* and *MA* operators. As indicated by Peña (2001), most outlier specifications in the literature (additive, innovational, level shifts, and transitory changes) are particular cases of equation (1) under proper parameterizations of the polynomial operator $v_i(L)$. Detailed definitions of the different intervention variables are included in table 3.

On the other hand, z_{jt} is a vector of exogenous variables that includes price and service changes of different tickets, and $\partial_j(L)$ represents the dynamic model for the j-th exogenous variable. Before going into the estimation details, the following are comments regarding the components of the z_{jt}.

First, rates of growth of deflated prices for all tickets are included in z_{jt}. It can be shown that price changes are identical for SMT and SBT, very similar for 10BT and 10MT, and exhibit a high correlation ($\rho = 0.81$) between price changes of 10MT and TC. This severe collinearity will have important consequences in assessing posterior robustness on parameter estimation. While it is well known that multicollinearity does not have any effect on forecasting, one should be cautious when interpreting the coefficients. However, the estimation results we present here are robust to alternative model specifications, except for 10MT, as a consequence of the high multicollinearity between price changes of 10MT and TC (see García-Ferrer et al. 2003 for details).

Second, service quality can be measured using several indicators: route length, number of stations, number of trains or buses, vehicle-kilometers, etc. Since all indicators are highly collinear, we decided to use route length as the main indicator of service quality. Apart from these indicators, there are no data available on other service quality indicators.

The estimation of the models was done by exact maximum likelihood using SCA statistical software. Estimation results are shown in tables 4 and 5, but we will discuss each table separately. This leads us to the first issue related to the choice of the estimation period for each variable. Although most of the

TABLE 3 Definitions of Intervention Variables

Name	Definition	Facts	Affected variables
EASTER	1 in Easter months, 0 otherwise	Easter effects	All
DAYS	Trading days per month	Trading day effects	All
MAR89	1 in 1989:3, 0 otherwise	Strike in bus mode and large increase in single-ticket prices	SMT, 10MT
MAR90	1 in 1990:3, 0 otherwise	Strike in bus mode	SMT, SBT, 10BT
APR90	1 in 1990:4, 0 otherwise	Strike in bus mode and large increase in single-ticket prices	10MT, 10BT
JAN91	1 in 1991:1, 0 otherwise	Strike in Metro mode and large increase in single-ticket prices	SMT
APR91	1 in 1991:4, 0 otherwise	Strike in Metro mode	SMT
FEB92	1 in 1992:2, 0 otherwise	Strike in bus mode	SMT, SBT, 10BT
MAR92	1 in 1992:3, 0 otherwise	Strike in bus mode	SBT, 10BT
APR92	1 in 1992:4, 0 otherwise	Consequences of bus strikes	SBT
JAN98	1 from 1998:1, 0 otherwise	Introduction of the Metrobus ticket	SMT, SBT
FEB93	1 in 1993:2, 0 otherwise	?	JTC
APR93	1 in 1993:4, 0 otherwise	Metro service interruption Temporary closing of certain lines	JTC
MAR94	1 in 1994:3, 0 otherwise	?	TC, JTC
JUL95	1 in 1995:7, 0 otherwise	?	JTC

Key: ? = no known explanation for this intervention variable.

database covers the period January 1987 to December 2002, some problems preclude using this whole period as a generalized database. First, monthly data on SBT and 10BT are not available before 1988. Second, the TCs from 1987 to 1990 include information that was inconsistent with the posterior 1991 to 2001 data. Due to the gradual introduction of different travel card options from 1987 to 1990, the data generation process cannot be considered identical before and after 1990. Third, in the case of JTC, the early period 1987 to 1990 database suffers from identical problems and, consequently, was discarded.

For the Metro and bus estimates in table 4, the following results are noteworthy:

1. All variables in the table are affected by the same difference operators as the endogenous variable, but prices are only affected by a ∇_{12} seasonal difference because they are already given as rates of growth. Therefore, by integrating out the seasonal difference on both sides of each equation, the estimated coefficients can be interpreted as elasticities.
2. The effect of trading days is positive for all tickets, ranging from 0.7% for SBT to 1.6% for the 10BT variable. The size of the effects on the Metro tickets is very similar.
3. On the contrary, the Easter effect is negative (as expected) for all tickets. However, the size of the effect is larger in the case of the 10-ride tickets (both for Metro and EMT).
4. The remaining intervention variables are related either to strikes or to large increases in public transport fares. Nevertheless, effects vary among different types of tickets. In the case of SMT, for instance, we found positive and statistically significant coefficients in APR91 and FEB92, which are associated with strikes in the competing mode (bus). The negative effect in JAN91 can be explained by a Metro strike, while the mixed effects in the MAR90 coefficients correspond, respectively, to a bus strike and to a large price increase in SMT the following month, which explains why this effect does not show up in the 10MT equation. The long bus network strike seen in

TABLE 4 Estimated Causal Models for the Metro and Bus Variables
In logs

	$\nabla\nabla_{12}SMT$	$\nabla\nabla_{12}10MT$	$\nabla\nabla_{12}SBT$	$\nabla\nabla_{12}10BT$
Constant	0.0004 (0.0022)	–0.0009 (0.0027)	0.0009 (0.0016)	0.0003 (0.0002)
DAYS	0.0096* (0.0017)	0.0151* (0.0030)	0.0073* (0.0022)	0.0164* (0.0021)
EASTER	–0.0479* (0.0101)	–0.0987* (0.0158)	–0.0310* (0.0127)	–0.1064* (0.0141)
MAR89	–0.0311 (0.0235)	–0.1064* (0.0451)		
MAR90	$0.0609^{*} -0.1265^{*}B -0.006^{*}B^{2}$ (0.0322) (0.0364) (0.0270)		$-0.6667^{*} -0.1482^{*}B$ (0.0759) (0.0631)	$-0.6491^{*} -0.2287^{*}B$ (0.0457) (0.0387)
APR90		0.0494 (0.0520)		
JAN91	–0.0998* (0.0272)			
APR91	0.0436* (0.0245)			
FEB92	0.0942* (0.0216)		$-1.6170^{*} -0.3406^{*}B -0.1301^{*}B^{2}$ (0.0323) (0.0417) (0.0381)	$-2.0973^{*} -0.4516^{*}B$ (0.0383) (0.0411)
JAN98	$-0.1190^{*} -0.0951^{*}B -0.0721^{*}B^{2}$ (0.0269) (0.0303) (0.0283)		$-0.1107^{*} -0.1010^{*}B$ (0.0360) (0.0349)	
$\Delta PSMT$	–1.2817* (0.1202)	0.7601* (0.1640)		
$\Delta P10MT$	–0.1878 (0.4350)	–2.4665* (0.6658)		
$\Delta PSBT$			–1.2776* (0.1600)	0.2205* (0.0679)
$\Delta P10BT$			0.7461 (0.5128)	–0.5178* (0.1838)
ΔPTC	0.6922 (0.5463)	2.6274* (0.8564)	–0.6958 (0.6393)	0.9075* (0.2516)
ϕ_1	–0.5577 (0.3319)		–0.2586* (0.0961)	
ϕ_2			–0.3086* (0.0908)	
θ_1	–0.5775 (0.3386)			0.4047* (0.0767)
θ_2				0.5842* (0.0774)
Φ_{12}			–0.5154* (0.0731)	
Θ_{12}	0.3145* (0.0830)	0.5629* (0.0651)		0.7979* (0.0871)
Θ_{24}				–0.2589* (0.0809)
$\hat{\sigma}_a$	0.035717	0.065537	0.040829	0.041271
R^2	0.89	0.60	0.98	0.99
LBQ (12, 24, 36)	5.4, 16.0, 27.2	9.1, 11.5, 22.3	14.6, 23.1, 25.0	7.8, 19.7, 36.2
Jarque-Bera (*p*-value)	2.43 (0.297)	0.289 (0.865)	2.11 (0.348)	6.36 (0.042)
White's heterosk (*p*-value)	22.82 (0.83)	6.08 (0.81)	4.35 (0.99)	2.11 (0.999)

Notes: Standard errors are in parenthesis. * significant at 5%; $\nabla = (1 - B)$ and $\nabla_{12} = (1 - B^{12})$: regular and seasonal differences; *B*: lag operator; ΔPT_i ($i = SMT$, $10BT$, SBT, $10BT$, TC): rates of growth of deflated prices for Metro and bus tickets; LBQ: Ljung-Box *Q* statistics; $\hat{\sigma}_a$ = residual standard error.

TABLE 5 Estimated Causal Models for TC and JTC
In logs

	$\nabla_{12}TC$	$\nabla\nabla_{12}\,JTC$
Constant	0.0544* (0.0073)	–0.0368* (0.0076)
DAYS	0.0018 (0.0024)	0.0026* (0.0010)
EASTER	–0.0535* (0.0106)	–0.0710* (0.0041)
FEB93		0.0579* (0.0136)
APR93		–0.0870* (0.0137)
MAR94	0.0368* (0.0165)	
JUL95		0.0492* (0.0085)
ΔPTC_i	0.1576 (0.3032)	–0.0134 (0.1387)
MRL_i	$\dfrac{\underset{(0.1344)}{0.2517^*}B^4 \underset{(0.1665)}{-0.4175^*}B^5}{(1+\underset{(0.0672)}{0.9897^*}B)}$	$\underset{(0.0861)}{0.1800^*}B^4 \underset{(0.0977)}{-0.2087^*}B^5$
θ_1	0.4557* (0.1316)	
θ_2	–0.5399* (0.0997)	0.4671* (0.1641)
θ_3		–0.5261* (0.1628)
Θ_{12}	–0.1285 (0.0939)	–0.4048* (0.0949)
ϕ_1	0.3240* (0.1042)	0.9753* (0.1044)
ϕ_2		–0.6448* (0.1997)
ϕ_3	0.2694* (0.0964)	0.3357* (0.1314)
ϕ_4		–0.3400* (0.0942)
$\hat{\sigma}_a$	0.02560	0.03035
R^2	0.70	0.86
LBQ (12, 24, 36)	11.2, 22.8, 45.6	10.6, 14.6, 21.9
Jarque-Bera (*p*-value)	2.75 (0.252)	2.69 (0.260)
White's heterosk (*p*-value)	6.68 (0.946)	5.86 (0.97)

Notes: Standard errors in parenthesis. * represents significant at 5%; $\nabla = (1-B)$ and $\nabla_{12} = (1-B^{12})$: regular and seasonal differences; *B*: lag operator; ΔPTC_i ($i = TC, JTC$): rates of growth of deflated price for TC tickets; *MRL*: Metro route length; LBQ: Ljung-Box *Q* statistics; $\hat{\sigma}_a$ = residual standard error.

the FEB92 variable shows temporary fluctuations of –80.2%, 28.9%, and 12.2% in February, March, and April 1992 with respect to the previous month in the same year and an even larger change (–82.7% in February and 36.3% in March) in the case of 10BT.[3] These figures show that fewer bus ticket purchases were only compensated by a mild increase in single-Metro ticket purchases (+9.4%) during that period, as a result of less Metro network coverage at that time. Finally, the JAN98 variable corresponds to the introduction of a new Metrobus ticket that negatively affected single-ticket sales permanently after its introduction.

5. The estimated coefficients associated with price increases are the corresponding elasticities. In general, users were highly sensitive to single- and 10-ride fares with own-price elasticities values ranging from –0.52 to –2.47. However, while the estimated elasticities in the cases of SMT, SBT, and 10BT proved to be robust to alternative model specifications, this was not the case for 10MT as a consequence of the high multicollinearity between price changes of 10MT and TC.[4] The remaining cross elasticities were in line with those obtained by García-Ferrer et al. (2003). The only significant evidence is found in the case of the two 10-trip tickets. For 10MT, positive cross effects were found for single tickets (0.76) and travel cards (2.63). Similar results were found for 10BT although the size of the cross effects was much smaller than in the Metro case: 0.22 for single tickets and 0.91 for travel cards.
6. The Ljung-Box statistics at 12- , 24- , and 36-month lags show no signs of residual autocorrelation in the estimated models.

For the travel card estimates, the following results are worth noting (see table 5):

1. The estimation period was shorter (108 observations), given the anomalies mentioned earlier at the beginning of both series.
2. For the model estimated for JTC, all variables were affected by the same difference operators as the endogenous variable, but prices were only affected by a ∇_{12} transformation, because they were already given as rates of growth. By integrating out the seasonal difference in both sides of each equation, the estimated coefficients can be interpreted as elasticities. Nevertheless, in the TC equation, this variable was affected only by a seasonal difference and the straightforward interpretation of the coefficients as elasticities was not possible.
3. The effect of trading days is almost negligible and only significant for JTC. This is a logical result given the (monthly card) characteristic of these tickets.
4. The Easter effect is statistically significant and has the expected negative signs for both variables. The quantitative impacts (–5.3% and –7.1%) are similar to the ones observed in single- and 10-trip tickets.
5. The remaining intervention variables are additive outliers that had not appeared earlier. The negative sign of APR93 is associated with service interruption and the temporary closing of certain lines. The CTM does not have any reasonable explanation for the positive coefficients associated with FEB93, MAR94, and JUL95.
6. Travel card tickets (TC and JTC) were not affected either by their own price increases or by the price increases of competing tickets. This result is hardly surprising given the CTM pricing policy, which held the price of travel cards constant (during a large part of the sample) since their introduction in 1987. Among the service variables included in the causal model's information set, the Metro route length (MRL_i) now becomes statistically significant for the two travel card variables.[5] The

[3] The estimated coefficients represent the variations in log y_t due to the strike. For instance, the estimated percentage fall in SBT in February 1992 (estimated coefficient –1.6170) is computed as $e^{-1.6170} - 1 = -0.8015$.

[4] As a matter of fact, when ΔPTC is removed from the equation, the own-price elasticity for 10MT goes down to –1.10. A word of caution applies equally to the cross-effects results between 10MT and TC.

[5] As a matter of fact, MRL_i is never statistically significant in the single- and 10-ride equations. It only becomes significant for TC and JTC when the estimation period ends in December 1999 and later, due to the abrupt changes in MRL_i during this last period.

estimated dynamic responses of this variable indicate a long-term effect that lasts up to five months, implying a very long and smooth demand response to changes in supply.

7. The Ljung-Box statistics at 12- , 24- , and 36-month lags do not indicate serial correlation problems in the estimated models.

Dynamic Harmonic Regression Model

The DHR model developed by Young et al. (1999) belongs to the Unobserved Component (UC) type and is formulated within the state-space framework. This model is based on a spectral approach, the hypothesis of which maintains that the observed time series can be decomposed into several DHR components whose variances are concentrated around certain frequencies. This hypothesis is appropriate if the observed time series has well-defined spectral peaks, which implies that its variance is distributed around narrow frequency bands. Basically, the method attempts to: 1) identify the spectral peaks, 2) assign a DHR component to each spectral peak, 3) optimize the hyper-parameters that control the spectral fit of each component to its corresponding spectral peak, and 4) estimate the DHR components using the Kalman Filter and the Fixed Interval Smoothing (FIS) algorithms.

In the univariate case, the DHR model can be written as a special case of the univariate UC model that has the general form:

$$y_t = T_t + S_t + e_t; \quad \mathrm{t} = 0, 1, 2, ..., \tag{2}$$

where y_t is the observed time series, T_t is the trend or low-frequency component, S_t is the seasonal component, and e_t is an irregular component normally distributed with zero mean value and variance σ_e^2. These types of models have been extensively studied in the literature (see, e.g., Kitagawa 1981; Harvey 1989; and West and Harrison 1997).

Equation (2) is appropriate for dealing with economic data exhibiting pronounced trend and seasonality as is the case with the monthly variables used in this paper. When set in state-space form, each component is modeled in a manner that allows y_t to be represented as a set of discrete-time equations that are the basis for recursive state-space estimation and forecasting. Let us introduce specific representations for T_t and S_t.

Trend Model

Together with its derivative D_t, the low-frequency component T_t can be described by the following second-order generalized random walk (GRW) model,

$$\begin{bmatrix} T_t \\ D_t \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ 0 & \gamma \end{bmatrix} \begin{bmatrix} T_{t-1} \\ D_{t-1} \end{bmatrix} + \begin{bmatrix} \eta_{1t} \\ \eta_{2t} \end{bmatrix}$$

where $\eta_t = \begin{bmatrix} \eta_{1t} \; \eta_{2t} \end{bmatrix}'$ is a white noise vector with zero mean and covariance matrix $\boldsymbol{Q}$. For simplicity, we assumed that $\boldsymbol{Q}$ is a diagonal matrix diag (q_{11}, q_{22}), with unknown elements q_{11} and q_{22}. This GRW model subsumes as special cases:

- the very well-known random walk
  $(RW;\ \alpha = 1,\ \beta = \gamma = 0,\ q_{22} = 0)$;
- the smooth random walk
  $(SRW;\ 0 < \alpha < 1,\ \beta = \gamma = 1,\ q_{11} = 0)$; and
- the integrated random walk
  $(IRW;\ \alpha = \beta = \gamma = 1,\ q_{11} = 0)$.

In the current example, the BGF algorithm (Bujosa et al. 2002) identifies an IRW model for all trends. In this case,

$$\begin{aligned} T_t &= T_{t-1} + D_{t-1} \\ D_t &= D_{t-1} + \eta_{2t} \end{aligned} \tag{3}$$

where T_t and D_t can be interpreted as the level and slope (derivative) of a time variable trend. The variance of η_{2t} (q_{22}) is the only unknown in equation (3) and can be estimated through the *noise variance ratio* (NVR), the ratio between q_{22} and the variance of the irregular component σ_e^2:

$$NVR_T = \frac{q_{22}}{\sigma_e^2} \tag{4}$$

When smoothing the series with the Kalman Filter in order to estimate the trend, the NVR_T works as a smoothing parameter. Very low NVR_T values are indicative of near deterministic linear trends. In the limit, when $NVR_T = 0$, the estimated trend is

TABLE 6 *NVR* Estimates of the Main Variables

Series	T	S^{12}	S^6	S^4	S^3	$S^{2.4}$	S^2	σ_e^2
SMT	0.091	0.359	0.117	0.109	0.343	0.071	0.052	0.00005
SBT	0.001	0.023	0.013	0.024	0.017	0.013	0.011	0.00038
10MT	0.029	0.948	0.434	0.302	0.357	0.490	0.301	0.00004
10BT	0.001	0.010	0.007	0.026	0.107	0.034	0.022	0.00036
TC	0.008	0.041	0.006	0.081	0.023	0.026	0.006	0.00021
JTC	0.050	3.779	1.818	1.199	0.278	0.026	0.012	0.00004

linear. On the other hand, for large NVR_T values the estimated trend mimics the original time series y_t.

Seasonal Model

We assume that the seasonal component in equation (2) can be represented by a

$$S_t = \sum_{j=1}^{N_s} \{ a_{j_t} \cos(\omega_j t) + b_{j_t} \sin(\omega_j t) \} \tag{5}$$

where the regression coefficients $a_{j_t}, b_{j_t}, j = 1 \ldots N_s$ are time variable to handle nonstationary seasonality. As in the previous trend model, time variation in a_{j_t} and b_{j_t} may follow any variant within the GRW framework. The BGF algorithm identifies that parameter variation in a_j and b_j is modeled as an RW process,

$$\begin{bmatrix} a_{j_t} \\ b_{j_t} \end{bmatrix} = \begin{bmatrix} a_{j_{t-1}} \\ b_{j_{t-1}} \end{bmatrix} + \begin{bmatrix} \xi_{1t} \\ \xi_{2t} \end{bmatrix}. \tag{6}$$

This assumption is very useful for time series with growing amplitude seasonality as the one found in most series in this paper.

The presence of important outliers, however, can easily affect both the estimation and the posterior identification and estimation stages of the DHR models. To avoid these effects, we implemented the following iterative process using our own Matlab code:

1. Initial identification and estimation of DHR models was proposed.
2. Using the Kalman Filter and FIS algorithms on the original series, outliers are treated as missing values and variance intervention was used (Young and Ng 1989) to handle level changes due to large variations in fares or the introduction of new types of tickets.
3. New series were reconstructed where each outlier was substituted by its estimated value and the level shifts were accounted for through variance intervention.
4. Finally, using the reconstructed series, we returned to step 1.

In all series in this paper, robust results were obtained after three iterations and the *NVR* estimation results, shown in table 6, correspond to the third iteration. Also, in all cases, the identification stage suggests an *IRW* trend component and an *RW* model for the seasonal component of period 12 and its harmonics (6, 4, 3, 2.4, and 2). Again, as in the causal models, the Ljung-Box statistics did not indicate any evidence of residual autocorrelation. Note that this iterative procedure was not necessary on the causal models, since specific intervention variables were properly introduced in the model's specification.

Finally, plots of the estimated trend and seasonal components for the six variables (not shown) indicate that, in all cases, the estimated NVR_T for the trends are different from zero, confirming the absence of deterministic linear trends but suggesting (apart from the breaks) very smooth long-term behavior. On the other hand, the estimated seasonal components are indicative of a clear changing seasonality pattern confirming the inappropriateness of deterministic seasonal schemes for this dataset.

FORECASTING

A large body of literature exists on the subject of travel demand forecasts (see, e.g., Gillen 1977; Lave 1994; Baumgartner 1995; Slavin 1996; and Lythgee and Wardman 2002, among others). The motivation for these types of forecasts is self-evident, because the most likely reason for making forecasts is to assess different policies. The level of detail and the forecasting horizons change considerably among different studies as a function of the travel demand model, the characteristics of the data, and the number of input variables. Broadly speaking, these external influences can be separated into two components: those related to demographic and economic changes as well as other external variables, and those directly related to the public transportation system. When long-term planning is the goal, both components need to be predicted and both involve economic assumptions about their future behavior. However, when using monthly data and when (as in our case) short- and medium-term forecasts are the objective, we can assume that the first set of factors are reasonably incorporated in the stochastic long-term trends. As we will show later, it is the second set of factors (basically abrupt changes in travel supply and changing seasonality patterns) that is responsible for large variations in public travel demand in the Madrid metropolitan area and the presence of forecast errors during certain months.[6]

Several periods-ahead forecasts for the six variables considered in this paper (using the causal and DHR models) were done for 2000, 2001, and 2002. The initial estimation sample ends in December 1999. We made 1-, 6-, 12-, and 24-months-ahead forecasts. We re-estimated all models adding one data point at a time and made new predictions. We proceeded in the same way until the last sample data point was available for estimation (November 2002 for 1-step-ahead forecasts, June 2002 for 6-steps-ahead forecasts, December 2001 for 12-steps-ahead forecasts, and December 2000 for 24-steps ahead forecasts). The forecasting for these periods was difficult given the large increase in Metro services over this time period, particularly during 1999. This difficulty meant a mixture of effects that influenced not only total demand but also temporary passenger shifts from Metro to EMT as a consequence of closing Metro stations and opening temporary new bus lines.

Measures of Forecast Accuracy

The literature repeatedly stresses the adoption of a variety of error measures (e.g., among others, Armstrong and Collopy 1992; Fildes 1992), because the selection of the best forecasting technique may depend on the choice of a particular accuracy measure. In this section, we computed different measures of forecast accuracy for each variable, model, and forecast horizon. Conflicts among them (although not desirable) only indicate different goals of alternative prediction exercises. The first measure is the root mean squared error (RMSE) defined for horizon h and model i as

$$RMSE_i(h) = \sqrt{\frac{1}{N}\sum_{t=T}^{T+N} e_{i,t+h}^2}$$

where N is the number of forecasts, t is the number of observations used for estimation, $e_{i,t+h}$ is the forecast error defined as true value minus the forecast produced with model i (either causal or DHR).

The second measure is the mean absolute error (MAE) defined as

$$MAE_i(h) = \frac{1}{N}\sum_{t=T}^{T+N} |e_{i,t+h}|$$

Also, to assess the forecast performance of both models, particularly in the medium term, we computed annual percentage errors (APE) and forecasted growth rates (FGR) and compare them with observed growth rates (OGR). Annual percentage errors for a certain year obtained with model i = *DHR*, *causal* and forecast horizon h are defined as

$$APE_i(h) = \frac{\sum_{n=1}^{12} y_{i,t+n|t+n-h} - \sum_{n=1}^{12} y_{t+n}}{\sum_{n=1}^{12} y_{t+n}}\% \qquad (7)$$

where n runs over the months within a year, t represents December of the previous year and $y_{i,t+n|t+n-h}$ is the forecast obtained with model i for month n with observations up to the previous h

[6] Among the variables associated with service changes, the Metro route length is primarily responsible for the demand changes experienced in the Madrid metropolitan area, particularly in 2000. This variable remains unchanged from 1988 to 1993 and changes very little (4%) from 1994 to 1997. However, it shows a huge increase (43%) during the last months of 1998 and throughout 1999, and zero growth during 2000 and 2001.

TABLE 7 Mean Absolute Value (MAE) and Root Mean Squared Error (RMSE) of 1-, 6-, 12-, and 24-Steps-Ahead Forecasts for Metro and Bus Tickets and Travel Cards for Alternative Models
Forecasting periods: 2000–2002

	SMT		10MT		SBT		10BT		TC		JTC	
	Causal	DHR	Causal	DHR	Causal	DHR	Causal	DHR	Causal	DHR	Causal	DHR
***h* = 1**												
MAE	0.031	0.035	0.039	0.041	0.039	0.046	0.027	0.041	0.025	0.031	0.029	0.034
RMSE	0.037	0.043	0.049	0.054	0.059	0.066	0.034	0.051	0.043	0.041	0.039	0.043
***h* = 6**												
MAE	0.046	0.054	0.065	0.055	0.085	0.090	0.029	0.038	0.036	0.044	0.026	0.040
RMSE	0.051	0.066	0.085	0.067	0.119	0.113	0.038	0.050	0.059	0.053	0.030	0.046
***h* = 12**												
MAE	0.078	0.085	0.103	0.075	0.115	0.112	0.027	0.040	0.027	0.043	0.022	0.029
RMSE	0.088	0.100	0.123	0.089	0.142	0.128	0.034	0.050	0.051	0.051	0.030	0.036
***h* = 24**												
MAE	0.156	0.233	0.130	0.107	0.140	0.268	0.032	0.064	0.073	0.074	0.075	0.049
RMSE	0.174	0.265	0.180	0.134	0.164	0.287	0.039	0.085	0.144	0.096	0.095	0.059

months. The annual forecast growth rates are defined as

$$FGR_i(h) = \frac{\sum_{n=1}^{12} y_{i,\,t+n|t+n-h} - \sum_{n=1}^{12} y_{t+n-1}}{\sum_{n=1}^{12} y_{t+n-1}}\% \qquad (8)$$

Although both RMSE and MAE are the usual yardsticks of forecast accuracy, for longer forecasting horizons (beyond the usual one-step-ahead period), their sole use may be not only inappropriate but misleading (García-Ferrer and Queralt 1997). In this case, we contend that FGR and APE become more relevant criteria. Table 7 presents the results for the RMSE and MAE, while the APE and FGRs are shown in table 8.

As might be expected, no model dominates the other under all the accuracy criteria and forecasting horizons. However, the following tentative conclusions can be drawn from these tables.

For the aggregate MAE and RMSE criteria (table 7), both indicators generally agree on which model performs better. As expected, both models' forecasting performance deteriorates as the forecast horizon grows. Only the causal model for 10BT shows similar MAE and RMSE values for different forecast horizons without signs of deterioration. For this variable, the DHR model deteriorates only in the 24-steps-ahead forecast horizon. When considering both criteria, all variables, and all forecast horizons (48 cases), the causal model emerges as a winner in 35 cases, while the DHR model is best in 12 cases. In one case (RMSE, h = 12, TC variable), the empirical results are identical for both models. However, as we will show later, some differences are not statistically significant when using the Diebold-Mariano test.

For the APE and FGR criteria, results change considerably among variables, forecast year, and forecasting horizons. In general, 1-step-ahead forecasts are excellent for both models, although the causal model outperforms the DHR in 24 out of 36 cases. For this forecast horizon, a large number of APE results are below 1%. The largest APE values are –2.77% for the causal model and 3.36% for the DHR model. For 6- and 12- steps-ahead forecasts, results are only computed for 2001 and 2002. In the first case (6-steps-ahead), the causal model outperforms its competitor in 15 out of 24 cases. Its median APE is 1.58%, while the one corresponding to the DHR is 2.66%. In the second case (12-steps-ahead), results indicate a similar performance (each model performing best in 12 cases), although the median APE is 2.83% for the causal model and 3.70% for the DHR model.

Finally, 24-steps-ahead results are only available for 2002. For this particular year, forecasts deterio-

TABLE 8 Annual Percentage Errors (APE), Forecasted Annual Growth Rates (FGR), Observed Annual Growth Rates (OGR) of *h* = 1-, 6-, 12-, and 24-Steps-Ahead Forecasts for Metro, Bus Tickets, and Travel Cards for Alternative Models
Forecasting Period: 2000–2002

	SMT		10MT		SBT		10BT		TC		JTC	
	Causal	DHR	Causal	DHR	Causal	DHR	Causal	DHR	Causal	DHR	Causal	DHR
2000, *h* = 1												
APE	–0.87	–0.17	0.30	1.82	–2.77	–0.30	–1.32	–1.63	–0.04	–1.43	–0.43	2.28
FGR	12.22	13.01	5.96	7.57	0.87	3.44	–3.75	–1.05	9.21	7.69	–4.06	–1.45
OGR	13.20		5.64		3.75		–2.47		9.26		–3.65	
2001, *h* = 1												
APE	0.97	1.23	1.15	1.30	0.25	0.88	0.25	0.40	0.15	3.36	–0.22	2.11
FGR	5.67	5.94	0.90	1.06	–1.61	–0.99	–3.12	–2.97	6.20	9.61	–3.10	–0.83
OGR	4.65		–0.24		–1.86		–3.36		6.04		–2.88	
2002, *h* = 1												
APE	0.84	–0.07	–0.78	–3.36	1.86	1.28	–0.80	–1.70	0.56	0.04	–1.17	0.42
FGR	–4.85	–5.71	6.15	3.39	–14.36	–14.84	–1.09	–1.98	5.01	4.47	–3.29	–1.73
OGR	–5.64		6.98		–15.90		–0.29		4.43		–2.15	
2001, *h* = 6												
APE	2.65	3.00	5.09	2.27	–0.19	4.25	–0.22	1.09	3.61	4.33	–0.62	2.96
FGR	7.42	7.79	4.84	2.02	–2.05	2.32	–3.57	–2.30	9.87	10.63	–3.48	0.00
OGR	4.65		–0.24		–1.86		–3.36		6.04		–2.88	
2002, *h* = 6												
APE	3.55	1.86	–5.76	–5.56	7.37	4.73	–1.51	–2.28	1.05	2.36	–1.61	1.42
FGR	–2.30	–3.89	0.82	1.03	–9.72	–11.94	–1.79	–2.56	5.52	6.89	–3.72	–0.76
OGR	–5.64		6.98		–15.90		–0.29		4.43		–2.15	
2001, *h* = 12												
APE	3.52	7.57	10.46	7.20	–4.38	7.10	–0.62	0.42	2.01	2.17	–0.07	2.48
FGR	8.33	12.57	10.19	2.02	–6.16	2.32	–3.96	–2.95	8.17	8.35	–2.95	–0.47
OGR	4.65		–0.24		–1.86		–3.36		6.04		–2.88	
2002, *h* = 12												
APE	9.91	9.69	–10.13	–6.13	17.79	15.62	–2.14	–2.80	1.19	4.62	2.12	1.35
FGR	3.71	3.50	–3.86	0.42	–0.96	–2.79	–2.42	–3.08	5.64	9.26	–0.07	–0.83
OGR	–5.64		6.98		–15.90		–0.29		4.43		–2.15	
2002, *h* = 24												
APE	16.94	28.68	10.66	7.00	10.41	34.00	–2.48	–2.20	6.17	6.39	–1.17	3.98
FGR	10.34	21.42	18.38	14.47	–7.17	12.67	–2.76	–2.48	17.17	11.11	–3.29	1.75
OGR	–5.64		6.98		–15.90		–0.29		4.43		–2.15	

rate considerably for SMT and SBT as a consequence of huge drops in their tickets sales during that year. Both models perform in a similar fashion, although the median APE is 8.30% for the causal model and 6.70% for the DHR case. One interesting byproduct of both APE and FGR criteria is the presence of signs that may be indicative of potential gains (by canceling errors of different signs) of combining forecasts over individual forecasts. According to table 8, for instance, there are, a priori, plausible forecasting gains for JTC (at all forecast horizons) and also for SBT, 10BT, and SMT for certain years and forecast horizons.

Statistical Comparison of Forecast Accuracy

In this subsection, we perform the Diebold-Mariano test (1995) to compare the forecast accuracy of the two models. (For a recent revision of this type of test see, e.g., Mariano 2002.) We will also apply its small sample version (Harvey et al. 1997). We concentrate on the mean squared error (MSE) loss function. Denote by $g(e_{i,t+h})$ the loss function for model $i = 1, 2$ and by $d_t = g(e_{1,t+h}) - g(e_{2,t+h})$ the loss differential between the two models. There is no statistical difference between the two forecasting procedures if the expected value of the loss differential is zero. Therefore, the null hypothesis is $H_0 : E(d_t) = 0$ versus the alternative $H_1 : E(d_t) \neq 0$. Diebold and Mariano (1995) use the following test statistic

$$DM = \frac{\bar{d}}{\left[\frac{\hat{f}}{N}\right]^{1/2}} \tag{9}$$

where

$$\bar{d} = \frac{1}{N} \sum_{t=T}^{T+N} g(e_{1,t+h}) - g(e_{2,t+h}) \tag{10}$$

and $\hat{f}$ is a consistent estimator of f, the asymptotic variance of $\sqrt{N} \cdot \bar{d}$ under the null. We will compute $\hat{f}$ as

$$\hat{f} = \sum_{k=-M}^{k=M} \hat{\gamma}_d(k) \tag{11}$$

with M being the lag truncation and $\hat{\gamma}_d(k)$ the sample autocovariance of order k of the loss differential series $\{d_t\}$. The asymptotic distribution of DM is $N(0, 1)$. Harvey et al. (1997) proposed a small sample modification of the previous test when g is the MSE function

$$DM^* = \frac{N^{1/2} DM}{\left[\frac{N + 1 - 2h + h(h-1)}{N}\right]^{1/2}} \tag{12}$$

where now the lag truncation is taken as $M = h - 1$ for the h-steps-ahead forecast (recall that errors in h-steps-ahead forecasts are a moving average process of order $h - 1$), and Harvey et al. (1997) compared the DM^* statistic with a t distribution with $N - 1$ degrees of freedom.

Table 9 presents the results of the two versions of Diebold and Mariano's test (DM and DM^*) of the squared loss function for 1-, 6-, and 12-steps-ahead forecasts. For 24-steps-ahead forecasts, the forecast errors could be distributed as $MA(23)$, so we do not have enough predictions (only 13) to properly estimate the variance, and the test cannot be applied. As table 9 shows, in 12 out of 18 cases the differences between the two forecasting methods are not statistically significant. Positive significant values indicate that the DHR model outperforms the causal model, which occurs only for 12-steps-ahead forecasts in the 10MT variable. On the contrary, negative significant values indicate that the causal model outperforms the DHR model, which occurs in six cases when using the DM test but is reduced to three when using its modified small sample version. The overall conclusion from the application of the test is that no model systematically dominates the other. This leads us to consider forecast combinations as a possible way to improve individual forecasts.

Forecast Combination

Forecast combination is discussed in the literature as a useful device to improve accuracy. Several types of linear combinations of forecasts can be computed: simple averages, weights based on the precision of forecasts, and regression methods, among others (see, e.g., Diebold 2003).

Let $y_{i,t+h|t}$ represent the h-steps-ahead forecast obtained with model $i = DHR, C$ (where C stands for causal) with information up to time t. An easy way to compute a forecast combination is to average the forecasts of the two competing models $y^{AVE}_{t+h|t}$

$$y^{AVE}_{t+h|t} = \frac{y_{DHR,t+h|t} + y_{C,t+h|t}}{2}. \tag{13}$$

Table 10 presents the RMSE for the DHR and causal models, as well as the combined forecast $y^{AVE}_{t+h|t}$ for all variables and horizons considered. From this table, we can conclude the following:

TABLE 9 Comparison of Forecast Accuracy: Results of the Diebold-Mariano (DM) and Modified Diebold-Mariano (DM*) Tests for the Square Loss Function and Several Forecast Horizons
Forecasting Period: 2000–2002

	SMT	10MT	SBT	10BT	TC	JTC
h **= 1**						
DM	–1.34	–0.57	–1.04	**–2.32**	0.23	–0.69
DM*	–1.32	–0.56	–1.02	**–2.30**	0.22	–0.68
h **= 6**						
DM	**–2.89**	1.80	0.37	**–2.08**	0.60	**–3.52**
DM*	**–2.38**	1.48	0.31	–1.71	0.49	**–2.90**
h **= 12**						
DM	–1.73	**2.78**	1.51	**–2.96**	–0.03	**–3.03**
DM*	–0.93	1.50	0.62	–1.56	–0.02	–1.64

Key: Bold indicates statistical difference between the two forecasting methods for alpha = 0.05. Positive (negative) values indicates that the DHR performs better (worse) than the causal model.

TABLE 10 Root Mean Squared Error (RMSE) of 1-, 6-, 12-, and 24-Steps-Ahead Forecasts for Metro and Bus Tickets and Travel Cards for the DHR, Causal, and Average Forecasts
Forecasting Period: 2000–2002

	SMT	10MT	SBT	10BT	TC	JTC
h **= 1**						
Causal	0.037	0.049	0.059	0.034	0.043	0.039
DHR	0.043	0.054	0.066	0.051	0.041	0.043
Average	0.037	0.042	0.059	0.035	0.037	0.033
h **= 6**						
Causal	0.051	0.085	0.119	0.038	0.059	0.030
DHR	0.066	0.067	0.113	0.050	0.053	0.046
Average	0.055	0.069	0.111	0.038	0.052	0.031
h **= 12**						
Causal	0.088	0.123	0.142	0.034	0.051	0.030
DHR	0.100	0.089	0.128	0.050	0.051	0.036
Average	0.091	0.101	0.126	0.036	0.045	0.026
h **= 24**						
Causal	0.174	0.180	0.164	0.039	0.144	0.095
DHR	0.265	0.134	0.287	0.085	0.096	0.059
Average	0.213	0.145	0.204	0.051	0.108	0.028

1. The combined forecast does not always improve the RMSEs of the individual forecasts.
2. Nevertheless, it is never worse than $y_{DHR,t+h|t}$ and $y_{C,t+h|t}$ at the same time.
3. In half the cases (12 out of 24), the RMSE of the combined forecast is between the RMSEs of the individual models.
4. In the remaining cases (12 out of 24), the combined forecast is better than or equal to the previous forecasts (DHR and causal).
5. The improvement is observed mainly for $h = 1$ (with the exception of the 10BT) and for the JTC variable, as the table 8 results suggest.

The rationale behind these results might be the following. Theoretically, for horizon 1, the forecast

TABLE 11 RMSE of the Different 1-Step-Ahead Forecasts of JTC
Forecast Sample: 2000–2002

	DHR	CAUSAL	AVE	REG	REG(12)*	VAR
RMSE	0.039	0.043	0.033	0.029	0.047	0.033

* The forecast sample in this case is 2001–2002.

errors are white noise, so the probability of these errors from the two models being a different sign is 0.5. When this occurs, the forecast average might be closer to the real value, because the average compensates for overprediction by one method with underprediction in the other. For longer forecast horizons, the theoretical forecast errors behave as $MA(h-1)$. If there is a turning point, it is possible that none of the models captures it, especially for a large h, and the probability of forecast errors of different signs might be lower without any possibility of compensation. Because the major gains are observed in the case of JTC, we concentrated on this variable to analyze other forecast combinations.

We analyzed two other forecasting combinations: linear regressions and variance-covariance methods. (For a recent review of these methods see, e.g., Newbold and Harvey 2002.) Consider the following regression

$$y_{t+h} = \beta_0 + \beta_1 y_{1,t+h|t} + \beta_2 y_{2,t+h|t} + e_{t+h} \quad (14)$$

where e_{t+h} can behave as $MA(h-1)$ and let

$$y^{REG}_{t+h|t} = \hat{\beta}_0 + \hat{\beta}_1 y_{1,t+h|t} + \hat{\beta}_2 y_{2,t+h|t} \quad (15)$$

where $\hat{\beta}_i$ is the OLS estimator of β_i, and the subindices (1 and 2) show the two competing models. By using the regression equation (15), problems such as severe multicollinearity are possible, especially if the processes to be forecast are not stationary. It is also important to have a large enough number of forecasts in order to estimate the regression coefficients, leaving the remaining h to make true ex-ante forecasts. Otherwise, the true values of the variables are used to estimate the coefficients, and the associated RMSEs are only in-sample (rather than out-of-sample) accuracy measures.

The error variance of the linear combination of the two forecasts

$$y^{VAR}_{t+h|t} = (1-\lambda) y_{1,t+h|t} + \lambda y_{2,t+h|t} \quad (16)$$

with $0 \le \lambda \le 1$, is minimized if

$$\lambda = \frac{\sigma_1^2 - \rho\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2} \quad (17)$$

where σ_1^2 and σ_2^2 are the variances of the forecast errors of the DHR and causal models, respectively, and ρ is the correlation coefficient. Note that this is the same as the regression method when $\beta_0 = 0$ and $\beta_1 + \beta_2 = 1$. When the correlation coefficient between the two forecast errors is zero, λ can be estimated from the data as

$$\hat{\lambda} = \frac{\sum_{t=T}^{T+N} e^2_{1,t+h}}{\sum_{t=T}^{T+N} e^2_{1,t+h} + \sum_{t=T}^{T+N} e^2_{2,t+h}} \quad (18)$$

Table 11 shows RMSE values for 1-step-ahead forecasts of JTC using both original and combined alternatives. In columns 5 and 6, we report the forecasts based on linear combinations formed with OLS coefficients as in equation (15). We tried two different alternatives: 1) using all the forecast sample to estimate the coefficients (REG in table 11), and 2) using the coefficients estimated with the data for 2000 (12 data points) to compute the linear combination forecasts for 2001 and 2002 (REG(12) in table 11). In column 7, we present the RMSE of the linear combination of the forecasts equation (16) when λ is estimated in equation (18). For comparison purposes, we also added in columns 2 to 4 the RMSE of the DHR, causal, and average forecasts. Notice that the smallest RMSE is given through the regression method REG, but recall that both REG and VAR forecasts are not truly ex-ante, because they use the future values of the variable in order to estimate the coefficients of the linear combinations. When we restrict the sample used in the regression equation (14) to the observations corresponding to 2000 and use these coefficients to compute the true ex-ante forecasts, the RMSE for 2001 and 2002 deteriorates to 0.047. We also tried to make true ex-ante forecasts using the restricted sample with the variables in first differences, but the RMSE deteriorated even more. Overall, the best forecasts were obtained by the average forecast of the DHR and causal models.

CONCLUSIONS

The recent experience of the Madrid metropolitan area shows that it is possible to reverse declining historical trends in public transport ridership. This was achieved through an integrated fare scheme based on low-cost travel tickets and improvements in the quality of service. To adequately plan for future public transportation facilities requires reliable predictions of public transport demand that take into account the users response to changes in prices and the characteristics of the service.

Using recent monthly data, we addressed the problem of forecasting the demand for a large number of tickets that are subject to multiple, complex calendar effects, changing supply service, and changing seasonality, in addition to superimposition of outliers. Two different approaches were used to deal with these issues. The first one is a causal model based on a transfer function dynamic model that allows the incorporation of intervention and exogenous variables in a flexible way. The other is the Dynamic Harmonic Regression model, a new variant of unobserved component models with time-varying parameters, that allows the trend and the seasonal components to adapt as soon as the new information becomes available.

Both methodologies are capable of dealing with the nonstationarity and strong seasonality features that characterize the database. Additionally, all series exhibit the presence of large outliers related to several causes—Easter and trading day effects, many strikes, and abrupt changes in public transport supply at the end of the sample—which considerably complicate the forecasting exercise. The estimation results indicate that the effects of these input variables have the expected signs and are highly significant from a quantitative point of view. However, their effects change considerably for different types of tickets and transportation modes.

Several periods-ahead forecasts (1, 6, 12, and 24 months) for the 6 variables considered in this paper were obtained for 2000, 2001, and 2002. The forecasting for this period was particularly difficult given the major increase in Metro services just before this period. This meant a mixture of effects that influenced not only total demand but, also, temporary passenger shifts among different modes.

We tried to make this forecasting exercise as complete as possible by adopting a large variety of error measures. Forecast accuracy was assessed using four different aggregate measures as well as variants of the so-called model-free testing procedures. Recent developments have shown that these tests are relevant in a wide variety of circumstances (nonquadratic and asymmetric loss functions, serially correlated errors, and non-Gaussian distributions) where previous tests were not applicable. As our forecasting exercise illustrates, they have been very valuable in assessing statistical significance of the differences found between the individual modeling alternatives. We also looked at the accuracy gain derived from pooling the individual forecasts. Although the gain is not uniform for all variables and forecasting horizons, it is almost uniform for all variables in the case of one-step-ahead forecasts and tends to decrease as the forecasting horizon grows. For the most important variables in this dataset (TC and JTC), however, the pooling alternative seems to work well at all forecasting horizons. Finally, we have shown how careful attention is needed in distinguishing between in-sample and truly out-of-sample comparisons when assessing forecast accuracy.

ACKNOWLEDGMENTS

The authors are very grateful to A. Martín Arroyo, three anonymous referees, and Guest Editors Keith Ord and Peg Young for very helpful comments and suggestions that have led to a much better paper. We are also grateful to the Consorcio de Transportes de Madrid for providing the data. This research was sponsored by Comunidad Autónoma de Madrid, contract grant 06-0170/2000 and the Spanish Ministery of Science and Technology, DGICYT contract grant BEC2002-00081.

REFERENCES

Armstrong, J. and F. Collopy. 1992. Error Measures for Generalizing About Forecasting Methods: Empirical Comparisons. *International Journal of Forecasting* 8:69–90.

Baumgartner, J.P. 1995. Beyond Transportation Forecasting. *International Journal of Transportation Economics* 22:3–14.

Box, G.E.P. and G.C. Tiao. 1975. Intervention Analysis with Applications to Economic and Environmental Problems. *Journal of the American Statistical Association* 70:70–79.

Bujosa, M., A. García-Ferrer, and P.C. Young. 2002. An ARMA Representation of Unobserved Component Models Under Generalized Random Walk Specifications: New Algorithms and Examples. Mimeo.

Carbajo, J.C. 1988. The Economics of Travel Passes. *Journal of Transport Economics and Policy* 22:153–173.

Diebold, F.X. 2003. *Elements of Forecasting.* Mason, OH: Thomson/South-Western.

Diebold, F.X. and R.S. Mariano. 1995. Comparing Predictive Accuracy. *Journal of Business and Economic Statistics* 13:253–263.

Fildes, R. 1992. The Evaluation of Extrapolative Forecasting Methods. *International Journal of Forecasting* 8:81–98.

FitzRoy, F. and I. Smith. 1999. Season Tickets and the Demand for Public Transport. *Kyklos* 52:219–238.

García-Ferrer, A., M. Bujosa, A. de Juan, and P. Poncela. 2003. Demand Forecast and Elasticities Estimation of Public Transport. Mimeo.

García-Ferrer, A. and R.A. Queralt. 1997. A Note on Forecasting International Tourism Demand in Spain. *International Journal of Forecasting* 13:539–549.

Gillen, D.W. 1977. Alternative Policy Variables to Influence Urban Transport Demand. *Canadian Journal of Economics* 10:686–695.

Harvey, A. 1989. *Forecasting Structural Time Series Models and the Kalman Filter,* 2nd edition. Cambridge, UK: Cambridge University Press.

Harvey, D.I., S.J. Leybourne, and P. Newbold. 1997. Testing the Equality of Prediction Mean Squared Errors. *International Journal of Forecasting* 13:281–291.

Kitagawa, G. 1981. A Nonstationary Time Series Model and Its Fitting by a Recursive Algorithm. *Journal of Time Series Analysis* 2:103–116.

Lave, C.A. 1994. A Behavioral Approach to Modal Split Forecasting. *The Economics of Transport,* vol. 1. Edited by H. Mohring. Aldershot, England: Elgar Publications. 215–232.

Lythgee, W.F. and M. Wardman. 2002. Demand for Rail Travel to and from Airports. *Transportation* 29(2):125–143.

Mariano, R.S. 2002. Testing Forecast Accuracy. *A Companion to Economic Forecasting.* Edited by M. Clements and D. Hendry. Oxford, England: Blackwell Publishers. 284–297.

Matas, A. and J.L. Raymond. 2003. Demand Elasticity on Tolled Motorways. *Journal of Transportation and Statistics* 6(2/3):91–108.

Newbold, P. and D.I. Harvey. 2002. Forecast Combination and Encompassing. *A Companion to Economic Forecasting.* Edited by M. Clements and D. Hendry. Oxford, England: Blackwell Publishers. 268–283.

Pucher, J. and S. Kurth. 1996. Verkehrsverbund: The Success of Regional Public Transport in Germany, Austria and Switzerland. *Transport Policy* 2:279–291.

Peña, D. 2001. Outliers, Influential Observations, and Missing Data. *A Course in Time Series Analysis.* Edited by D. Peña, G.C. Tiao, and R.S. Tsay. New York, NY: John Wiley. 136–170.

Slavin, H. 1996. An Integrated Dynamic Approach to Travel Demand Forecasting. *Transportation* 23(3):313–350.

West, M. and J. Harrison. 1997. *Bayesian Forecasting and Dynamic Linear Models,* 2nd edition. New York, NY: Springer-Verlag.

Young, P.C. and C. Ng. 1989. Variance Intervention. *Journal of Forecasting* 8:399–416.

Young, P.C., D. Pedregal, and W. Tych. 1999. Dynamic Harmonic Regression. *Journal of Forecasting* 18:369–394.

APPENDIX A: BGF ALGORITHM

In the DHR model, T_t and S_t consist of a number of DHR components, $s_t^{p_j}$, with the general form

$$s_t^{p_j} = a_{j_t}\cos(\omega_j t) + b_{j_t}\sin(\omega_j t),$$

where p_j and $\omega_j = 2*\pi/p_j$ are the period and the frequency associated with each j-th DHR component respectively. T_t is the zero frequency term $(T \equiv s_t^{\infty} = a_{0t})$, while the seasonal component is

$$S_t = \sum_{j=1}^{R} s_t^{p_j}$$

where $j = 1,\ldots,R$ are the seasonal frequencies. Hence, the complete DHR model is

$$y_t^{dhr} = \sum_{j=0}^{R} s_t^{p_j} + e_t =$$

$$\sum_{j=0}^{R} \{ a_{j_t}\cos(\omega_j t) + b_{j_t}\sin(\omega_j t)\} + e_t$$

The oscillations of each DHR component, $s_t^{p_j}$, are modulated by a_{j_t} and b_{j_t}, which are stochastic time-variable parameters that follow an AR(2) stochastic process with at least one unit root

$$\left.\begin{aligned}(1-\alpha_j B)(1-\beta_j B)a_{j_t} &= \xi_{j_{t-1}} \\ (1-\alpha_j B)(1-\beta_j B)b_{j_t} &= \xi_{j_{t-1}}\end{aligned}\right\} \; 0 \le \alpha_j, \beta_j \le 1,$$

$$\{\xi_j\} \sim \text{w.n. } N(0, \sigma_j^2)$$

therefore, nonstationarity is allowed in the various components.

This DHR model can be considered a straightforward extension of the classical harmonic regression model, in which the gain and phase of the harmonic components can vary as a result of estimated temporal changes in the parameters a_{j_t} and b_{j_t}.[7]

The method for optimizing the hyper-parameters of the model (i.e., the variances

$\left[\boldsymbol{\sigma}^2_{dhr} = \sigma_0^2, \sigma_1^2, \ldots, \sigma_R^2\right]'$

of the processes ξ_j, $j = 0 \ldots, R$, and the variance σ_e^2 of the irregular component) was formulated by Young et al. (1999) in the frequency domain and is based on expressions for the pseudospectrum of the full DHR model:

$$f_{dhr}(\omega, \boldsymbol{\sigma}^2) = \sum_{j=0}^{R} \sigma_j^2 S_j(\omega) + \sigma_e^2;$$

$$\boldsymbol{\sigma}^2 = \left[\boldsymbol{\sigma}^2_{dhr}, \sigma_e^2\right]'$$

where $\sigma_j^2 S_j(\omega)$ are the pseudospectra of the DHR components s^{p_j}, and σ_e^2 is the variance of the irregular component. The optimization processes seek the vector $\boldsymbol{\sigma}^2$ that minimizes[8]

$$\min_{[\sigma^2] \in \mathbb{R}^{R+1}} \left\| f_y(\omega) - f_{dhr}(\omega, \boldsymbol{\sigma}^2) \right\|, \tag{19}$$

[7] The main difference between the DHR model and related techniques, such as Harvey's structural model (Harvey 1989), lies in the formulation of the UC model for the periodic components and the method of optimizing the hyper-parameters.

[8] Young et al. (1999) modified the problem using the residual variance $\hat{\sigma}^2$ from the fitted AR model as an estimation of σ_e^2, and then dividing by $\hat{\sigma}^2$, so to seek the vector $\mathbf{NVR} = \left[1, NVR_0, \ldots, NVR_R\right]$, where $NVR_j = \sigma_j^2 / \hat{\sigma}^2$.

where $f_y(\omega)$ is the spectrum of the observed time series. The DHR components follow nonstationary ARMA processes; therefore, in order to find an ordinary least squares (OLS) solution for equation (19), the unit modulus AR roots of $f_{dhr}(\omega, \boldsymbol{\sigma}^2)$ need to be eliminated. The DHR components $s_t^{p_j}$ are stochastic processes of the form

$$\varphi_j(\boldsymbol{B}) s_t^{p_j} = \theta_j B \xi_{j_{t-1}}, \qquad \{\xi_{j_t t}\} \sim \text{w.n. } N\left(0, \sigma^2_{\xi_{j_t}}\right)$$

and the pseudospectrum of the complete DHR model is

$$f_{dhr}(\omega, \boldsymbol{\sigma}^2) = \sum_{j=1}^{R} \sigma_j^2 \frac{\theta_j(e^{-i\omega})\theta_j(e^{i\omega})}{\varphi_j(e^{-i\omega})\varphi_j(e^{i\omega})} + \sigma_e^2.$$

If equation (19) is multiplied by the function

$\Psi(\omega) = \Phi(e^{-i\omega})\Phi_j(e^{i\omega})$,

where $\Phi(B)$ is the minimum order polynomial with all unit modulus AR roots of the complete DHR model, the algorithm minimizes

$$\min_{\sigma^2 \in \mathbb{R}^{R+2}} \left\| \Psi(\omega) f_y(\omega) - \Psi\omega \cdot f_{dhr}(\omega, \sigma^2) \right\| \tag{20}$$

but, because in equation (20) all the unit modulus AR roots cancel, the minimization problem can be solved by OLS.

Finally, $f_y(\omega)$ can be substituted in equation (20) by the estimated AR spectrum

$$\hat{f}_y(\omega) = \hat{\sigma}^2\left(\phi_y(e^{-i\omega})\phi_y(e^{i\omega})\right)^{-1},$$

where $\phi_y(B)$ is an AR polynomial fitted to the series, and $\hat{\sigma}^2$ is the residual variance of the fitted AR model. The size, shape, and location of the spectral peaks of $\hat{f}_y(\omega)$ are used to identify the models of each of the DHR components s^{p_j}.

Once the models of the DHR components have been identified, and the hyper-parameters

$\sigma^2_{dhr} = \left[\sigma_0^2, \sigma_1^2, \ldots, \sigma_R^2\right]'$

have been optimized by OLS, the DHR components T_t, S_t, and e_t can be estimated using the Kalman Filter and Fixed Interval Smoothing.

Predicting and Monitoring Casualty Numbers in Great Britain

ROBERT RAESIDE

Transport Research Institute
Napier University
Sighthill Court
Edinburgh EH11 4BN Scotland
Email: r.raeside@napier.ac.uk

ABSTRACT

This paper presents analyses and forecasts of trends related to road traffic and pedestrian casualties and fatalities in Great Britain. For people killed and seriously injured, these forecasts are based on extrapolation of the absolute number of casualties. For casualties classified as slight, forecasts are made of the rate of casualties per 100 million vehicle-kilometers. Forecasts, using autoregressive models, are then compared with government targets and show that at the aggregate level it is unlikely that, for the numbers who are killed or seriously injured, these targets will be achieved.

INTRODUCTION

One of the key performance measures of the safety of a nation's transport system is the number of people who are killed or seriously injured in road accidents. Apart from the human tragedy, estimates show that each fatal road accident in Great Britain costs £1,447,490 (approximately $2,665,000) while a serious casualty costs £168,260 (approximately $310,000) (DETR 2002). National governments provide targets that traffic managers, infrastructure designers, vehicle manufacturers, and the legal system strive to achieve. The latest set of targets for

KEYWORDS: National road safety trends, statistical forecasting, casualty and fatality prediction.

Great Britain is for the year 2010. Compared with the average of 1994 to 1998, hoped-for reductions are as follows:

- the number of fatal or serious injuries in road accidents by 40%,
- the number of children killed or seriously injured by 50%, and
- slight injuries per 100 million vehicle-kilometers by 10%.

This paper reports on an exploration of recent casualty[1] time series in Great Britain and forecasts these series up to 2010 to determine if it is likely that targets set by the government of Great Britain are achievable. As will be discussed, most forecasting approaches predict the casualties per 100 million vehicle-kilometers; however, this paper attempts to conduct the straightforward forecasting of the time series of the number of casualties. These are then compared with the casualty reduction targets. It is hoped this will be useful to those involved with road safety to determine if targets for the number of those killed or seriously injured in vehicle accidents can be met or if more effort is required.

For the previous targets, which were set in 1987, the aim was to reduce deaths and serious injuries by one-third by 2000 compared with the average for 1981 to 1985. This target was surpassed; in fact, road deaths fell by 39% and serious injuries by 45%. The success in Great Britain has come about through legislative changes aimed at altering driver behavior and improving infrastructure and vehicle crash protection.

The number of casualties is of concern throughout Europe, where there are over 40,000 deaths and 1.7 million people injured per year, directly costing some 160 billion Euros—and the young are most affected (European Commission 2003).

In 2000, the European Commission initiated the European Road Safety action program with the intention of halving the number of those killed or seriously injured in road accidents by 2010. It took 30 years for the previous halving of rates, so this must be seen as ambitious. The focus of the program is on:

[1] *Casualties* are fatal, serious, or slight injuries.

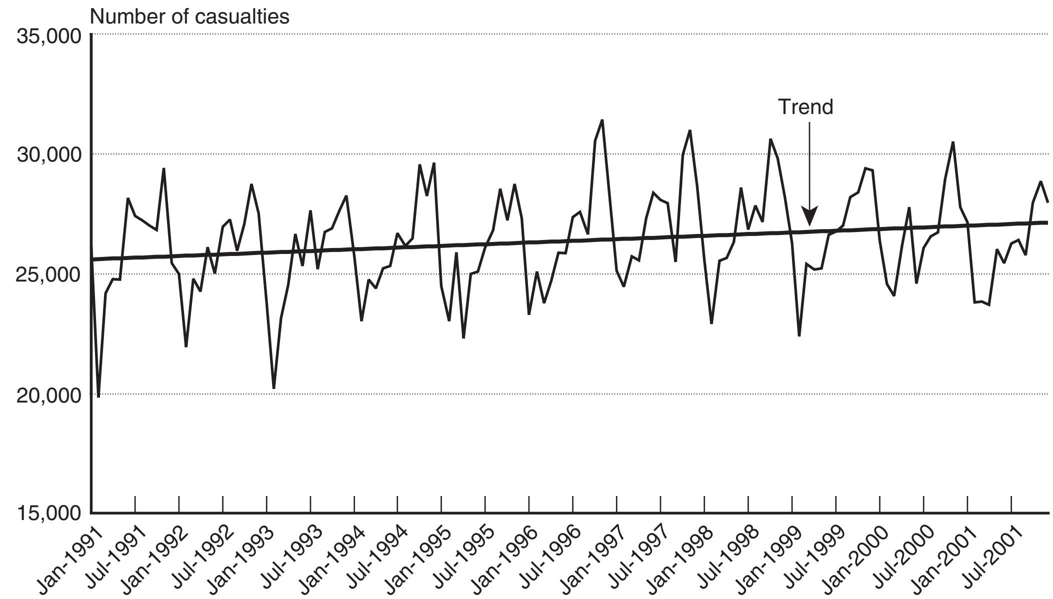
1. Improving driver behavior through education and enforcement of speed limits, use of safety belts, and penalties for drinking/drug use while driving.
2. Encouraging improvements in vehicle design and technological advances to enhance both occupant protection and pedestrian survivability.
3. Encouraging improvement of the road infrastructure.
4. Reducing the risk to people during the transport of commercial goods and while using public transport.
5. Improving emergency services and care for road accident victims.
6. Promoting research into transport safety and improving accident data collection and dissemination.

To enact these policies, the European Commission has produced a European Road Safety Charter as an exemplar of good practice. Government agencies are required to sign and have their compliance with the charter monitored and publicized—hence the need to predict and monitor the casualty series.

Great Britain is now one of the safest of the European countries in terms of road traffic injuries and compares favorably with all Organization for Economic Cooperation and Development (OECD) countries (OECD 2003). However, there is still room for improvement, especially for child pedestrian fatality rates per 100,000 people. As of 2001, Great Britain lagged behind many similar western European countries (Scottish Executive 2002). Of particular concern is that since 1991 the total number of road traffic casualties in Great Britain has shown a slight but significant upward trend. In figure 1, the slope of the fitted trend line indicates an increase of 11.69 each month, which is statistically significant with a P-value of 0.013.

Figure 2, which is broken down by type of casualty, presents a somewhat different picture from the total. The numbers of fatal and serious casualties have decreased markedly over the period, whereas the number of slight injuries has increased. (This last change may reflect improved reporting systems, especially as insurance companies require a police report if anyone is injured.) Given that in Great

FIGURE 1 Casualty Trends in Great Britain: 1991–2001



Britain since 1991 the number of vehicle-kilometers traveled has increased by more than 15%, the risk of fatal or serious injury has been substantially reduced (by 24.5% and 28.1%, respectively, over the period 1991 to 2001).

REVIEW OF PREVIOUS WORK

An often-used approach to forecasting killed and seriously injured (KSI) casualties has been to take a time series of annual rates and extrapolate a fitted negative exponential model. Sometimes allowance is made through the use of disturbance terms for special events, such as the introduction of legislation to make seat-belt wearing compulsory, but in general the models are univariate in nature and incorporate little in the way of explanatory variables. A good example is the work of Broughton (1991) who fitted extensions of the model

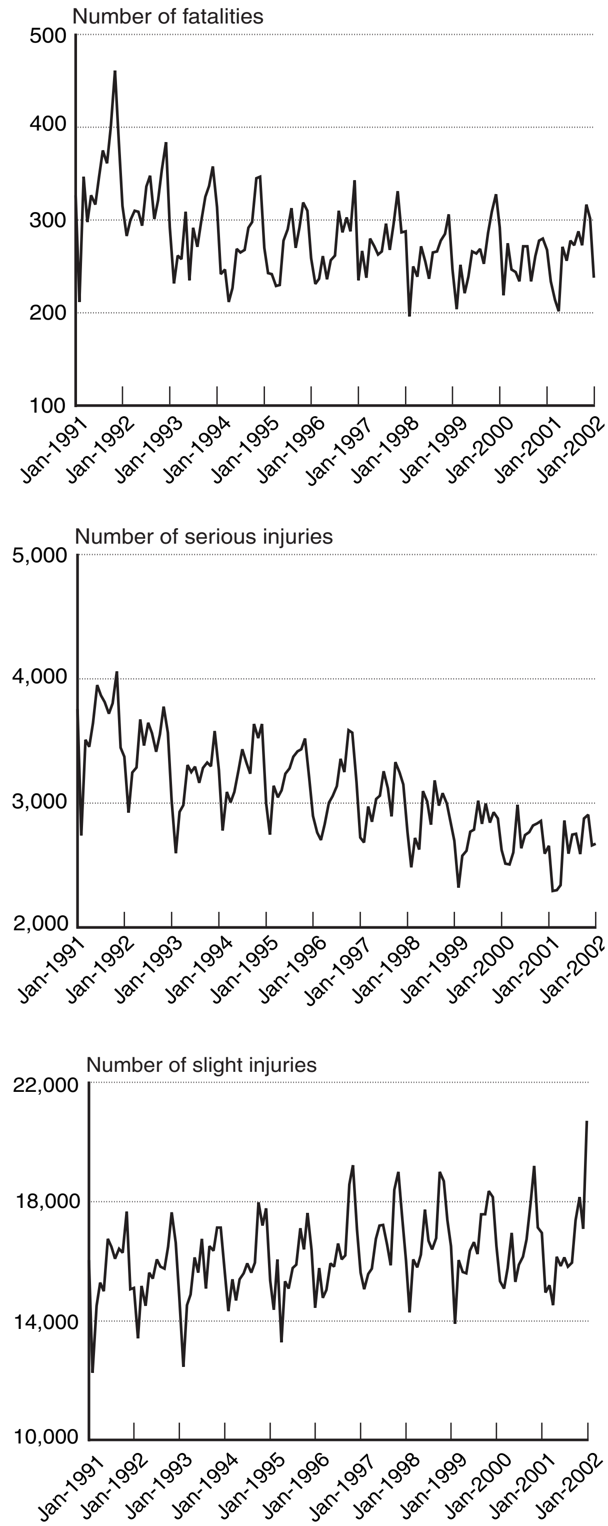
$$\mathit{log}(\mathit{casualties}/\mathit{traffic\ volume}) = a + b^{*}\mathit{year} + \mathit{an\ intervention\ term} \quad (1)$$

to data on Great Britain road casualties from 1949 to 1989. This model gave good forecasts of the number of fatal casualties in 2000: 3,312 with a 90% prediction interval of 2,892 to 3,826. There were, in fact, 3,409 fatalities. Broughton's forecasts of KSI casualty numbers and all casualty numbers underpredicted by 4.8% and 24.8%, respectively. This approach requires the number of vehicle-kilometers driven to be forecast ahead. Thus, either official forecasts of the amount of kilometers driven must be made, or the growth of this series will need to be modeled, perhaps using a sigmoid model, as Oppe (1989) suggested.

Although supported by Nilsson (1997), the use of kilometers as a denominator is problematic because most casualty accidents occur relatively close to the place of residence of the person or persons involved (Petch and Henson 2000; Scottish Executive 2002). Also, the certainty of estimation of a nation's annual driving is not clear. However, an argument can be made that using kilometers as a denominator would serve as a proxy for the number of people driving.

Harvey and Durbin (1986) applied structural time series methods using ARIMA models with an intervention term to the monthly data series of the numbers killed and seriously injured in Great Britain from January 1969 to December 1984. The results demonstrated the effectiveness of the introduction of seat belt legislation. Raeside and White (2004) used ARIMA models for the monthly series of KSIs from 1991 to 2001. From these models, they projected the numbers of fatal and serious casualties in 2010 to show that targets set by the government of Great Britain may be met. But using ARIMA models to predict eight years ahead from

FIGURE 2 Casualty Trends by Type



this short time series must be regarded as speculative, and this is reflected in the relatively wide prediction intervals of these models. A more precise means of assessing progress to targets is required.

Haight (1991), in his editorial commentary on a special issue of *Accident Analysis and Prevention,* advocated models for predicting fatalities with year and traffic volume as variables.

$$Fatalities = a \times b^{year} \times Traffic^{c} \qquad (2)$$

The models were transformed and fitted using Poisson regression. Brude (1995) successfully applied a version of these models to forecast the number of fatalities in Sweden to the year 2000 from data covering 1977 to 1991. Guria and Mara (2001) incorporated this type of modeling into a control chart to give the "probability of achieving the target given the past outcomes of the year." They highlighted the importance of day and month effects on the variability of the casualty series.

Lassarre (2001) extended Smeed's (1949) work to develop a family of structural time series models using Harvey's (1989) approach to estimate fatality time series across 10 European countries. He found that, since 1962, fatalities have decreased at an average annual rate of 6%. Balkin and Ord (2001) applied a stochastic structural equation modeling approach to predict the effect of speed limit changes on the number of fatal crashes on both urban and rural interstate highways in the United States. Seasonal influences can be accounted for in their approach, and comparisons between states were made. They found that the view that higher speeds means more fatalities could not be universally supported.

Page (2001) modeled safety trends in OECD countries from 1980 to 1994 and constructed a safety index comprising population variables, numbers of buses and coaches, employment rates, and rates of alcohol consumption. Page then used these variables in regression models to demonstrate that fatality rates per billion vehicle-kilometers have generally decreased. The OECD countries with the highest index (safest) were Sweden, the Netherlands, Norway, the United Kingdom, and Switzerland, and the lowest were Greece, Belgium, the United States, Portugal, and Spain.

Much discussion of the improving casualty trends has appeared in the literature, some of which is summarized in Raeside and White (2004). These trends appear to be the result primarily of improvements in road infrastructure and the crashworthiness of vehicles. Behavioral and legislative influence appear, from the literature, to be of second order. Figure 2 provides some support for this, as the total number of accidents seems to have remained about the same but fatal and serious accidents have decreased, which indicates greater levels of personal protection. Accident rates have also improved as a result of changes in behavior, especially reductions in rates for child pedestrians and bicycling (DiGuiseppi et al. 1997; Stone and Broughton 2003).

Little use has been made of explanatory models in predicting casualties, with the notable exception of Brannas (1995), who used a Poisson regression model. He based his work on that of Zeger (1988) to successfully forecast road accidents in Vasterbotten County in Sweden using variables representing exposure and weather, plus daylight variables. The model Brannas considered is as follows:

$$\Pr(y_t) = \frac{\lambda_t^{y_t} e^{-\lambda_t}}{\lambda_t!}(t = 1......,T) \tag{3}$$

where

$$\lambda_t = e^{x_t \beta}$$

and x_t is a 1 x k vector of covariates representing weather and daylight. This, however, is more suited to forecasting for micro areas than for national forecasts. The same is true for the numerous Poisson models developed by civil engineers. They are similar to equation (2) but contain variables representing the geometry of the junction, the nature of the conflict, traffic volumes, and major road features, and they can be used to predict accidents at particular road junctions (Maher and Summersgill 1995).

Unfortunately, little has been done to use the predictive models to assess the probability of meeting casualty reduction targets. To achieve this aim, this paper employs simple models based on Broughton's (1991) approach to produce predictive distributions based on numbers rather than rates per 100 million vehicle-kilometers driven. The models do not incorporate traffic volumes (except for slight injuries).

FORECASTS OF CASUALTIES

The annual data series of the numbers of fatal, serious, and slight casualties were taken from table 9.10 of *Road Accidents Great Britain 2002* (DETR 2002). The series were then modeled using autoregressive and linear trend terms. Natural logarithms were used for fatalities and for the serious casualties, so that the models would be negative exponential in nature and similar to that of Broughton (1991), but using the numbers rather than rates per billion vehicle-kilometers and not employing an intervention term. For slight casualties, results of the natural logarithm of the casualty rate per 100,000 vehicle-kilometers allowed comparisons with the official target. The trend variable was formed by subtracting 1970 from the year. The autoregressive models were fitted using SPSS; the exact maximum likelihood method was used. The models for fatal and serious injuries may be written as:

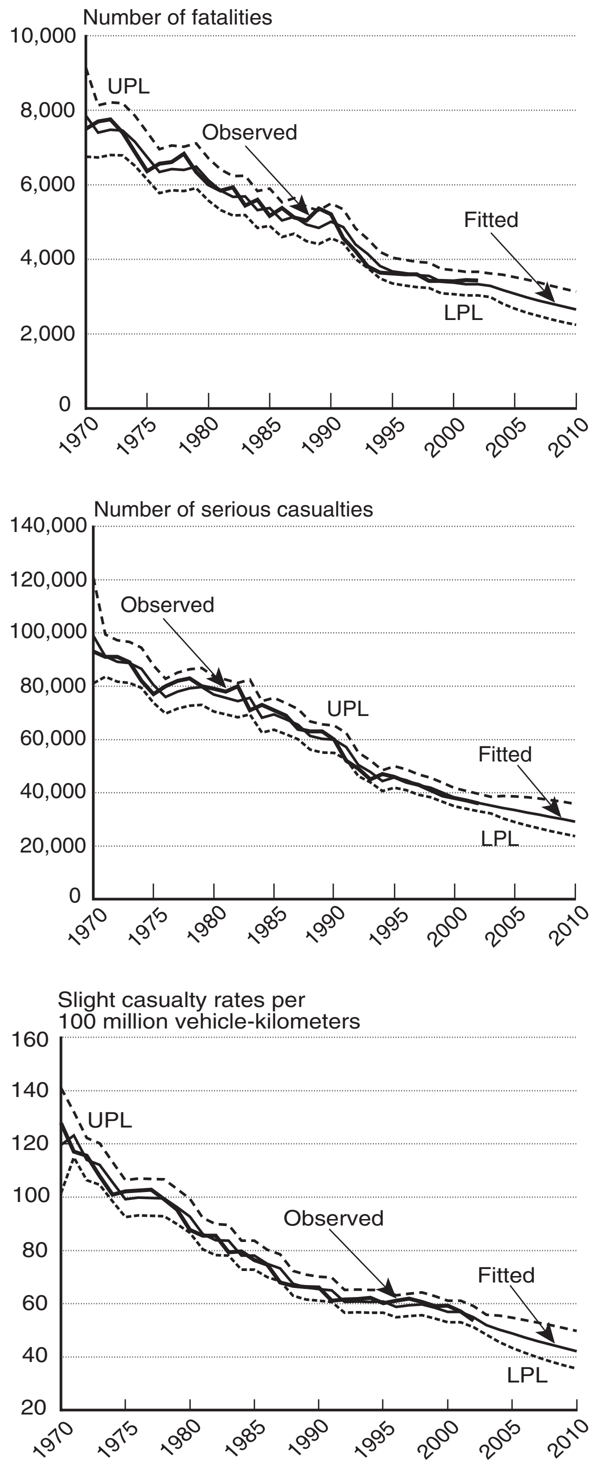
$$ln(casualties\ in\ year\ t) = a + b^*(year–1970) + c^*ln(casualties\ in\ year\ t–1) \tag{4}$$

For slight injuries, casualties per 100 million vehicle-kilometers were used instead of casualty numbers. Table 1 shows the coefficients and fit parameters of

TABLE 1 Coefficients and Fit of Models for Forecasting the Natural Logarithm of the Numbers of Fatal and Serious Casualties and of the Rate of Slight Casualties

	Number of fatalities		Number of serious casualties		Slight casualty rates per 100 million vehicle-kilometers	
Parameter	Coefficient	Standard error	Coefficient	Standard error	Coefficient	Standard error
Autoregressive term	0.684	0.130	0.831	0.093	0.840	0.088
Trend	–0.027	0.002	–0.030	0.003	–0.026	0.003
Constant	8.996	0.043	11.534	0.065	0.205	0.054
Standard error	0.045		0.042		0.034	
Adjusted R^2	0.965		0.966		0.966	

FIGURE 3 Forecasts of Fatalities, Serious Casualties, and Slight Casualties



KEY: LPL = lower prediction level; UPL = upper prediction level

the models. The models all fitted well, and it should be noted that all models show a significant downward trend.

Figure 3 presents the forecasts generated by each of these models, with the lower and upper prediction level limits (LPL and UPL) also displayed. The forecast of fatalities for 2010 ranges over a 95% prediction interval from 2,246 to 3,142 with an expectation of 2,656 deaths. This is a 26% reduction from the 1994 to 1998 average of 3,578, which is used as the government's baseline for measuring improvements. This is disappointing for planners and policymakers—if the predictive distribution for the log of causalities is considered approximately normal, then the chance of attaining or exceeding the 40% reduction target is less than 14%. By adding the forecasts of the number of serious casualties to the forecasts of fatalities, we get a 95% prediction interval for KSI casualties of 26,002 to 38,989, with an expectation of 31,839. The expected value of KSI casualties is only 33% less than the baseline figure of 47,656. The probability of meeting or exceeding the target of 40% is only 31.2%. Thus, the attainment of this target is unlikely.

The government target for slight injuries was a 10% reduction from the 1994 to 1998 baseline of 46.30 slight injuries per 100 million vehicle-kilometers. The forecasts show that a reduction of just over 9% is expected. The probability of achieving or surpassing the 10% target is 0.476. Thus it appears that in Great Britain, the road casualty improvement targets for slight injuries may not be reached. However, the prospects are more optimistic than for the KSI series if injuries are assumed to be a function of the number of trips and not the number of vehicle-kilometers. This "optimism" may well be the consequence of increased traffic volumes rather than improved safety.

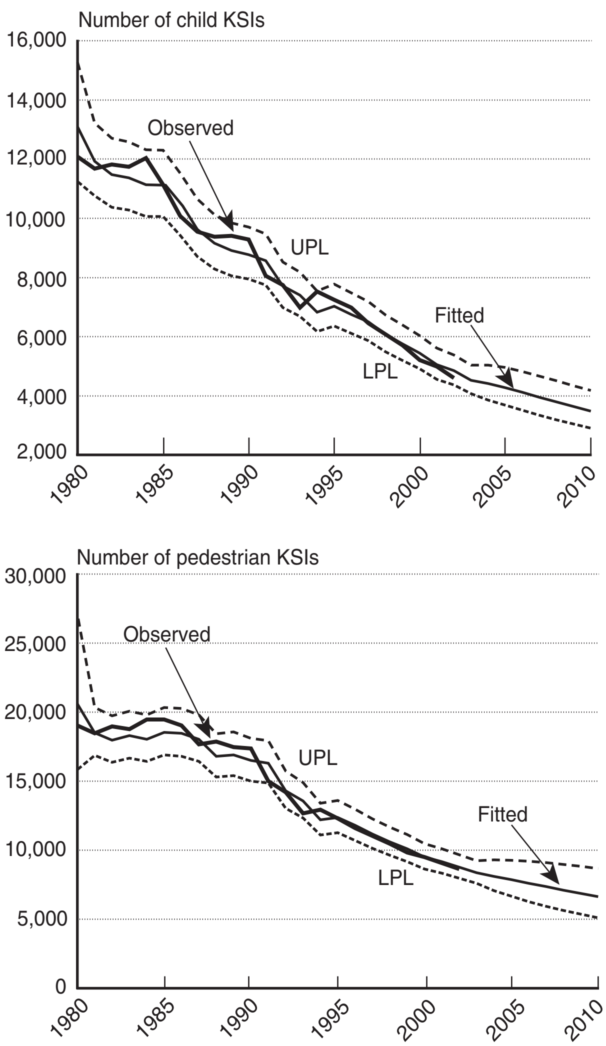
The other important target of halving the number of children who are killed or seriously injured by 2010 and the pedestrian casualty series will be examined next. Table 2 presents coefficients of the model of the natural logarithm for child and pedestrian KSIs. Again the models fitted well and displayed a significant downward trend. The forecasts along with the LPLs and UPLs produced from these models are displayed in figure 4.

Child KSI casualties are forecasted to fall to 3,482 with a 95% prediction interval of 2,899 to 4,182, a reduction of just over 50% from the 1994 to 1998 baseline. Although this is close to the target

TABLE 2 Coefficients and Fit of Models for Forecasting the Natural Logarithm of the Number of Child and Pedestrian KSI Casualties

	Child casualties		Pedestrian casualties	
Parameter	Coefficient	Standard error	Coefficient	Standard error
Autoregressive term	0.618	0.180	0.877	0.086
Trend	–0.044	0.003	–0.037	0.006
Constant	9.966	0.074	10.340	0.138
Standard error	0.047		0.044	
Adjusted R^2	0.948		0.950	

FIGURE 4 Forecasts of Child and Pedestrian KSIs



KEY: LPL = lower prediction level; UPL = upper prediction level

of 50%, the probability of meeting or exceeding the target is 0.530. Pedestrian casualties are projected to be reduced by 43%, with the number of KSI pedestrians falling from 11,667 to 6,652. As no target is given for pedestrian casualty reduction, no probabilities of attainment can be computed. Table 3 shows the probabilities of attaining the targets where targets are available.

TABLE 3 Probabilities of Attaining Road Accident Improvement Targets by 2010

Accident series	Percentage reduction target for 2010	Probability of attainment
Fatalities	40	0.133
Killed or seriously injured (KSI)	40	0.312
Slight	10	0.476
Child KSI	50	0.530

CONCLUSIONS

Governments of many developed countries set periodic road safety targets. The latest targets in Great Britain are for 2010 and relate to the number of people killed or seriously injured on Britain's roads and the rate of slight injuries per 100 million vehicle-kilometers driven. For effective use of resources it is important to monitor progress to these targets. This paper presents a methodology for forecasting casualty trends and monitoring progress toward targets.

The paper presents trends in casualty numbers for fatal, serious, and slight injuries, as well as those involving pedestrians (with a separate category for children). Progress in improving casualty numbers seems promising for children and for slight injuries, but attaining the reduction targets for 2010 is uncertain. For the killed and seriously injured category, the probability of attaining a 40% reduction is fairly slim, and a greater effort is needed to ensure convergence on this target. One possibility for reducing casualties is to apply and enforce measures to reduce traffic levels in Great Britain. While targets should be aspirational rather than set at easily attainable levels, the issue of road traffic accidents is politically contentious. Accounting for the marked seasonality of the data may provide targets that are more likely to be attainable. This is the subject of future research.

ACKNOWLEDGMENTS

I wish to gratefully acknowledge the suggestions for improvements and changes to the paper made by the referees and editors.

REFERENCES

Balkin, S. and J.K. Ord. 2001. Assessing the Impact of Speed-Limit Increases on Fatal Interstate Crashes (with Discussion). *Journal of Transportation and Statistics* 4(1):1–26.

Brannas, K. 1995. Prediction and Control for a Time-Series Count Data Model. *International Journal of Forecasting* 11:263–270.

Broughton, J. 1991. Forecasting Road Accident Casualties in Great Britain. *Accident Analysis and Prevention* 23(5):353–362.

Brude, U. 1995. What Is Happening to the Number of Fatalities in Road Accidents? A Continuous Monitoring of Development up to the Year 2000. *Accident Analysis and Prevention* 27(3):405–410.

Department of the Environment, Transport and the Regions (DETR). 2002. *Road Accidents Great Britain 2002: The Casualty Report.* Norwich, UK: Department of Transport, The Stationery Office.

DiGuiseppi, C., I. Roberts, and L. Li. 1997. Influence of Changing Travel Patterns on Child Death Rates from Injury: Trend Analysis. *British Medical Journal* 314(7082):710–717.

European Commission. 2003. *Halving the Number of Road Accident Victims in the EU by 2010: A Shared Responsibility.* Available at http://eurpoa.eu.nt/comm/dgs/energy_transport/index_fr.html, as of May 2004.

Guria, J. and K. Mara. 2001. Predicting Performance of Annual Safety Outcomes. *Accident Analysis and Prevention* 33(3):387–392.

Haight, F.A. (ed). 1991. Theoretical Models for Traffic Safety. *Special Issue—Accident Analysis and Prevention* 23(1):1–5.

Harvey, A.C. 1989. *Forecasting Structural Time Series Models and the Kalman Filter.* Cambridge, UK: Cambridge University Press.

Harvey, A.C. and J. Durbin. 1986. The Effects of Seat Belt Legislation on British Road Casualties: A Case Study in Structural Time Series Modelling (with Discussion). *Journal of the Royal Statistical Society, Series A* 149:187–227.

Lassarre, S. 2001. Analysis of Progress in Road Safety in Ten European Countries. *Accident Analysis and Prevention* 33(6):743–751.

Maher, M. and I. Summersgill. 1995. A Comprehensive Methodology for the Fitting of Predictive Accident Models. *Accident Analysis and Prevention* 28(3):281–296.

Nilsson, G. 1997. *Methods and Necessity of Exposure Data in Relation to Accident and Injury Statistics: Future Development of IRTAD,* IRTAD special report by the Swedish Road and Transport Research Institute. Available at http://www.bast.de/htdocs/fachthemen/irtad/english/abstschweden.html, as of September 2004.

Oppe, S. 1989. Macroscopic Models for Traffic and Traffic Safety. *Accident Analysis and Prevention* 21(3):225–232.

Organization for Economic Cooperation and Development (OECD). 2003. *International Road Traffic Accident Database* (IRTAD). Available at http://www.bast.de/htdocs/fachthemen/irtad/index.htm, as of May 2004.

Page, Y. 2001. A Statistical Model To Compare Road Mortality in OECD Countries. *Accident Analysis and Prevention* 33(3):371–385.

Petch, R.O. and R.R. Henson. 2000. Child Road Safety in the Urban Environment. *Journal of Transport Geography* 8:197–211.

Raeside, R. and D. White. 2004. Predicting Casualty Numbers in Great Britain, paper No. 04-3684. Proceedings of the 83rd Annual Meeting of the Transportation Research Board, January 11–15, 2004. Washington, DC.

Scottish Executive. 2002. *Road Accidents Scotland.* Edinburgh, Scotland.

Smeed, J.R. 1949. Some Statistical Aspects of Road Safety Research. *Journal of the Royal Statistical Society, Series A* 1:1–34.

Stone, M. and J. Broughton. 2003. Getting Off Your Bike: Cycling Accidents in Great Britain in 1990–1999. *Accident Analysis and Prevention* 35(4):549–556.

Zeger, S.L. 1988. A Regression Model for Time Series of Counts. *Biometrika* 75:621–629.

Estimating the Impact of Recent Interventions on Transportation Indicators

KEITH ORD[1,*]
PEG YOUNG[2]

[1] The McDonough School of Business
Georgetown University
37th and "O" Streets, NW
Washington, DC 20057

[2] Bureau of Transportation Statistics
U.S. Department of Transportation
400 Seventh Street, SW, Room 3430
Washington, DC 20590

ABSTRACT

Whenever an unusual event disrupts the structural patterns of a time series, one of the aims of a forecaster is to model the effects of that event, with a view to establishing a new basis for forecasting. Intervention analysis has long been the method of choice for such adjustments, but it is often represented as a procedure for dealing with events in the middle of the time series rather than for the most recent observations. In this paper, we develop a method, termed the three-intervention approach, to provide a flexible solution to this problem. We examine its application for a number of transportation series that were disrupted by the tragic events of September 2001. Analyses of the series using up to six months of post-event data show good agreement with results based on longer post-event series, and suggest that the proposed method will often provide adequate modifications to a series in a timely manner. The method is applicable to most economic time series, but has been tested only for transportation series.

INTRODUCTION

The time that elapses between the occurrence of events and the production of the statistical records

Email addresses:
* Corresponding author—ordk@georgetown.edu
P. Young—peg.young@bts.gov

KEYWORDS: Airline traffic, forecasting, intervention analysis, process control, structural time series, transportation indicators.

that describe them will always be too long. The production speed of monthly series will continue to improve in the area of transportation as elsewhere, but time lags in production are inevitable. Thus, when new data become available, it is important to remove any distortions caused by recent events so that we can both understand what has happened and predict future developments. This quality will be jeopardized whenever the latest values do not provide a clear indication of the true situation. Common examples include the disruption of travel patterns because of extreme weather conditions or a loss of service due to a labor dispute. We refer to such effects as interventions in the time series, which may return to its previous level rapidly, slowly, or not at all. In such circumstances, the data may be misleading and require adjustment before underlying trends can be discerned.

When an intervention occurs some months in the past, data are available on either side of the affected month(s), and we may use the more conventional methods of intervention analysis to make adjustments (see, e.g., DeLurgio 1998, chapter 12 or Harvey 1989, section 7.6). In these circumstances, the nature of the intervention can usually be identified and it remains only to estimate the model parameters.

A rather different problem arises when the intervention has just taken place. The same general methods are appropriate but the amount of data available to describe the event is necessarily very limited. Further, the nature of the change (e.g., permanent or temporary) may be uncertain. Nevertheless, we wish to ascertain the nature of the change and to estimate its impact as quickly as possible so that series predictability may be restored. This paper develops such an early response system and tests its performance empirically. In the next section we describe the basic ideas, and the following section describes their implementation in a structural modeling framework. After that we present the analysis of a single series to illustrate ideas. The general approach and results for a number of transportation series are then summarized and discussed. The paper concludes with final comments on the proposed form of intervention analysis.

AN INTERVENTION ANALYSIS FRAMEWORK

The basic ideas for monitoring a process over time are central to statistical process control (SPC). The use of time-series modeling in SPC follows from the seminal work of Alwan and Roberts (1988, 1995). In SPC, we conventionally distinguish two sources of variation (see Alwan 2000, pp. 217–220):

- **Common cause variation** reflects the natural variation inherent in the process, and
- **Special (or assignable) cause variation** is any variation in the process introduced by a recognizable factor (e.g., a worn tool or a poorly trained operator).

In the present context, we are interested in identifying recent changes, and the possible types of assignable cause need to be identified more clearly. Thus, it is useful to divide assignable cause variation into three categories, which we may examine by different means:

- **Additive outlier (AO)**—A factor has a short-term temporary impact on the series, which is resolved within a single observational period. The series then returns to its original state. For example, the effects of a blizzard on airplane traffic would typically be of this nature.
- **Temporary change (TC)**—A factor has a relatively short-term impact on the series, which returns to its previous state over a number of time periods. For example, a prolonged strike in an industry will reduce production, which gradually recovers over the next several months.
- **Level shift (LS)**—A factor causes the series to shift to a new level, and the series stays at that new level. For example, a change in the law on seat belts will lead to a shift in the number of fatalities.

The emphasis in these three types of intervention is on a sudden change in the series, and such changes are the basis of our present study. By contrast, it is possible to observe slowly changing conditions that may lead to fundamental changes in the series of interest. For example, improved engine design might produce greater fuel efficiency ratings for automobiles, but such an effect would be seen only very gradually in an aggregated series on aver-

TABLE 1 Statistical Definitions of Assignable Causes in Time Series

Assignable cause	Statistical definition	Value of Y_t: $t < T$	$t = T$	$t > T$
Additive outlier (AO)	$X_t = 1$, *when* $t = T$; $X_t = 0$, *otherwise*	μ	$\mu + I$	μ
Temporary change (TC)	$X_t = d^{t-T}$, *when* $t \geq T$; $X_t = 0$, $t < T$	μ	$\mu + I$	$\mu + Id^{t-T}$
Level shift (LS)	$X_t = 1$, *when* $t \geq T$; $X_t = 0$, $t < T$	μ	$\mu + I$	$\mu + I$

age miles per gallon. Such changes will be incorporated into the trend terms of our models and so are not identified directly. Finally, we note that the seasonal pattern in the series may vary over time. For example, airlines may alter their seasonal pricing strategies, which would lead to a shift in travel patterns. Both these effects are important for longer term predictability, but they are less critical in the shorter term.

We make the three assignable causes operational in accordance with the definitions in table 1, where we assume that the event takes place at time T, and X_t is an indicator (or dummy) variable that indicates the timing of the event. Thus, in the simplest case of a series that has a constant mean, μ, over time with random disturbances ε, except for interventions, the series y_t would be modeled as:

$$y_t = \mu + X_t I + \varepsilon_t \quad (1)$$

where I measures the magnitude of the intervention.

Figures 1 through 3 provide graphical examples of the AO, TC, and LS, respectively. A major question with the TC is the rate of adjustment. Indeed, the AO may be viewed as a TC with adjustment factor, $d = 0$, or sufficiently small to disappear within one time period. Likewise, an LS may be viewed as a TC with d approaching 1.0. Keep in mind that we will typically have a very limited amount of data with which to estimate these effects, and that the direct estimation of adjustment rates is difficult even when we have a considerable number of observations after the intervention (see Box and Tiao 1975). Although many interventions are unique in nature, the notion of the time taken to recover from the effect is usually quite well understood by those in the industry. Thus, subject matter specialists can sometimes provide reasonable estimates of the half-

FIGURE 1 Example of Additive Outlier (AO)

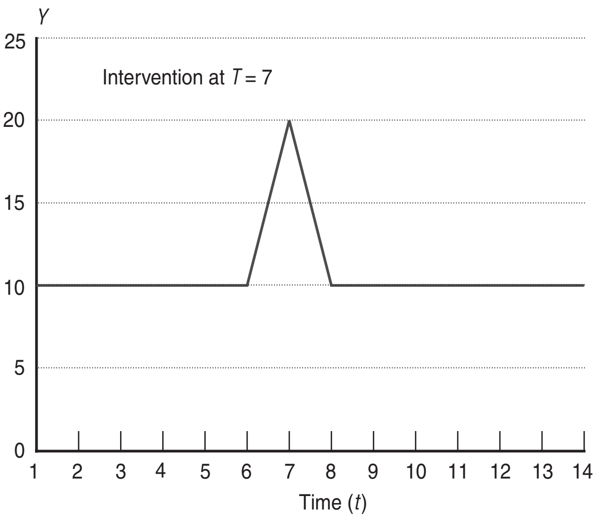
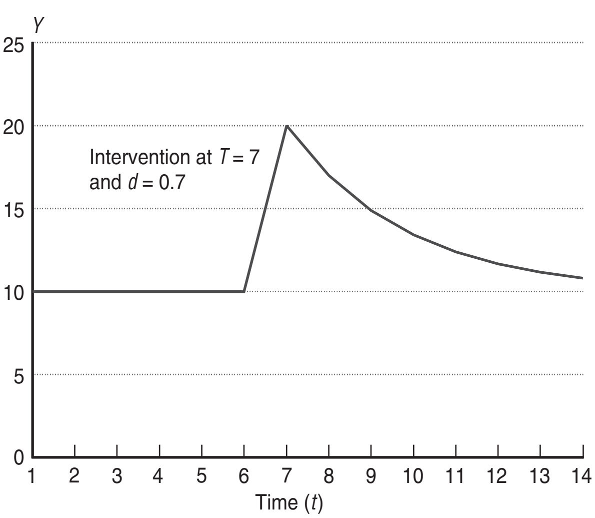


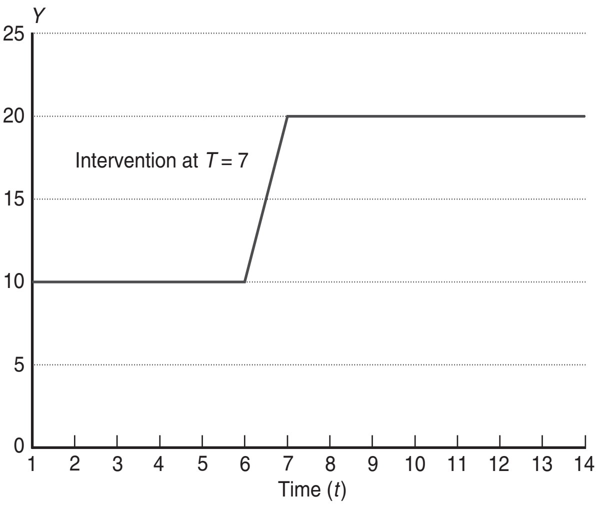
FIGURE 2 Example of Temporary Change (TC)



life for a new intervention, even though the magnitude of the effect cannot be reliably assessed in advance.

Chen and Liu (1993) recommend $d = 0.7$ as a convenient choice, which results in a half-life of the TC of about three months. That is, in months 1, 2,

FIGURE 3 Example of Level Shift (LS)



and 3, the weights assigned to the TC are 1.0, 0.7, and 0.49, reducing to about half the starting weight. Likewise, a value of d = 0.8 corresponds to a half-life of about four months and d = 0.9 to a half-life of just over seven months. A value of d greater than 0.9 becomes almost indistinguishable from a level shift in the short term. Conversely, a value appreciably less than 0.7 may be represented by a one- or (at most) two-period AO.

Based both on the argument of Chen and Liu (1993) and observations about likely recovery times from transportation analysts, we used d = 0.7 in our empirical study of monthly series. As a general matter, an investigator should pay attention to both the phenomenon under study and the frequency with which the data are recorded.

THE STRUCTURAL TIME-SERIES MODEL

Although terms such as "trend" and "seasonal" are intuitively appealing, they are mental constructs because we cannot observe them directly. Therefore, we use a structural modeling approach that treats them as *unobserved components* (Harvey 1989; Harvey and Shephard 1993). In the empirical work, we used the STAMP (Structural Time Series Analyser, Modeller, and Predictor) software in conjunction with GiveWin (for details, see Koopman et al. 2000).

The trend is the long-run component in the series; it designates the general direction in which the series is moving. The trend consists of two parts: the level (which is the current value of the trend) and the slope (which represents the change in the level from one period to the next). Both the level and the slope may be either fixed or evolve over time. A slope may or may not be present depending on the nature of the phenomenon being studied. The seasonal component represents variations over the year, such as increased traffic during the summer. Again, a seasonal component may or may not be present and, if present, may be fixed or evolve over time. The irregular component represents the unexplained variation in the series. We define the components at time t as follows: level = μ_t; slope = β_t; seasonal component = γ_t; and irregular component = ε_t. We assume that the process is observed at unit time intervals (t, $t+1$,...) and that there are s such intervals in a year (e.g., s = 12 for monthly data). We then allow each component to evolve over time according to the specifications:

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \tag{2}$$

$$\beta_t = \beta_{t-1} + \zeta_t \tag{3}$$

and

$$\gamma_t = -\gamma_{t-1} - \gamma_{t-2} - \dots - \gamma_{t-s+1} + \omega_t \tag{4}$$

Equation (4) provides the dummy variable form of the seasonal component (the reader is directed to Koopman et al. (2000) for the trigonometric formulation of the seasonal component).

The quantities η_t, ζ_t, and ω_t represent zero mean, random shifts in the corresponding component. We assume such shifts to be independent of one another and uncorrelated over time; we also assume that they are independent of the "irregular" component, ε_t, seen in equation (5) below. Equations (2) through (4) are known as the *state* or *transition* equations, because they describe the underlying states of the process or the transition of the components from one time period to the next.

Equations (2) and (3) together provide a general framework for describing the evolution of the trend. If the process being modeled does not require all of these components, they can be dropped from the specification. The components are tested in sequential fashion as follows (Harvey 1989, pp. 248–256):

- Does the slope disturbance term have positive variance? (Zero variance corresponds to the slope being fixed over time.)

- If the slope disturbance term has a zero variance, does the slope parameter estimate significantly differ from zero? (An insignificant slope coefficient having a slope disturbance term with a zero variance indicates that the slope term should be dropped from the model.)
- Does the level disturbance have positive variance? (Zero variance corresponds to the mean level being fixed over time.)

If all three statistical tests produced negative outcomes, the overall trend term would be reduced to a constant.

When the time series is seasonal, we check the following:

- Does the seasonal disturbance term have positive variance? (Zero variance corresponds to a stable seasonal pattern.)
- If the seasonal disturbance term has a zero variance, are the seasonal components significantly different from zero? (Is there any seasonal pattern? Should seasonality be dropped from the model?)

If the seasonal disturbance term has zero variance but the seasonal components are significantly different from zero, we are left with a "classical" model with fixed seasonal components. If the seasonal pattern is rejected completely, we reduce the model purely to its trend components.

The observed series is related to the state of the system by the *observation (or measurement) equation:*

$$y_t = \mu_t + \gamma_t + \varepsilon_t \tag{5}$$

where ε_t denotes the irregular component. The irregular component has zero mean and is assumed to be serially uncorrelated (i.e., not predictable) and independent of the disturbances in the state equations.

Estimation proceeds by maximum likelihood (Harvey 1989, pp. 125–128). Operational details are provided in Koopman et al. (2000, section 8.3). The key parameters are the four variances corresponding to the disturbance terms $[\sigma_\varepsilon^2, \sigma_\eta^2, \sigma_\zeta^2, \text{ and } \sigma_\omega^2]$. Note that we assume these variances are constant over time; the time series may need to be transformed to justify this assumption, at least to a reasonable degree of approximation. The four variance terms control the form of the model, allowing each component of level, slope, and seasonality to be stochastic or fixed; slope and seasonal elements may be present or absent. Table 2 illustrates the principal

TABLE 2 Some of the Principal Models in the Structural Framework

Type of model	σ_ε^2	σ_η^2	σ_ζ^2	σ_ω^2
Level only				
Constant mean	yes	0	0	0
Local level	yes	yes	0	0
Random walk	0	yes	0	0
Trend only				
Deterministic	yes	0	0	0
Local level with fixed slope	yes	yes	0	0
Random walk with fixed drift	0	yes	0	0
Local linear trend (Holt)	yes	yes	yes	0
Smooth trend	yes	0	yes	0
Second difference	0	0	yes	0
Seasonal (with selected trend)				
Fixed seasonals	(yes or 0)	(yes or 0)	(yes or 0)	0
Varying seasonals	(yes or 0)	(yes or 0)	(yes or 0)	yes
Basic Structural Model	yes	yes	yes	yes

Source: Based on Koopman et al. (2000), p. 141.

variations. If fixed components are included in a model, the corresponding terms appear in the state equations (e.g., fixed seasonal coefficients), but the variance term is zero. If the components are stochastic, the same terms appear in the model, but the variance is strictly positive. The most general form is the Basic Structural Model (BSM), in which all components are stochastic. The BSM forms the starting point for the model development process and is the standard form employed in STAMP. The program then "tests down" to eliminate any components that are not required.

ANALYSIS OF LATE ARRIVAL OF SCHEDULED FLIGHTS

By way of illustration, we consider an example that has received considerable publicity in recent years, namely late arrival of flights, or airline delays. The particular series examined in this section describes the monthly percentage of scheduled flights for major U.S. air carriers not arriving on time, or the Late Arrivals time series. A plot of this series, for the period September 1987–February 2002, is shown in figure 4. The tragic events of September 2001 changed many lives in fundamental ways and also had a serious effect on the level of activity in the airline industry. Therefore, we will analyze the series initially only up to August 2001 and consider the aftermath of the terrorist attacks in the next section.

An initial set of interventions (prior to September 2001) was identified using information provided by the Bureau of Transportation Statistics (published in a report known as *Transportation Indicators*), combined with an initial analysis using the AUTOBOX software (produced by Automatic Forecasting Systems[1]).

Two significant pulses, or AOs, were found for this series within the time period of September 1987 through August 2001: January 1996 and December 2000. These two interventions were weather-related and were incorporated into the STAMP modeling process. Our analysis of the series, using STAMP, revealed that the most appropriate model included a stochastic level, no slope, and fixed seasonal components. This model yields the outputs shown in figures 5 through 8. Figure 5 shows the smoothed

[1] Information on AUTOBOX software can be found at http://www.autobox.com.

FIGURE 4 Percentage of Scheduled Flights for Major U.S. Air Carriers Not Arriving on Time: September 1987–February 2002

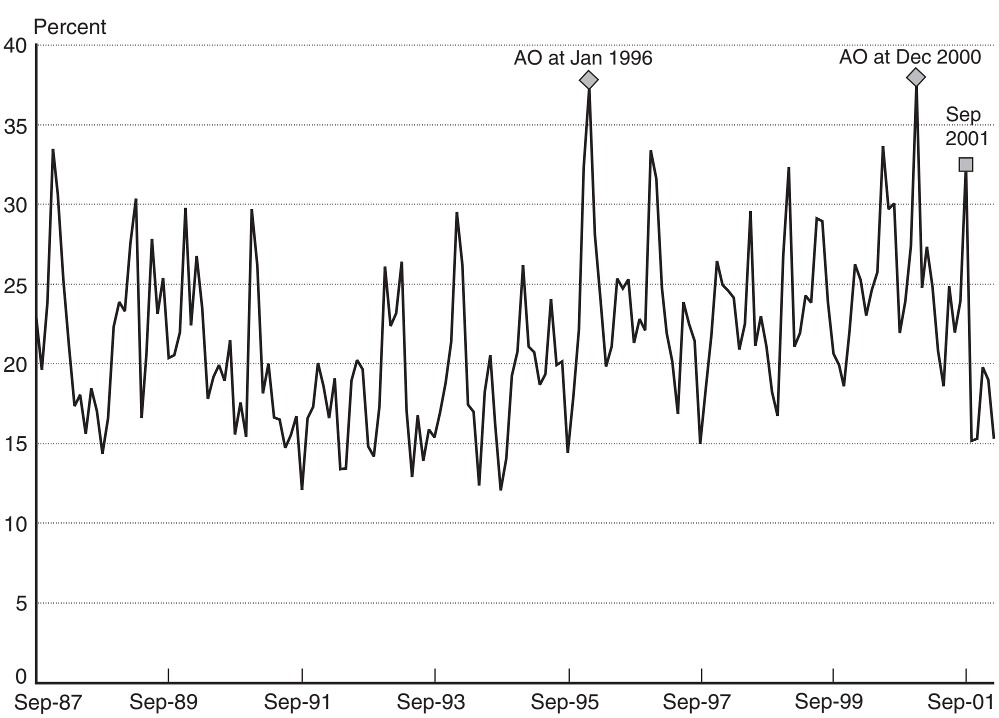
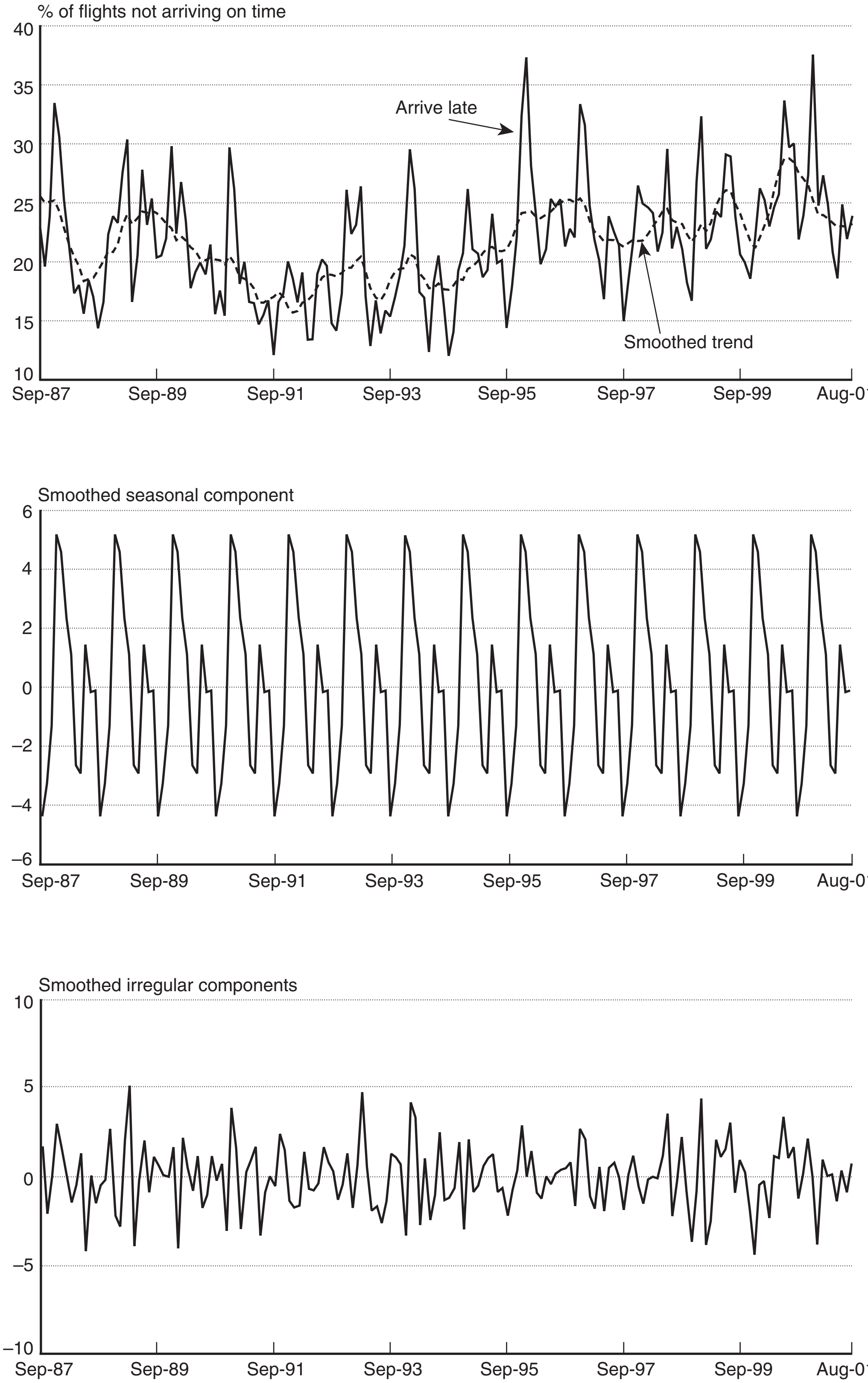


FIGURE 5 Smoothed Components of the Late Arrivals Series
Generated by STAMP



Note: The first panel shows the observed series and the fitted trend, the second shows the seasonal component, and the third shows the irregular component.

trend, the seasonal components and the irregular component. The *smoothed* versions use the entire series to construct the trend, seasonal, and irregular components; it is a better choice for gaining a perspective on the evolution of the series, because the estimates use observations both before and after the time period in question.

When these plots are compared with the *filtered* (or *forecast*) components in figure 6, the increased roughness of the trends in the latter set become evident. The seasonal pattern in the filtered series also changes over time, because the filtered components use only the observations up to the current time in each set of calculations. Thus, the filtered components are directly useful for prediction purposes, and we use only these components in subsequent analyses.

As noted earlier, our initial analysis was based on the data from September 1987 through August 2001. Now that the model has been specified, the holdout sample of data from September 2001 through February 2002 is placed back into the dataset, and same model is fitted onto the full set of data. Figure 7 shows the standardized residuals for the full fitted series and highlights the impact of post-September 11, 2001. The horizontal lines correspond to ± 2 standard deviations, and the plot may be thought of as a Shewhart chart.[2] As expected, the chart indicates a sharp rise in the percentage of late arrivals in September, followed by a major decline in late arrivals in October 2001, due primarily to reduced traffic levels.

The automated analysis in STAMP suggests an outlier in September 2001 and a level shift in October 2001. Using these interventions, the final fitted trend for the full set of data is shown in figure 8. Although the overall performance appears to be satisfactory, the further declines in the trend after October 2001 seem inconsistent with the model and suggest the need for further analysis. This problem is reflected more clearly in later analyses (seen in figures 13 and 14). Since data for such a modeling exercise are necessarily very limited, we need to use our judgment on likely future developments, as illustrated in the next section when we use the framework developed earlier.

[2] Shewhart charts are widely used in statistical process control to identify out-of-control conditions from which we would seek to identify assignable causes. The center line in figure 7 (equal to zero here because we are plotting regression residuals) and the vertical axis represent the number of standard deviations (SD) that an observation lies above or below the mean. In conventional use, a single observation that is more than three SD from the mean, or two successive observations more than two SD from the mean, is said to signal an out-of-control condition, in contrast to a state of statistical control, i.e., a stable system of random variation.

GENERAL APPROACH TO INTERVENTION MODELING

We applied the framework developed earlier within the following context. After September 11, 2001, the airlines experienced massive disruptions in their schedules. In October and later months, the overall operating system gradually returned to normal, but passenger traffic resumed at lower levels than prior to the attack. This sequence of events may be represented by the following set of interventions:

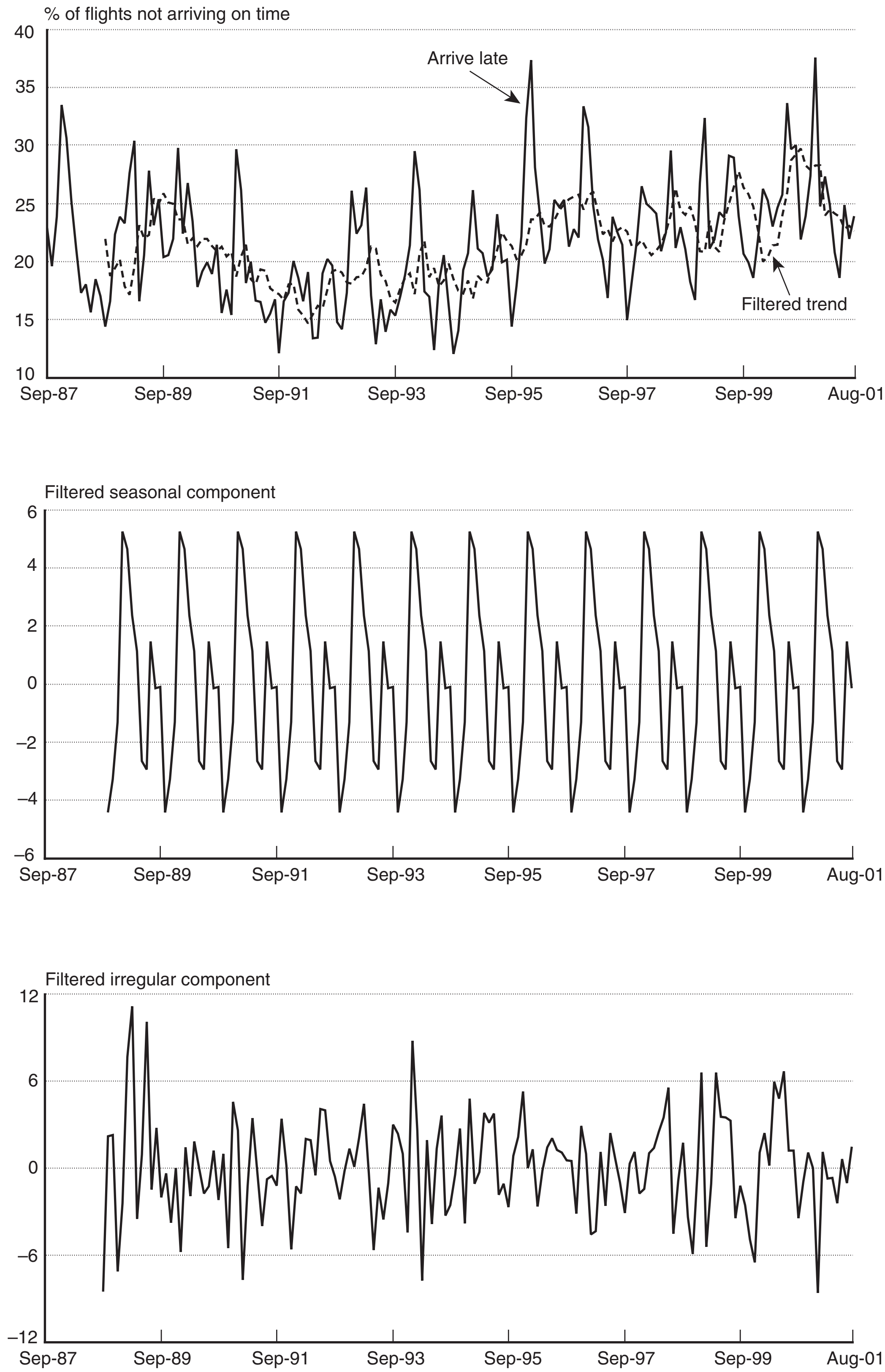
- A purely transient effect (AO) relating to the month of September only.
- A temporary change or shift (TC) that started in October 2001 and gradually disappeared. We could have started this effect in September, but felt that October provided a simpler interpretation. As noted earlier, we used $d = 0.7$ in all cases.
- A permanent effect (LS) that changed all mean values of the series from November 2001 on. Again, note that we could have started this factor in September or October, but we felt that the present construction affords a simpler interpretation by separating out the start dates of the three interventions. Provided all three interventions are retained in the model, the particular choice of starting dates will not affect the fitted or forecast values in the series.

Data

We considered the following five series, primarily selected from the air transportation sector, since this was the mode of transportation most affected:

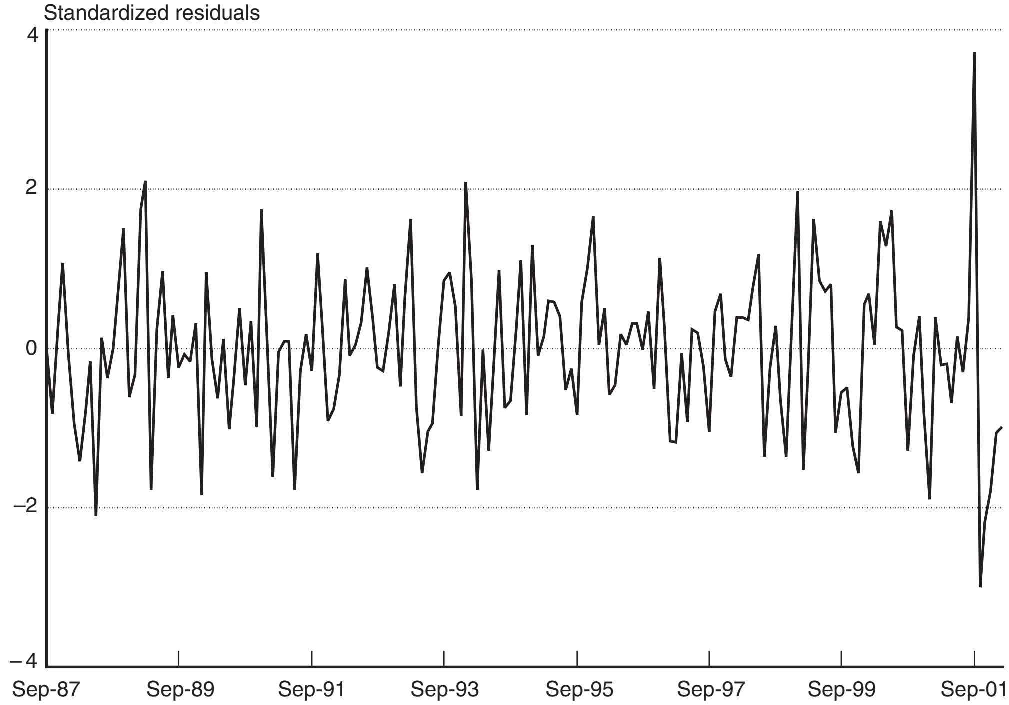
- **Late arrivals**—percentage of scheduled flights by major U.S carriers not arriving on time.
- **Cancellations**—percentage of scheduled flights by major U.S. carriers that were canceled.

FIGURE 6 Filtered Components of the Late Arrivals Series: September 1987–August 2001
Generated by STAMP



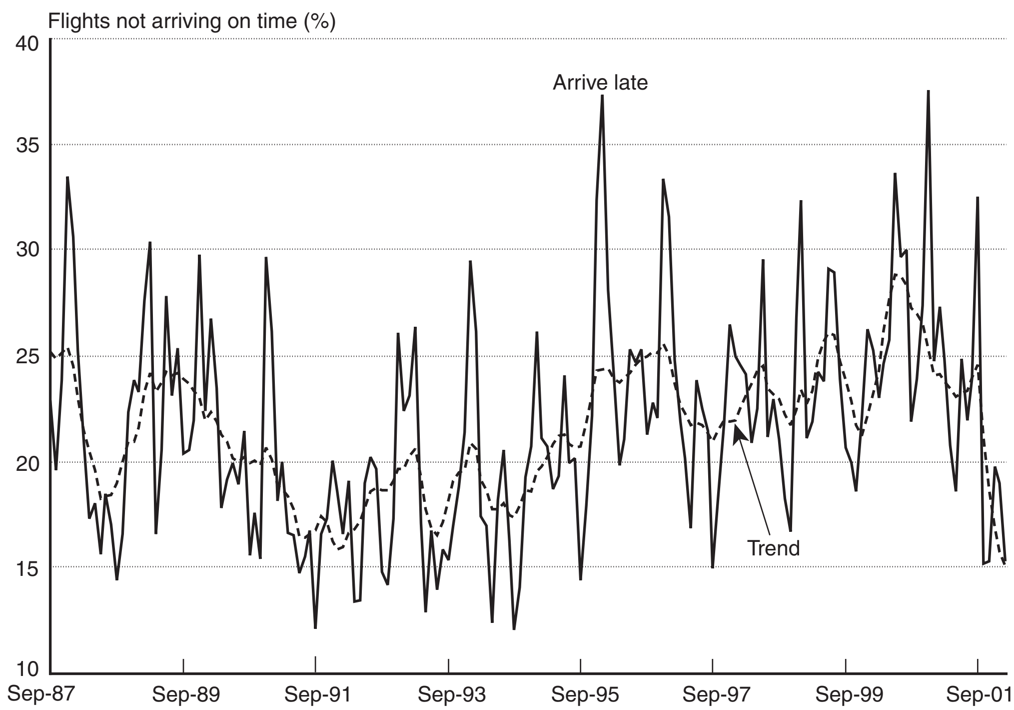
Note: The first panel shows the observed series and the fitted trend, the second shows the seasonal component, and the third shows the irregular component.

FIGURE 7 Shewhart Chart of Standardized Residuals of Late Arrivals Series: September 1987–February 2002



Note: The upper and lower grid lines are two standard deviations away from the mean.

FIGURE 8 Trend for Late Arrivals Series: September 1987–February 2002



- **Domestic enplanements**—number of passengers boarding domestic aircraft (millions).
- **Air revenue passenger-miles**—revenue-earning miles flown by passengers on major U.S. carriers (billions).
- **Rail revenue passenger-miles**—revenue-earning passenger-miles carried by Amtrak and the Alaska Railroad (millions).

The late arrivals series was illustrated in figure 4; figures 9 through 12 provide graphs of the other series being examined. In all cases, the data are available on the U.S. Department of Transportation website (http://www.dot.gov). The first four series are collected by the Office of Airline Information in the Bureau of Transportation Statistics (also available at http://www.bts.gov), and the fifth is produced by the Federal Railroad Administration.

In our original analysis of these data (Young and Ord 2002), we were able to use only a small number of observations post-intervention (table 3). In this paper, we report those initial analyses recomputed using the data as later revised by the agencies. It should be noted that the different time periods used in the analyses reflect data availability at the times the analyses were completed (May 2002 and January 2003). These inevitable delays serve to underscore the importance of timely and reliable adjustments to series after interventions.

FIGURE 9 Percentage of Scheduled Flights Canceled by Major U.S. Carriers: September 1987–February 2002

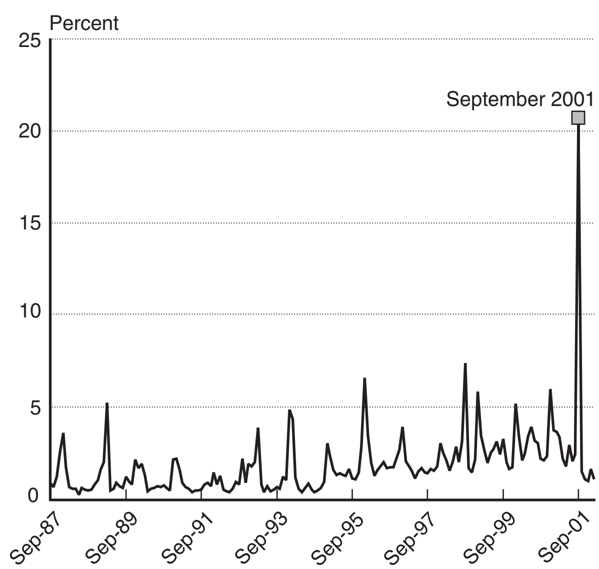


FIGURE 10 Number of Enplanements on Domestic Aircraft: September 1987–February 2002

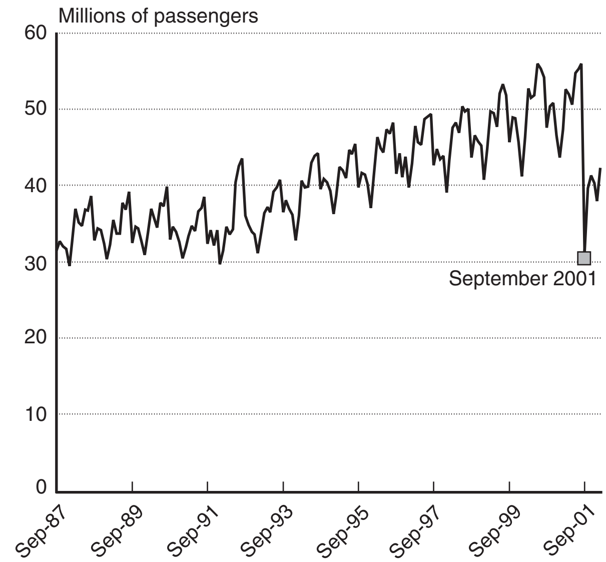
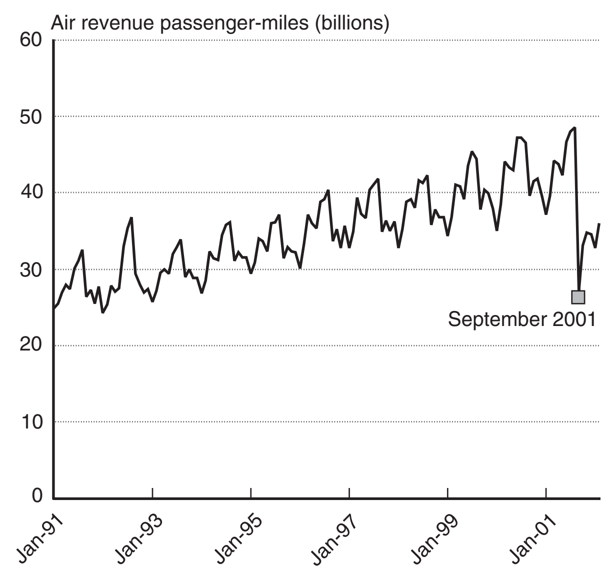


FIGURE 11 Revenue-Earning Miles Flown by Passengers on Major U.S. Carriers: January 1991–February 2002



Model Development

Our procedure was as follows:

1. Develop a model for the series up to August 2001, incorporating AO and LS outliers where needed (as for late arrivals in our earlier analysis).

FIGURE 12 Revenue-Earning Passenger-Miles Carried by AMTRAK and the Alaska Railroad: January 1979–November 2002

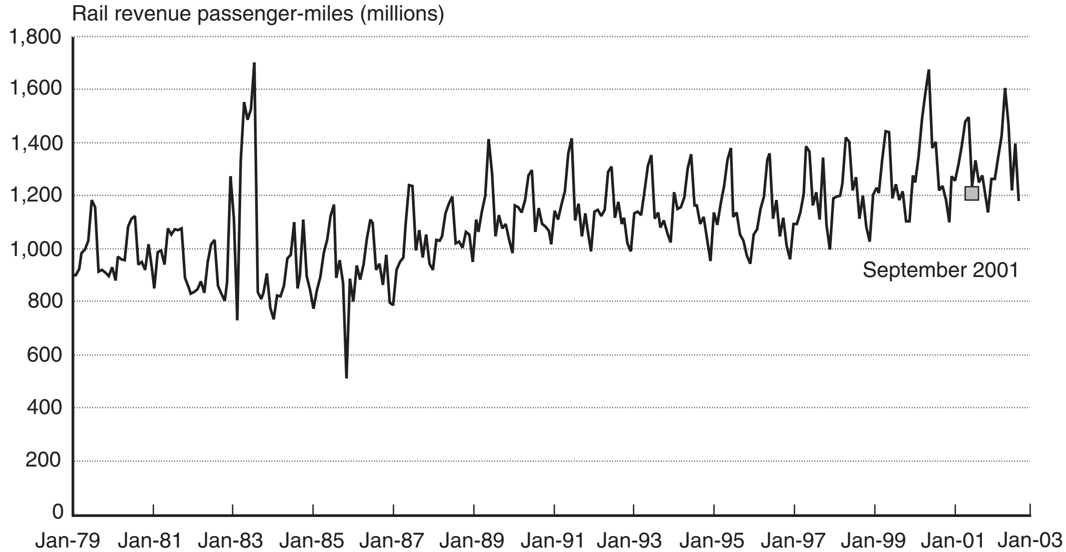


TABLE 3 Summary of Data Structure and Changes Since Earlier Analyses

Time series	Revisions	Latest month available as of May 2002	Latest month available as of January 2003
Late arrivals	None	February 2002	October 2002
Cancellations	None	February 2002	October 2002
Enplanements	Very minor	December 2001	June 2002
Air revenue passenger-miles	01/99–12/99: minor changes	December 2001	June 2002
Rail revenue passenger-miles	None	February 2002	November 2002

2. Using the data available as of May 2002 (table 3), run the same model with AO, LS, and TC components as specified above and "test down" to eliminate insignificant coefficients. This analysis was performed initially in early June 2002 (Young and Ord 2002) and minor differences in the results are reported only to the extent that changes occurred in the reported series after that time.
3. Using the data available as of January 2003 (table 3), run the same model with AO, LS, and TC components as specified above and test down to eliminate insignificant coefficients.
4. Use the models developed in steps 2 and 3 to generate successive one-step-ahead forecasts for the most recent data to see if the earlier analysis (step 2) provided an adequate description of the structural changes in the series.

The original models are summarized in table 4. The data revisions noted above did not lead to any changes in specification; the changes in the estimated coefficients were minor in all cases. The entries in table 4 are to be interpreted as in the following example for late arrivals:

State equations (stochastic level, no slope, fixed seasonals):

$$\mu_t = \mu_{t-1} + \eta_t \tag{2a}$$

and

$$\gamma_t + \gamma_{t-1} + \ldots + \gamma_{t-s+1} = 0, \text{ or } \gamma_s \textit{ fixed} \tag{4a}$$

Measurement equation:

$$y_t = \mu_t + \gamma_s + X_1 I_1 + X_2 I_2 + \varepsilon_t \tag{5a}$$

where s denotes the number of seasons, γ_j denotes the parameter for the fixed seasonal effect in period

TABLE 4 Original Models for Each Series

Time series	Time period	Level	Slope	Seasonal	Outliers
Late arrivals	9/1987–8/2001	Stochastic	None	Fixed	Outliers: 1/1996; 12/2000
Cancellations	9/1987–8/2001	Stochastic	Stochastic	Stochastic	Outliers: 3/1989; 3/1993; 1/1994; 2/1994; 1/1996; 9/1998; 1/1999; 12/2000
Enplanements	1/1991–8/2001	Stochastic	Stochastic	Stochastic	Outlier: 11/1996 Levels: 6/1992; 10/1992
Air revenue passenger-miles	1/1991–8/2001	Stochastic	Stochastic	Stochastic	Outlier: 11/1996 Levels: 6/1992; 10/1992
Rail revenue passenger-miles	1/1987–8/2001	Stochastic	None	Fixed	Outliers: 7/1989; 12/1997 Level: 11/2000

j, and (X_1, X_2) denote the AO interventions at January 1996 and December 2000. Numerical details are omitted in the interests of space.

Empirical Results

Because we had only between four and six observations in the initial study (from September to December or February depending on the series), it was only to be expected that the LS and TC estimates would be highly correlated, and the TC dropped out in three of the five series. What is remarkable is that when the analyses were re-run with the later observations included, the changes were minor in all cases. The results are given in table 5.

Several conclusions may be drawn from table 5:

- In all cases, large adverse effects were identified in September, as expected. Rail traffic was reduced as well as air traffic, because people were reluctant to travel at all.
- In all cases, the estimates based on the first analysis seem to provide adequate adjustments to the series.
- Cancellations and late arrivals showed negative level shifts reflecting the reduced amount of traffic in subsequent months. As airports gradually resumed normal operations, we might have expected these series to have resumed their earlier levels, but the initial estimates seem to have provided a reasonable assessment of the reactions. These effects are probably the result of less congestion as the result of lower traffic volumes.
- The temporary effects for enplanements and air revenue passenger-miles are about five times the size of the permanent level shift. All these effects were negative, indicating the adverse effect on the airline industry. Both the larger temporary effect and the smaller final impact seem to have been adequately recognized in the first analysis.
- The rail revenue passenger-miles series shows no change after the first month, which is consistent with the data and reflects the relative independence of the two markets.
- Although the details are not reproduced here, the diagnostics for each series indicated that the descriptions were consistent with the data available. Since the same set of three interventions was applied to each series, this provides some evidence that the descriptions are reasonable, although further data are clearly needed to validate that claim.

The results shown in table 5 indicate that the parameter estimates of the three-intervention terms based on limited data proved to be comparable with those estimates with more data points after the intervention. But does this model fit also imply comparable results when forecasting?

In order to study this issue, we chose to compare the quality of forecasts with the three-interventions (AO, TC, and LS) and without the three-interventions in the model. Figures 13 and 14 illustrate graphically the forecasts of air revenue passenger-miles based on the models without (figure 13) and

TABLE 5 Results of Intervention Analyses for the Five Series

Time series	# obs. after 8/2001	AO (Sept.)	TC (Oct.)	LS (Nov.)	*Rs*-squared
Late arrivals	6	14.28	–3.11	–5.52	0.404
		(0.0000)	(0.3334)	(0.0431)	
	14	13.91	–4.64	–4.94	0.402
		(0.0000)	(0.1264)	(0.0648)	
Cancellations	6	18.12	–0.06	–1.56	0.947
		(0.0000)	(0.8978)	(0.0001)	
	14	18.00	0.72	–1.80	0.947
		(0.0000)	(0.2450)	(0.0009)	
Enplanements	4	–17.42	–11.99	–2.25	0.813
		(0.0000)	(0.0000)	(0.0082)	
	10	–17.31	–11.18	–2.75	0.798
		(0.0000)	(0.0000)	(0.0017)	
Air revenue passenger-miles	4	–13.94	–10.02	–1.60	0.800
		(0.0000)	(0.0000)	(0.0245)	
	10	–13.82	–9.33	–2.08	0.787
		(0.0000)	(0.0000)	(0.0036)	
Rail revenue passenger-miles	4	–9.97	–0.18	3.38	0.564
		(0.0002)	(0.9460)	(0.1335)	
	11	–9.71	1.71	1.46	0.554
		(0.0002)	(0.4480)	(0.3801)	

Note: The first two rows for each series provide the estimated coefficients and their *p*-values based on the first analysis; the next two rows provide the same information for the second analysis.

with (figure 14) the three-intervention coefficients (forecast based on the model fitted on data ending December 2001).

If no interventions are incorporated in the model, the forecast for air revenue passenger-miles would be a continuing downward trend, whereas the incorporation of interventions identifies the initial downturn and the subsequent gradual though partial recovery. This pattern would, of course, eventually be identified without the intervention analysis. However, the modified trend is identified much more quickly and reliably when the three-intervention model is applied.

The "no-interventions" forecasts will provide a basis of comparison for the three-interventions forecasts with differing forecast origins (December 2001, March 2002, and June 2002). To measure the forecast accuracy of each model, we chose to calculate the Mean Absolute Percentage Error (MAPE) values

$$MAPE = \frac{\sum |e_t / y_t|}{n}\text{, where} \qquad (6)$$

e_t is the forecast error for time t, y_t is the observed value at time t, and n is the number of forecasts. The *MAPE* values were calculated for forecasts based on the two types of models (with and without interventions) from the three forecast origins (December 2001, March 2002, and June 2002). The summations cover the period from the forecast origin to the latest observation available (table 6). For each forecast origin, the *MAPE* values for the two types of models are then compared by creating a "Relative Value" (or RV) of those two *MAPES*:

$$\text{RV} = MAPE_{\text{no interventions}} / MAPE_{\text{three-interventions}} \qquad (7)$$

The results of the RV calculations for the different forecast origins are shown in table 6.

We need to refer back to table 4 in order to understand the results in table 6. Late arrivals and rail revenue passenger-miles have fixed seasonal patterns, so that the *model estimates* for the seasonal components of these series are not affected by the intervention. The other three series have stochastic seasonals and the three-interventions model identi-

FIGURE 13 Air Revenue Passenger-Miles Forecast (as of December 2001) with No Interventions for September 2001: January 1991–August 2003

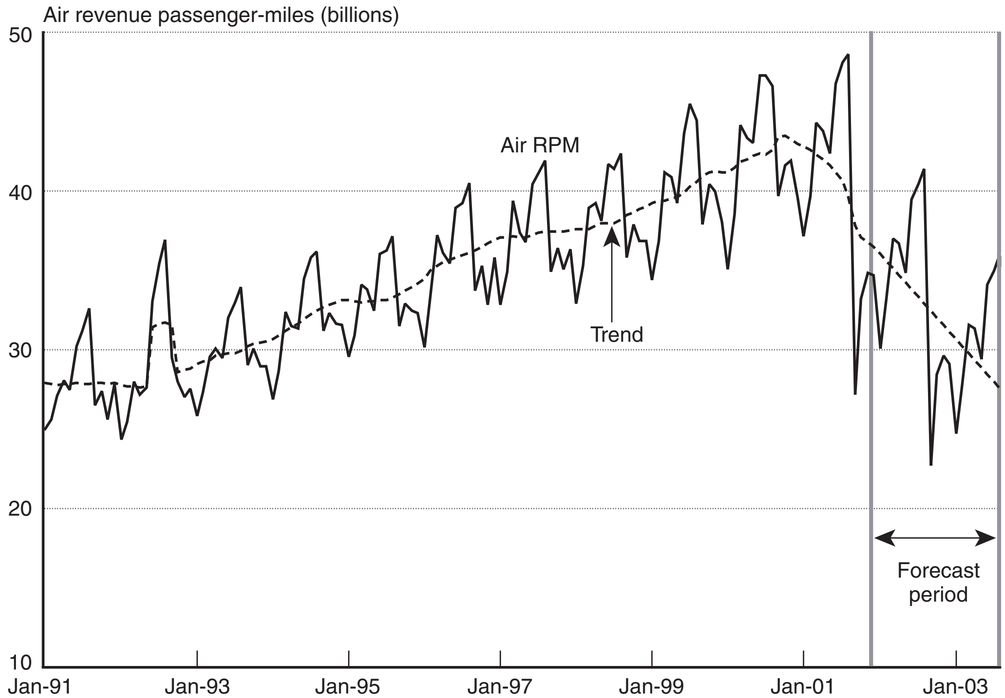
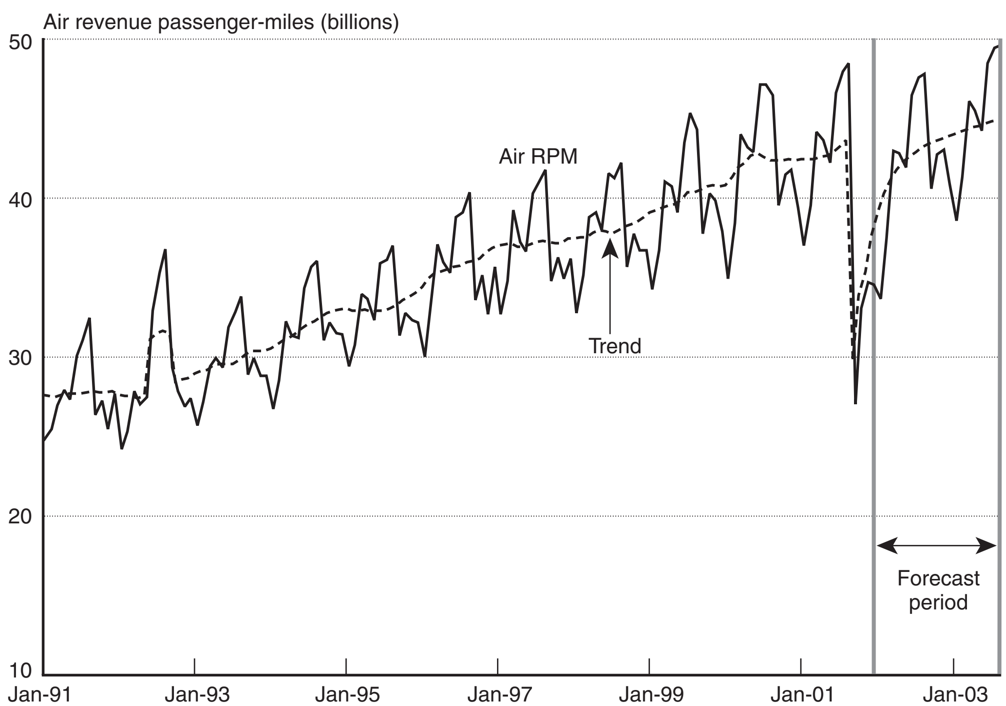


FIGURE 14 Air Revenue Passenger-Miles Forecast (as of December 2001) with Three-Interventions for September 2001: January 1991–August 2003



fies the disruptions in their seasonal patterns more accurately than the no-interventions model, especially for the cancellations series (figure 9).

All five series have stochastic levels so that the intervention is gradually incorporated into the model structure over a period of months. Consequently, for the RVs for the forecast horizon with the least number of time periods after September 2001 (December 2001), the intervention approach provided more accurate forecasts for each of the five

TABLE 6 Relative Values (RV) of *MAPE* Using Different Forecast Origins

Series	Last observation	Forecast origin: December 2001	March 2002	June 2002
Late arrivals	October 2002	1.06	1.14	1.03
Cancellations	October 2002	7.21	6.65	21.80
Enplanements	December 2002	1.51	0.74	1.26
Air revenue passenger-miles	December 2002	1.85	0.76	1.13
Rail revenue passenger-miles	November 2002	1.11	0.97	1.00

Note: RV values greater than one indicate that the interventions provided greater forecast accuracy.

series. However, as more data values are obtained, those series that contain local level and fixed seasonal components seem to have corrected quickly and therefore do not show any longer term improvement through the intervention approach. Such effects have been noted in other forecasting studies (e.g., Makridakis and Hibon 2000) and reflect the adaptive nature of the models used.

The structural modeling approach does allow the model to adapt itself to changes in the data, but the incorporation of the interventions allows the model to react more quickly. For our five sets of data, the series seem to be tracked reasonably by the three-interventions model after three to four months, whereas it takes about six months or longer (especially for cancellations) for the no-interventions model to self-adjust. The importance of the proposed procedure lies in the ability to reduce the time required to discern the underlying trends after the intervention has occurred.

FINAL COMMENTS

A key requirement in forecasting is that adjustments should be made for unusual events so that the series can be forecast on its new trajectory. In addition, we also seek to make changes that are reliable, yet quick to take effect. These requirements are especially important when the series has been subjected to a major intervention and we wish to identify the newly emerging trend.

The results of this study suggest that the flexible use of the three-interventions approach we have described provides adequate adjustments by three to four time periods after the event. This contrasts with a six-month or longer delay in self-adjustment even for a flexible model, and probably a longer time if a model-based procedure is not used at all. In addition, the three-interventions method enables us to provide an initial partitioning of the effects into short-term, transient, and permanent shifts, which is important for planning purposes. Furthermore, the structural technique employed to calculate these models can be easily updated as new data become available, so that the previous month's assessment of the shifts can be compared with the latest results and an assessment made of how quickly the system is returning to a new stable level after the intervention. We are cautiously optimistic that the proposed approach offers a way forward in dealing more expeditiously with interventions at the end of a time series.

REFERENCES

Alwan, L.C. 2000. *Statistical Process Analysis*. Boston, MA: Irwin/McGraw-Hill.

Alwan, L.C. and H.V. Roberts. 1988. Time-Series Modeling for Statistical Process Control. *Journal of Business and Economic Statistics* 6(1):83–95.

______. 1995. The Problem of Misplaced Control Limits. *Applied Statistics* 44(3):269–278.

Box, G.E.P. and G.C. Tiao. 1975. Intervention Analysis with Applications to Economic and Environmental Problems. *Journal of the American Statistical Association* 70:70–79.

Chen, C. and L.-M. Liu. 1993. Joint Estimation of Model Parameters and Outlier Effects in Time Series. *Journal of the American Statistical Association* 88:284–297.

DeLurgio, S.A. 1998. *Forecasting Principles and Applications*. Boston, MA: Irwin/McGraw-Hill.

Harvey, A.C. 1989. *Forecasting, Structural Time Series Models, and the Kalman Filter*. Cambridge, UK: Cambridge University Press.

Harvey, A.C. and N. Shephard. 1993. Structural Time Series Models. *Handbook of Statistics, Volume 11*. Edited by G.S. Maddala, C.R. Rao, and H.D. Vinod. Amsterdam, Netherlands: Elsevier Science.

Koopman, S.J., A.C. Harvey, J.A. Doornik, and N. Shephard. 2000. *Stamp: Structural Time Series Analyser, Modeller, and Predictor.* London, UK: Timberlake Consultants.

Makridakis, S. and M. Hibon. 2000. The M3 Competition. *International Journal of Forecasting* 16(4):451–476.

Young, P. and K. Ord. 2002. Monitoring Transportation Indicators, and an Analysis of the Effects of September 11, 2001, paper presented at the 22nd International Symposium on Forecasting, Dublin, Ireland.

Airline Networks: An Econometric Framework to Analyze Domestic U.S. Air Travel

DIPASIS BHADRA*
PAMELA TEXTER

Center for Advanced Aviation System Development (CAASD)
The MITRE Corporation
7515 Colshire Avenue
McLean, VA 22102
USA

ABSTRACT

In this paper, we examine the U.S. domestic airline network. Using an exhaustive definition of the airline network and a cross-section pooled time series dataset for 35 consecutive quarters covering 1995:Q1 to 2003:Q3, we analyzed domestic scheduled air transportation. Results suggest the existence of increased vertical disintegration of market segments following the events of September 11, 2001 (9/11). The effects of 9/11 have affected all network classes, with the largest impact on the point-to-point variants. The expansion of Southwest Airlines affected all variants of the network positively, with a proportionately larger impact on the point-to-point over the hub-and-spoke variants. The results of this study are expected to help inform both operational decisionmaking and policymaking. Results may also be useful to manufacturers in projecting the size and mix of the aircraft fleet that are expected to be compatible with the evolving network.

INTRODUCTION

Events beginning with the recession in spring 2001 and the terrorist attacks on September 11, 2001 (9/11), have destabilized the U.S. aviation industry. The accumulated net income from the second half of the 1990s ($22.8 billion from 1995

E-mail addresses:
* Corresponding author—dbhadra@mitre.org
P. Texter—ptexter@mitre.org

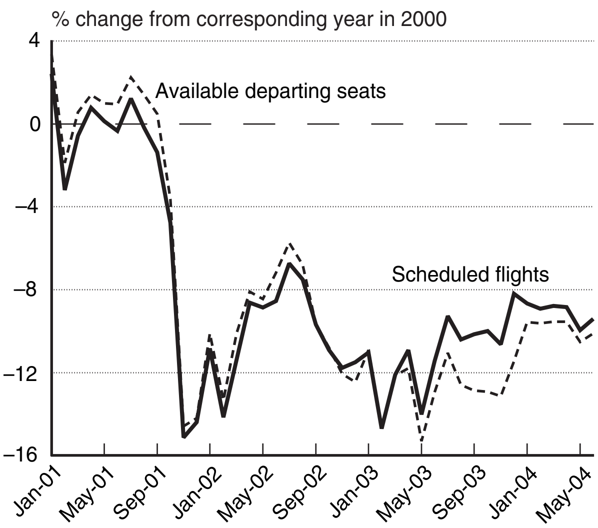
KEYWORDS: Airline networks, transportation forecasting, econometric modeling, commercial airlines, regional jets. JEL Code L93: Air Transportation.

to 2000) was wiped out completely by the $24.8 billion in losses incurred during the subsequent 11 quarters (2001:Q2 to 2003:Q4) (ATA 2004; *Airline Monitor 2004*), despite a U.S. government-provided cash grant of $5 billion and a loan program totaling $10 billion soon after the events of 9/11. Due to a significant hike in the price of jet fuel (more than 40¢ per gallon) that began at the end of 2002 and has continued well into 2004, the industry is expected to lose $2 billion to $3 billion in 2004. Without a jet fuel price increase, the industry would most likely have returned a small profit in 2004 due to improved traffic conditions and a slightly better fare environment.

The events of 2001[1] led to a massive restructuring of the airline industry that addressed weak basic business practices. The most significant changes were in capacity reductions in the number of available seat-miles and the number of flights (figure 1). These necessary adjustments reflect a drop in demand, a decline in business travel, and the availability of internet booking. In addition, a realigned fare structure narrowed the gap between premium and walkup fares and leisure fares. Finally, renegotiations of labor and other contracts, and simplification of the network structure, have also played key roles in the restructuring of the industry.

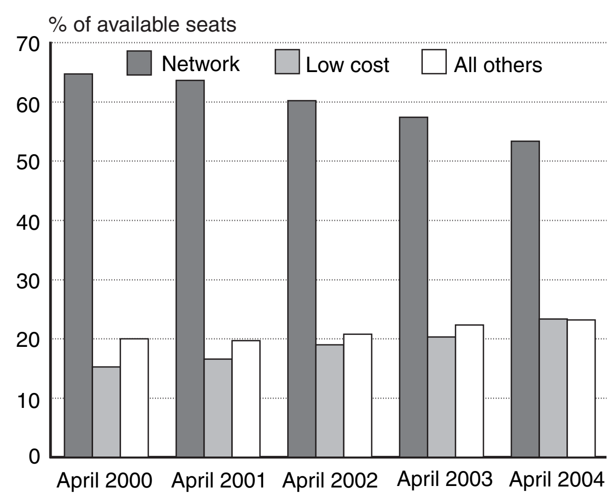
The overall downward capacity adjustments affected industry participants differently, with network carriers affected more than low-cost carriers (LCCs).[2] LCCs and regional carriers (carriers specializing in regional jets—RJ carriers)[3] appear to have increased their capacity as the network carriers' capacity shrank. While there was an overall fall in demand immediately after 9/11, LCCs and RJ carriers fared better than the network carriers (figure 2).

FIGURE 1 U.S. Airline Industry Capacity Adjustments Since 2000



Source: M.R. Dayton, "Trends and Demand in Aviation Markets," presentation at the ATCA/FAA/Nav Canada Technical Symposium, Office of Inspector General, U.S. Department of Transportation, 2004.

FIGURE 2 Airline Market Share by Type of Carrier



Note: *All others* is primarily regional jet carriers but may include a small percentage of scheduled charter carriers.

Source: M.R. Dayton, "Trends and Demand in Aviation Markets," presentation at the ATCA/FAA/Nav Canada Technical Symposium, Office of Inspector General, U.S. Department of Transportation, 2004.

[1] These events include the economic recession starting in spring 2001, the increasing use of internet technology in airline booking, and the tragic events of 9/11. While the first two are associated with cyclical and secular parts of the time series, the events of 9/11 are not.

[2] Here we define network carriers as American, Continental, Delta, Northwest, United, and US Airways. They generally run their operations through a system of hub-and-spoke airports. Some LCCs, Air Tran in particular and JetBlue to some extent, are also following the hub-and-spoke network model. In 2003, network carriers accounted for about 73% of total revenue passenger-miles (RPM) and provided 72% of available seat-miles (ASM) (*Airline Monitor 2004*), two standard measures of airline output.

We define the LCCs (in order of importance with respect to shares in ASM and RPM) as Southwest, America West, ATA, JetBlue, Air Tran, and Frontier. It is important to recognize that LCC markets are continuously evolving, both in terms of their market shares and the number of participants.

[3] These carriers use small jets and generally supply service for other airlines. Some examples of these carriers are Air Wisconsin partnering with Air Tran, American Eagle partnering with American and Delta, and Cape Air partnering with Continental.

Southwest Airlines was the nation's largest LCC in 2003, with almost half the total LCC market (48% of available seat-miles (ASM) and 45% of revenue passenger-miles (RPM), representing supply and demand, respectively) (*Airline Monitor 2004*). This airline now ranks second, after Delta Airlines, in terms of U.S. domestic passenger enplanements, accounting for about 10% of overall domestic ASM and RPM.[4] Throughout the last decade, Southwest Airlines expanded its level of activity, although activity accelerated somewhat after 9/11. By aggressively gaining market share, they became a major force in U.S. air travel. Given what appears to be an increasingly distributed network structure (Berry 2004) with less emphasis on hubs and connections,[5] Southwest appears to have adopted a point-to-point or distributed network variant.

Southwest Airlines has shown that offering lower fares induces increased overall air travel demand (Bennet and Craun 1993; Morrison 2001), which we will call the *Southwest effect* in this paper. In addition to its primary focus on serving larger metropolitan areas through secondary airports, Southwest also flies from airports designated as large hubs based on their traffic levels (e.g., Baltimore-Washington (BWI), Phoenix (PHX), Las Vegas (LAS), and Midway (MDW)).[6] Therefore, an expanding Southwest may also have a positive impact on hub-to-hub travel and on spokes that connect to these hub airports through inducing demand in those networks. This process may be further enhanced if other LCCs entering the market follow Southwest's network structure.

Continuous restructuring of the industry, characterized by capacity realignment, changing market shares, and an evolving network, has affected the fleet mix as well. Regional or feeder carriers increasingly take up the markets from which network carriers have retreated. Essentially, network carriers increasingly outsource some of their markets to regional carriers. This vertical disintegration of markets once held by network carriers (i.e., market fragmentation), leads to a greater number of scheduled flights flown by regional jets. Consequently, the number of segments and aircraft operations may go up significantly even though the market, as a whole, may be smaller than before. Figure 3 presents this process of substitution from markets served by network carriers (using larger jets) to markets served by regional carriers (flying regional jets), or fragmentation of markets.

These changes have had a profound impact on the industry as a whole. Ed Greenslet, a long-time aviation industry analyst, summed this up recently (*Aviation Week & Space Technology* 2002, p. 52): "…the domestic airline landscape is changing before our eyes, and the consequences for the tradi-

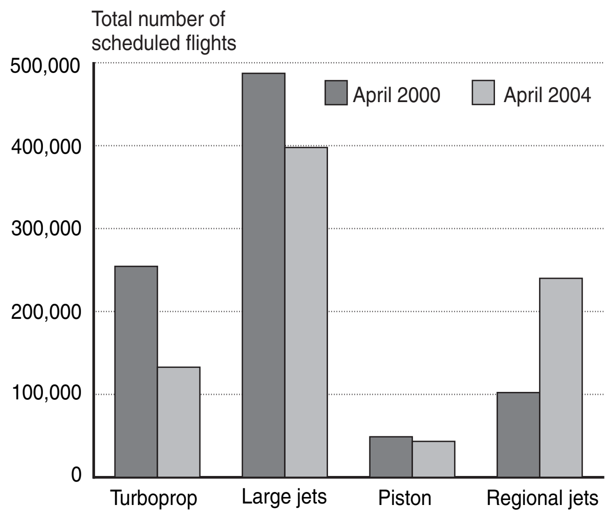
[4] Southwest ranks third with respect to total passenger enplanements, following American Airlines and Delta Airlines. However, it ranks second when evaluated in terms of domestic enplanements. In 2003, Delta Airlines handled 78 million enplanements compared with Southwest's 74 million. American Airlines carried the most (both domestic and international), about 83 million enplanements with a higher share of international enplanements than Delta. Southwest does not fly any international routes currently.

[5] The term *distributed network* is used here to represent situations where airlines distribute their operations among more modestly sized airports for traffic traveling between two ends of the network, e.g., among Midway, Nashville, and St. Louis airports for east-west traffic in Southwest's network. This is in sharp contrast to using one or two airports heavily as their main hubs (e.g., Chicago and Denver for American and United) to serve a similar purpose. The term is also used to represent situations where schedules are distributed more uniformly throughout the day as opposed to schedules that have sharp peaks and offpeaks as is often the case in hub-and-spoke airports. Distributed networks (i.e., networks with distributed traffic and distributed schedules) have been found to complement point-to-point networks more than hub-and-spoke networks (Berry et al. 2004). In this sense, distributed networks align more with point-to-point networks.

[6] Airport hubs in this paper use the U.S. Department of Transportation, Federal Aviation Administration definition. There are four categories, based on the percentage of total national enplanements (i.e., physical counts): large hubs (≥1% of total enplanements), medium hubs (0.25%–0.999% of total enplanements), small hubs (0.05%–0.249% of total enplanements), and nonhubs (<0.05% of total enplanements). These are physical hubs.

There is a second "operational" definition that categorizes airports as a hub if inbound flights are scheduled to arrive from multiple origins within a short period of time, thus creating a "bank" of passengers. The coordinated arrival and departure banks together form a wave of activities and lead to peaks in airlines schedules. At present, some physical hubs are also operational hubs. However, an airport can be an operational hub without being a large physical hub (e.g., small airports primarily serving connecting passengers), while a physical hub may exist without being an operational hub (e.g., large airports primarily serving origin and destination passengers).

FIGURE 3 Number of Scheduled Flights by Type of Aircraft: April 2000 and April 2004



Notes: Turboprops include aircraft such as the ATR-42/72, Aerospatiale, Beechcraft 1900, Dornier 328 Turbo, and JETST-31 BAE. Pistons are typically fixed wing Cessnas and Pipers, e.g., the CES-150-185, AERO-200, and PA 12-32.

Source: M.R. Dayton, "Trends and Demand in Aviation Markets," presentation at the ATCA/FAA/Nav Canada Technical Symposium, Office of Inspector General, U.S. Department of Transportation, 2004.

tional airlines are only beginning to be felt. That's because the route networks of low-cost, low-fare airlines have grown large enough to make alternative service available in almost all of the large business markets." The bankruptcy declaration of United Airlines (Dec. 6, 2002), following US Airways' earlier bankruptcy (Aug. 11, 2002), appears to bring that speculation one step closer to reality.

If market shares of airlines are changing and affecting network structure, this is an important phenomenon to be addressed by both analysts and policymakers. The expected transition is certain to have an impact on almost all areas of the National Airspace System and will affect operations as well.

Understanding the emerging network is key to foreseeing what is driving future operational issues as well. A changing network will also have a significant impact on airframe manufacturers. For example, Airbus has aggressively marketed the very large A380 model aircraft over the last few years. As of the third quarter of 2004, Airbus had 139 firm orders for the aircraft. In order for production to be economically viable, Airbus requires about 250 orders. Boeing, on the other hand, recently abandoned the Sonic Cruiser program in favor of the more traditional fuel-efficient 7E7 jets. As of April 2004, Boeing had received its first order for 50 7E7s from Japan's All Nippon Airways. The size and speed of these two aircraft and the types of markets for which they are suited indicate that they are expected to serve clearly different niches within the network: the A380 appears to continue with the assumption that the long-haul hub-to-hub network (e.g., international long-haul routes) will anchor air transportation and be enhanced by feeder routes (i.e., hub-and-spokes), while the 7E7 is designed primarily to serve more point-to-point traffic.

In light of these phenomena, this paper is an attempt to understand the evolutionary nature of the U.S. airline network. In particular, we address and quantify three empirical issues:

- how the changing role of Southwest Airlines affected the network structure,
- how the increasing use of regional jets affected the network structure, and
- how the events of 9/11 were a catalyst for changes in the network structure.

Addressing these issues may provide some important insights that could lead to improved policymaking in a changing environment. It may also allow us to forecast the structure of the network. The paper is organized as follows: the next section presents our definition of an airline network; we then discuss the empirical framework and the data; next, we present our methodology and empirical results; and we conclude with policy suggestions and areas for further research.

NETWORK DEFINITION

The airline network is a dynamic environment that has numerous variants. As the business models of participating airlines change, so will the airline network. The market environment facing the network carriers, those with substantial hub-to-hub and hub-to-spoke operations in selected airports, has become increasingly competitive. A complex web linking declining average yield[7] with a narrowing margin

[7] Average annual yield (i.e., itinerary fare/passenger-miles flown) has declined by 2% annually over the last two decades following deregulation in 1978. Over the next two decades, analysts predict this rate of decline will slow down to 0.9% a year (SED-F 2003; USDOT FAA 2003).

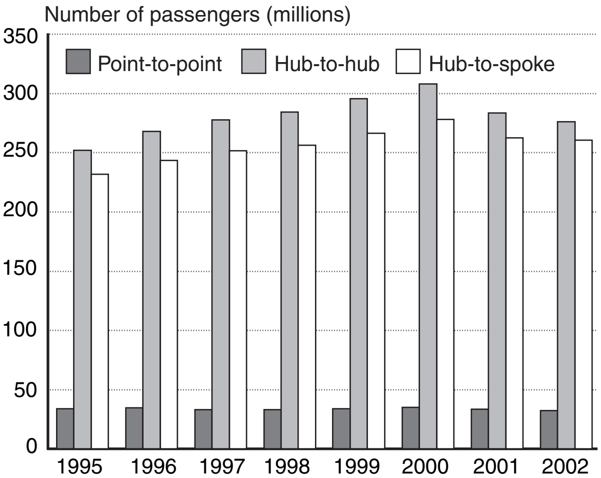
between premium and walkup versus competitive fares is forcing network carriers to undertake painful cost-cutting measures.

Acknowledging this dynamism and recognizing the present structure of the airline network, we defined the network based on its physical characteristics. Using the U.S. Department of Transportation (USDOT), Federal Aviation Administration (FAA) definition (see footnote 6), our network consists of 35 hub airports—a combination of 31 large hubs and 4 medium hubs. Although we used the physical definition for a hub, many of these airports are also operational hubs for both network carriers and LCCs. The hubs are listed in appendix A.[8] These airports together accounted for 73% of total scheduled enplanements and 69% of total scheduled aircraft operations in 2002. Although most of these airports qualify for the FAA definition of large hubs (> 1% of national enplanements), four other airports, Cleveland Hopkins (OH), Washington Reagan (DC), Memphis (TN), and Portland (OR), were included to maintain consistency with the FAA's Operational Evolution Plan (OEP) airports. Finally, appendix A provides information on which airlines are the primary and secondary air carriers at these airports.[9]

We defined the three variants of the airline network as follows:

- **Point-to-point (PP) variant** covers air travel that takes place between non-OEP airports (e.g., Teterboro (NJ) Airport to Hagerstown (MD) Regional Airport). Any travel outside OEP airports as listed in appendix A represents the point-to-point variant of the network.
- **Hub-to-hub (HH) variant** covers air travel that takes place between two major hubs (i.e., travel between OEP airports; e.g., Atlanta Hartsfield to Boston Logan).

FIGURE 4 Annual Passenger Enplanements at Airports, by Type of Network Variant



- **Hub-to-spoke (outbound) and spoke-to-hub (inbound) (HS)** covers air travel for which either the origin or the destination (but not both) is a major hub (i.e., travel between non-OEP airports and OEP airports; e.g., Atlanta Hartsfield to Teterboro).

In order to measure variants of network activities, we used two variables, the number of passenger enplanements and the number of actual aircraft departures performed. As figure 4 shows, the number of passenger enplanements in the PP variant was dwarfed by the number of enplanements under both the HH and HS variants. While the HH and HS variants together accounted for around 93% to 95% of the total enplanements, the PP share of the overall network has been in the range of 5% to 7%.

Because Southwest Airlines concentrates its operations in PP markets, it has a higher percentage of the PP variant of the airline network (ranging between 62% and 70% of total enplanements) than the HH (about 4% to 5.5%) and HS (10% to 17%) variants.

THE FRAMEWORK: RESEARCH QUESTIONS, DATA, AND METHODOLOGY

Research Questions

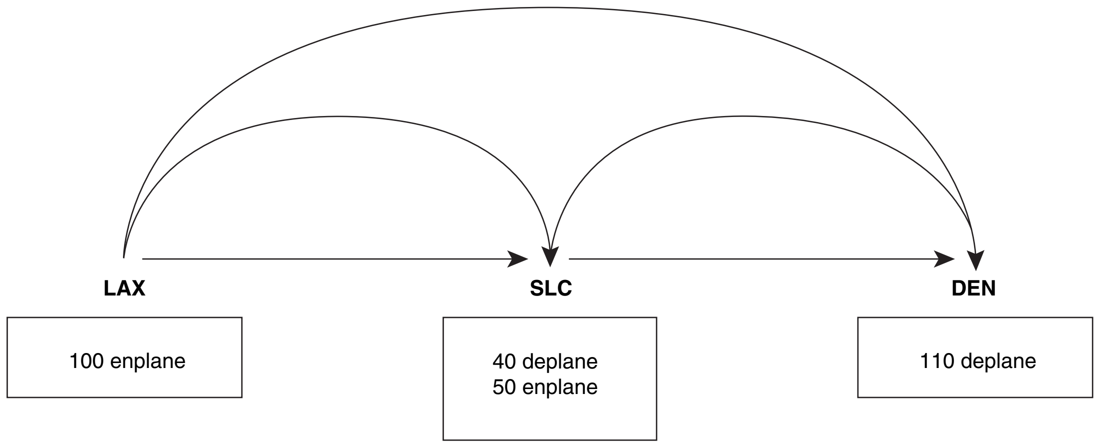
Here, we formulate three empirical issues for testing:

1. How has the expansion of Southwest Airlines affected different variants of the networks?
2. How has the changing share of regional jets affected parts of the network differently?
3. How have the events of 9/11 affected the overall network and different parts of the network?

[8] These 35 airports are also known as Operational Evolution Plan (OEP) airports. The OEP is a major FAA initiative to meet emerging air transportation needs for the next 10 years. For more details, see http://www.faa.gov/programs/oep/index.htm.

[9] The line between primary and secondary carriers is somewhat arbitrary. We provide information (reported by the *Official Airline Guide*) on air carriers that are conducting hub operations in these airports, irrespective of the magnitude.

FIGURE 5 Illustration of T100 data



Key: DEN = Denver International Airport; LAX = Los Angeles International Airport; SLC = Salt Lake City International Airport.

Notes: **Nonstop segments** are represented by straight arrows, i.e., number of passengers transported between points (between takeoff and landings).
LAX to SLC: 100 passengers transported
SLC to DEN: 110 passengers transported
Onflight markets are represented by curved lines, i.e., where passengers are enplaned and deplaned on a flight.
LAX to SLC: 40 passengers
LAX to DEN: 60 passengers
SLC to DEN: 50 passengers
For a one-stop flight, the number of passengers would be the same as under segment and market.

Source: U.S. Department of Transportation, Office of the Secretary, *O&D Survey Reporting Regulations for Large Air Carriers: Code of Federal Regulations Part 241, Section 19-7* (Washington, DC: 1992).

Answering these questions may provide some important insights into the process by which the network is undergoing changes. Furthermore, understanding these changes may also allow us to forecast the structure of the network into the future.

Data

Data for this exercise come from the Bureau of Transportation Statistics (BTS), DOT T100 schedule. T100 is the transportation schedule for Form 41 data that every major airline is required to submit to BTS each quarter. T100 is divided into two parts: T100 market segment data (T100M), which cover on-flight origin and destination (O&D) or direct markets; and the T100 segment data (T100S), which contains data for market segments. In particular, T100S is the Data Bank 28DS of Form 41 that provides segment traffic (i.e., the number of passenger enplanements, freight ton-miles, and departures scheduled and performed) by scheduled air carriers for freight and mail by service class and type of aircraft equipment, capacity (i.e., available freight ton-miles and available passenger seat-miles), and performance indicators (i.e., ramp-to-ramp elapsed time and airborne elapsed time) by month and year. The data are reported by major air carriers operating between airports located within the boundaries of the United States and its territories (see CFR 2001 for more details). The data cover January 1995 to September 2003.[10] For our empirical analysis, we used T100 domestic segment quarterly data for the period covering 1995:Q1 to 2003:Q3 (data for 35 continuous quarters).

Figure 5 shows the nonstop segment data (T100S) for a hypothetical flight from Los Angeles (LAX) to Salt Lake City (SLC) and then to Denver (DEN). The data for the LAX-SLC segment thus includes not only the O&D traffic within that segment (i.e., people originating in LAX and destined for SLC), but also the passengers who are originating at LAX, stopping at SLC, and then flying on to Denver. The T100M market data, on the other hand, for LAX-SLC includes only those people originating in LAX and destined for SLC. Unfortu-

[10] See http://www.transtats.bts.gov and click on the aviation data link for T100 domestic data segments in the Form 41 traffic file.

nately, however, T100M is limited to fewer variables: number of passengers by O&D, freight, mail, carriers, distance, month, and year.

Each segment reported in T100S is unique, distinctively defined by air carrier and the type of equipment flown. Therefore, the LAX-SLC flights shown in figure 5 will be reported twice, for example, if a carrier flew the segment using two equipment types, holding all other factors constant. The total number of segments can be aggregated over the same O&Ds to provide a logical basis for defining the network. For example, there were 70,127 distinct segments in 2003:Q3. These unique segments reduced to 11,179 O&D segments when summarized by the same O&D, thus providing the basis for our estimation. Summed over 1,695,848 distinct segments for 1995:Q1 to 2003:Q3, we had 228,129 observations. These observations were used to estimate our econometric model.

Methodology and Results

Estimation

For both total passenger enplanements and aircraft departures performed, we specified econometric models in natural logs[11] by variants of the network as follows:

$ln(Pax_{ij;\,k}) = F$ [seasonal dummy, share of Southwest Airlines in total passenger, share of regional jets in total passenger, dummy representing 9/11, ln(one-quarter lag of passenger)] (1)

$ln(A/C\ Dep_{ij;\,k}) = F$ [seasonal dummy, share of Southwest Airlines in total departures, share of regional jets in total departures, dummy representing 9/11, ln(one-quarter lag of departures)] (2)

where

ln = natural log,

i = origin,

j = destination,

k = type of network.

The three variants of the network, k = 0, 1, 2, are defined as point-to-point (PP; k = 0), hub-to-hub (HH; k = 1), and hub-to-spoke, including both hub-to-spoke and spoke-to-hub traffic (HS; k = 2). The two endogenous variables, $ln(Pax_{ij;k})$ and $ln(A/C\ Dep_{ij;k})$ are the natural logs of the number of total passenger enplanements and total aircraft departures performed, respectively, aggregated within the i-j O&D market for the k-th variant of the network.

It is important to understand that both passenger enplanements and aircraft departures performed are generally determined by economic factors (e.g., fares and income), demographic factors (e.g., population and age distribution in O&D markets), and the quality of services (e.g., schedule choices and types of aircraft) (Bhadra 2003). Notice also that whether hubs connect directly to other airports (HS) or via other hubs (HH) depends on market features such as the size and composition of the market, fares, connection possibilities, and so forth (see Shy (2001, 215–231) for an analytical discussion of airline networks; and Bhadra and Hechtman (2004) for an empirical analysis). In our present dataset, however, not all such information is available.[12] Nonetheless, the independent variables specified above may capture the trends in passenger enplanements and performed departures quite substantially and well enough, as discussed earlier.

In particular, we postulate that seasons affect both passenger enplanements and departures performed, and those variations may differ depending on the type of network. Empirically speaking, air travel goes through cyclical variations, peaking during spring and summer (i.e., April-September; season dummy = 1) and hitting its trough during fall and winter (i.e., October-March; season dummy = 0). Thus, we designed a seasonal dummy variable to capture this cycle.

[11] We used natural logs, as opposed to levels, for two reasons. First, transforming the endogenous variables into their natural logs eliminated heteroskedasticity from the dataset. Second, estimated coefficients of log-transformed models have clearer intuitive appeal than using level variables.

[12] The list of variables available in T100S and T100M has been given above. Many factors, fares in particular, are reported in what is commonly known as the Origin and Destination Survey or the DB1B. That database, while containing useful information such as fares and quarterly passengers in an O&D market, does not include information on aircraft equipment types and other performance indicators (for more details on types of data and variables, see http://www.transtats.bts.gov/Databases.asp?Mode_ID=1&Mode_Desc=Aviation&Subject_ID2=0).

Following our earlier discussion, we formulated a variable that accounts for Southwest Airlines' share of total passenger enplanements and aircraft departures performed.[13] This share variable (i.e., the percentage share of Southwest Airlines in total enplanements) has been designed to capture the impact of the airline on different variants of the network (i.e., k = 0, 1, 2) defined over i-j segments. Similarly, the share of regional jets[14] has been formulated to capture their impact on totals. A dummy variable representing the beginning of the effect following 9/11 (i.e., 2001:Q3) was formulated to examine the effects of 9/11 on two endogenous variables defined for the network type.[15] In other words, this dummy variable assumes a value of 0 for the period 1995:Q1 to 2001:Q2 and a value of 1 for the period 2001:Q3 to 2003:Q3. Finally, an autoregressive term, the log of both enplanements and aircraft departures performed lagged one quarter, was used to capture the time series component of this time series pooled cross-section sample. It is important to note here that we postulate that both enplanements and aircraft departures performed are driven by the same set of explanatory variables and are determined simultaneously as a system.

Given the interdependency among enplanements and departures performed, it is likely that the error structures of the equations may be linked to each other. Although each equation in the system above appears to be independent and unrelated, they might be linked to each other through errors. Thus, this type of system is also called "disturbance-related" or "error-related" regression equations.

Under this circumstance, econometricians often recommend the use of the seemingly unrelated regression (SUR) technique for estimation (Pindyck and Rubinfeld 1991, p. 308). SUR is used when a system consists of two or more equations where errors may be correlated across equations. SUR is considered to be appropriate when all the right-hand side regressors are assumed to be truly exogenous and the errors satisfy the following conditions:

1. $\varepsilon_{ij;\,k}$ (i.e., error terms) have zero means and finite variances,
2. the variances of errors may differ, and,
3. there is a presumed correlation between $\varepsilon_{ij;1}$ and $\varepsilon_{ij;2}$.

Given that 1–3 are likely to be true for the dataset we used, we adopted the SUR methodology for estimation (SAS 1993).

Results

Table 1 presents the results of the estimation of the two equations specified linearly. Quite a few interesting features underlying the data and findings deserve special attention.

First, it is important to note that we make a distinction between the number of observations (N) used and the number read (i.e., the last two columns in table 1). The difference between N read and N used arises because of the unavailability of data, accounted for primarily by the lack of lagged variation in the one-quarter lag of passenger enplanements and departures performed on a particular segment. In other words, quarter t did not have a corresponding quarter t-1 observation. Reviewing these numbers across rows, it is evident that the PP network has relatively less continuity over time than the other two types of networks. In particular, almost half of the segments (46%) that were observed at quarter t for the PP network did not have lagged entries for quarter t-1. Hence only 39,880 observations were used from a total of 73,438 observations.

In comparison, the HS network, including both HH and HS routes, seems to have more continuity over time, and hence observations used are far closer to the available total number of observations. Given that we used 35 continuous quarters in our

[13] We define the share of passengers as: (*Southwest's passengers/total passengers*)*100. Thus, this and other share variables are expressed in 100th of units and not in a fraction, 0 to 1.

[14] Here we define RJs as the following: Canadair RJ-100/R, Canadair RJ145-200, Embraer EMB-135, Embraer EMB-145, Embraer EMB-140, Avroliner RJ85, BAE-146-3, and Do328JET. There may be other RJs outside of this definition.

[15] Other factors, e.g., slowdown of the economy, internet booking, etc., were also taking place around this time (see footnote 1). However, the purpose of including a dummy variable for 2001:Q3 and later is to capture the sudden unsystematic effect that took place in this quarter and its impact. In other words, this dummy variable should be interpreted as representing the impact of 2001:Q3 as a catalyst for all that changed in the time series before and after this event.

TABLE 1 System of Scheduled Aviation Activities in the United States: Number of Passengers and Aircraft Departures (in natural logs)

Network variants (1)	System equations (2)	Intercept (3)	Seasonal dummy (4)	Share of Southwest Airlines (5)	Share of RJ carriers (6)	9/11 dummy (7)	*Ln* (one quarter lag) (8)	Adj. R^2 (9)	*N* (used) (10)	*N* (read) (11)
PP	$ln(Pax_{ij})$	0.5039*	0.0958*	0.0032*	–0.0011*	–0.1083*	0.9107*	0.8903	39,880	73,438
	$ln(A/C\ Dep_{ij})$	0.1928*	0.0367*	0.0017*	–0.0006*	0.0465*	0.9215*	0.8865		
HH	$ln(Pax_{ij})$	–0.8040*	0.1205*	0.0009*	–0.0004*	–0.0161*	0.9203*	0.9039	34,192	35,827
	$ln(A/C\ Dep_{ij})$	0.4363*	0.0407*	0.0011*	0.0013*	–0.0110*	0.9263*	0.9155		
HS	$ln(Pax_{ij})$	–0.7568*	0.0863*	0.0012*	0.0006*	–0.0309*	0.9128*	0.8485	97,171	118,864
	$ln(A/C\ Dep_{ij})$	0.4290*	0.0006	0.0007*	0.0012*	–0.0028	0.9151*	0.8591		

Key: * = 99% level of significance.
Note: First row in each network variant represents scheduled passengers and the second row represents the number of aircraft departures.

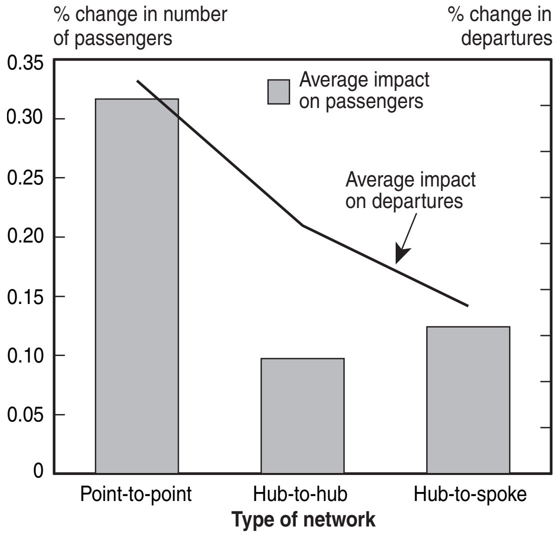
analysis, the number of observations under the HH network variant represents, on average, 1,000 segments per quarter. Segments under the HS variant, on the other hand, are almost three times the size of the HH variant per quarter. The PP segments fall in between those of HH and HS.

Second, the estimated system model appears to have a very good fit. In particular, all the explanatory variables describe the two endogenous variables very well, resulting in a high adj. R^2. Almost 90% of the variation in the dependent variables, across the different networks, is explained. Third, almost all the variables, with the exception of the seasonal and the 9/11 dummy variables in the case of the HS network, are statistically significant at the 99% level. Furthermore, the estimated parameters of the simultaneous system appear to confirm the expected signs for the empirical hypotheses for most variables.

In particular, the seasonal dummy variable confirms the hypothesis that in spring and summer both passenger enplanements and aircraft departures performed go up, increasing the most for the HH network, followed by the PP and HS variants. On average, the spring and summer quarters add about 12% to passenger enplanements and 4% to departures performed on the HH network variant, while adding 10% more passenger enplanements and 4% more departures to the PP network variant. The HS variant gains about 9% more passenger enplanements during the peak travel season, while departures in this network do not show significant seasonal changes, thus suggesting excess capacity. Examination of passenger data for 2000 for major carriers (i.e., those who use the HH and HS network variants primarily) indicates that, on average, passenger enplanements increased by 6.3 million per quarter during the spring and summer, or about 12% more than the overall quarterly average (ATA 2004).

As anticipated, Southwest Airlines impacts all variants of the network positively and in varying degrees. This further confirms the already empirically established *Southwest effect*. Figure 6 summarizes the findings from table 1. Notice that the effects of Southwest Airlines (column 5 in table 1) are captured by two variables in equations 1 and 2: Southwest Airlines' share of total passenger enplanements and departures performed for equations 1 and 2, respectively. Estimated parameters, multiplied by 100 (to account for the share expression in units of 100), thus represent the effect of a one percentage point increase in Southwest Airlines' market share on the percentage change in passenger enplanements and departures performed (i.e., natural logs of these two variables). Thus, a one percentage point increase in Southwest Airlines' market share in total segment passenger enplanements will add 0.32% more passengers each quarter (i.e., in the short run, holding all other factors constant) to the overall PP network. Similarly, a one percentage point Southwest expansion adds 0.09% and 0.012% more passengers overall for the HH and HS

FIGURE 6 Impact of a One Percentage Point Change in Southwest Airlines' Market Share on Different Variants of the Network



variants, respectively, for each representative quarter in the short run.

Compared with other research (Bennett and Craun 1993; Morrison 2001), the estimated values in our model represent a much smaller *Southwest effect*. In past studies, the Southwest effect has been estimated using the impact of the entry of Southwest Airlines on airline fares in different markets, and the impact of falling airfares, in turn, on passenger demand (e.g., Morrison 2001). For example, the effect of Southwest Airlines on fares was estimated to be in the range of 6% to 46% for every one percentage point increase in Southwest's market share, and the effects of those falling fares on passenger demand were estimated to be in the range of 5% to 10% for each percentage point decline in fares (Bennett and Craun 1993). Thus, the total Southwest effect (of a one percentage point increase in market share on the percentage increase in overall market share) has been estimated to be in the range of 30% to 460%. Bhadra (2003) estimated elasticities of demand for the overall U.S. domestic air markets in the range of 0.55 to 1.8, which would yield lower values for the Southwest effect in the range of 3.3% to 82.8%. As is apparent, our estimated effects are smaller still.

There are several reasons for this major difference in results. The earlier studies examined the Southwest Airlines phenomenon when the carrier was much smaller in size and for a particular year. We, however, studied the effects of the presence of Southwest on all networks and estimated these effects over a time series. The long-term accumulated effects of this expansion, as will become evident later, are not small. As noted earlier (see figure 4), Southwest Airlines' strongest presence is in the PP network. About 70% of all passengers in the PP network (32.34 million in 2002) flew Southwest Airlines (i.e., 22.65 million). In comparison, about 5% and 14% of the passengers in the HH and HS networks, respectively, flew Southwest. Between 1995 and 2002, about 35 million passengers flew in the PP network annually, which had an average annual growth rate of about 0.6%. During this period, however, Southwest Airlines expanded its PP network market share from 62% to 70%, or from 21 million passengers in 1995 to 22.65 million in 2002. Southwest Airlines grew, on average, twice as fast (1.13% per year) as the average annual growth rate of the entire PP network. Similar magnitudes of scale and growth also follow for the HH and HS networks. Clearly, these rates suggest much smaller expansion than seen in earlier studies.[16]

Second, the specification of the model may also be responsible for the results. Unlike earlier studies, we specified both passenger enplanements and departures performed as endogenous variables simultaneously determined via a common set of variables. If some of the independent variables assumed to be exogenous are actually endogenous (e.g., RJ carrier shares), this incorrect specification of the true model may seriously underestimate the magnitude of the coefficient on all the other exogenous variables, including the Southwest effect.

Although they appear to be rather small, the magnitude of the estimated parameters translate to considerable changes in the number of enplanements as Southwest Airlines expands its market share. It is obvious that more departures will have to be performed in order to accommodate these additional flows of enplanements due to a substan-

[16] An expansion in the magnitude of 30% to 460% (mentioned earlier) from the base of 1995 would require total passengers for Southwest Airlines under the PP network to be between 27 million and 118 million annually. This is unlikely given that the total size of the network is about 35 million passengers annually.

tial increase in Southwest Airlines' market share. Hence, departures would increase by 0.17% to accommodate greater passenger enplanements (0.32%) under the PP variant of the network; 0.11% to accommodate additional passenger enplanements (0.09%) under the HH variant; and 0.07% to accommodate the 0.12% increase in passenger enplanements under the HS variant of the network (see figure 6).

It is interesting to note that the Southwest expansion adds proportionately more enplanements and departures to the PP variant than to the other two types of networks. Therefore, the expansion of Southwest Airlines will enhance the PP-type network while still positively influencing both the HH and HS networks. This result simultaneously confirms the expanding distributed or PP network of Southwest Airlines and the Southwest effect.

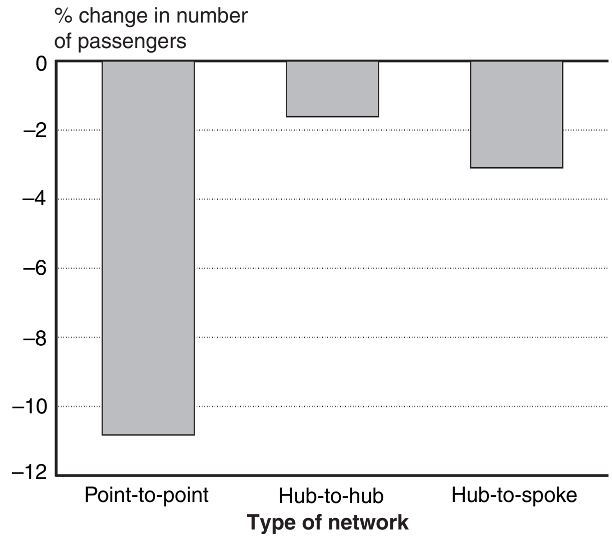
To the extent that the presence of RJs is truly exogenous,[17] RJs' percentage of shares of total enplanements and performed departures have interesting implications. The expansion of RJ carriers is associated with a reduction in traffic in the PP network, and it encourages and even accommodates an increase in the HS feeder operations network (see column 6 in table 1).

For the HH network, RJ expansion is accompanied by fewer passenger enplanements and more departures. Another way to look at this is that increasing the shares of RJs is apparently accompanied by more departures under the HH network to compensate for the smaller number of passengers carried by each aircraft. These findings seem to establish the point that RJs are an important tool for feeding hubs in hub-and-spoke networks.

Because RJs are generally considered to represent an improvement in the quality of service offered on feeder routes (i.e., in the HS network), the expansion of RJs is expected to enhance both the HS network and the HH network fed by HS routes. However, this expansion may negatively affect passenger travel on HH routes, where RJs are more likely to be perceived as a reduction in the quality of service as compared with larger jets (e.g., B737s) that previously served those routes. In the PP network, the negative impact of RJs on enplanements and departures may reflect (as in the HH network) a decline in traffic caused by broader economic factors not included in our model. It may also reflect higher fares required by the greater unit costs of RJs compared with larger B737-type jets. Finally, since Southwest Airlines operates B737s *exclusively*, the presence of RJs on a route is likely to be highly collinear with the *absence* of Southwest Airlines on the route, so that the RJ share may be picking up part of the explanatory power of the Southwest share.

The events following 9/11 had a sizeable impact on all types of networks (see column 7, table 1). As figure 7 demonstrates, the PP network suffered the most, losing, on average, 11% of its passengers but performing, interestingly, about 5% more departures per quarter following 9/11. The HS network, on the other hand, lost approximately 3% of its enplanements and 0.28% of its departures per quarter, which can be attributed to factors relating to the events of 9/11. The HH network lost about 1.6% of its passenger enplanements per quarter and 1.1% of its departures performed each quarter. The reductions in passenger enplanements in the PP network, combined with increased aircraft operations, suggest that large network carriers were outsourcing their operations on these routes to smaller regional carriers flying smaller aircraft. This vertical disintegration of markets appears to be stronger in the PP

FIGURE 7 Effect of 9/11 on Different Variants of the Network



[17] We acknowledge the limitation that in a well-specified model of an airline network, the share of RJs may be endogenous as well.

network post 9/11 than it was on the HH and HS networks, where enplanements and aircraft operations declined together.

Finally, past enplanements and departures performed have proven to be robust explanatory variables for both the endogenous variables in the system. For example, a 1% increase in past quarter passenger enplanements and departures performed would increase both current enplanements and current departures by about 0.9% (see column 8, table 1). This coefficient has further implications. Together with other structural variables, the long-term effects can be separated by the accumulated effects from all short-term variables as follows: $\Sigma b_i/(1-c)$, where b_i's are the estimated structural parameters of the two equations, and c is the estimated parameter of the log-lagged value of the endogenous variables. Notice here that while Σb_i captures all the short-term effects, $\Sigma b_i/(1-c)$ captures the long-term accumulated effects of all the variables on the endogenous variables. Thus, although a one percentage point expansion in Southwest Airlines' market share increases the passenger volume in the PP network by 0.32% in the short run, its long-term accumulated effect on total passenger enplanements is 3.58%, that is (0.32/(1–0.9107)). Applying this formulation to the sum of the effects of all variables would yield, for example, a cumulative effect of 2.21% and 2.50% on passenger enplanements and departures performed, respectively, for the PP network. In other words, the impact of all short-term structural parameters (i.e., Σb_i) accumulates at the rate of 0.197% per quarter for passenger enplanements and 0.196% per quarter for departures performed and eventually yields the long-term impacts. Long-term accumulated effects for passenger enplanements and departures performed were calculated, respectively, as 1.98% and 3.70% for the HH network and 2.76% and 2.26% for the HS network. These calculated rates correspond fairly well to those observed from actual data.

These findings may have important uses. First, they may be important for understanding the impact of Southwest Airlines and other LCCs on the airline network. For example, despite the smaller magnitudes and associated issues related to misspecification, our findings indicate that the expansion of Southwest Airlines in particular, and perhaps LCCs in general, would increase the flow of traffic in the overall network substantially. While this expansion would have a proportionately greater effect on the PP network, the HH and HS networks would be positively affected as well. Using the estimated coefficients underlying these segments, we can now compute the effect on segment traffic in any market in which Southwest Airlines operates. This means a clear isolation of the Southwest impact on segment passenger traffic and departures by network can be performed.

Second, some of these estimated coefficients can also be used to partially estimate and forecast, keeping other variables constant, the effect on individual markets that can be expected if Southwest Airlines enters those markets, for example, the Philadelphia market beginning in May 2004. The estimated parameters can be used to forecast traffic at the Philadelphia airport with the expansion of Southwest Airlines, and hence may aid airport infrastructure planning not only for Philadelphia but also for those airports linking with it. A similar reasoning applies to the RJ expansion. In addition, partial coefficients with respect to the 9/11 dummy variable can also be used to estimate financial losses in different segments incurred by different airlines.

Finally, the simultaneous system model as a whole can be used to forecast enplanements and departures performed for segment traffic under different types of network structures. The estimated model can provide a foundation for forecasting segment traffic given anticipated and projected values of the explanatory variables for different types of networks, especially with given market share levels for Southwest Airlines and RJ aircraft. Since the segment traffic is the primary measure of traffic flow dealt with by the air traffic control (ATC) system regularly, both at en route centers and at airports, the ability to better project segment traffic in a way that takes explicit account of airline networks may improve the ATC system. Better traffic projections may also help planners better allocate resources to improve the country's critical air transportation infrastructure.

The estimated model has been found to be robust and stable. However, the model is still somewhat limited. It can be argued that some of the explanatory variables, Southwest and RJ expansion in particular, are not truly exogenous. In fact, they result from interactions in the complex market processes and hence cannot be posited truly as independent variables. Furthermore, the dependent variables may depend on market and economic factors (e.g., fares and income), demographic factors, and quality of service. While incorporating these variables may improve the model, the marginal benefit in the overall model fit (e.g., improvement in the adj. R^2) may be somewhat limited. Finally, the model in its present form may be misspecified, as indicated by the comparison of estimated coefficients with that of other studies (e.g., Southwest effects). While we acknowledge these limitations, we believe that the proposed analytical framework and estimated model can provide a strong foundation for both policy analysis and forecasting.

CONCLUSIONS, POLICY OBSERVATIONS, AND FUTURE RESEARCH

In this paper, we examined the U.S. domestic airline network. By defining hub-to-hub, hub-and-spoke, and point-to-point as the three essential network components, we have identified domestic scheduled air transportation under each of these networks. Using cross-section pooled time series data for 35 consecutive quarters for all scheduled carriers in the United States between 1995:Q1 and 2003:Q3, we were able to estimate a simultaneous system comprising passenger enplanements and aircraft departures performed in types of networks and aggregated within O&D markets.

Our findings indicate the existence of increased vertical disintegration of market segments following the events of 9/11. Second, seasonality tends to play an important role in determining segment traffic, peaking during the spring and summer. Third, we found evidence that the expansion of Southwest Airlines affects all networks positively, with a proportionately larger impact on the point-to-point network than on the hub-and-spoke network. Fourth, regional jets have been found to affect the network in mixed ways, with negative impacts on the point-to-point network and positive impacts on the hub-and-spoke network. Fifth, effects from 9/11 have been generally negative on all three types of networks, with the largest impact falling on point-to-point passenger traffic, followed by passenger traffic on the hub-and-spoke and the hub-to-hub networks. Increased use of departures to accommodate a lower number of passenger enplanements, in the case of the point-to-point network, provides additional indirect evidence of market segmentation. Finally, auto-regressive terms for both enplanements and departures performed have proven to be robust explanatory variables.

These are meaningful results and may provide better insight into the nature of the U.S. airline network and the factors that explain its evolution over time. The overall statistical fit may also justify using the model to forecast the trends in scheduled network activities by segments. Furthermore, estimated coefficients may be used both to guide policy decisions and to facilitate the planning of the air transportation infrastructure. In particular, empirical results tend to demonstrate that the future airline network may be more distributed as Southwest Airlines or similar airlines expand their operations.

Our analysis is the first systematic effort, as far as we are aware, to empirically establish the existence of an emerging distributed network in the U.S. air transportation system. The infrastructure that served the hub-and-spoke operations in the past may be required to change to further accommodate the needs of Southwest Airlines and other LCCs. Additional attention should therefore be given to the segment traffic and the network structure that is emerging as the industry undergoes serious structural changes.[18]

The model we propose to capture these emerging trends, however, is somewhat limited due to the lack of other important explanatory variables: fare, income, and population at segment end points, and the quality characteristics influencing these services. Treating Southwest Airlines and RJ expansion as explanatory variables may also limit our understanding of the potentially endogenous character of

[18] The FAA's recent initiative (USDOT FAA 2004) examines this issue by looking further and beyond the infrastructure needs of the 35 traditionally large OEP airports.

these factors and may limit the model's applicability in forecasts. Due to these important omissions, our model may also suffer from misspecification. Improvement of the model could be approached by incorporating explanatory factors and determining which factors are truly endogenous. These are tasks for future research.

ACKNOWLEDGMENTS

An earlier version of this paper was presented at the Forecasting in Transportation Session of the 23rd International Symposium on Forecasting, Mérida, Mexico, June 15–18, 2003. The authors would like to thank four anonymous referees and the Guest Editors for their extensive comments and suggestions that led to significant improvement in the present version. We would also like to thank our colleagues, Tom Berry, Dr. Gerald Dorfman, Dr. Glenn Roberts, and Jackie Kee for their comments, help, and assistance with this paper.

The contents of this document reflect the views of the authors and The MITRE Corporation and do not necessarily reflect the views of the Federal Aviation Administration (FAA) or the Department of Transportation (DOT). Neither FAA nor DOT makes any warranty or guarantee, expressed or implied, concerning the content or accuracy of these views.

REFERENCES

Air Transport Association (ATA). 2004. Available at http://www.airlines.org/public/home/default1.asp.

Airline Monitor. 2004. A Review of Trends in the Airline and Commercial Jet Aircraft Industries, monthly issues. Available at http://airlinemonitor.com.

Aviation Week & Space Technology. 2002. International Air Transport: Restructure or Die. November 18, pp. 52–69.

Bennett, R. and J.M. Craun. 1993. The Airline Deregulation Evolution Continues: The Southwest Effect. Office of Aviation Analysis, U.S. Department of Transportation. May.

Berry, T. 2004. Airline Network Characteristics and Connectivity, paper presented at the Today's Aviation Capacity Issues and Solutions session of the 83rd Annual Meetings of the Transportation Research Board, Jan. 11–15, 2004, Washington, DC.

Berry, T., D. Bhadra, J. Gentry, and G. Nelson. 2004. Airlines Entry and Exit and the Impact on Air Traffic Management: An Analytical Framework for Zero-Sum and Positive-Sum Outcomes. *Journal of the Transportation Research Forum,* Fall 2004 (forthcoming).

Bhadra, D. 2003. Demand for Air Travel in the United States: Bottom-Up Econometric Estimation and Implications for Forecasts by O&D Pairs. *Journal of Air Transportation* 8(2):19–56, September.

Bhadra, D. and D. Hechtman. 2004. Determinants of Airport Hubbing in the U.S.: An Econometric Framework. *Public Works Management and Policy* 9(1):26–50, July.

Code of Federal Regulations (CFR). 2001. Ch II (Jan. 1, 2001, Edition), Pt. 241. U.S. Department of Transportation.

Morrison, S.A. 2001. Actual, Adjacent, and Potential Competition: Estimating the Full Effect of Southwest Airlines. *Journal of Transport Economics and Policy* 35(2):239–256.

Pindyck, R.S. and D.L. Rubinfeld. 1991. *Econometric Models and Economic Forecasts,* 3rd Edition. New York, NY: McGraw-Hill, Inc.

SAS Institute. 1993. *Applications Guide 2: Econometric Modeling, Simulation, and Forecasting, Version 6, First Edition.* SAS/ETS Software.

Socio-Economic Demand Forecasting (SED-F) Team. 2003. Socio-Economic Demand Study, Ch. 2 supporting a National Plan to Transform the National Air Transportation System, an internal working paper prepared for the Joint Planning and Development Office, an initiative of the U.S. Department of Transportation, Federal Aviation Administration (FAA) submitted by staff of FAA, the Logistics Management Institute, GRA Inc., The Volpe National Transportation Systems Center, and the National Aeronautics and Space Administration.

Shy, O. 2001. *The Economics of Network Industries.* Cambridge, MA: Cambridge University Press.

U.S. Department of Transportation (USDOT), Federal Aviation Administration (FAA). 2003. Aerospace Forecasts: Fiscal Years 2003–2014. Washington, DC

_____. 2004. *Capacity Needs in the National Airspace System: An Analysis of Airport and Metropolitan Area Demand and Operational Capacity in the Future.* Washington, DC.

APPENDIX A Activities at 35 Operational Evolution Plan (OEP) Airports in the United States: 2002

Airport code	Airport	Primary and secondary air carriers	City	State	Hub type	Number of annual scheduled passengers	Percentage of total passengers	Number of annual scheduled A/C ops	Percentage of total A/C ops
ATL	Atlanta Hartsfield Intl	Delta, Airtran	Atlanta	GA	L	36,321,239	5.79	642,727	4.87
ORD	Chicago O'Hare Intl	United, American	Chicago	IL	L	31,026,878	4.94	612,553	4.64
LAX	Los Angeles Intl	United, American, Southwest	Los Angeles	CA	L	26,323,259	4.19	447,170	3.39
DFW	Dallas-Ft Worth Intl	American, Delta	Dallas-Ft Worth	TX	L	24,148,619	3.85	493,772	3.74
PHX	Phoenix Sky Harbor Intl	America West, Southwest	Phoenix	AZ	L	16,930,419	2.70	370,247	2.80
DEN	Denver Intl	United, Frontier	Denver	CO	L	16,544,458	2.64	330,825	2.50
LAS	Las Vegas McCarran Intl	Southwest, America West, United	Las Vegas	NV	L	16,540,417	2.64	317,700	2.41
IAH	George Bush Intercontinental	Continental, American	Houston	TX	L	15,889,349	2.53	299,903	2.27
MSP	Minneapolis-St Paul Intl	Northwest, American	Minneapolis	MN	L	15,553,423	2.48	326,974	2.48
DTW	Detroit Metropolitan Wayne County	Northwest	Detroit	MI	L	15,124,490	2.41	337,816	2.56
JFK	John F Kennedy Intl	American, Jet Blue, Delta, United	New York	NY	L	15,050,456	2.40	245,475	1.86
SFO	San Francisco Intl	United, American	San Francisco	CA	L	14,856,842	2.37	260,501	1.97
EWR	Newark Intl	Continental, American	Newark	NJ	L	14,073,453	2.24	282,849	2.14
MIA	Miami Intl	American, Continental	Miami	FL	L	13,889,275	2.21	304,863	2.31
MCO	Orlando Intl	Delta, Southwest, American	Orlando	FL	L	12,631,347	2.01	201,203	1.52
SEA	Seattle-Tacoma Intl	Alaska, United	Seattle	WA	L	12,570,572	2.00	217,352	1.65
STL	Lambert-St Louis Intl	American, Southwest	St Louis	MO	L	12,412,120	1.98	295,148	2.23
PHL	Philadelphia Intl	US Airways, American	Philadelphia	PA	L	11,631,738	1.85	267,402	2.02
CLT	Charlotte Douglas Intl	US Airways	Charlotte	NC	L	11,589,824	1.85	239,173	1.81
BOS	Boston Logan Intl	American, US Airways, Delta	Boston	MA	L	10,665,476	1.70	207,138	1.57
LGA	La Guardia	US Airways, American, Delta	New York	NY	L	10,416,041	1.66	207,915	1.57
CVG	Cincinnati-Northern Kentucky Intl	Delta	Covington-Cincinnati, OH	KY	L	9,879,246	1.57	150,943	1.14

continued on next page

APPENDIX A Activities at 35 Operational Evolution Plan (OEP) Airports in the United States: 2002 *(Continued)*

Airport code	Airport	Primary and secondary air carriers	City	State	Hub type	Number of annual scheduled passengers	Percentage of total passengers	Number of annual scheduled A/C ops	Percentage of total A/C ops
BWI	Baltimore-Washington Intl	Southwest, US Airways	Baltimore	MD	L	9,489,296	1.51	210,349	1.59
HNL	Honolulu Intl	Aloha, Hawaiian, ATA	Honolulu	HI	L	9,266,615	1.48	174,544	1.32
PIT	Pittsburgh Intl	US Airways, Delta	Pittsburgh	PA	L	9,212,335	1.47	188,154	1.42
SLC	Salt Lake City Intl	Delta, Southwest	Salt Lake City	UT	L	8,908,798	1.42	151,121	1.14
FLL	Ft Lauderdale-Hollywood Intl	Continental, Delta, Southwest, American, US Airways, Spirit, Jet Blue	Ft Lauderdale	FL	L	7,983,704	1.27	147,874	1.12
IAD	Washington Dulles Intl	United, Delta, American	Washington	DC	L	7,936,618	1.26	130,351	0.99
MDW	Chicago Midway	American, Southwest	Chicago	IL	L	7,573,932	1.21	161,468	1.22
TPA	Tampa Intl	Southwest, Delta, US Airways, Continental, Air Tran, American	Tampa	FL	L	7,544,284	1.20	145,968	1.11
SAN	San Diego Intl-Lindburgh Field	Southwest, American, United	San Diego	CA	L	7,224,573	1.15	143,298	1.08
PDX	Portland Intl	Alaska, United, Southwest	Portland	OR	M	5,970,960	0.95	122,407	0.93
DCA	Washington Reagan Natl	US Airways, Delta	Washington	DC	M	5,311,436	0.85	121,456	0.92
CLE	Cleveland Hopkins Intl	Continental, American	Cleveland	OH	M	5,007,767	0.80	86,572	0.66
MEM	Memphis Intl	Northwest	Memphis	TN	M	4,784,135	0.76	237,385	1.80
Aggregate of all 35 airports						**460,283,394**	73.34	**9,080,596**	68.75
Scheduled air transportation in the nation as a whole						**627,600,000**	100.00	**13,208,600**	100.00

Key: A/C = aircraft; L = large hubs; M = medium hubs.

Notes: Only scheduled passenger and A/C operations were considered to make the data compatible with T-100 segment data.

Airport hubs in this paper use the U.S. Department of Transportation, Federal Aviation Administration definition. There are four categories of total enplanements (i.e., physical counts): large (> 1% of total enplanements), medium (0.25%–0.999% of total enplanements), small hubs (0.05%–0.249% of total enplanements), and nonhub (< 0.05% of total enplanements). These are physical hubs.

There is a second definition that categorizes airports as a hub where inbound flights are scheduled to arrive from multiple origins within a short period of time thus creating a bank of passengers. The coordinated arrival and departure banks together form a wave of activities and lead to peaks in airlines schedules. Some physical hubs are, thus, operational hubs. However, an airport can be an operational hub without being a physical hub (i.e., airports primarily serving connecting passengers), and a physical hub may exist without being an operational hub (i.e., airports primarily serving origin and destination passengers.

Sources: U.S. Department of Transportation (USDOT), Federal Aviation Administration (FAA), *Terminal Area Forecast (TAF)* (Washington, DC: 2003); USDOT, FAA, *Aerospace Forecasts, 2003–2014* (Washington, DC: U.S. Government Printing Office, 2003); and *Official Airline Guide* (OAG), available from http://www.oagflights.com.

JOURNAL OF TRANSPORTATION AND STATISTICS

Guidelines for Manuscript Submission

Please note: Submission of a paper indicates the author's intention to publish in the *Journal of Transportation and Statistics* (JTS). Submission of a manuscript to other journals is unacceptable. Previously published manuscripts, whether in an exact or approximate form, cannot be accepted. Check with the Managing Editor if in doubt.

Scope of JTS: JTS has two primary areas of focus. We publish articles that look at transportation from a statistical perspective. We also publish articles that do not use or discuss statistical techniques but present information on areas of transportation research that are of interest to JTS readers. Papers that do not use or discuss statistical techniques may be considered for the Research Notes section.

Manuscripts must be double spaced, including quotations, abstract, reference section, and any notes. All figures and tables should appear at the end of the manuscript with each one on a separate page. Do not embed them in your manuscript.

Because the JTS audience works in diverse fields, **please define** terms that are specific to your area of expertise.

Electronic submissions via email to the Managing Editor are strongly encouraged. We can accept PDF, Word, Excel, and Adobe Illustrator files. If you cannot send your submission via email, you may send a disk or CD by overnight delivery service or send a hardcopy by the U.S. Postal Service (regular mail; see below). Do not send disks or CDs through regular mail.

Hardcopy submissions delivered to BTS by the U.S. Postal service are irradiated. Do not include a disk in your envelope; the high heat will damage it.

The cover page of your manuscript must include the title, author name(s) and affiliations, and the telephone number and surface and email addresses of all authors.

Put the **Abstract** on the second page. It should be about 100 words and briefly describe the contents of the paper including the mode or modes of transportation, the research method, and the key results and/or conclusions. Please include a list of keywords to describe your article.

Graphic elements (figures and tables) must be called out in the text. Graphic elements must be in black ink. We will accept graphics in color only in rare circumstances.

References follow the style outlined in the *Chicago Manual of Style*. All non-original material must be sourced.

International papers are encouraged, but please be sure to have your paper edited by someone whose first language is English and who knows your research area.

Accepted papers must be submitted electronically in addition to a hard copy (see above for information on electronic submissions). Make sure the hard copy corresponds to the electronic version.

Page proofs: As the publication date nears, authors will be required to return article page proofs to the Managing Editor within 48 hours of receipt.

Acceptable software for text and equations is limited to Word and LaTeX. Data behind all figures, maps, and charts must be provided in Excel, Word, or DeltaGraph. Data in American Standard Code for Information Interchange (ASCII) text will be accepted but is less desirable. Acceptable software for graphic elements is limited to Excel, DeltaGraph, or Adobe Illustrator. If other software is used, the file supplied must have an .eps or .pdf extension. We do not accept PowerPoint.

Maps are accepted in a variety of Geographic Information System (GIS) programs. Please save your GIS files using .eps or .pdf extensions. If this is not possible, maps may be provided in ArcView3.x project files (.apr), ArcGIS8.x map files (.mxd), or Arc/Info metafiles (.map). Send your files using the .apr or .mxd extensions on a CD-ROM via overnight delivery service.

Send all submission materials to:
Marsha Fenn
Managing Editor, *Journal of Transportation and Statistics*
Bureau of Transportation Statistics
U.S. Department of Transportation
400 7th Street, SW
Room 7412, K-15
Washington, DC 20590
Email: marsha.fenn@bts.gov

BUREAU OF TRANSPORTATION STATISTICS UNITED STATES DEPARTMENT OF TRANSPORTATION

JOURNAL OF TRANSPORTATION AND STATISTICS

Volume 7 Number 1, 2004
ISSN 1094-8848

CONTENTS