Volume 6 Numbers 2/3, 2003
ISSN 1094-8848

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 6 Numbers 2/3, 2003
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
or a BTS Product Catalog:

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 Semi-truck on a Canadian highway
© 2004/Getty Images

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 6 Numbers 2/3
2003

Contents

Monthly Output Index for the U.S. Transportation Sector

KAJAL LAHIRI
University at Albany, SUNY

HERMAN STEKLER
George Washington University

WENXIONG YAO
University at Albany, SUNY

PEG YOUNG
Bureau of Transportation Statistics
U.S. Department of Transportation

ABSTRACT

In this paper, we develop a monthly output index for the U.S. transportation sector from January 1980 through April 2002, covering air, rail, water, truck, transit, and pipeline activities. Separate indexes for freight and passenger are also constructed. Our total transportation output index matches very well with the annual transportation output figures produced by the Bureau of Labor Statistics and the Bureau of Economic Analysis. The strong cyclical movements of transportation output appear to be more synchronized with the growth slowdowns in the U.S. economy than full-fledged recessions. Our index led the turning points of the six National Bureau of Economic Research-defined growth cycles over the period with an average lead time of six months at peaks and five months at troughs.

INTRODUCTION

In this paper, we develop an index of monthly economic activity for the transportation sector of the U.S. economy. In contemporary business cycle analysis, output is one of the four coincident economic indicators of the overall economy. Output refers to the physical quantity of items produced, as distinct from the sales value, which combines

KEYWORDS: transportation output, Fisher Ideal Index, business cycles, growth cycles, freight transportation.

quantity and price. In our context, transportation output measures freight movement and passenger travel by different transportation modes. Prior to our work, there was no unique indicator to measure the output of the transportation sector on a monthly basis. The Bureau of Economic Analysis (BEA) and the Bureau of Labor Statistics (BLS) of the U.S. government produce output measures for the transportation sector on an annual basis. The Federal Reserve Board does not produce an index of production for the manufacturing sector but does for service industries.

Even though there has been considerable development of National Bureau of Economic Research (NBER)-type indicator analysis for the whole economy, little work has been done in developing sectoral indicators. While Layton and Moore (1989) have developed leading indicators for the service sector, no monthly indexes of output for particular service industries exist.

In order to construct a monthly index of output for the transportation sector, it is first necessary to determine the constituent parts of the industry. We do that in the next section. Then we discuss the output data available for each of these components of the transportation sector. We also explore possible uses of the output index in business and growth cycle analysis. The newly developed output index is then compared against the annual transportation output figures produced by BEA and BLS.

COMPONENTS OF THE TRANSPORTATION SECTOR

We base our definition of the industry on the North American Industrial Classification System (NAICS). This definition also conforms to the Transportation Satellite Accounts (TSAs) associated with the National Income and Product Accounts (NIPA).

Although transportation activities generally include Household Production of Transportation Services (HPTS) in owner-operated automobiles and in-house as well as for-hire transportation by commercial establishments, in this study we consider only for-hire commercial activities for lack of available monthly data on the other two components. Official data on transportation services, defined in either the Standard Industrial Classification codes or NAICS, are confined to establishments that provide passenger and/or freight transportation services for a fee; neither in-house transportation nor HPTS are counted.[1] Although market activities by NAICS-defined establishments do not cover all transportation activities, for-hire is nevertheless the most informative component of the transportation sector.

For-hire transportation includes six subsectors: air, rail, water, truck, transit and ground passenger transportation, and pipeline. Even though these subsectors are representative of economic activity in the transportation industry and are closely associated with the sectors in the satellite NIPA, a problem must be noted. These series do not include all of the subsectors in the for-hire portion of the transportation sector. The subsectors included in NAICS but excluded here are: scenic and sightseeing transportation, support activities for transportation, postal service, and couriers and messengers. The industries included correspond to NAICS codes 481 to 486 and cover 89.7% to 93.9% of total transportation between 1980 and 2000, according to the "Gross Product by Industry" table in the November 2001 issue of the *Survey of Current Business*.

A useful monthly index of economic activity in the transportation sector can be derived from the available series, because the subsectors they represent constitute a significant portion of the entire industry. Moreover, the transportation subsectors that we used to construct the index of transportation output account for a substantial portion of U.S. gross domestic product (GDP). The aggregate value of for-hire transportation accounted for 3.1% and 3.0% of GDP in 1992 and 1996, respectively[2] (Fang et al. 1998 and 2000). Given the critical role that transportation plays in facilitating economic activity between sectors and across regions, an index of its output can be an important indicator

[1] Han and Fang (2000) and Chen et al. (2003) have shown the importance of in-house and household components, respectively, but their estimates are currently annual. Arguably, these two components should be included as part of transportation output when their monthly measures are developed.

[2] These numbers and other measures of the importance of transportation were derived from the value added of the industry. Using different concepts of the scope of the transportation industry would yield different measures of its importance, varying anywhere from 3.09% (transportation GDP) to 16.50% (transportation-driven GDP). See Han and Fang (2000).

for either the current or future level of general economic activity (see Ghosh and Wolf 1997).

DATA

The total Transportation Output Index was developed from eight series. Five of these series measure the level of freight activity, and the remaining three measure the level of passenger services. The series used to measure the freight component were trucking tonnage, air revenue ton-miles, rail revenue ton-miles,[3] a waterway tonnage indicator, and pipeline movements of petroleum products and natural gas. Similarly, the passenger output index was constructed from three series: air revenue passenger-miles, rail revenue passenger-miles,[4] and national transit ridership.[5] The sources and characteristics of all of these series are provided in appendix 1 (pages 16–23).

With the exception of pipeline, all data were available from January 1980 to April 2002. The pipeline data were available starting in January 1985 going to April 2002. The series that we used to measure pipeline transportation is constructed from data on movements of crude oil and petroleum products, consumption of natural gas, and the field production in Alaska.

Crude oil and petroleum products are moved between different Petroleum Administration for Defense Districts (PADDs), while natural gas is delivered to final users. The Alaska field production of crude oil and petroleum products is added, because it almost never enters the PADD system.[6] This addition accounts for the movement within Alaska along the Trans-Alaska Pipeline from the North Slope to the port of Valdez. However, movements of crude oil and petroleum and natural gas are measured in different units. The first is measured in millions of barrels per day while natural gas is measured in cubic feet. It is possible to combine them by converting both to tons (or Btu) with conversion factors.[7] Then the converted tonnage of petroleum and natural gas are added together as the measure of total movement by pipelines. Just as with the other series, these figures are converted into index number form with 1996 equal to 100.

In constructing the index, the weights were adjusted for the years in which the pipeline data were not available. Each series was then seasonally adjusted using the Census X-11 program.[8] We used the econometric software *EViews* (version 3.1) for this purpose. Because all of these series measure real quantities, no price deflation was required.

INDEX CONSTRUCTION

Weights for the Components Series

The total output of the transportation industry is an aggregate of real output generated by each of the components, and thus data from the eight series were used to construct the Transportation Output Index. Each series, representing the output quantity of a transportation subsector, was converted into index number form with 1996 equal to 100.

In order to construct the Transportation Output Index, I_m^A (A denotes "aggregate" and m denotes the month), for the entire transportation sector, the subsector indexes were combined by assigning weights to each of the components. The weights measure the relative importance of each subsector to the entire sector. They can also be interpreted as the "price" of services provided by different modes in quantity indexes.

While there are several different ways of measuring the relative importance of each subsector, we used value-added weights from the NIPA. Here, the value-added weights are more appropriate than

[3] The monthly rail revenue ton-miles data were obtained by interpolating the quarterly figures. We are now working on weekly railroad data on carloads and intermodal traffic to construct a monthly series. These figures will be used to update the index.

[4] Due to a change in data-collection procedure, rail revenue passenger-mile (RPM) values from January 1980 to December 1985 were unusable. The RPM values for these months were backcasted based on a regression of rail RPM on rail Revenue Passengers (RP), $Rail\_RPM = -27991243.120 + 51725.329*Rail\_RP - 0.485*Rail\_RP^2$, estimated over January 1986 to April 2002. Adjusted $R^2 = 0.562$.

[5] The transit data are monthly but are available only on a quarterly basis.

[6] Alaskan petroleum used to be mostly consumed within Alaska or other PADD regions due to an export ban. This ban was lifted in the early 1990s, and now most of it is exported to Japan.

[7] The conversion factors were obtained from the U.S. Department of Energy (DOE) and they are presented in appendix 1. DOE has two types of conversion factors, one based on Btu and one based on mass; both yield similar estimates.

[8] The X-11 program was originally developed by Shiskin et al. (1967).

gross output, because transportation is an intermediate sector whose economic contribution is calculated as the difference in the values of goods being transported. This definition conforms to the concept of GDP.

Weights were obtained from the annually updated "Gross Product by Industry" table published in the *Survey of Current Business* (November 2001). We disaggregated airline and railroad weights into their respective freight and passenger components by using the ratio of their operating revenues for the particular year.

Figure 1 shows the historic annual weights for each component of the Transportation Output Index. Since 1981, air passenger transportation, which dominates the airline industry, has an increasing weight relative to other subsectors, and railroad freight, which dominates rail transportation, has a decreasing weight. From 1980 to 2000, airline industry and railroad transportation weights changed from 18.8% to 33.0% and 21.5% to 8.1%, respectively.

Trucking maintained the greatest weight among all subsectors throughout the period, always in excess of 40.0%. The weights for rail passenger, air freight, pipelines, water transportation, and public transit were always below 8.0% and changed little over this period. The graph also reflects a less freight-intensive contemporary economy in that the total weight for freight movement relative to total transportation activity has steadily shrunk from 72.3% to 61.1% between 1980 and 2000.

Fisher Ideal Index

Given the weights, component series are aggregated into one single index using different index methods. Economic theory indicates that the preferred measure of quantity change is a geometric mean of the Laspeyres index and the Paasche index. This results in the so-called Fisher Ideal Index. The Fisher Ideal Index is one of the "superlative" aggregate indexes, which means current-weighted, while the other two are fixed-weighted using weights in a single period. The use of fixed-weighted measures for a quantity index, such as those derived from the Laspeyres quantity index, may result in "substitution bias" that overstates output growth for periods after the base year and understates growth for periods before the base year (see Landefeld and Parker (1995) for further explanation).

The tendency of substitution bias reflects the fact that those commodities for which output grows rapidly tend to be those for which prices change less proportionately. Although this bias may be small enough to be safely ignored for shorter sample periods, the output measures derived from a fixed-weighted index can become increasingly subject to "weighting effects" as the time between the weighting period and the current period lengthens. A similar but opposite problem occurs with the other type of fixed-weighted index, the Paasche quantity index, which uses current period prices as weights.

The Fisher Ideal Index, which is a chain index, registers changes that fall between those of the Laspeyres and the Paasche indexes. Because of its many advantages, BEA has used this new methodology since 1996 to publish the NIPA (Landefeld and Parker 1995). The Board of Governors of the Federal Reserve Board (FRB) has also adopted the Fisher Ideal formula in constructing the Industrial Production Index since the mid-1990s (Corrado et al. 1997). Conceptually, our transportation output measure is very similar to FRB's Industrial Production Index in the sense that both measure the physical production of a sector.

The new formula for the growth of monthly transportation indexes is given by

$$\frac{I_m^A}{I_{m-1}^A} = \sqrt{\frac{\sum_j I_{jm} P_{jy(m-6)}}{\sum_j I_{jm-1} P_{jy(m-6)}} * \frac{\sum_j I_{jm} P_{jy(m+6)}}{\sum_j I_{jm-1} P_{jy(m+6)}}} \quad (1)$$

where

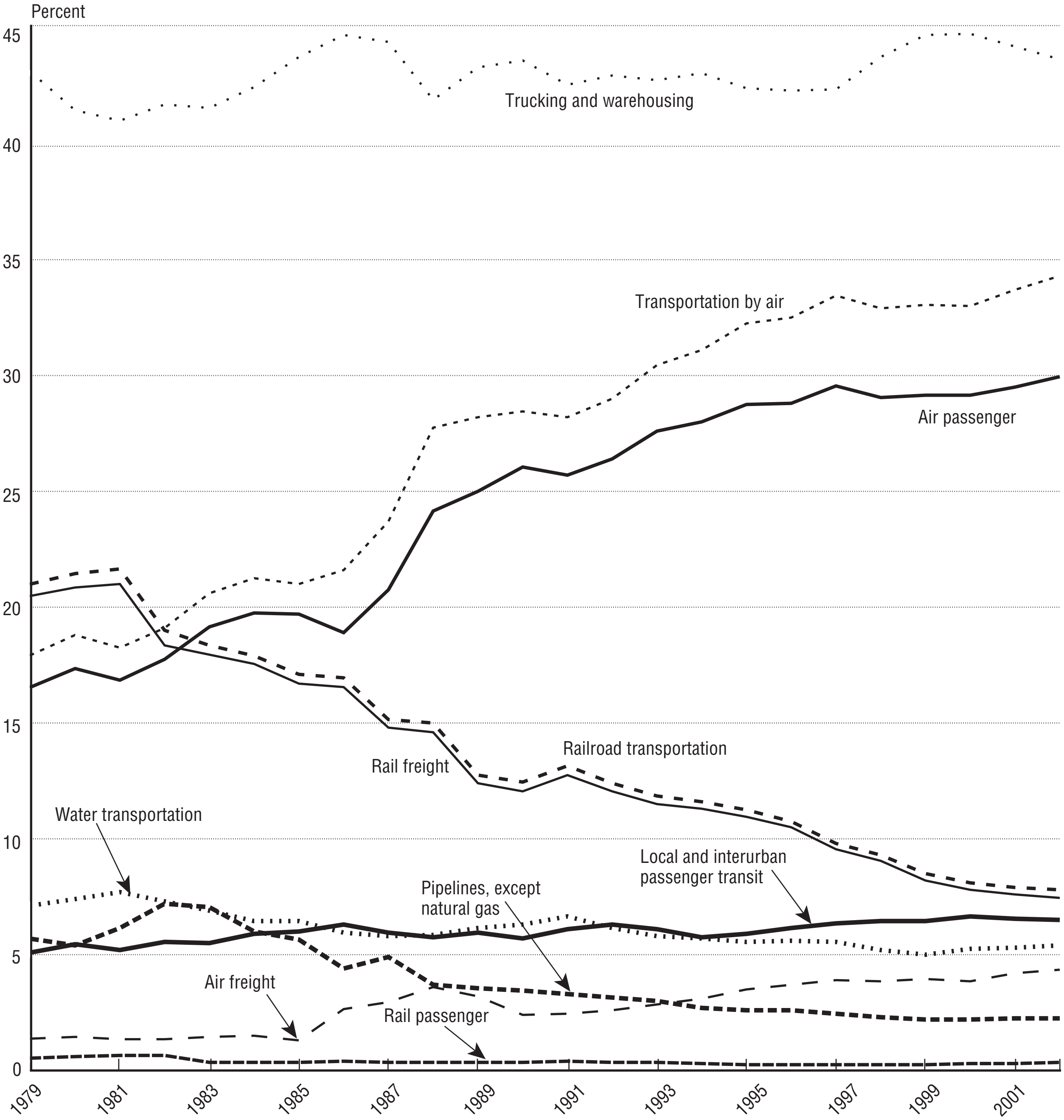
I_{jm} is the output index in subsector *j* in month *m;*

$P_{jy(m)}$ is the value-added weight for subsector *j* in year *y;*

y(m) is the year containing the month *m*.

The Transportation Output Index (Fisher Ideal) uses annual outputs weighted by previous, current, and next year prices. To compute the output quantity index as a chain-typed annually-weighted Fisher Index, we required the unit value added for both the current and the next year. While the "Gross Product by Industries" table is usually published in the November issue of the *Survey of Current Business*, the estimates for recent periods were obtained in

FIGURE 1 Annual Weights for the Aggregation of Transportation



Source: "Gross Product by Industry" table, *Survey of Current Business,* November 2001.

two steps. First, the industry producer price index (PPI) for each subsector of transportation (for transit, we used the consumer price index for intracity transportation, because PPI is not available for this subsector) that BLS produces on a monthly basis were extrapolated to obtain the annual averages for the current year (2002) and the next year (2003). Second, the unit value-added measures were extrapolated based on these annual averages of industry PPI. The Transportation Output Index, as well as its freight and passenger component subtotals, is computed as the cumulative product of a monthly series of these growth estimates from January 1980 onward. For $I_0^A = 100$ in the base year,

$$I_m^A = \frac{I_m^A}{I_{m-1}^A} \times \frac{I_{m-1}^A}{I_{m-2}^A} \times \cdots \times \frac{I_1^A}{I_0^A} \times 100 \qquad (2)$$

Figure 2 compares the Fisher Ideal Index of total transportation output with its alternative index

FIGURE 2 Total Transportation Index: Linked Laspeyes vs. Fisher Ideal

computed from the linked Laspeyres.[9] They are found to be almost identical. Any difference would arise from the weights used. As seen earlier in figure 1, the weight on the largest component, trucking, has been pretty stable in the sample period, which limits any potential substitution bias. FRB also found a similar result when they recomputed their Industrial Production Index using the Fisher Ideal Index[10] (Corrado et al. 1997). However, because of its potential advantages, the transportation indexes derived from the Fisher Ideal Index were used for our analysis in this paper.

THE CHARACTERISTICS OF THE INDEX

Classic Business Cycles

The monthly values of the resulting indexes for January 1980–April 2002 are tabulated in appendix 2 (pages 24–27). The Total Transportation Output Index, the Freight Transportation Output Index, and the Passenger Transportation Output Index are presented in figures 3a to 3c. Dark shaded areas represent the NBER-defined recessions in the U.S. economy and lightly shaded areas represent the

TABLE 1 Final Weight for Transportation Indexes (Linked Laspeyres)

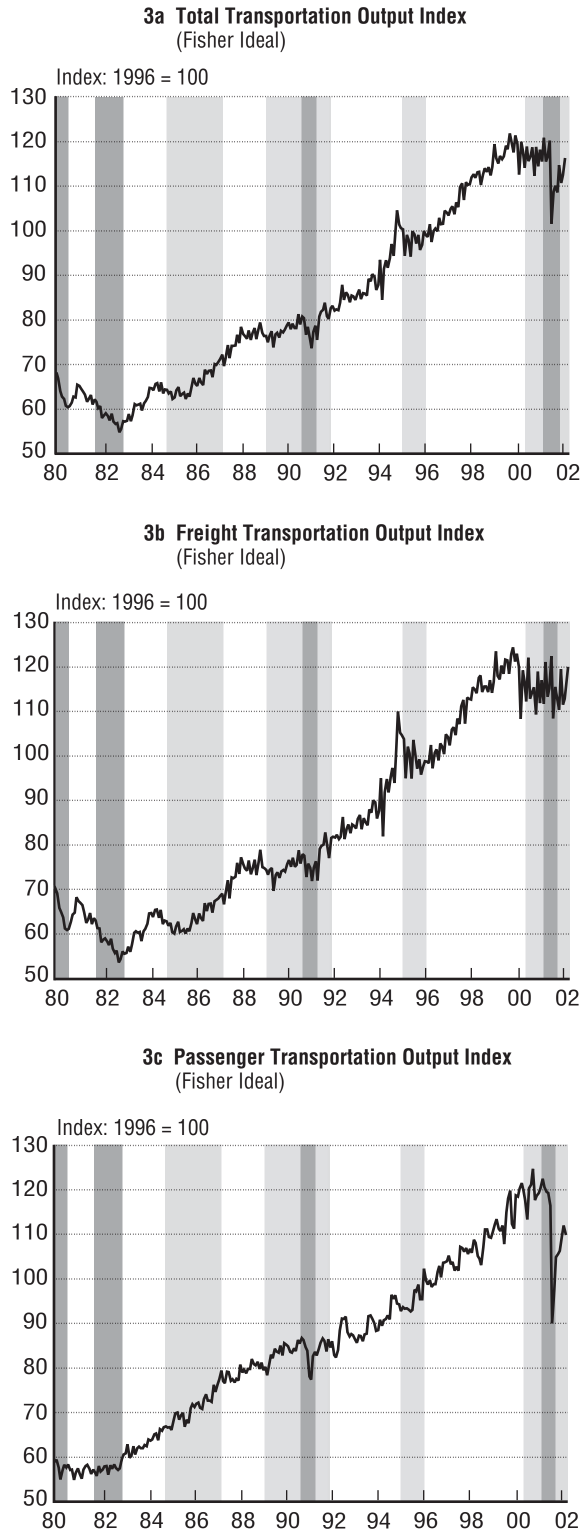
Subsector of transportation	1996 Transportation Satellite Accounts (adjusted)	
Rail	17.3%	
Passenger		0.8%
Freight		16.5%
Truck	42.2%	
Water	4.7%	
Air	24.7%	
Passenger		21.3%
Freight		3.4%
Pipeline	9.7%	
Transit	1.4%	
Total	**100.0%**	

Source: Adapted from Fang et al. (2000).

[9] The standard formula for the linked Laspeyres quantity index is $I_M^A = \Sigma I_m \cdot p_0 / \Sigma I_0 \cdot p_0$ where p_0 is the price in the base period. (Note that we set $I_0 = 100$.) It shows changes in physical movements in the transportation sector with prices held fixed at base year values, which is 1996 here (Corrado et al. 1997). Because the public transit subsector is often supported by public subsidies, its value-added figures are sometimes negative. As a result, we had to calculate the weight assigned to this sector as the average of the ratio of its output to the total transportation industry output for 1996. For airlines and railroads, we determined the relative amount of operating revenue obtained from transporting passengers and freight to disaggregate the weight into passenger and freight. The weights for the Laspeyres index are obtained from the 1996 TSA (Fang et al. 2000) and presented in table 1.

[10] We thank Professor Ariel Pakes of Harvard University for an illuminating discussion on this finding.

FIGURE 3 Three Transportation Output Indexes: Seasonally Adjusted



Note: Dark shaded areas represent the NBER-defined recessions in the U.S. economy; lightly shaded areas represent the NBER-defined growth cycle recessions in the economy (the trough for the latest growth slowdown has not been determined).

NBER-defined growth cycle recessions. These indexes are based on the seasonally adjusted component series that are individually graphed in appendix 1.

Certain characteristics of these indexes should be noted. First, all of them show strong upward trends, with the Total Transportation Output Index showing a compounded annual growth rate of 2.65% between January 1980 and August 2001. Both the passenger and freight indexes also grew over this period, with rates of 3.19% and 2.56%, respectively. (We compared the growth rates through August 2001, because the terrorist attacks of September 11, 2001, drastically affected the passenger component of the transportation sector.) The indexes also show declines in their values, reflecting the economic recessions of July 1981–December 1982, July 1990–March 1991, and March 2001–November 2001. Sharp downward movement also occurred in both the freight and passenger indexes after September 11 and was most pronounced in the passenger index. Overall, the cyclical movement of the freight index dominates that in the Total Transportation Output Index.

The peak (trough) occurs when the Transportation Output Index reaches the highest (lowest) point of its cyclical fluctuations, which would exclude from consideration some temporary positive (negative) irregular disturbances. We followed the NBER dating algorithm described in Bry and Boschan (1971, chapter 2) to identify each of the peaks and troughs. The algorithm uses a series of rules to distinguish the real peaks and troughs from spurious ones. For instance, a movement from a peak to a trough (phase) cannot be shorter than 6 months and a complete cycle must be at least 15 months long. Using these criteria, the cyclical turning points of the Total Transportation Output Index together with the NBER business and growth cycle chronologies are reported in table 2.

Table 2 shows that cyclical peaks in the Transportation Output Index occurred prior to the economic recessions of July 1981–December 1982, July 1990–March 1991, and March 2001–November 2001. In the case of the July 1990–March 1991 recession, we defined the peak in the index to have occurred in February 1988, nearly 29 months prior to the beginning of the economic recession. After

TABLE 2 Lead and Lag Analysis Between Transportation and the Economy

NBER-defined chronologies of economy[1]				Business cycle of Transportation Output Index					
						Lead and lag of transportation vs.			
Recessions		Growth cycle		Chronology		Recessions of economy		Growth cycle of economy	
P	T	P	T	P	T	P	T	P	T
–	Jul–80	–	Jul–80	–	Jul–80	–	0	–	0
Jul–81	Nov–82	Jul–81	Dec–82	Feb–81	Oct–82	–5	–1	–5	–2
–	–	Sep–84	Jan–87	Aug–84	Sep–85	–	–	–1	–16
Jul–90	Mar–91	Jan–89	Dec–91	Feb–88	Mar–91	–29	0	–11	–9
–	–	Jan–95	Jan–96	Dec–94	Jul–95	–	–	–1	–6
Mar–01	Nov–01	Jun–00	–	Nov–99	Sep–01	–16	–2	–7	–
					Mean	–17	–1	–5	–7
					Median	–16	–0.5	–5	–6

[1] Business cycle chronologies are taken from http://www.nber.org/; growth cycle chronologies are taken from Zarnowitz and Ozyildirim (2002).
Key: P = peak; T = trough.

February 1988, index growth stagnated, but surged in December 1988, followed by a period of steady decline. Following the Bry-Boschan censoring rule of identifying real peaks, we regard December 1988 as a temporary disturbance. The transportation sector started to recover in July 1989, but its growth was interrupted in August 1990, which is one month after the beginning of the economic recession. The Index started to move up at about the same time as the economic recovery after March 1991.

The Total Transportation Output Index clearly peaked 16 months prior to the beginning of the latest recession. It appears that the Index started to move up in June 2001, but the events of September 11 have distorted the data. September 2001 also marks the lowest point in aggregate transportation activity since its last peak in November 1999 and is roughly coincident with the recently announced trough of November 2001 for the latest economic recession. The Index has been recovering since then, albeit with interruptions.

Overall, the Transportation Output Index led the three peaks with a considerable lead time (median 16 months);[11] the signals for recovery were almost contemporaneous. The index would have given two false signals for economic recession in August 1984 and December 1994. However, they were not false in the sense that these peaks were followed by recessions in the growth cycle. Hence, the strong cyclical changes in transportation output appear to be more synchronized with growth slowdowns rather than full-fledged recessions of the U.S. economy. This also suggests that the cyclical movement in these indexes foreshadows the growth cycles of the economy more consistently than the business cycles. Thus, the newly constructed Transportation Output Index can be very useful in monitoring the fluctuations in general economic activity from the perspective of transportation.

When we look at the freight and passenger transportation indexes separately in figures 3b and 3c, we find that the cyclical movements in the Total Transportation Output Index are mostly determined by freight movement. The freight index reached its peak and trough during the same months as the total index during the July 1981–November 1982 recession. The passenger index, on the other hand, did not have the corresponding cyclical movement during this period. Freight activities dominated the transportation sector in the early 1980s.

During the economic recession of July 1990–March 1991, the freight index peak occurred two months before that of the total index, while the pas-

[11] Between 1953 and 1982, the average lead time of the composite index of 11 leading indicators relative to the NBER-defined reference cycles is 9.7 months at peaks and 4.6 months at troughs (see table 11.4 in Zarnowitz 1992).

senger index started to decline in September 1990, which is one month after the peak of the economy. A similar phenomenon occurred during the latest recession. The freight index peak occurred at about the same time as the total index, but with a much deeper amplitude. The passenger index reached its peak 12 months later. Furthermore, September 11 had a more profound impact on passenger transportation than on freight transportation. As a result, the total index mimics the movement in the passenger index more closely during this recessionary episode than on previous occasions.

The sequence of peaks and troughs in these indexes and their relationship to business cycles in the economy may reflect some interesting underlying linkages. Freight movement adjusts early to the demand or supply shocks in the economy; these adjustments or fluctuations across different sectors can eventually lead to a full-fledged recession or be limited to sectoral cycles. On the other hand, passenger transportation activities are affected when the state of the overall economy has changed due to demand shocks, especially in a recession. The last two recessions seem to follow this stylized scenario. Because every recession is caused by a mixture of different demand and supply factors, the relative changes in the passenger and freight indexes may not always follow the above sequence. Overall, turning points in the total index stay between those of its two components, but tend to be closer to those of the freight index.

Growth Cycles

In a growth cycle, the economy undergoes alternating periods of deceleration and acceleration that may not develop into a full-fledged recession (see Zarnowitz 1992, chapters 7 and 8; and Zarnowitz and Ozyildirim 2002). Growth cycles are less well known compared with classic business cycles, and they usually cover both full-fledged business cycles and growth slowdowns. Technically, the growth cycle refers to the cyclical component of a typical time series, which is the deviation of a seasonally adjusted series from its estimated trend. Over our sample period, there were six such episodes in the overall economy, four of which included the recessions of the period. They are all clearly discernable with major downswings in the Total Transportation Output Index in figures 3a to 3c.

Depending on the method of estimation of the trend from a time series, growth cycles could differ. The conventional NBER algorithm to estimate the secular trend and identify the growth cycles is the Phase Average Trend (PAT) method (Boschan and Ebanks 1978). The PAT starts by determining preliminary turning points based on the deviation from a 75-month moving average (first approximation) of a deseasonalized time series. Then, values at the turning points are averaged to obtain phase averages (each phase is defined on two turning points). The three-item moving averages of these phase averages are subsequently computed to obtain the so-called "triplets." The midpoints of the triplets are connected, and the connected level series is further adjusted to match the level of the original series. Then a 12-month moving average (second approximation) of the adjusted series yields the estimated secular trend.[12]

Using the estimated trend, the NBER growth cycles are defined based on the deviation of the deseasonalized series from the PAT. We then compare the growth cycles of the Transportation Output Index obtained using the PAT with the NBER growth cycle chronology. The growth cycles of the

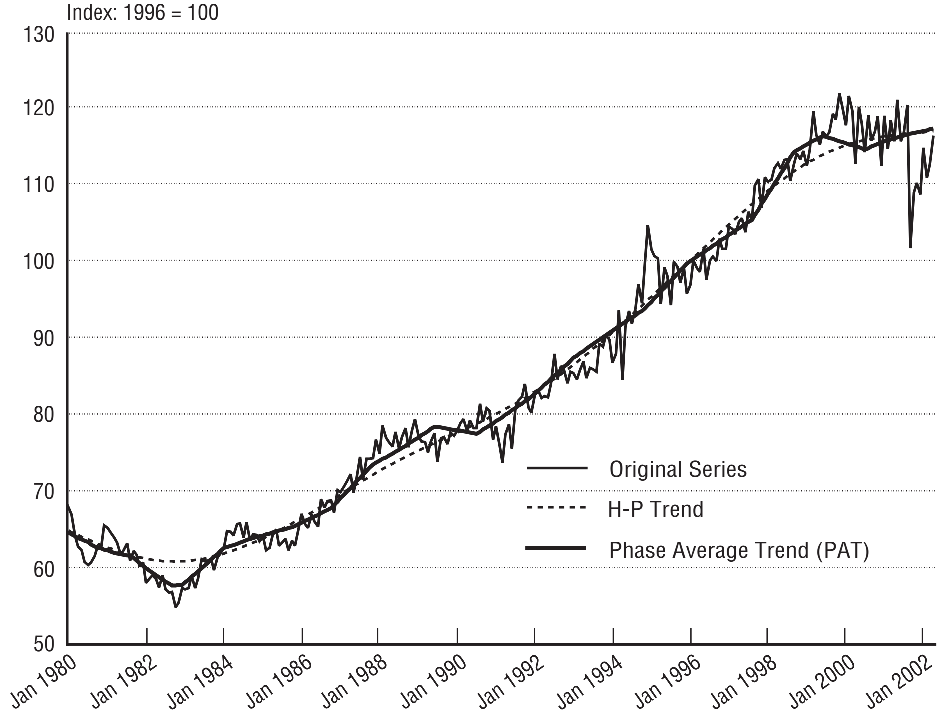
[12] Since the calculation of the PAT can be tedious, a good alternative would be the use of the H-P filter (Hodrick and Prescott 1997). The H-P filter chooses the trend value s_t of the deseasonalized data y_t to minimize

$$\sum_{t=1}^{T}(y_t - s_t)^2 + \lambda \sum_{t=2}^{T-1}((s_{t+1} - s_t) - (s_t - s_{t-1}))^2 .$$

The penalty λ parameter controls the smoothness of the series. The larger the value of λ is, the smoother the trend. Currently, the H-P filter can be implemented using most econometric software (e.g., *EViews*).

Zarnowitz and Ozyildirim (2002) point out that the selection of the trend is inevitably associated with considerable arbitrariness, which has long been a source of confusion in the literature of growth cycles. However, they found that estimated trends are generally similar for the PAT and the H-P filter when the value of λ is around 108,000 for monthly data, and the PAT is superior to its alternatives in the matter of details. Consistent with their finding, with the value of λ = 108,000, the two estimated trends based on the PAT and the H-P filter were very similar, as depicted in figure 4. By its very nature, however, the PAT attributes a somewhat bigger part of the cyclical movements to trend.

FIGURE 4 Trends in the Transportation Output Index



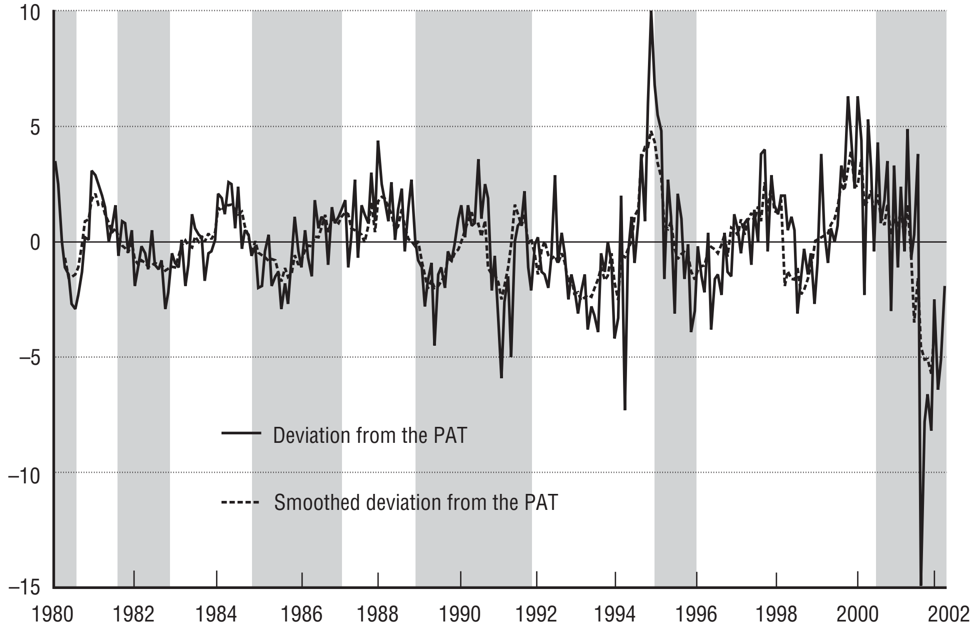
Transportation Output Index together with its smoothed version are compared with the NBER-defined growth cycles for the overall economy in figure 5. The smoothing was done using a filter developed by *Statistics Canada* (Hertzberg and Beckman 1989). We found that the Total Transportation Output Index led the growth cycle consistently with *average* lead times of six months at peaks and five months at troughs. Only for the economic slowdown of January 1995–January 1996 was the Transportation Output Index roughly coincident both at the peak and the trough. Figure 5 also reveals slowdowns in the transportation sector from July 1992–August 1993 (mainly due to a sharp decline in air passenger travel at that time) and October 1997–August 1998 (a short and shallow slowdown compared with others), which were not followed by corresponding slowdowns in the overall economy. Except for these caveats, our Transportation Output Index gave correct signals for all economy-wide slowdowns of the period. A look at the freight and passenger indexes suggests that the classic business and growth cycle characteristics of transportation output are mainly due to the freight component, and the passenger component does not show a consistent lead-lag relationship with reference to the economic cycle.

We should, however, point out that the lead time analysis presented above does not take into account either the lag involved in obtaining the data necessary to construct the series or the necessity of employing a filter rule that by its very nature involves a delay in identifying changes. It is necessary to develop some filter rule (e.g., a three consecutive decline rule for signaling a downturn) that would enable analysts, in real time, to distinguish between the irregular movements and the true signals of cyclical turns.[13] After all, a leading indicator is only as good as the filter rule that interprets its movements. These rules typically involve tradeoffs of accuracy for timeliness and missed signals for false alarms, see Lahiri and Wang (1994). We have so far identified the peaks and troughs of the indexes from an ex post perspective. Further analysis is needed to establish the ex ante predictive ability of the Transportation Output Index. In future research, we plan to develop filter rules that would enable us, in real time, to distinguish between the irregular movements and the true signals of cyclical turns.

---

[13] For a discussion of alternative rules for forecasting the cyclical movements of the Composite Index of Leading Indicators for the economy, see Stekler (1991, pp. 169–181).

FIGURE 5 Growth Cycles in the Transportation Output Index



Note: Shaded areas represent the NBER-defined growth slowdown in the U.S. economy (the trough for the latest growth slowdown has not been determined).

COMPARISON WITH ALTERNATIVE OUTPUT MEASURES

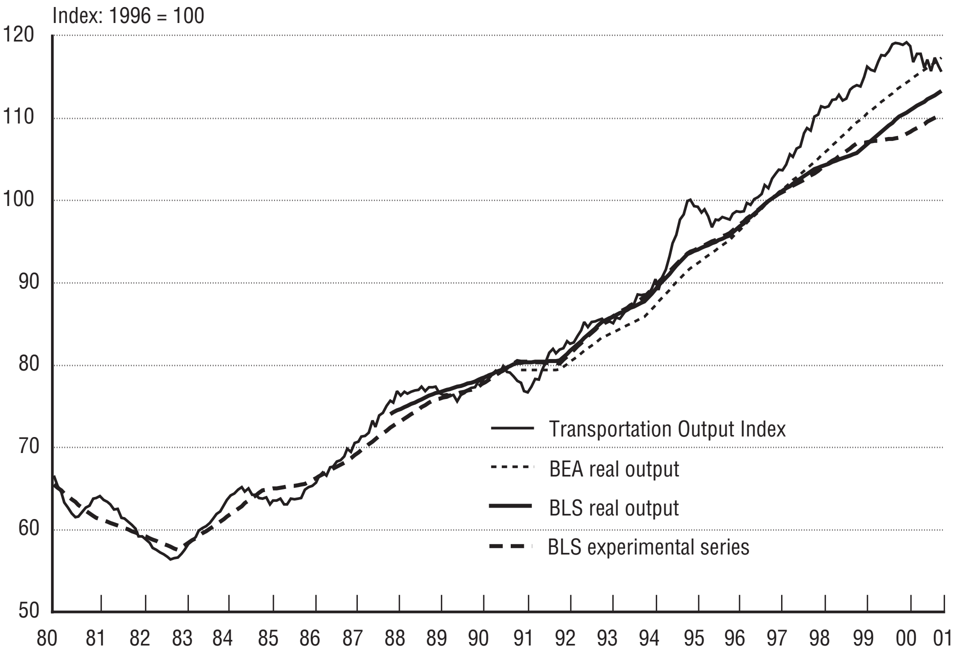
It is also possible to compare our Total Transportation Output Index with annual data produced by BEA and BLS on the gross output of the transportation sector. Gordon (1992) and, more recently, Bosworth (2001) and Yuskavage (2001) have provided valuable insights into the different methodologies and data that BEA and BLS use to construct the output. The Office of Productivity and Technology of BLS maintains an annual series on transportation output that begins at 1987.

Gullickson and Harper (2002) present an analysis using experimental BLS output data based on a multifactor economic growth model that goes back to 1947. Since BEA went through a major overhaul in generating gross output data in the 1980s, and after 1991 it switched to using BLS's Producer Price Index to compute the price deflator, we plotted the BEA series obtained from the *Survey of Current Business (*November 1997) only after 1991.

Figure 6 shows that even though the four transportation output series are derived using widely different approaches, remarkably similar trends are exhibited (values of all series were normalized at 1996 = 100). The average values of the four series are also very similar. The BEA series, which has more comprehensive coverage and is benchmarked to the five-year economic census, followed our Transportation Output Index closely throughout the 1990s, whereas the BLS series seems to have slowed down since 1998. More importantly, it appears that while the three alternative annual output measures reflect the long-term trends, our monthly transportation output measure is superior to them in reflecting cyclical movements in this sector. In figure 6, our Transportation Output Index deviates temporarily from the other three series whenever there are recessions and growth slowdowns in the economy.

Following Gordon (1992) and Bosworth (2001), in table 3, we present alternative estimates of output growth in the transportation sector and in its three major subsectors—trucking, railroads, and airlines—between 1980 and 2000. For this comparison, we did not include the BLS real output series because it is available only after 1987 and it is very similar to the BLS experimental series. The growth rates are also reported separately for 1980–1991 and 1992–2000. In computing these rates, we converted our monthly values to annual figures. For the total output, the growth rates of our index fall between the BEA and BLS rates in all periods. The same is true for trucking except that our index has a

FIGURE 6 Comparison of Monthly Transportation Index with Annual BEA and BLS Outputs



higher growth rate than both BEA and BLS during 1992–2000. For railroads, our index has higher rates of growth than that of BEA and BLS for the overall period and in the 1990s. However, during 1980–1991, the rail growth rate of our index was between the BEA and BLS values. For airlines, our index is almost the same as that of the BLS index, whereas the BEA figures are somewhat higher.

Interestingly, we found that our monthly index has a lot more cyclical variation than the other three series. This is not surprising in view of the fact that the BEA and BLS values are annual and are benchmarked to five-year economic surveys. Given that we constructed the Total Transportation Output Index using monthly data on a series of eight related factors, most of which were not previously used, it is heartening to note the level of agreement in the three series. The advantage of our approach, however, is that the index can be made available on a monthly basis such that the health of the transportation sector can be monitored in real time.

CONCLUSIONS

In this paper, we developed a monthly output index of the U.S. transportation sector for January 1980 through April 2002, covering air, rail, water, truck, transit, and pipeline activities. The included industries cover from 89.7% to 93.9% of the total for-hire transportation GDP during 1980 to 2000. We use both linked Laspeyres and Fisher Ideal Index methods to construct the indexes. These two series were found to be very similar. Separate indexes for freight and passenger transportation were also constructed, and freight was found to be the dominant component in the Total Transportation Output Index. The index closely follows the annual transportation output figures produced by BLS and BEA, even though our monthly index displays more pronounced cyclical movements. Thus, our approach to measuring output in the transportation sector can be useful for measuring productivity in the sector and can be extended to other nonmanufacturing sectors as well.

We also examined the characteristics of the transportation output measure in relation to the classical business and growth cycles of the overall economy. The transportation output cycles are studied using the Phase Average Trend and Hodrick-Prescott filter. The strong cyclical movements in transportation output appear to be more synchronized with the growth slowdowns rather than the full-fledged recessions of the U.S. economy. Based on the cycles generated from the PAT, we found that the index led the NBER-defined growth cycles with an average lead time of six months at peaks and five months at

TABLE 3 Comparisons of Alternative Measures of Output Growth in the Transportation Sector
Compound annual rate

Output measures	1980–2000	1980–1991	1992–2000
Trucking			
BEA real output	4.8%	4.8%	3.9%
BLS experimental real output	2.3%	1.3%	2.8%
Transportation Output Index	3.4%	1.7%	4.5%
Railroads			
BEA real output	1.8%	1.5%	1.7%
BLS experimental real output	1.8%	0.8%	2.6%
Transportation Output Index	2.2%	1.0%	3.3%
Airlines			
BEA real output	5.4%	5.7%	4.6%
BLS experimental real output	5.0%	4.9%	4.4%
Transportation Output Index	5.0%	4.9%	4.4%
Total			
BEA real output	4.2%	4.1%	3.9%
BLS experimental real output	2.3%	1.3%	2.8%
Transportation Output Index	3.0%	1.9%	3.7%

Sources: BEA output data are from U.S. Department of Commerce, Bureau of Economic Analysis, "Gross Output by Detailed Industry," table. See Gullickson and Harper (2002) for the BLS experimental output series.

troughs with almost no false signals. Admittedly, the lead-lag analysis reported here is retrospective. In future research, we would like to develop ex ante filter rules that would enable us, in real time, to distinguish between true cyclical turns and irregular movements of the transportation series. Further analysis is needed to establish the ex ante predictive value of the Transportation Output Index.

While we believe the Total Transportation Output Index yields a valid measure of output in the industry, we recognize there are some data problems and that refinements in the indexes may be necessary to improve it in the future. First, this index only measures output in the services sector of the industry. The activity involved in the production of transportation equipment is not included, nor is the activity involved in the construction of transportation infrastructure.

Second, within the services sector only for-hire transportation is included. The activity involved in intrafirm (in-house) and household transportation (HPTS) has been excluded. To the extent that for-hire and these two transportation activities display different trends, the current index will not yield a precise picture of economic activity in the industry. Han and Fang (2000) estimated that in-house and for-hire components of total transportation activity constituted nearly 1.97% and 3.16%, respectively, of total GDP in 1997. Furthermore, Chen et al. (2003) estimated the magnitude of HPTS to be about 1.9 times that of all for-hire transportation industries between 1991 and 2000. Inclusion of both in-house and HPTS components would increase the contribution of transportation services to the total GDP from 3.16% to 11.0%, if based on TSA 1997 data. In the future, it will be useful to incorporate these two components as part of our Transportation Output Index once monthly data are available. In addition, the index excludes activity in some of the minor for-hire subsectors like scenic and sightseeing, support activities, postal service, and couriers and messengers.

Third, the waterborne component of the index only includes internal waterway traffic. It does not include deep seas, Great Lakes, coastal trade, or cruise travel. Again, if the trends in the excluded items differ from the data included, the results would be imprecise. Monthly data on some of these excluded items are currently being developed by the U.S. Army Corps of Engineers and can be easily integrated in our analysis as soon as they are available.

Finally, monthly data on national transit ridership are available on a quarterly basis and lag by four months. Other monthly data are sometimes available with a lag of one to three months. For the purpose of releasing the output index with a usual lag of one to two months, some of the latest monthly data must be forecasted on a provisional basis using methods discussed in McGuckin et al. (2001). Fortunately, however, the major components of the series (trucking, air, and rail freight) are available quickly, and hence monthly figures for the total transportation sector can be reported soon after release with confidence.

Despite these caveats and suggestions for refining the indexes, as presently constructed they can provide sufficiently accurate estimates of the level of economic activity in the transportation sector.

ACKNOWLEDGMENTS

We thank the following individuals for numerous insightful discussions and help in identifying relevant data sources: Mazhar Ali Awan, Terry Branson, David Chesser, Clyde Crimmel, Bob Costello, Christie Dawson, Stan Ellis, Robert Finkelstein, Robert Ganz, Bill Gullickson, Frank Hardesty, William Jeffers, Paul Kern, Brian Moyer, Sandia Porter, Paul Posey, Jeff Potter, Brian Sloboda, Miranda Stuart, Thea Thomas, Sharon Trench, and Jay Wieriman.

This study was supported by a grant from the Bureau of Transportation Statistics (BTS), U.S. Department of Transportation, and benefited from the comments from BTS staff and many others, including Anirvan Banerji, Bill Bannister, Charlie Han, Rick Kowalewski, Ashish Sen, Jack Wells, and Victor Zarnowitz. This paper also benefited from the insightful comments of three anonymous referees and the Editor-in-Chief of the journal.

The contents of this paper reflect the views of the authors and not necessarily those of the Bureau of Transportation Statistics.

REFERENCES

Boschan, C. and W.W. Ebanks. 1978. The Phase-Average Trend: A New Way of Measuring Economic Growth. *1978 Proceedings of the Business and Economic Statistics Section.* Washington, DC: American Statistical Association.

Bosworth, B. 2001. Output and Productivity in the Transportation Sector: An Overview. The Brookings Institution.

Bry, G. and C. Boschan. 1971. Cyclical Analysis of Time Series: Selected Procedures and Computer Programs, NBER Technical Paper 20. National Bureau of Economic Research.

Chen, D., B. Fang, X. Han, and B. Sloboda. 2003. Household Production of Transportation Services Through Automobiles: An Expansion of the Transportation Satellite Accounts. U.S. Department of Transportation, Bureau of Transportation Statistics, draft.

Corrado, C., C. Gilbert, and R. Raddock. 1997. Industrial Production and Capacity Utilization: Historical Revision and Recent Developments. *Federal Reserve Bulletin.* February.

Fang, B., X. Han, A.M. Lawson, and S.K.S. Lum. 1998. U.S. Transportation Satellite Accounts for 1992. *Survey of Current Business* 78(4), April.

Fang, B., X. Han, S. Okubo, and A.M. Lawson. 2000. U.S. Transportation Satellite Accounts for 1996. *Survey of Current Business* 80(5), May.

Ghosh, A.R. and H.C. Wolf. 1997. Geographical and Sectoral Shocks in the U.S. Business Cycle, NBER Working Paper #6180. National Bureau of Economic Research. September.

Gordon, R.J. 1992. Productivity in the Transportation Sector. *Output Measurement in the Service Sectors.* Edited by Z. Griliches. Chicago, IL: University of Chicago Press. pp. 371–428.

Gullickson, W. and M.J. Harper. 2002. Bias in Aggregate Productivity Trends Revisited. *Monthly Labor Review* 125(3), March.

Han, X. and B. Fang. 2000. Four Measures of Transportation's Economic Importance. *Journal of Transportation and Statistics* 3(1):15–30.

Hertzberg, M.P. and B.A. Beckman. 1989. Business Cycle Indicators: Revised Composite Indexes. *Business Conditions Digest* 29(1), January.

Hodrick, R.J. and E.C. Prescott. 1997. Postwar U.S. Business Cycles: An Empirical Investigation. *Journal of Money, Credit, and Banking* 29(1):1–16.

Lahiri, K. and J. Wang. 1994. Predicting Cyclical Turning Points with Leading Index in a Markov Switching Model. *Journal of Forecasting* 13:245–263.

Landefeld, J.S. and R.P. Parker. 1995. Preview of the Comprehensive Revision of the National Income and Product Accounts: BEA's New Featured Measures of Output and Prices. *Survey of Current Business,* July.

Layton, A. and G.H. Moore. 1989. Leading Indicators for the Service Sector. *Journal of Business and Economic Statistics* 7(3):379–386.

McGuckin, R.H., A. Ozyildirim, and V. Zarnowitz. 2001. Composite Index of Leading Economic Indicators: How to Make It More Timely, NBER Working paper #8430. National Bureau of Economic Research.

Shiskin, J., A.H. Young, and J.C. Musgrave. 1967. The X-11 Variant of the Census Method II Seasonal Adjustment Program, Technical Paper 15. U.S. Department of Commerce, Bureau of the Census.

Stekler, H.O. 1991. Forecasting Cyclical Turning Points. *Leading Economic Indicators: New Approaches and Forecasting Records*. Edited by K. Lahiri and G.H. Moore. Cambridge, England: Cambridge University Press.

Yuskavage, R.E. 2001. Issues in the Measurement of Transportation Output: The Perspective of the BEA Industry Accounts, Brookings Workshop. May.

Zarnowitz, V. 1992. *Business Cycles: Theory, History, Indicators, and Forecasting*. Chicago, IL: University of Chicago Press.

Zarnowitz, V. and A. Ozyildirim. 2002. Time Series Decomposition and Measurement of Business Cycles, Trends and Growth Cycles, NBER Working Paper #8736. National Bureau of Economic Research.

Authors' Addresses: Corresponding Author: Kajal Lahiri, Economics Department, University at Albany, SUNY, Albany, NY 12222. Email: klahiri@albany.edu.

Herman Stekler, Department of Economics, George Washington University, 2201 G Street, NW, Washington, DC 20052. Email: hstekler@gwu.edu.

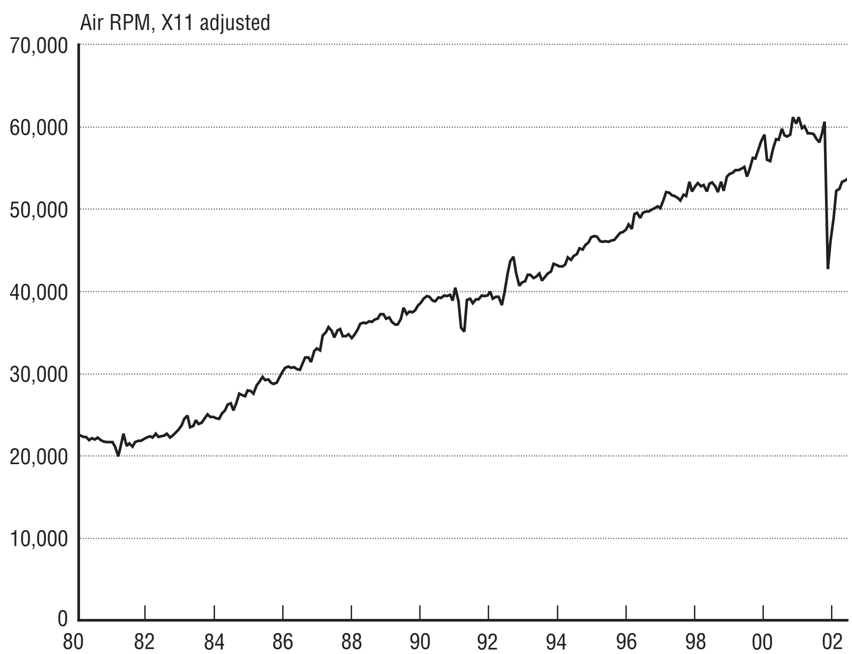
Wenxiong Yao, Economics Department, University at Albany, SUNY, Albany, NY 12222.
Email: wy7098@albany.edu.

Peg Young, Bureau of Transportation Statistics, U.S. Department of Transportation, 400 Seventh St., SW, Room 3430, Washington, DC 20590. Email: peg.young@bts.gov.

Appendix 1
Documentation of the Data Series

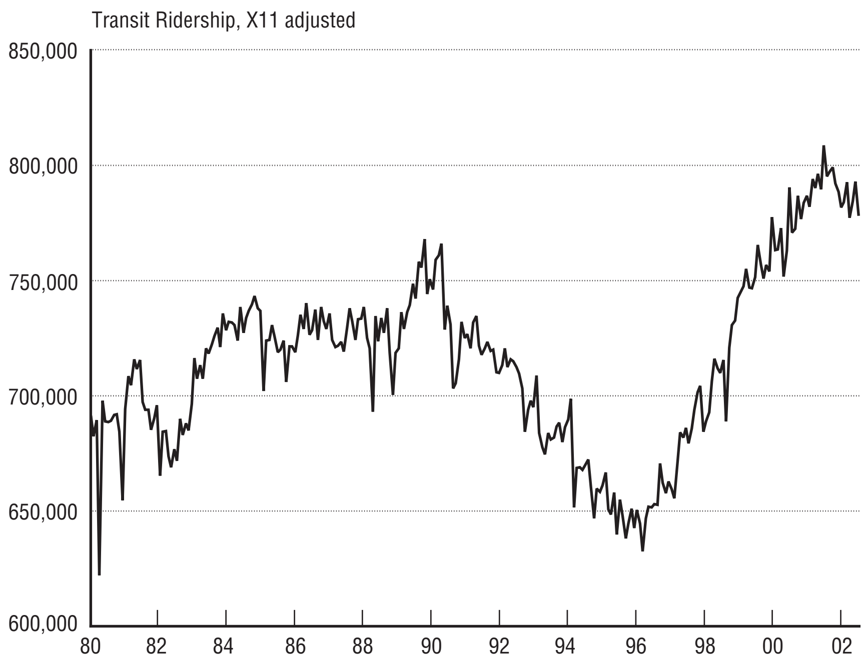
1. Air Revenue Passenger-Miles (RPM)

Name of series	Air Revenue Passenger-Miles (RPM)
Explanation	One revenue passenger transported one mile
Source	U.S. Department of Transportation, Bureau of Transportation Statistics, Office of Airline Information, *Air Carrier Traffic Statistics Monthly*, available at http://www.bts.gov/oai, January 1992
Data format	Preliminary data; seasonally adjusted (in thousands)
Publication date	Available at the end of the month for the 2 previous months
Revisions	The latest 12 months of data are preliminary
Comments	Based on BTS Form 41 filed by large certificated air carriers



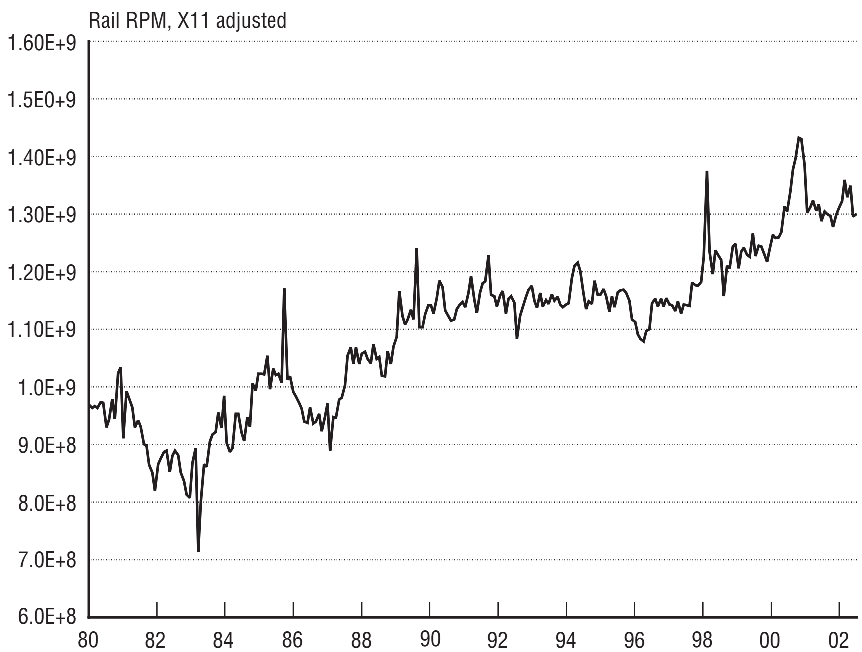
2. National Transit Ridership

Name of series	National Transit Ridership
Explanation	Estimated unlinked passenger trips
Source	American Public Transportation Association (APTA), *APTA Quarterly Transit Ridership Report,* available at http://www.bts.gov since January 1992
Data format	Preliminary data; seasonally adjusted (in thousands of riders)
Publication date	Available in the first day of each quarter for the 2 previous quarters
Revisions	The latest 3 years of data are preliminary
Comments	Includes ridership of commuter rail, heavy rail, light rail, and others



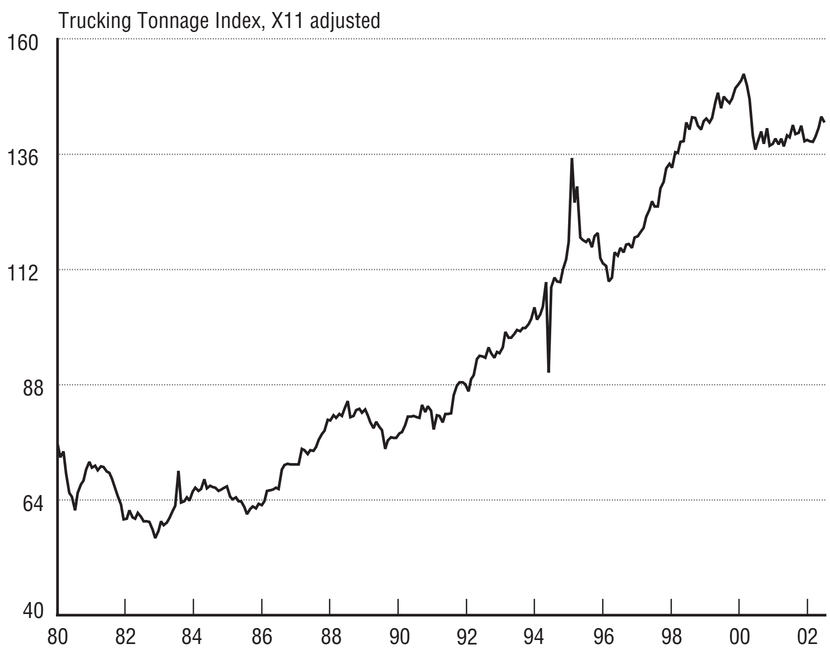
3. Rail Revenue Passenger-Miles (RPM)

Name of series	Rail Revenue Passenger-Miles (RPM)
Explanation	RPMs carried by Amtrak and Alaska Railroads
Source	U.S. Department of Transportation, Federal Railroad Administration (FRA), Office of Safety Analysis, *FRA Accident/Incident Bulletin,* available at http://safetydata.fra.dot.gov/OfficeofSafety/Default.asp)
Data format	Preliminary data; seasonally adjusted (in millions of riders)
Publication date	Beginning of each month for previous 2 months
Revisions	The latest 12 months of data are preliminary
Comments	RPM for January 1980–December 1985 were estimated from data of revenue passengers, because empty trains were counted into RPM before that time



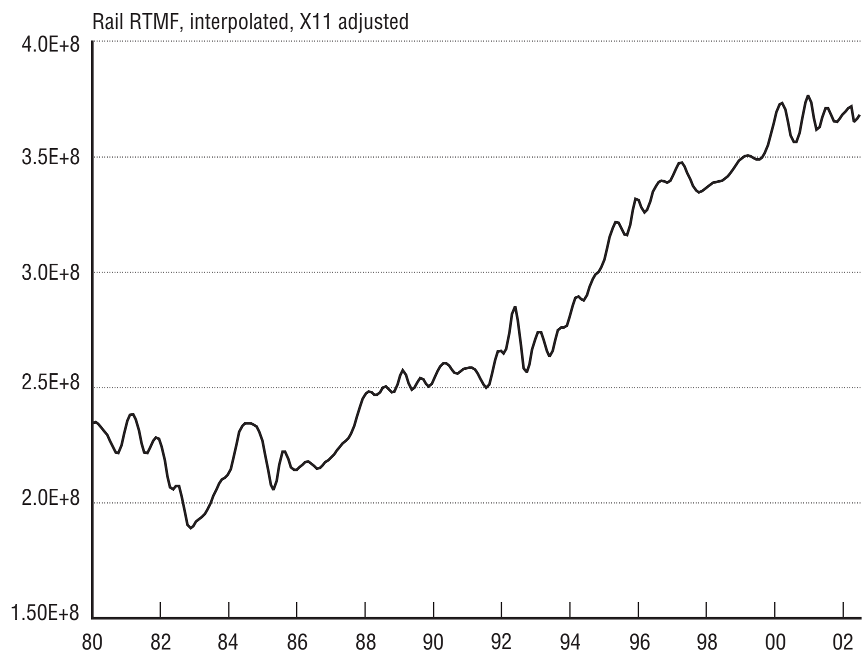
4. Trucking Tonnage Index (TTI)

Name of series	Trucking Tonnage Index (TTI)
Explanation	Truck loads
Source	American Trucking Association (ATA), *Monthly Trucking Report*
Data format	Index: 1996 = 100; monthly, seasonally adjusted and unadjusted
Publication date	3rd of each month for the previous 2 months
Revisions	The latest monthly data are preliminary
Comments	Estimated from tonnage reported by ATA's members in 50 states



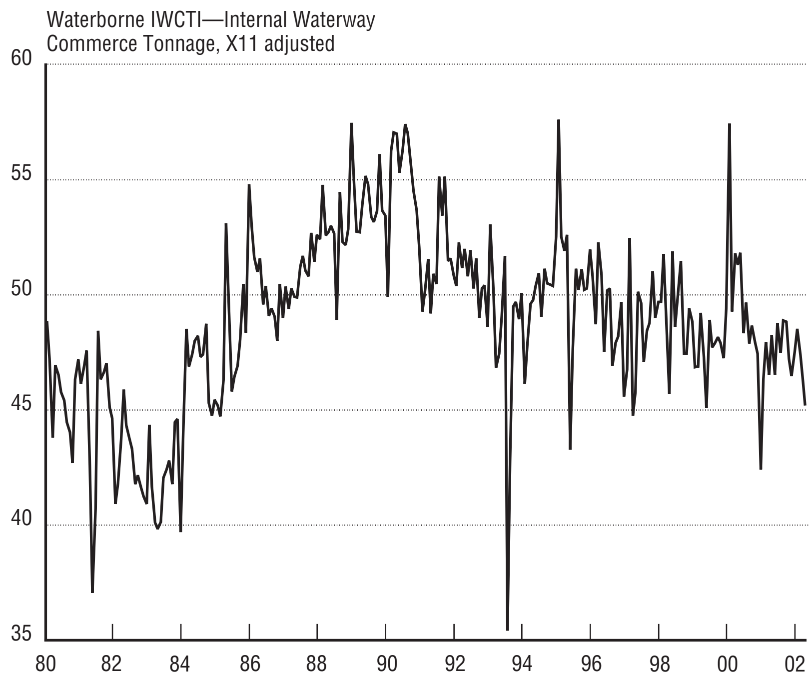
5. Railroads Revenue Ton-Miles of Freight (RTMF)

Name of series	Railroads Revenue Ton-Miles of Freight (RTMF)
Explanation	Carloads of 20 railroads (total containers and trailers) in the United States
Source	Association of American Railroads, *Weekly Railroad Traffic,* available at http://www.bts.gov since the 1st week of 1996
Data format	Preliminary data; quarterly; seasonally adjusted (in billions)
Publication date	Second month of each quarter for the 2 previous quarters
Revisions	The latest 12 months of data are preliminary
Comments	Monthly data were not available. We interpolated from the quarterly data; however, we expect to work with the monthly series soon.



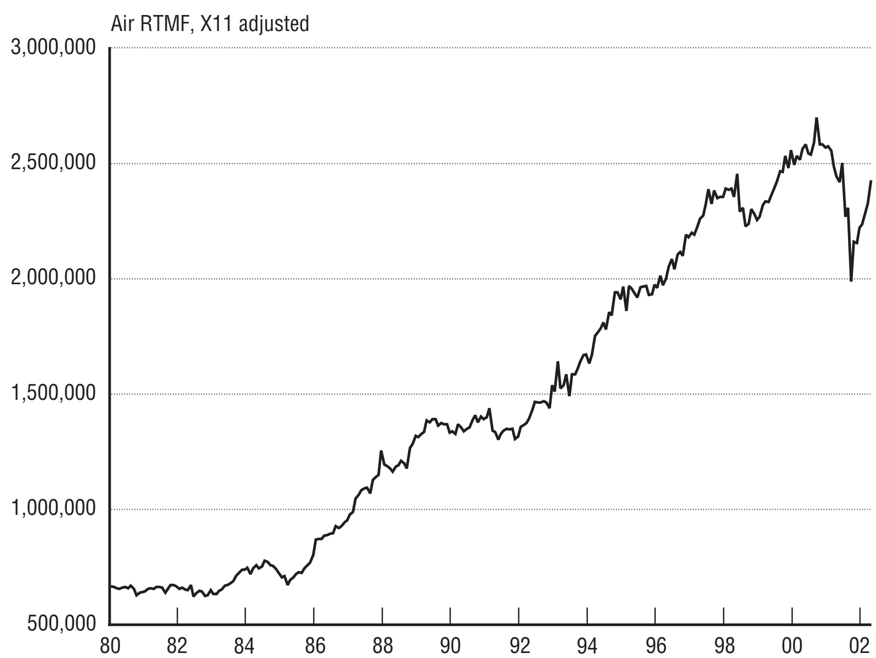
6. Total Internal Commerce Tonnage Indicator (TICTI)

Name of series	Total Internal Commerce Tonnage Indicator (TICTI), all commodities
Explanation	Internal waterway tonnage of coal, petroleum and chemicals, and food and farm products; estimated from 11 key locks on 9 rivers
Source	U.S. Army Corps of Engineers, Waterborne Commerce Statistics Center, available at http://www.iwr.usace.army.mil/ndc/monthlyindicators.htm, since January 1994
Data format	Preliminary data; seasonally adjusted (in millions of short tons)
Publication date	The beginning of each month for the 2 previous months
Revisions	The latest 12 months of data are preliminary
Comments	The data do not include great lakes, coastal and deep-sea waterborne traffic, which are currently not available



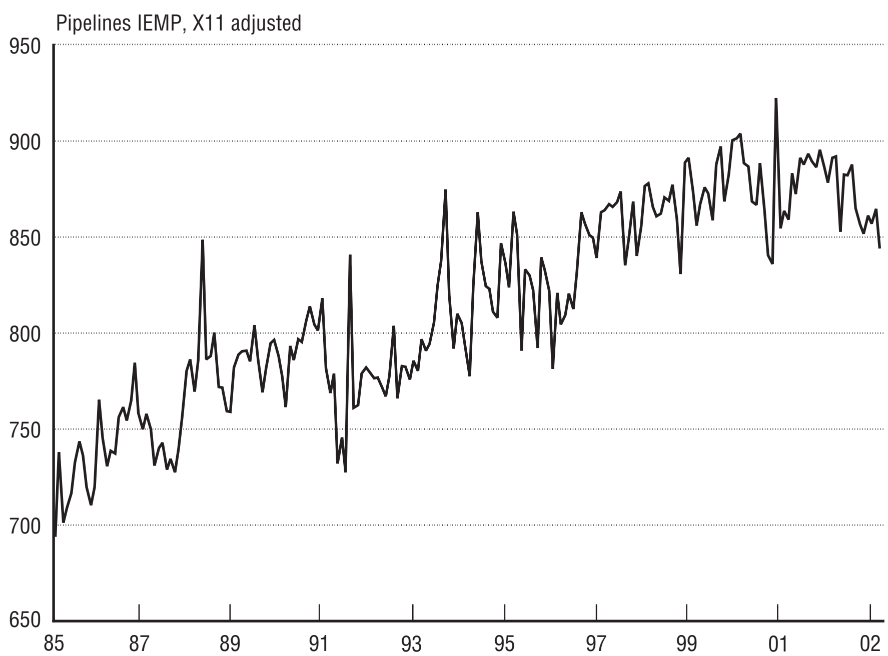
7. Air Revenue Ton-Miles of Freight and Mails (RTMFM)

Name of series	Air Revenue Ton-Miles of Freight and Mail (RTMFM)
Explanation	Ton-miles of freight and express mail transported by the air industry
Source	U.S. Department of Transportation, Bureau of Transportation Statistics, Office of Airline Information, *Air Carrier Traffic Statistics Monthly,* available at http://www.bts.gov/oai since January 1992
Data format	Preliminary data; seasonally adjusted (in thousands)
Publication date	The end of the month for the 2 previous months
Revisions	The latest 12 months of data are preliminary
Comments	Based on BTS Form 41 filed by large certificated air carriers



8. Index of Energy Movements by Pipeline (IEMP)

Name of series	Index of Energy Movements by Pipeline (IEMP)
Explanation	Movements of crude oil and petroleum products between PADDs; Alaska field production and consumption of natural gas
Source	U.S. Department of Energy, Energy Information Administration, *Petroleum Supply Monthly* (for movements of crude oil and petroleum products) and *Monthly Energy Review* (for natural gas and Alaska field production)
Data format	Final data; seasonally adjusted (in millions of tons)
Publication date	23rd–26th of each month for the 2 previous months
Revisions	No revision
Comments	Before January 1985, movements of crude oil between PADDs were not included in the total. In constructing IEMP, crude oil and petroleum products are in million barrels per day and natural gas is in cubic feet and are converted into tons using conversion factors. Conversion factors: 1 cubic foot of natural gas = 1,020 Btu (heat unit); 1 million Btu = 0.025 tons of oil equivalent; 1 barrel of petroleum products = 5.326 millions of Btu (heat unit).



Appendix 2

Monthly Values of Transportation Indexes

Time	Total Transportation Index	Freight Transportation Index	Passenger Transportation Index
Jan-80	68.2	70.9	58.9
Feb-80	66.9	69.1	59.4
Mar-80	64.1	65.8	57.7
Apr-80	62.8	65.0	55.0
May-80	62.3	63.7	57.0
Jun-80	60.8	61.2	58.3
Jul-80	[T]60.4	60.9	57.7
Aug-80	60.7	61.1	58.4
Sep-80	61.5	62.7	57.1
Oct-80	62.8	64.3	57.2
Nov-80	62.6	64.8	54.9
Dec-80	[P]65.5	68.2	56.5
Jan-81	65.2	67.4	57.5
Feb-81	64.6	67.0	56.3
Mar-81	63.9	66.5	55.2
Apr-81	63.3	64.9	57.2
May-81	61.7	62.5	58.0
Jun-81	62.3	63.2	58.3
Jul-81	63.2	64.7	57.4
Aug-81	61.0	62.2	56.2
Sep-81	62.2	63.5	57.2
Oct-81	61.7	63.0	56.6
Nov-81	60.2	61.3	55.8
Dec-81	60.7	61.2	58.0
Jan-82	58.0	58.1	56.8
Feb-82	58.6	58.7	57.2
Mar-82	59.1	59.0	57.9
Apr-82	58.5	58.3	58.0
May-82	57.5	57.5	56.1
Jun-82	59.0	58.9	58.2
Jul-82	57.2	56.7	57.4
Aug-82	56.7	55.7	58.4
Sep-82	56.9	56.2	57.6
Oct-82	[T]54.8	53.6	57.1
Nov-82	55.5	54.3	57.5
Dec-82	57.3	56.1	59.6
Jan-83	57.2	55.5	60.6
Feb-83	57.4	55.7	60.9
Mar-83	58.9	57.1	62.9
Apr-83	57.3	56.0	59.9
May-83	58.6	57.7	60.2
Jun-83	61.1	60.2	62.5
Jul-83	60.9	60.6	60.6
Aug-83	61.0	60.2	62.0
Sep-83	61.2	60.4	62.4
Oct-83	59.7	58.4	62.0
Nov-83	61.3	60.6	62.0
Dec-83	61.8	61.1	62.7
Jan-84	62.7	62.3	62.5
Feb-84	64.8	64.6	64.0
Mar-84	64.7	64.7	63.7
Apr-84	64.2	63.8	64.0
May-84	65.7	65.4	65.1
Jun-84	65.8	65.5	65.3
Jul-84	64.0	63.6	63.8
Aug-84	[P]66.0	65.4	66.3
Sep-84	63.5	62.2	65.9
Oct-84	64.4	63.0	66.7
Nov-84	64.3	62.8	66.9
Dec-84	63.5	61.8	66.6
Jan-85	64.0	62.5	66.7
Feb-85	62.3	60.4	66.1
Mar-85	62.6	60.1	68.2
Apr-85	64.2	61.7	69.8
May-85	65.0	62.7	70.0
Jun-85	62.9	60.4	68.4
Jul-85	63.4	60.7	69.5
Aug-85	63.7	61.0	69.9
Sep-85	[T]62.3	60.1	66.8
Oct-85	63.5	61.2	68.4
Nov-85	62.8	60.6	67.6
Dec-85	65.1	62.5	71.0
Jan-86	67.0	64.7	71.9
Feb-86	65.7	63.2	71.3
Mar-86	65.2	62.2	72.0
Apr-86	67.0	64.6	72.2
May-86	66.0	63.6	71.1
Jun-86	65.4	62.9	70.8
Jul-86	68.9	66.9	73.0

Key: P = peak; T = trough.

(continues)

Monthly Values of Transportation Indexes *(continued)*

Time	Total Transportation Index	Freight Transportation Index	Passenger Transportation Index
Aug-86	67.9	65.1	74.1
Sep-86	68.6	66.7	72.7
Oct-86	68.7	66.8	72.7
Nov-86	67.2	64.8	72.4
Dec-86	70.1	67.4	76.2
Jan-87	69.9	67.6	74.6
Feb-87	70.6	68.0	76.3
Mar-87	71.3	68.5	77.5
Apr-87	72.2	69.0	79.4
May-87	69.7	66.6	77.0
Jun-87	71.5	69.2	76.6
Jul-87	74.4	72.1	79.2
Aug-87	71.4	68.0	79.1
Sep-87	74.2	72.7	77.1
Oct-87	74.2	72.5	77.5
Nov-87	74.3	72.9	76.8
Dec-87	76.7	76.0	77.6
Jan-88	74.8	73.5	77.2
Feb-88	P 78.5	77.2	80.8
Mar-88	76.9	75.7	79.0
Apr-88	76.4	74.7	79.6
May-88	75.8	74.2	78.8
Jun-88	77.7	76.5	79.8
Jul-88	75.5	73.4	79.8
Aug-88	77.0	74.6	82.0
Sep-88	78.2	76.7	80.9
Oct-88	75.7	73.2	80.6
Nov-88	77.9	76.1	81.4
Dec-88	79.4	78.9	79.8
Jan-89	77.1	75.0	81.2
Feb-89	76.4	74.8	79.5
Mar-89	76.4	74.4	80.2
Apr-89	75.0	73.3	78.4
May-89	76.4	74.4	80.4
Jun-89	77.5	74.8	83.0
Jul-89	73.8	69.7	82.2
Aug-89	76.8	73.2	84.3
Sep-89	77.1	73.8	83.7
Oct-89	76.1	72.6	83.4
Nov-89	77.6	74.0	85.0
Dec-89	77.2	74.4	82.8
Jan-90	77.8	74.0	85.6
Feb-90	78.8	75.7	85.2
Mar-90	79.4	76.6	84.9
Apr-90	77.9	75.1	83.6
May-90	79.2	77.1	83.4
Jun-90	78.2	75.1	84.4
Jul-90	78.2	75.5	83.8
Aug-90	81.3	78.8	86.3
Sep-90	79.0	75.8	85.5
Oct-90	80.8	77.9	86.7
Nov-90	80.4	77.6	86.3
Dec-90	76.7	72.8	84.7
Jan-91	78.5	75.8	83.9
Feb-91	75.9	74.8	78.1
Mar-91	T 73.7	71.9	77.4
Apr-91	77.3	74.8	82.6
May-91	78.7	76.3	83.6
Jun-91	75.5	72.0	82.8
Jul-91	80.8	79.1	84.3
Aug-91	81.8	79.8	85.7
Sep-91	82.3	80.1	86.6
Oct-91	83.9	82.8	86.1
Nov-91	80.9	79.5	83.7
Dec-91	80.2	77.0	86.5
Jan-92	82.6	81.6	84.6
Feb-92	83.2	81.8	86.0
Mar-92	82.1	81.7	82.9
Apr-92	82.4	82.3	82.4
May-92	82.2	81.3	84.1
Jun-92	83.9	81.8	88.4
Jul-92	87.9	86.3	91.1
Aug-92	84.5	81.3	91.0
Sep-92	86.2	83.6	91.5
Oct-92	85.7	84.6	88.1
Nov-92	84.0	82.6	87.1
Dec-92	85.5	84.6	87.6
Jan-93	85.3	84.3	87.4
Feb-93	84.5	83.7	86.4
Mar-93	85.7	85.8	85.7
Apr-93	86.8	86.7	87.3
May-93	84.7	83.5	87.2
Jun-93	86.0	86.0	86.0
Jul-93	85.9	85.1	87.5
Aug-93	85.5	84.4	87.7
Sep-93	89.2	88.0	91.6

Key: P = peak; T = trough.

(continues)

Monthly Values of Transportation Indexes *(continued)*

Time	Total Transportation Index	Freight Transportation Index	Passenger Transportation Index
Oct-93	88.7	87.3	91.8
Nov-93	90.3	90.0	91.1
Dec-93	89.7	89.6	90.0
Jan-94	86.7	85.9	88.3
Feb-94	87.9	87.7	88.5
Mar-94	93.5	95.0	90.8
Apr-94	84.5	81.9	89.5
May-94	91.4	91.8	90.7
Jun-94	93.4	94.8	91.0
Jul-94	91.7	91.8	91.8
Aug-94	93.8	95.3	90.9
Sep-94	97.0	97.3	96.4
Oct-94	94.4	93.9	95.5
Nov-94	99.7	101.5	96.3
Dec-94	P 104.6	110.0	94.5
Jan-95	101.4	105.2	94.3
Feb-95	100.6	104.7	92.8
Mar-95	100.3	103.8	93.7
Apr-95	94.4	95.0	93.3
May-95	99.1	102.1	93.4
Jun-95	98.0	100.6	93.1
Jul-95	T 94.2	94.9	92.7
Aug-95	99.9	103.6	93.0
Sep-95	99.2	100.1	97.6
Oct-95	97.1	97.1	97.2
Nov-95	99.1	99.3	98.7
Dec-95	95.7	95.8	95.4
Jan-96	96.9	97.7	95.4
Feb-96	100.0	98.8	102.3
Mar-96	99.0	98.7	99.6
Apr-96	98.6	98.4	98.7
May-96	101.5	102.5	99.7
Jun-96	97.5	97.2	98.2
Jul-96	100.0	100.7	98.8
Aug-96	100.5	101.5	98.8
Sep-96	99.9	99.0	101.6
Oct-96	102.9	102.8	102.9
Nov-96	101.5	102.2	100.3
Dec-96	101.6	100.4	103.6
Jan-97	104.5	104.9	103.7
Feb-97	104.1	104.2	104.0
Mar-97	103.4	102.4	105.2
Apr-97	105.0	105.9	103.3
May-97	105.5	106.5	103.8
Jun-97	103.7	104.7	101.8
Jul-97	106.4	108.4	102.7
Aug-97	105.3	107.2	101.8
Sep-97	109.8	111.2	107.2
Oct-97	110.6	112.6	106.9
Nov-97	106.9	107.3	106.1
Dec-97	110.9	113.1	106.9
Jan-98	110.3	112.9	105.6
Feb-98	110.5	112.7	106.7
Mar-98	112.0	115.4	105.8
Apr-98	112.7	114.9	108.8
May-98	111.9	114.1	108.1
Jun-98	113.1	117.6	105.1
Jul-98	113.2	118.1	104.5
Aug-98	110.3	114.4	103.1
Sep-98	112.6	115.6	107.1
Oct-98	114.0	115.6	111.1
Nov-98	113.2	114.4	111.1
Dec-98	114.3	117.0	109.4
Jan-99	112.4	114.2	109.2
Feb-99	114.4	116.7	110.5
Mar-99	119.4	123.5	112.2
Apr-99	116.3	118.0	113.4
May-99	115.1	117.2	111.3
Jun-99	116.8	120.3	110.7
Jul-99	116.1	118.5	112.0
Aug-99	116.7	121.7	107.8
Sep-99	119.1	121.5	114.9
Oct-99	118.3	118.5	118.0
Nov-99	P 121.8	122.9	119.8
Dec-99	120.0	124.4	112.3
Jan-00	117.6	121.2	111.3
Feb-00	121.4	123.0	118.7
Mar-00	119.4	120.0	118.5
Apr-00	112.6	108.3	120.2
May-00	120.0	119.3	121.5
Jun-00	117.8	116.7	119.8
Jul-00	114.1	112.2	117.6
Aug-00	118.9	122.1	113.3
Sep-00	115.6	112.9	120.3
Oct-00	116.8	114.4	121.1
Nov-00	118.7	115.4	124.7

Key: P = peak; T = trough.

(continues)

Monthly Values of Transportation Indexes *(continued)*

Time	Total Transportation Index	Freight Transportation Index	Passenger Transportation Index
Dec-00	112.3	109.3	117.7
Jan-01	118.8	119.0	118.6
Feb-01	114.5	111.9	119.3
Mar-01	118.2	117.0	120.4
Apr-01	115.5	111.7	122.5
May-01	120.9	121.1	120.6
Jun-01	115.5	113.3	119.5
Jul-01	116.6	115.2	119.3
Aug-01	120.2	122.5	116.4
Sep-01	^{T}101.6	108.4	90.0
Oct-01	108.8	115.5	97.2
Nov-01	110.1	113.1	104.9
Dec-01	108.6	110.4	105.6
Jan-02	114.7	119.5	106.3
Feb-02	110.7	111.5	109.6
Mar-02	112.5	112.9	112.0
Apr-02	116.3	120.0	109.9

Key: T = trough.

The Importance of Transportation in the Canadian Economy

JEFF HARRIS
Transport Canada

ABSTRACT

This paper uses direct and indirect demands for transportation as a proportion of final demand to assess the relative share of transportation-related demand in Canadian gross domestic product (GDP) from 1971 to 1996. The data are derived from the Canadian input/output tables. Three trends are highlighted over the time period studied: a growing share of transportation-related trade in GDP, a decline in transportation investment, and a decline in the transportation margins associated with the distribution of commodities. Overlying these trends, as a major determinant of transportation as a share of GDP, is the volatility of transport fuel prices. Transportation as a share of GDP has been fairly steady over the time period studied, representing 20.7% of GDP in 1996, with a peak of 21.1% in 1981 corresponding to the peak in fuel prices, and a low of 19.1% during the recession of the early 1990s.

INTRODUCTION

This paper uses the Canadian system of national accounts (CNA) to estimate the proportion of Canadian gross domestic product (GDP) that is transport-related, with transportation final demand

KEYWORDS: transportation, economic analysis, national accounts, GDP, consumption, investment, imports, exports, direct and indirect demand, domestic demand.

as the measure of GDP instead of value added or income. Final demand is used here because it allows for a broader definition of transportation, notably including the use of transportation equipment, fuel, and infrastructure traditionally not considered to be part of the transportation industries, but instead considered private transportation demand by industry and consumers, and public transportation demand by government.

The paper draws on the work of Han and Fang (1998) on the final demand-based methodology used by the U.S. Department of Transportation and expands on that methodology in two principal ways:

1. by broadening the definition of transportation-related demand to include both transportation fuels (e.g., gasoline, diesel oil) and transportation margins,[1] as well as transportation equipment, infrastructure, and industries;
2. by including not only direct transportation-related demand but also indirect demand, which is embedded in nontransportation-related final demand. For example, shoes that are consumed as part of final demand require transportation as an intermediate input in their production, thus some part of the final cost of shoes to consumers can be considered as indirect transportation-related final demand.

The data used in this paper are from the CNA input-output (IO) tables, using the industry and commodity classification system based on the Canadian Standard Industrial Classification (SIC) system. The SIC has been replaced by the North American Industry Classification System (NAICS). However, at the time this paper was written, NAICS-coded data were not available for the years assessed (1971–1996). Statistics Canada's IO division provided all data described in this paper.[2]

It should be noted that this paper does not develop a satellite account for transportation in the same sense as the U.S. satellite account for transportation described in Fang et al. (2000).[3] Transportation satellite accounts typically develop an estimate for a new industry, for example, private trucking, by disaggregating and then reaggregating data derived from other industries. For example, trucking employment, transportation equipment investment, or transportation fuel use contained in other industries, such as the retail or wholesale trade industries, will form the basis for estimating the private trucking industry.

This paper neither disaggregates nor reaggregates data, but uses the existing rows and columns contained in the CNA, while imposing a definition of transportation-related demand that encompasses both existing industries (e.g., transportation industries) and commodities (e.g., transportation equipment, fuel). The same methodology used in this paper could be applied to transportation satellite account IO tables where, for example, private trucking was reconstituted as a separate industry column, thereby increasing the accuracy of the description of transportation within the national economy. A more detailed description of the methodology used here is provided later in this paper.

The paper is structured in three sections: the CNA, methodology, and results. The paper traces the evolution of transportation-related demand from 1971 to 1996, at five-year intervals based on the Canadian census years, where 1996 is the most current IO table available. These time periods were also selected to allow for observations relative to the business cycle, with 1971 and 1976 straddling the

[1] Transportation margins (TMs) are a concept unique to the CNA and represent an estimate of the transportation costs incurred in the distribution of commodities. They form, along with trade margins and indirect taxes, the principal difference between commodities at factor prices and final prices as the sum of final demand categories in the CNA. In order to fully account for transportation in the CNA, the margins must either be included separately in transportation demand or disaggregated and merged with other appropriate transportation commodities (e.g., freight trucking) as a form of satellite account. Because this paper is not developing a satellite account, the TMs were included as a separate commodity.

[2] A small amount of suppression for confidentiality of certain entries in the IO tables was undertaken by Statistics Canada prior to providing the tables, typically accounting for 1%–2% of the total values contained in the tables. Estimates of the suppressed data were generated based on comparisons of the provided actual row and column total values, and comparisons of the values in different years with unsuppressed data.

[3] It can be noted that the development of a similar satellite account for private trucking in Canada is proceeding, and when completed, can be used with the same methodology described in this paper to develop a more precise estimate of the share of transport demand in GDP.

TABLE 1 Description of the w- and l-Level Matrices

	Rows	Columns	Years
l-level			1961–1996
Final demand	476 commodities	122 categories	
Use	476 commodities	167 industries	
Make	476 commodities	167 industries	
w-level			1986–1996
Final demand	679 commodities	138 categories	
Use	679 commodities	244 industries	
Make	679 commodities	244 industries	

1973 OPEC oil crisis, 1981 and 1991 representing recession years, and 1986 and 1996 representing years of recovery and positive growth.

CANADIAN NATIONAL ACCOUNTS

The CNA, as represented by the IO tables, are structured into the make (M), use (U), and final demand (FD) matrices. The M table is a commodity by industry matrix that indicates which commodities are made by which industry. The U table is a commodity and primary inputs by industry matrix that indicates the amount of commodities and primary inputs used by industry. The FD table is a commodity and primary inputs by FD category matrix that indicates the amount of commodities and primary inputs that form part of FD. FD categories are broadly classified as consumption (C), investment (I), government spending (G), imports (IM), and exports (EX) where

$FD = GDP = C + I + G - IM + EX$, and

FDD (final domestic demand) $= C + I + G$.

In the CNA, detailed commodities are more numerous than either industry or final demand classifications, in order to account for industries producing more than one commodity or joint production, thus leading to rectangular matrix forms. Commodities are goods or services while primary commodities represent returns to primary factors (e.g., labor, capital), as well as variables such as indirect taxes. There are four levels of detail for which IO tables are available in the CNA: s, m, l, and w. S corresponds to small (in terms of numbers of rows and columns), m to medium, and w to working (the most detailed level available). L refers to an historical link series, which provides a consistent classification of industries and commodities going back over time to 1961. This paper uses the Canadian SIC-based system used at Statistics Canada for the 1986 to 1997 IO tables, which has now been replaced by the NAICS.[4] Table 1 illustrates the dimensions of the two most detailed levels available (l, w) based on the SIC, where the l (or historical link) tables are used in this paper to represent a consistent classification over time, though at some loss of detail.

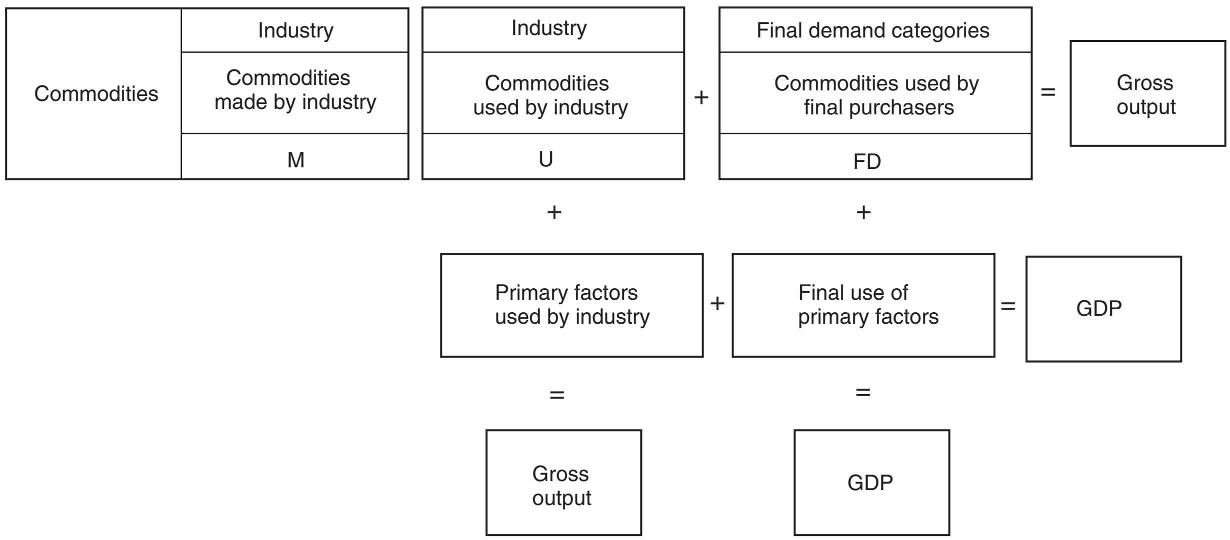
Figure 1 shows the relationship between the three IO tables and illustrates two common means of calculating GDP—by the sum of primary inputs and the sum of final demands. The third common means of calculating GDP is value added, which is done by subtracting the U from the M matrix, leaving a commodity by industry matrix of value added.

In terms of how supply and demand for transportation is represented in the CNA tables, the M matrix indicates the total supply of transportation (where M shows the value and type of commodities produced by each industry, or gross output). The U and FD matrices represent, respectively, the intermediate and final demands for commodities, where U indicates which commodities are used as inputs in production and FD indicates those commodities that are allocated to final demand categories and thus form GDP.

In terms of transport demand, the direct demand for transportation (e.g., automobiles, rail investment) is found in the FD tables, while indirect demand is derived from transportation commodities (e.g., domestic freight) used as inputs in the U matrix. Indirect demand refers to the proportion of

[4] A NAICS-based l-level historical series is currently being developed by Statistics Canada, with data for 1961–1999 expected to be available by the end of 2004.

FIGURE 1 The Canadian Input/Output Tables



Source: Statistics Canada, 1989.

transportation commodities embedded in nontransport final demand. For example, shoes consumed as part of final demand require transportation as an intermediate input in their production; thus a proportion of the final cost of shoes to consumers can be considered as indirect transportation-related final demand. The definition of the industries and commodities classified as transportation-related demand in this paper is provided in the following section.

One difficulty, or opportunity, in accounting for transportation in the CNA is commercial transportation, which is attributed to "fictive" industries, notably transportation margins (TMs). Margins are created to account for the difference between factor prices, or price at the factory gate, and final prices, or prices charged to the consumer, with the two main types of margins being transportation and trade margins.[5] TMs appear as both columns and rows in the *M* and *U* matrices and as a row in the *FD* matrix, and can be interpreted as the distribution costs imputed to transportation industries. The use of TMs means that transportation industry commodities are allocated between: 1) transportation used in distribution, or transportation costs as a portion of the difference between the factor price and final price, and 2) transportation used in production, or transportation costs as a share of factor prices.

The fictive TMs industry uses solely for-hire commercial freight transportation industry commodities (e.g., trucking, rail) as inputs in the *U* matrix, while its commodity output is represented as a single row in both the *U* and *FD* matrices. This means that components of the TMs (e.g., trucking, rail) cannot be attributed to specific industries, although they can be classified in aggregate through the inputs of the *U* matrix. Table 2 lists the composition, or inputs, to the TMs industry from 1971 to 1996. The trends show the most growth in trucking as a share of the TMs, going from 41.1% in 1971 to 60.6% in 1996, with the largest growth from 1991 (52.3%) to 1996 (60.6%), possibly reflecting trade trends and the ongoing movement to just-in-time distribution. Over the same period, rail declined from 41.7% to 27.3% and water transport from 11.9% to 3.6% of the TMs.

These pronounced trends are an early indication of this paper's theme—the observation that many of the most interesting trends in commercial freight transportation are actually found within the evolution of the TMs. As an indication of the importance of the TMs in this area, in 1996 the rail and truck inputs to the TMs accounted for 81.9% and 62.6%, respectively, of the gross outputs of the two commercial transportation industries, as table 3 illustrates. However, caution must be taken when interpreting the TMs, particularly for smaller

5 The other main component of the difference between factor prices and final prices are indirect taxes (e.g., excise taxes) and subsidies.

TABLE 2 Composition of the Transportation Margins: 1971–1996
Millions of Canadian dollars

Transportation margins	1971 Dollars	%	1976 Dollars	%	1981 Dollars	%	1986 Dollars	%	1991 Dollars	%	1996 Dollars	%
Air	66	1.9	128	1.9	590	5.2	508	3.3	512	3.2	512	2.5
Other services incidental to transportation	44	1.2	236	3.6	534	4.7	994	6.4	1,065	6.7	1,481	7.2
Water	421	11.9	986	14.8	1,053	9.2	1,004	6.5	1,250	7.9	735	3.6
Services incidental to water transportation	147	4.1	296	4.5	380	3.3	588	3.8	520	3.3	258	1.3
Railway	1,479	41.7	2,396	36.1	4,162	36.4	4,838	31.1	4,761	29.9	5,605	27.3
Truck	1,457	41.1	2,729	41.1	5,292	46.3	8,143	52.3	8,320	52.3	12,446	60.6
Total inputs, transportation margins	3,548		6,643		11,421		15,567		15,916		20,525	

TABLE 3 Allocation of Commercial Freight Industries' Gross Output to Transportation Margins, Own Uses,[1] and Commercial Transportation at Factor Prices: 1976–1996

	1976 TMs	1976 Own use	1976 Factor price[2]	1986 TMs	1986 Own use	1986 Factor price	1996 TMs	1996 Own use	1996 Factor price
Water	88.4%	24.1%	–12.5%	55.6%	20.6%	23.8%	30.8%	7.9%	61.3%
Services incidental to water transportation	61.3%	35.9%	2.8%	48.0%	14.6%	37.4%	17.5%	2.1%	80.4%
Railway	81.3%	2.1%	16.7%	78.2%	1.1%	20.8%	81.9%	0.8%	17.4%
Truck	70.2%	8.6%	21.2%	66.0%	2.3%	31.7%	62.6%	3.8%	33.6%

[1] Own use refers to own commodities used by the commodity-producing industry as inputs, for example, trucking commodities used by the trucking industry.

[2] The negative value reflects a negative entry in commodity final demand generated by the imbalance of imports over exports and little direct consumer demand.

freight industries (e.g., air, marine freight) as they are an artificial industry created within the CNA to distinguish between factor and final prices. The TMs also do not include private or in-house freight transportation.

METHODOLOGY

The first step in determining the importance of transportation is the definition of direct transport demand in terms of the commodities (rows) and final demand categories (columns) of the *FD* matrix. An important point to note is that no disaggregation or reaggregation of the existing l-level rows or columns were used in this paper, as this would represent the development of a satellite account for transportation, a partial example of which is provided in the U.S. national accounts with the satellite account for private truck transport.

In terms of commodities, all rows associated with transportation equipment (including tires), transportation industries (including pipelines), as well as the TMs were used. Selected commodities were also chosen to represent transportation fuels, transportation construction, and other transportation services, three areas where the level of aggregation at the l level of the CNA leads to somewhat incomplete datasets.

In the case of transportation fuels, diesel and aviation fuel are aggregated with heating oil as one commodity (or row), and thus transport demand for this commodity will be slightly overestimated. However, this commodity is included as part of transport fuels, along with motor gasoline.

In terms of construction, the level of detail does not allow for disaggregating transportation-related expenditures from several commodities, notably construction repair, other engineering construction, and railways and communications. The only row in construction that is unequivocally transportation-related is roads, highways, and airport construction. A similar problem of inadequate detail also occurs with other service commodities, principally repair services and trade margins as well as indirect taxes. The only other services commodity that is unequivocally transportation-related is car and truck rentals.

Thus, in order to develop a more complete picture of transportation-related demand, certain *FD* categories or columns were also classified as direct transport demand, notably all columns in consumption unequivocally related to transportion demand, as well as equipment and construction investment by transportation industries. Therefore, construction commodities (e.g., railway track) or service commodites (e.g., trade margins, repair services), as well as indirect taxes found in these *FD* columns will also be classified as direct transport demand. This should limit, but not completely account for, the underestimation of transportion-related demand due to transportation-related expenditures embedded within construction, other services, or indirect tax commodities.

Given this definition, the direct demand for transportation (FD_t) can be classified as a matrix of $c * f$ proportions where *c* refers to commodities, *f* to final demand categories, and all nontransport-related commodities or columns are specified as zeros. Summing FD_t by commodities generates the breakdown of direct transportation-related demand presented in this paper.

Indirect Demand

Indirect demand refers to the demand for transport embedded in the factor price of nontransport-related commodities and *FD* columns, where transport-related commodities and columns are as defined above. Determining the indirect demand for transportation requires all three matrices, the *M*, *U*, and *FD*.

The first step in calculating the indirect demand is to generate a proportional matrix of *U*, where all commodity entries are divided by the total gross inputs to generate a $c * i$ (or industry) matrix of proportions, here called the *UP* matrix. Each individual entry in this matrix indicates the proportion of industry gross inputs accounted for by that commodity.

The second step is multiplying the *UP* matrix by the transpose of the *M* matrix (M^t) where the transpose is required due to the rectangular format of the CNA.[6] Thus

$$UP * M^t = C$$

where *C* is a $c * c$ matrix indicating the commodity values used in the production of other commodities. Converting the *C* matrix to proportions in a similar manner to the *U* matrix by dividing the commodity rows by the commodity gross outputs generates the *CP* matrix, a $c * c$ matrix where each entry indicates the proportion of each commodity used in the production of all other commodities.

In the third step, we return to *FD*, and in order to avoid double counting, subtract FD_t or

$$FD - FD_t = FD_{nt}$$

where FD_{nt} represents a $c * f$ matrix with zero entries and the direct transportation commodities and columns as defined above are located. This step ensures that transportation commodities used in the production of other such commodities (e.g., transportation industries used to produce transportation equipment) are not double counted.

In the fourth step, the FD_{nt} matrix is reduced to a $c * 2$ matrix, where the two columns represent indirect domestic demand and exports, here called FD_{ID}. Indirect domestic demand (*IDD*) is calculated as $C + I + G - I_p = IDD$, where I_p refers to the proportion of imports that are attributed to final demand. This commodity import proportion is cal-

[6] The transpose is required given the rectangular matrix forms of the CNA in order to ensure that the conformability condition for matrix multiplication applies, specifically that the column dimension of the lead matrix (*UP*) is equal in number to the row dimension of the lag matrix (M^t). An alternate, if more cumbersome, procedure would be to initially merge the various commodity rows to industry dimensions, thus generating a square set of *U* and *M* matrices.

culated as the relative share of final demand by commodity as a proportion of commodity gross output, multiplied by total imports. Thus, *IDD* is an estimate of the nontransportation-related domestic demand met by domestic producers.

The final step requires:

$CP * FD_{ID} = ID$, where *ID* (indirect demand) refers to a $c * 2$ matrix listing the commodity values used in the production of the FD_{nt} commodities. Using the same definition of transport commodities as listed above in FD_t allows for the calculation of the indirect share of transport in FD_{ID}.

RESULTS

Direct Transportation Demand

Consumption

Transportation-related consumption has remained at a fairly steady proportion of total consumption, varying between a high of 16.2% in 1981, at the height of the fuel price spike brought about by the OPEC cartel, to a low of 14.7% in the recession of the early 1990s (table 4). In 1996, the largest component of transport consumption was "other transportation services" associated with equipment sales and use, that is, trade margins and repairs. The "other transportation services" category shows a low of 4.0% in 1981 and a high of 4.4% in 1996.

The second largest category of consumption in 1996 was transportation equipment (at factor price), varying from a high of 4.5% in 1986, in the growth period following the recession of the early 1980s, to a low of 3.7% in the recession of the early 1990s. Automobiles constitute the largest segment of transportation equipment, although a slight decrease in consumption from 3.1% in 1971 to 2.5% in 1996 is seen. However, increases in the truck category, from 0.2% in 1971 to 0.9% in 1996, compensate for the decrease in the automobile category.

Indirect taxes (e.g., sales tax, excise tax) were the third largest component, accounting for 2.8% of consumption in both 1991 and 1996. This represents an increasing trend from a low of 2.1% of consumption in 1981, with the share of indirect taxes also lower in the 1970s relative to the 1990s.

Transportation industries, or commercial transportation, show a steady level of consumption at approximately 2% in all years. Within commercial transportation, two slight trends are evident, an increasing share of air transportation, from 0.7% in 1971 to 1.1% in 1996, and a declining trend for surface passenger transportation, from 0.8% in 1971 to 0.5% in 1996. The TMs show a more pronounced decline, from 1.0% of consumption in 1971 to 0.5% in 1996, possibly associated with trucking deregulation and price competition, and the corresponding increased trend toward truck freight use in the TMs.

The most volatile of transportation commodities are transportation fuels, swinging from a high of 3.0% of consumption in 1981 to a low of 1.2% in 1996, after beginning at 1.5% in 1971. Although accounting for a relatively low share of consumption, the influence of fuel price swings is evident, with the 1981 fuel price peak generating the most atypical year of all years assessed.

Investment

Investment can be classified into three separate categories to distinguish between transportation industries, other industries, and government. Note that all investment undertaken by transportation industries is classified as transportation-related, but only investment in commodities that are transportation-specific is counted for government and business.

The components of investment by transportation industries experienced a major swing from 1971 to 1996, with an increasing share of transportation equipment—29.0% in 1971 to 59.5% in 1996—and a declining share of transportation construction—63.3% in 1971 to 32.9% in 1996 (table 5). This reflects the reduced level of construction investment in railways as they consolidated their construction capital stock, which is consistent with their declining market share in freight transportation. It also reflects a surge in investment in other equipment, particularly between 1991 and 1996, possibly reflecting a more recent swing from investment in transportation equipment to information and communications technology. As of 1996, "other equipment" was the largest component in the transportation equipment category, ahead of aircraft and other complete equipment (e.g., trailers and semi-trailers). As expected, annual investment is rather volatile, particularly aircraft investment; nevertheless, we can see a slight upward trend for aircraft and trucks and a slight downward trend for railroad equipment.

TABLE 4 Transportation-Related Consumption as a Proportion of Total Consumption
Millions of Canadian dollars

	1971		1976		1981		1986		1991		1996	
Commodities	**Dollars**	**%**	**Dollars**	**%**	**Dollars**	**%**	**Dollars**	**%**	**Dollars**	**%**	**Dollars**	**%**
Transportation equipment												
Aircraft, parts, and repairs	3	0.0	7	0.0	12	0.0	21	0.0	24	0.0	31	0.0
Automobiles, including passenger vans	1,698	3.1	3,114	2.9	5,146	2.7	9,098	3.1	10,082	2.5	11,926	2.5
Trucks, road tractors, and chassis	123	0.2	541	0.5	791	0.4	1,735	0.6	2,357	0.6	4,534	0.9
Other complete equipment	289	0.5	609	0.6	932	0.5	1,264	0.4	1,630	0.4	2,080	0.4
Motor vehicle parts, including bodies and tires	225	0.4	435	0.4	694	0.4	1,030	0.4	736	0.2	944	0.2
Shipbuilding and ship repair	2	0.0	3	0.0	4	0.0	6	0.0	7	0.0	8	0.0
Total	**2,340**	**4.2**	**4,709**	**4.4**	**7,579**	**4.0**	**13,154**	**4.5**	**14,836**	**3.7**	**19,523**	**4.0**
Transportation fuels												
Motor gasoline	439	0.8	1,385	1.3	3,986	2.1	3,483	1.2	4,072	1.0	4,728	1.0
Diesel and fuel oil, aviation fuel	382	0.7	950	0.9	1,729	0.9	1,119	0.4	948	0.2	1,101	0.2
Total	**821**	**1.5**	**2,335**	**2.2**	**5,715**	**3.0**	**4,602**	**1.6**	**5,020**	**1.3**	**5,829**	**1.2**
Transportation construction												
Road, highway, and airport runway construction	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Highway and bridge maintenance	33	0.1	46	0.0	66	0.0	55	0.0	112	0.0	95	0.0
Total	**33**	**0.1**	**46**	**0.0**	**66**	**0.0**	**55**	**0.0**	**112**	**0.0**	**95**	**0.0**
Transportation industries												
Air	392	0.7	947	0.9	2,141	1.1	2,911	1.0	4,459	1.1	5,261	1.1
Railway	60	0.1	90	0.1	165	0.1	207	0.1	183	0.0	161	0.0
Water	19	0.0	42	0.0	150	0.1	203	0.1	333	0.1	406	0.1
Truck	72	0.1	124	0.1	288	0.2	357	0.1	361	0.1	417	0.1
Surface passenger	452	0.8	733	0.7	1,273	0.7	1,867	0.6	2,332	0.6	2,579	0.5
Other transportation services	2	0.0	3	0.0	5	0.0	7	0.0	7	0.0	6	0.0
Pipeline	70	0.1	146	0.1	275	0.1	609	0.2	638	0.2	884	0.2
Total	**1,067**	**1.9**	**2,085**	**1.9**	**4,297**	**2.2**	**6,161**	**2.1**	**8,313**	**2.1**	**9,714**	**2.0**
Transportation margins	553	1.0	1,049	1.0	1,553	0.8	2,137	0.7	2,136	0.5	2,264	0.5
Other transportation services												
Trade margins	1,429	2.6	2,366	2.2	4,417	2.3	7,523	2.6	9,912	2.5	12,347	2.6
Other services (e.g., repairs)	877	1.6	1,994	1.8	3,163	1.7	4,703	1.6	6,951	1.7	8,345	1.7
Rental of automobiles and trucks	60	0.1	103	0.1	156	0.1	303	0.1	462	0.1	323	0.1
Total	**2,366**	**4.3**	**4,463**	**4.1**	**7,736**	**4.0**	**12,529**	**4.3**	**17,325**	**4.3**	**21,015**	**4.4**
Total, indirect taxes	**1,260**	**2.3**	**2,622**	**2.4**	**3,924**	**2.1**	**7,455**	**2.6**	**11,037**	**2.8**	**13,517**	**2.8**
Total, transportation-related consumption	**8,440**	**15.3**	**17,309**	**16.0**	**30,870**	**16.2**	**46,093**	**15.9**	**58,779**	**14.7**	**71,957**	**14.9**
Total, consumption	55,073		108,121		191,116		289,559		399,933		482,058	

TABLE 5 Investment by Transportation Industries
Millions of Canadian dollars

Commodities	1971 Dollars	1971 %	1976 Dollars	1976 %	1981 Dollars	1981 %	1986 Dollars	1986 %	1991 Dollars	1991 %	1996 Dollars	1996 %
Transportation equipment												
Aircraft, parts, and repairs	98	8.3	43	2.5	907	17.6	451	12.3	2,014	28.4	1,095	15.9
Automobiles, including passenger vans	13	1.1	9	0.5	21	0.4	23	0.6	28	0.4	40	0.6
Trucks, road tractors, and chassis	22	1.9	49	2.9	50	1.0	315	8.6	292	4.1	283	4.1
Other complete equipment	19	1.6	90	5.3	209	4.1	181	4.9	182	2.6	764	11.1
Motor vehicle parts, including bodies and tires	5	0.4	10	0.6	4	0.1	39	1.1	34	0.5	18	0.3
Railroad equipment and parts	109	9.3	188	11.0	359	7.0	277	7.6	114	1.6	448	6.5
Shipbuilding and ship repair	23	2.0	75	4.4	284	5.5	77	2.1	217	3.1	249	3.6
Other equipment	52	4.4	153	8.9	381	7.4	265	7.2	400	5.7	1,209	17.5
Total	**341**	**29.0**	**617**	**36.1**	**2,215**	**42.9**	**1,628**	**44.4**	**3,281**	**46.3**	**4,106**	**59.5**
Transportation construction												
Road, highway, and airport runway construction	8	0.7	5	0.3	12	0.2	49	1.3	29	0.4	452	6.5
Other construction	737	62.6	964	56.3	2,697	52.3	1,711	46.6	3,375	47.7	1,816	26.3
Total	**745**	**63.3**	**969**	**56.6**	**2,709**	**52.5**	**1,760**	**48.0**	**3,404**	**48.1**	**2,268**	**32.9**
Transportation margins	**9**	**0.8**	**7**	**0.4**	**17**	**0.3**	**17**	**0.5**	**43**	**0.6**	**24**	**0.3**
Trade margins	**30**	**2.5**	**63**	**3.7**	**102**	**2.0**	**133**	**3.6**	**239**	**3.4**	**366**	**5.3**
Indirect taxes	**53**	**4.5**	**56**	**3.3**	**117**	**2.3**	**128**	**3.5**	**113**	**1.6**	**138**	**2.0**
Total, transportation	**1,174**	**100.0**	**1,712**	**100.0**	**5,160**	**100.0**	**3,666**	**99.9**	**7,080**	**100.0**	**6,902**	**100.0**
Total investment by transportation industries	1,174		1,711		5,159		3,668		7,080		6,902	

Transportation investment by other businesses accounted for 7.2% of total investment in 1996, with a low of 5.0% in 1991 and a high of 7.5% in 1986 (table 6). This investment is almost exclusively in transportation equipment where a pronounced business cycle trend is evident, with lower investment in recessions and higher investment in recoveries. As of 1996, transportation equipment accounted for 6.8% of total investment, up from a low of 4.4% in the recession year of 1991. The largest share of business investment was in automobiles (4.4% in 1996) and trucks (1.5% in 1996), with a slight increase in automobiles and a slight decline in trucks, possibly reflecting increased outsourcing from private to commercial transportation.

Government investment is primarily road-related construction, with total transportation investment accounting for 25.9% of government investment in 1996—with a low of 23.8% in 1991 and a high of 32.2% in 1971 (table 7). Government transportation investment exhibits a declining trend, with road investment declining from 29.3% of government investment in 1971 to a low of 21.3% in 1991, followed by a slight upturn to 23.4% in 1996.

TABLE 6 Transportation-Related Investment by Businesses Other Than Transportation Industries
Proportion of total business investment; millions of Canadian dollars

	1971		1976		1981		1986		1991		1996	
Commodities	**Dollars**	**%**	**Dollars**	**%**	**Dollars**	**%**	**Dollars**	**%**	**Dollars**	**%**	**Dollars**	**%**
Transportation equipment												
Aircraft, parts, and repairs	34	0.2	47	0.1	325	0.4	374	0.4	33	0.0	410	0.3
Automobiles, including passenger vans	311	1.9	778	2.0	1,613	2.2	3,207	3.7	2,451	2.3	5,227	4.4
Trucks, road tractors, and chassis	351	2.1	1,073	2.8	1,531	2.1	1,818	2.1	1,761	1.7	1,769	1.5
Other complete equipment	150	0.9	426	1.1	457	0.6	340	0.4	367	0.3	663	0.6
Railroad equipment and parts	32	0.2	28	0.1	9	0.0	2	0.0	26	0.0	8	0.0
Shipbuilding and ship repair	38	0.2	72	0.2	111	0.2	39	0.0	33	0.0	21	0.0
Total	**916**	**5.5**	**2,424**	**6.2**	**4,046**	**5.5**	**5,780**	**6.7**	**4,671**	**4.4**	**8,098**	**6.8**
Transportation construction												
Road, highway, and airport runway construction	76	0.5	194	0.5	408	0.6	316	0.4	270	0.3	232	0.2
Total	**76**	**0.5**	**194**	**0.5**	**408**	**0.6**	**316**	**0.4**	**270**	**0.3**	**232**	**0.2**
Transportation margins	**84**	**0.5**	**206**	**0.5**	**353**	**0.5**	**412**	**0.5**	**397**	**0.4**	**320**	**0.3**
Total, transportation	**1,076**	**6.5**	**2,824**	**7.3**	**4,807**	**6.6**	**6,508**	**7.5**	**5,338**	**5.0**	**8,650**	**7.2**
Total, business Investment	16,560		38,819		73,024		86,867		105,727		119,643	

Summing the three categories of transportation investment generates the result that the highest level of transportation-related investment as a proportion of total investment was in 1971 (16.4%—table 8). A contradictory business cycle movement can be observed, with investment by transportation industries becoming the largest component of transport investment in recessions (1981, 1991) while other business forms the largest component in recoveries (1986, 1996). The share of government exhibits a steady decline, from the largest share in 1971 (6.2%) to 3.4% of total investment in 1996.

Trade: Exports and Imports

Transportation commodities account for a large share of Canadian exports, 30.0% as of 1996, with a low of 26.0% in the fuel-price-generated recession of 1981 and a high of 34.8% in the recovery year of 1986 (table 9). The main component of exports is transportation equipment, accounting for 23.3% of exports in 1996, with a high of 27.2% in 1986 and a low of 17.4% in 1981. In all years, the main components of transportation equipment exports are automobiles, motor vehicle parts, and trucks.

The second largest component of exports are the TMs, accounting for 2.8% of final export prices in 1996, down from 5.1% in 1971, with a steady decline in all years. Again, this steady decline correlates with an increasing share of trucking in the TMs, and a reorientation and concentration of trade with the United States. Commercial transportation (transportation industries) exhibits a steadier share of exports, also accounting for 2.8% in 1996, with a high of 3.0% in 1991 and a low of 1.4% in 1976. The upward trend in commercial transportation is generated by increasing exports of air, truck, and pipeline services.

TABLE 7 Transportation-Related Investment by Government
Proportion of government investment; millions of Canadian dollars

Government	1971 Dollars	1971 %	1976 Dollars	1976 %	1981 Dollars	1981 %	1986 Dollars	1986 %	1991 Dollars	1991 %	1996 Dollars	1996 %
Transportation equipment												
Aircraft, parts, and repairs	16	0.4	27	0.4	87	0.8	60	0.4	31	0.2	0	0.0
Automobiles, including passenger vans	40	1.0	87	1.3	63	0.6	98	0.7	110	0.6	137	0.7
Trucks, road tractors, and chassis	27	0.6	46	0.7	82	0.8	129	0.9	156	0.8	139	0.7
Other complete equipment	17	0.4	32	0.5	24	0.2	73	0.5	64	0.3	90	0.5
Motor vehicle parts, including bodies and tires	2	0.0	5	0.1	3	0.0	4	0.0	6	0.0	0	0.0
Railroad equipment and parts	0	0.0	63	0.9	0	0.0	0	0.0	0	0.0	0	0.0
Shipbuilding and ship repair	9	0.2	25	0.4	89	0.8	129	0.9	74	0.4	70	0.4
Total	**111**	**2.6**	**285**	**4.3**	**348**	**3.3**	**493**	**3.5**	**441**	**2.3**	**436**	**2.3**
Transportation construction												
Road, highway, and airport runway construction	1,232	29.3	1,764	26.3	2,805	26.3	3,610	25.6	4,082	21.3	4,463	23.4
Total	**1,232**	**29.3**	**1,764**	**26.3**	**2,805**	**26.3**	**3,610**	**25.6**	**4,082**	**21.3**	**4,463**	**23.4**
Transportation margins	**9**	**0.2**	**16**	**0.2**	**29**	**0.3**	**41**	**0.3**	**48**	**0.3**	**35**	**0.2**
Total, transportation	**1,352**	**32.2**	**2,065**	**30.8**	**3,182**	**29.8**	**4,144**	**29.4**	**4,571**	**23.8**	**4,934**	**25.9**
Total, government investment	4,205		6,696		10,665		14,089		19,208		19,066	

TABLE 8 Transportation-Related Investment as a Proportion of Total Investment
Millions of Canadian dollars

	1971 Dollars	1971 %	1976 Dollars	1976 %	1981 Dollars	1981 %	1986 Dollars	1986 %	1991 Dollars	1991 %	1996 Dollars	1996 %
Transportation industries	1,174	5.4	1,712	3.6	5,160	5.8	3,666	3.5	7,080	5.4	6,902	4.7
Other industries	1,076	4.9	2,824	6.0	4,807	5.4	6,508	6.2	5,338	4.0	8,650	5.9
Government	1,352	6.2	2,065	4.4	3,182	3.6	4,144	4.0	4,571	3.5	4,934	3.4
Total, transportation (% total investment)	**3,602**	**16.4**	**6,601**	**14.0**	**13,149**	**14.8**	**14,318**	**13.7**	**16,989**	**12.9**	**20,486**	**14.1**
Total, investment (% final demand)	**21,942**	**27.9**	**47,226**	**30.0**	**88,848**	**31.0**	**104,624**	**25.9**	**132,015**	**25.0**	**145,611**	**21.5**
Final demand	78,594		157,624		286,236		403,783		527,967		678,222	

TABLE 9 Transportation Commodities as a Share of Exports
Millions of Canadian dollars

Commodities	1971 Dollars	1971 %	1976 Dollars	1976 %	1981 Dollars	1981 %	1986 Dollars	1986 %	1991 Dollars	1991 %	1996 Dollars	1996 %
Transportation equipment												
Aircraft, parts, and repairs	326	1.5	527	1.2	2,055	2.1	3,208	2.2	5,186	3.0	7,501	2.3
Automobiles, including passenger vans	2,091	9.9	3,794	8.6	5,535	5.7	17,297	12.1	15,455	9.0	33,827	10.5
Trucks, road tractors, and chassis	554	2.6	1,305	2.9	3,020	3.1	5,253	3.7	7,027	4.1	10,645	3.3
Other complete equipment	165	0.8	203	0.5	558	0.6	819	0.6	1,049	0.6	3,168	1.0
Motor vehicle parts, including bodies and tires	1,354	6.4	3,134	7.1	5,076	5.2	11,485	8.0	9,321	5.4	18,173	5.7
Railroad equipment and parts	35	0.2	117	0.3	463	0.5	567	0.4	563	0.3	1,422	0.4
Shipbuilding and ship repair	4	0.0	230	0.5	208	0.2	153	0.1	37	0.0	117	0.0
Total	**4,529**	**21.5**	**9,310**	**21.0**	**16,915**	**17.4**	**38,782**	**27.2**	**38,638**	**22.4**	**74,853**	**23.3**
Transportation fuels												
Motor gasoline	0	0.0	26	0.1	232	0.2	392	0.3	919	0.5	1,252	0.4
Diesel and fuel oil, aviation fuel	85	0.4	268	0.6	1,114	1.1	864	0.6	1,350	0.8	1,775	0.6
Total	**85**	**0.4**	**294**	**0.7**	**1,346**	**1.4**	**1,256**	**0.9**	**2,269**	**1.3**	**3,027**	**0.9**
Transportation industries												
Air	5	0.0	24	0.1	746	0.8	1,040	0.7	1,198	0.7	2,210	0.7
Railway	67	0.3	90	0.2	208	0.2	226	0.2	208	0.1	231	0.1
Water	142	0.7	248	0.6	724	0.7	638	0.4	914	0.5	1,528	0.5
Truck	79	0.4	143	0.3	132	0.1	561	0.4	1,237	0.7	2,552	0.8
Surface passenger	0	0.0	0	0.0	82	0.1	174	0.1	207	0.1	308	0.1
Other transportation services	12	0.1	24	0.1	247	0.3	397	0.3	542	0.3	641	0.2
Pipeline	21	0.1	79	0.2	487	0.5	424	0.3	888	0.5	1,557	0.5
Total	**326**	**1.5**	**608**	**1.4**	**2,626**	**2.7**	**3,460**	**2.4**	**5,194**	**3.0**	**9,027**	**2.8**
Transportation margins	**1,067**	**5.1**	**2,048**	**4.6**	**4,270**	**4.4**	**5,872**	**4.1**	**6,202**	**3.6**	**9,061**	**2.8**
Rental of automobiles and trucks	0	0.0	0	0.0	102	0.1	256	0.2	293	0.2	477	0.1
Total, transportation	**6,007**	**28.5**	**12,260**	**27.7**	**25,259**	**26.0**	**49,626**	**34.8**	**52,596**	**30.6**	**96,445**	**30.0**
Total, exports	21,109		44,291		97,027		142,757		172,159		320,988	

Transportation commodities account for a slightly smaller share of Canadian imports, 25.9% as of 1996, with a low of 24.7% in the recession of 1981 and a high of 32.6% in the recovery year of 1986 (table 10). As it is for exports, the main component of imports is transportation equipment, accounting for 22.9% of imports in 1996, with a high of 29.6% in 1986 and a low of 21.7% in

TABLE 10 Transportation Commodities as a Share of Imports
Millions of Canadian dollars

	1971		1976		1981		1986		1991		1996	
Commodities	Dollars	%	Dollars	%	Dollars	%	Dollars	%	Dollars	%	Dollars	%
Transportation equipment												
Aircraft, parts, and repairs	–299	1.6	–425	0.9	–2,396	2.6	–3,111	2.3	–3,745	2.2	–4,973	1.8
Automobiles, including passenger vans	–1,217	6.4	–2,685	6.0	–5,451	5.9	–12,766	9.5	–12,501	7.4	–14,206	5.2
Trucks, road tractors, and chassis	–325	1.7	–840	1.9	–1,429	1.6	–3,071	2.3	–3,111	1.8	–5,618	2.1
Other complete equipment	–138	0.7	–438	1.0	–648	0.7	–895	0.7	–1,634	1.0	–2,847	1.1
Motor vehicle parts, including bodies and tires	–2,507	13.1	–5,782	12.9	–9,624	10.5	–19,091	14.3	–17,849	10.6	–33,275	12.3
Railroad equipment and parts	–50	0.3	–99	0.2	–205	0.2	–476	0.4	–392	0.2	–934	0.3
Shipbuilding and ship repair	–18	0.1	–68	0.2	–197	0.2	–179	0.1	–111	0.1	–95	0.0
Total	**–4,554**	**23.9**	**–10,337**	**23.0**	**–19,950**	**21.7**	**–39,589**	**29.6**	**–39,343**	**23.3**	**–61,948**	**22.9**
Transportation fuels												
Motor gasoline	–20	0.1	–1	0.0	–128	0.1	–419	0.3	–562	0.3	–787	0.3
Diesel and fuel oil, aviation fuel	–193	1.0	–228	0.5	–688	0.7	–1,171	0.9	–936	0.6	–1,019	0.4
Total	**–213**	**1.1**	**–229**	**0.5**	**–816**	**0.9**	**–1,590**	**1.2**	**–1,498**	**0.9**	**–1,806**	**0.7**
Transportation industries												
Air	–8	0.0	–19	0.0	–826	0.9	–1,234	0.9	–2,224	1.3	–3,136	1.2
Railway	0	0.0	0	0.0	–205	0.2	–229	0.2	–292	0.2	–237	0.1
Water	–134	0.7	–543	1.2	–327	0.4	–307	0.2	–411	0.2	–464	0.2
Truck	–28	0.1	–89	0.2	–215	0.2	–209	0.2	–500	0.3	–1,589	0.6
Surface passenger	0	0.0	0	0.0	–208	0.2	–312	0.2	–540	0.3	–639	0.2
Other transportation services	–12	0.1	–13	0.0	–23	0.0	–24	0.0	–24	0.0	–4	0.0
Total	**–182**	**1.0**	**–664**	**1.5**	**–1,804**	**2.0**	**–2,315**	**1.7**	**–3,991**	**2.4**	**–6,069**	**2.2**
Rental of automobiles and trucks	0	0.0	0	0.0	–107	0.1	–160	0.1	–281	0.2	–333	0.1
Total, transportation	**–4,949**	**26.0**	**–11,230**	**25.0**	**–22,677**	**24.7**	**–43,654**	**32.6**	**–45,113**	**26.8**	**–70,156**	**25.9**
Total, imports	–19,067		–44,899		–91,940		–133,937		–168,606		–270,870	

1981. In all years, the main components of transportation equipment imports are motor vehicle parts and automobiles, reflecting Canada's role in transportation equipment manufacturing.

The second largest share of transportation imports is commercial transportation (transportation industries), accounting for 2.2% of imports in 1996, with a high of 2.4% in 1991 and low of 1.0% in 1971. The increasing share of commercial transportation relative to the 1970s derives primarily from greater imports of air and trucking services. There are no TMs in imports, as import factor prices are calculated from the border rather than from the factory gate.

Indirect Transportation Demand

Indirect transportation demand refers to demand for transportation commodities, which is embedded in the price of the nontransportation commodities that make up part of final demand. For example, some portion of the price of shoes is accounted for by the transportation commodities used in the domestic production and distribution of shoes. Indirect demand can be divided into two categories, indirect domestic demand (*IDD*) and exports.

Government spending on transportation, as opposed to investment, is contained within *IDD*. In the CNA, government is treated as an industry that primarily produces services. Thus, government services commodities are listed primarily as a single entry in *FD*, under the column government spending. In order to produce services, the government uses inputs such as transportation commodities (table 11). Detailed government inputs are listed in the *U*, not the *FD* matrix, and thus transportation commodities used to produce government services will appear in *IDD* and not direct *FD*.

Indirect Transportation in Domestic Demand

Indirect transportation-related domestic demand (*ITDD*) accounted for 2.6% of domestic demand (*DD*) in 1996, the lowest level of all years assessed, with a high of 3.6% in 1981, again reflecting the price spike in fuels (table 12). As of 1996, transportation industries represented the largest component of *ITDD* at 1.0% of *DD*, with a surprisingly even trend over time that extends through all categories of commercial transportation. The two largest components of commercial transportation were air and surface passenger transportation, with surface passenger transportation consisting mainly of taxicab use by businesses and ambulance and school bus transport as part of government services. As with direct demand, the most interesting trends related to commercial transportation are found in the TMs, which accounted for 0.4% of *DD* in 1996, a steady decline from 0.9 % in 1971.

As of 1996, the second largest component of *ITDD* was transportation equipment, primarily motor vehicle parts, at 0.6% of *ITDD*, with again a relatively constant trend over time. Fuel was the most volatile component of *ITDD*, accounting for 0.4% in 1996 and 1971, with a high of 1.1% in 1981.

Indirect Exports

In terms of exports, TMs represent the largest share of indirect transportation demand, accounting for 0.8% of exports in 1996, a steady decline from 1.2% in 1971 (table 13). Commercial transportation accounts for 0.5%, a slight decline from 0.7% in 1971, with pipeline transportation as the leading component. Transportation equipment, again primarily motor vehicle parts, accounts for 0.3% of indirect demand from exports, with a relatively steady trend. Indirect demand for fuel exports is volatile, ranging from a high of 0.9% in 1981 to a low of 0.3% in 1996.

Total Transport Demand as a Share of GDP

The previous sections have provided a detailed assessment of the share of direct and indirect transportation as a proportion of the relevant components of final demand. This section aggregates the commodities presented earlier to generate estimates of transportation as a share of GDP and domestic demand. The detailed descriptions of the different components of transportation are aggregated to transportation equipment, fuel, construction, industries, margins, other transportation services (trade margins, repairs, automobile rental services), and indirect taxes.

The share of transportation in GDP has been relatively stable over the time periods selected, accounting for 20.7% of GDP in 1996 and 1971, with a high of 21.1% associated with the fuel price peak of 1981, and low of 19.1% in the restructuring recession of 1991 (table 14).

Several broad trends can be discerned, the most important of which is the increasing trade related to transportation. In 1996, transportation-related exports were the largest single component of transport demand at 14.2% of GDP, going from a low of 7.6% of GDP in 1971. Particularly strong growth occurred from 1991 (10.0%) to 1996, and from 1981 (8.8%) to 1986 (12.3%), indicating a business cycle trend. Imports also grew, but at a slower pace, from –6.3% in 1971 to –10.3% in 1996, which indicates an increasing trade surplus. While the transportation equipment category dominates trade, transportation industries also show an increasing level of trade, with exports accounting for 1.3% of GDP in 1996, having moved steadily

TABLE 11 Government Spending on Transportation as a Share of Total Inputs
Millions of Canadian dollars

Government spending	1971 Dollars	1971 %	1976 Dollars	1976 %	1981 Dollars	1981 %	1986 Dollars	1986 %	1991 Dollars	1991 %	1996 Dollars	1996 %
Transportation equipment												
Aircraft, parts, and repairs	89	0.4	167	0.3	531	0.6	980	0.7	695	0.4	663	0.3
Other complete equipment	12	0.1	72	0.1	94	0.1	144	0.1	165	0.1	255	0.1
Shipbuilding and ship repair	73	0.3	32	0.1	109	0.1	368	0.3	711	0.4	304	0.1
Total	**174**	**0.7**	**271**	**0.5**	**734**	**0.8**	**1,492**	**1.1**	**1,571**	**0.8**	**1,222**	**0.6**
Transportation fuels												
Motor gasoline	44	0.2	97	0.2	260	0.3	224	0.2	221	0.1	233	0.1
Diesel and fuel oil, aviation fuel	61	0.3	252	0.5	584	0.7	445	0.3	397	0.2	366	0.2
Total	**105**	**0.4**	**349**	**0.7**	**844**	**0.9**	**669**	**0.5**	**618**	**0.3**	**599**	**0.3**
Transportation construction												
Highway and bridge maintenance	22	0.1	48	0.1	133	0.1	148	0.1	238	0.1	216	0.1
Total	**22**	**0.1**	**48**	**0.1**	**133**	**0.1**	**148**	**0.1**	**238**	**0.1**	**216**	**0.1**
Transportation industries												
Air	3	0.0	4	0.0	42	0.0	90	0.1	106	0.1	80	0.0
Water	10	0.0	11	0.0	23	0.0	21	0.0	37	0.0	24	0.0
Railway	2	0.0	11	0.0	3	0.0	1	0.0	1	0.0	2	0.0
Truck	52	0.2	105	0.2	191	0.2	256	0.2	286	0.1	254	0.1
Surface passenger	156	0.7	274	0.5	723	0.8	1,001	0.7	1,543	0.8	1,695	0.8
Other transportation services	0	0.0	1	0.0	1	0.0	2	0.0	14	0.0	6	0.0
Pipeline	12	0.1	37	0.1	61	0.1	154	0.1	163	0.1	177	0.1
Total	**235**	**1.0**	**443**	**0.9**	**1,044**	**1.2**	**1,525**	**1.1**	**2,150**	**1.1**	**2,238**	**1.0**
Transportation margins	**24**	**0.1**	**54**	**0.1**	**91**	**0.1**	**145**	**0.1**	**151**	**0.1**	**158**	**0.1**
Rental of automobiles and trucks	40	0.2	100	0.2	169	0.2	212	0.2	300	0.2	349	0.2
Total, transportation	**600**	**2.5**	**1,265**	**2.5**	**3,015**	**3.4**	**4,191**	**3.1**	**5,028**	**2.6**	**4,782**	**2.2**
Total, government inputs	23,694		50,091		89,697		133,658		195,471		213,483	

up from a low of 0.4% in 1971, with comparable figures for imports being –0.9% and –0.2%. The direct export associated TMs have maintained a fairly steady level over time, accounting for 1.3% of GDP in 1996, as have the indirect exports TMs (0.4% in 1996). This indicates that the decreases in TMs as a share of exports over time, as indicated in earlier sections on exports and indirect exports, are compensated by the increasing volume of exports.

All of these trade-related trends are consistent with the growth in importance of trade in the Canadian economy, particularly that associated with the advent of the North American Free Trade Agreement. Table 15 presents exports and imports as a share of GDP for 1971 to 1996, showing

TABLE 12 Transportation Commodities as an Indirect Share of Domestic Demand
Millions of Canadian dollars

Indirect transport, domestic demand	1971 Dollars	1971 %	1976 Dollars	1976 %	1981 Dollars	1981 %	1986 Dollars	1986 %	1991 Dollars	1991 %	1996 Dollars	1996 %
Transportation equipment												
Aircraft, parts, and repairs	87	0.1	163	0.1	508	0.2	923	0.2	674	0.1	659	0.1
Automobiles, including passenger vans	0	0.0	0	0.0	0	0.0	0	0.0	1	0.0	0	0.0
Trucks, road tractors, and chassis	1	0.0	6	0.0	9	0.0	0	0.0	13	0.0	3	0.0
Other complete equipment	13	0.0	72	0.0	89	0.0	135	0.0	161	0.0	303	0.0
Motor vehicle parts, including bodies and tires	378	0.5	601	0.4	851	0.3	1,290	0.3	1,594	0.3	2,232	0.4
Railroad equipment and parts	4	0.0	11	0.0	11	0.0	25	0.0	18	0.0	13	0.0
Shipbuilding and ship repair	70	0.1	35	0.0	115	0.0	349	0.1	663	0.1	286	0.0
Total	**553**	**0.7**	**887**	**0.6**	**1,584**	**0.6**	**2,721**	**0.7**	**3,124**	**0.6**	**3,495**	**0.6**
Transportation fuels												
Motor gasoline	132	0.2	348	0.2	1,527	0.5	1,139	0.3	1,281	0.2	1,251	0.2
Diesel and fuel oil, aviation fuel	179	0.2	649	0.4	1,447	0.5	1,060	0.3	1,179	0.2	1,193	0.2
Total	**312**	**0.4**	**998**	**0.6**	**2,974**	**1.1**	**2,199**	**0.6**	**2,460**	**0.5**	**2,444**	**0.4**
Transportation construction												
Highway and bridge maintenance	21	0.0	46	0.0	129	0.0	146	0.0	231	0.0	208	0.0
Total	**21**	**0.0**	**46**	**0.0**	**129**	**0.0**	**146**	**0.0**	**231**	**0.0**	**208**	**0.0**
Transportation industries												
Air	200	0.3	396	0.3	791	0.3	1,141	0.3	1,574	0.3	2,095	0.3
Railway	42	0.1	77	0.0	104	0.0	131	0.0	135	0.0	182	0.0
Water	44	0.1	76	0.0	118	0.0	127	0.0	129	0.0	147	0.0
Truck	182	0.2	329	0.2	527	0.2	560	0.1	668	0.1	775	0.1
Surface passenger	207	0.3	380	0.2	827	0.3	1,183	0.3	1,849	0.4	1,994	0.3
Other transportation services	22	0.0	53	0.0	161	0.1	234	0.1	407	0.1	667	0.1
Pipeline	118	0.2	219	0.1	279	0.1	539	0.1	562	0.1	618	0.1
Total	**815**	**1.1**	**1,529**	**1.0**	**2,806**	**1.0**	**3,915**	**1.0**	**5,324**	**1.0**	**6,478**	**1.0**
Transportation margins	**690**	**0.9**	**1,314**	**0.8**	**1,965**	**0.7**	**2,550**	**0.6**	**2,453**	**0.5**	**2,495**	**0.4**
Rental of automobiles and trucks	146	0.2	391	0.2	713	0.3	1,039	0.3	1,332	0.3	1,522	0.2
Total, transportation	**2,538**	**3.3**	**5,166**	**3.3**	**10,171**	**3.6**	**12,570**	**3.2**	**14,923**	**2.8**	**16,643**	**2.6**
Total, domestic demand	76,552		158,232		281,149		394,963		524,414		628,104	

TABLE 13 Transportation Commododities as an Indirect Share of Exports
Millions of Canadian dollars

Indirect transport, exports	1971 Dollars	1971 %	1976 Dollars	1976 %	1981 Dollars	1981 %	1986 Dollars	1986 %	1991 Dollars	1991 %	1996 Dollars	1996 %
Transportation equipment												
Aircraft, parts, and repairs	3	0.0	4	0.0	10	0.0	17	0.0	30	0.0	85	0.0
Automobiles, including passenger vans	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Trucks, road tractors, and chassis	1	0.0	1	0.0	1	0.0	0	0.0	1	0.0	3	0.0
Other complete equipment	7	0.0	10	0.0	1	0.0	1	0.0	3	0.0	163	0.1
Motor vehicle parts, including bodies and tires	38	0.2	89	0.2	195	0.2	277	0.2	310	0.2	561	0.2
Railroad equipment and parts	3	0.0	6	0.0	6	0.0	20	0.0	14	0.0	12	0.0
Shipbuilding and ship repair	6	0.0	12	0.0	25	0.0	27	0.0	41	0.0	52	0.0
Total	**57**	**0.3**	**122**	**0.3**	**238**	**0.2**	**342**	**0.2**	**399**	**0.2**	**875**	**0.3**
Transportation fuels												
Motor gasoline	50	0.2	91	0.2	299	0.3	241	0.2	258	0.2	337	0.1
Diesel and fuel oil, aviation fuel	96	0.5	248	0.6	611	0.6	374	0.3	455	0.3	660	0.2
Total	**146**	**0.7**	**339**	**0.8**	**909**	**0.9**	**615**	**0.4**	**713**	**0.4**	**997**	**0.3**
Transportation construction												
Highway and bridge maintenance	0	0.0	0	0.0	1	0.0	2	0.0	2	0.0	3	0.0
Total	**0**	**0.0**	**0**	**0.0**	**1**	**0.0**	**2**	**0.0**	**2**	**0.0**	**3**	**0.0**
Transportation industries												
Air	10	0.0	23	0.1	74	0.1	109	0.1	142	0.1	252	0.1
Railway	9	0.0	17	0.0	34	0.0	44	0.0	45	0.0	59	0.0
Water	11	0.1	21	0.0	35	0.0	50	0.0	70	0.0	118	0.0
Truck	53	0.3	84	0.2	115	0.1	162	0.1	248	0.1	442	0.1
Surface passenger	3	0.0	6	0.0	17	0.0	30	0.0	48	0.0	69	0.0
Other transportation services	1	0.0	3	0.0	16	0.0	24	0.0	41	0.0	87	0.0
Pipeline	68	0.3	109	0.2	179	0.2	325	0.2	387	0.2	643	0.2
Total	**156**	**0.7**	**264**	**0.6**	**470**	**0.5**	**744**	**0.5**	**980**	**0.6**	**1,670**	**0.5**
Transportation margins	**254**	**1.2**	**454**	**1.0**	**911**	**0.9**	**1,289**	**0.9**	**1,433**	**0.8**	**2,608**	**0.8**
Rental of automobiles and trucks	30	0.1	62	0.1	108	0.1	177	0.1	222	0.1	403	0.1
Total, transportation	**644**	**3.1**	**1,240**	**2.8**	**2,638**	**2.7**	**3,168**	**2.2**	**3,750**	**2.2**	**6,558**	**2.0**
Total, exports	21,109		44,291		97,027		142,757		172,159		320,988	

TABLE 14 Transportation Demand as a Proportion of GDP and Domestic Demand
Millions of Canadian dollars

	1971		1976		1981		1986		1991		1996	
Commodities	**Dollars**	**%**	**Dollars**	**%**	**Dollars**	**%**	**Dollars**	**%**	**Dollars**	**%**	**Dollars**	**%**
Transportation consumption												
Transportation equipment	2,340	3.0	4,709	3.0	7,579	2.6	13,154	3.3	14,836	2.8	19,523	2.9
Transportation fuel	821	1.0	2,335	1.5	5,715	2.0	4,602	1.1	5,020	1.0	5,829	0.9
Transportation construction	33	0.0	46	0.0	66	0.0	55	0.0	112	0.0	95	0.0
Transportation industries	1,067	1.4	2,085	1.3	4,297	1.5	6,161	1.5	8,313	1.6	9,714	1.4
Transportation margins	553	0.7	1,049	0.7	1,553	0.5	2,137	0.5	2,136	0.4	2,264	0.3
Other transportation services	2,366	3.0	4,463	2.8	7,736	2.7	12,529	3.1	17,325	3.3	21,015	3.1
Indirect taxes	1,260	1.6	2,622	1.7	3,924	1.4	7,455	1.8	11,037	2.1	13,517	2.0
Total	**8,440**	**10.7**	**17,309**	**11.0**	**30,870**	**10.8**	**46,093**	**11.4**	**58,779**	**11.1**	**71,957**	**10.6**
Transportation investment												
Transportation equipment	1,368	1.7	3,326	2.1	6,609	2.3	7,901	2.0	8,393	1.6	12,640	1.9
Transportation construction	2,053	2.6	2,927	1.9	5,922	2.1	5,686	1.4	7,756	1.5	6,963	1.0
Transportation margins	102	0.1	229	0.1	399	0.1	470	0.1	488	0.1	379	0.1
Other transportation services	26	0.0	63	0.0	102	0.0	133	0.0	239	0.0	366	0.1
Indirect taxes	53	0.1	56	0.0	117	0.0	128	0.0	113	0.0	138	0.0
Total	**3,602**	**4.6**	**6,601**	**4.2**	**13,149**	**4.6**	**14,318**	**3.5**	**16,989**	**3.2**	**20,486**	**3.0**
Transportation exports												
Inventory and scrap	–47	–0.1	170	0.1	992	0.3	1,247	0.3	–1,019	–0.2	–1,238	–0.2
Transportation equipment	4,529	5.8	9,310	5.9	16,915	5.9	38,782	9.6	38,638	7.3	74,853	11.0
Transportation fuel	85	0.1	294	0.2	1,346	0.5	1,256	0.3	2,269	0.4	3,027	0.4
Transportation industries	326	0.4	608	0.4	2,626	0.9	3,460	0.9	5,194	1.0	9,027	1.3
Transportation margins	1,067	1.4	2,048	1.3	4,270	1.5	5,872	1.5	6,202	1.2	9,061	1.3
Other transportation services	0	0.0	0	0.0	102	0.0	256	0.1	293	0.1	477	0.1
Total	**6,007**	**7.6**	**12,260**	**7.8**	**25,259**	**8.8**	**49,626**	**12.3**	**52,596**	**10.0**	**96,445**	**14.2**
Transportation imports												
Transportation equipment	–4,554	–5.8	–10,337	–6.6	–19,950	–7.0	–39,589	–9.8	–39,343	–7.5	–61,948	–9.1
Transportation fuel	–213	–0.3	–229	–0.1	–816	–0.3	–1,590	–0.4	–1,498	–0.3	–1,806	–0.3
Transportation industries	–182	–0.2	–664	–0.4	–1,804	–0.6	–2,315	–0.6	–3,991	–0.8	–6,069	–0.9
Transportation margins	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Other transportation services	0	0.0	0	0.0	–107	-0.0	–160	-0.0	–281	–0.1	–333	–0.0
Total	**–4,949**	**–6.3**	**–11,230**	**–7.1**	**–22,677**	**–7.9**	**–43,654**	**–10.8**	**–45,113**	**–8.5**	**–70,156**	**–10.3**

(continues)

TABLE 14 Transportation Demand as a Proportion of GDP and Domestic Demand *(Continued)*
Millions of Canadian dollars

	1971		1976		1981		1986		1991		1996	
Commodities	**Dollars**	**%**	**Dollars**	**%**	**Dollars**	**%**	**Dollars**	**%**	**Dollars**	**%**	**Dollars**	**%**
Indirect transportation domestic demand												
Transportation equipment	553	0.7	887	0.6	1,584	0.6	2,721	0.7	3,124	0.6	3,495	0.5
Transportation fuel	312	0.4	998	0.6	2,974	1.0	2,199	0.5	2,460	0.5	2,444	0.4
Transportation construction	21	0.0	46	0.0	129	0.0	146	0.0	231	0.0	208	0.0
Transportation industries	815	1.0	1,529	1.0	2,806	1.0	3,915	1.0	5,324	1.0	6,478	1.0
Transportation margins	690	0.9	1,314	0.8	1,965	0.7	2,550	0.6	2,453	0.5	2,495	0.4
Other transportation services	146	0.2	391	0.2	713	0.2	1,039	0.3	1,332	0.3	1,522	0.2
Total	**2,538**	**3.2**	**5,166**	**3.3**	**10,171**	**3.6**	**12,570**	**3.1**	**14,923**	**2.8**	**16,643**	**2.5**
Indirect transportation exports												
Transportation equipment	57	0.1	122	0.1	238	0.1	342	0.1	399	0.1	875	0.1
Transportation fuel	146	0.2	339	0.2	909	0.3	615	0.2	713	0.1	997	0.1
Transportation construction	0	0.0	0	0.0	1	0.0	2	0.0	2	0.0	3	0.0
Transportation industries	156	0.2	264	0.2	470	0.2	744	0.2	980	0.2	1,670	0.2
Transportation margins	254	0.3	454	0.3	911	0.3	1,289	0.3	1,433	0.3	2,608	0.4
Other transportation services	30	0.0	62	0.0	108	0.0	177	0.0	222	0.0	403	0.1
Total	**644**	**0.8**	**1,240**	**0.8**	**2,638**	**0.9**	**3,168**	**0.8**	**3,750**	**0.7**	**6,558**	**1.0**
Total, transportation	**16,234**	**20.7**	**31,515**	**20.0**	**60,402**	**21.1**	**83,368**	**20.6**	**100,905**	**19.1**	**140,695**	**20.7**
Final demand	78,594		157,624		286,236		403,783		527,967		678,222	
Total, domestic transportation	**12,141**	**15.9**	**24,419**	**15.4**	**45,920**	**16.3**	**62,272**	**15.8**	**75,462**	**14.4**	**92,202**	**14.7**
Domestic demand	76,552		158,232		281,149		394,963		524,414		628,104	

exports growing from 26.9% of GDP in 1971 to 47.3% in 1996, with particularly high growth from 1991 to 1996.

A second major trend is the decline in transportation-related investment as a share of GDP, down from 4.6% of GDP in 1971 and 1981 to a low of 3.0% in 1996 (see table 14). This is due to a decline in transportation construction as a share of GDP, going from 2.6% of GDP in 1971 to 1.0% in 1996. This results from the ongoing consolidation of railway construction capital stock and a lower level of government investment in roads as a share of GDP, possibly indicating a mature transportation infrastructure. This relative decline may also be associated with increased investment in information and communications technology (ICT), as investment flows from mature industries, such as transportation, to new and growing industries such as ICT. Alternatively, the decline in transportation investment may point to an infrastructure investment deficit, particularly for road infrastructure, where the government may not have been investing sufficiently to meet the increased demand for road infrastructure, as indicated by the growth in consumption and investment in road transportation equipment as well as trucking.

A third major trend is the decline in the TMs associated with domestic demand, both direct and indirect. In consumption, TMs show a steady

TABLE 15 Exports and Imports as a Percentage of GDP: 1971–1996
Millions of Canadian dollars

	1971		1976		1981		1986		1991		1996	
	Dollars	**%**	**Dollars**	**%**	**Dollars**	**%**	**Dollars**	**%**	**Dollars**	**%**	**Dollars**	**%**
Exports	21,109	26.9	44,291	28.1	97,027	33.9	142,757	35.4	172,159	32.6	320,988	47.3
Imports	19,067	24.3	44,899	28.5	91,940	32.1	133,937	33.2	168,606	31.9	270,870	39.9
GDP	**78,594**		**157,624**		**286,236**		**403,783**		**527,967**		**678,222**	

Source: Statistics Canada, National Income and Expenditures Accounts.

decline from 0.7% of GDP in 1971 to 0.3% in 1996, while for indirect domestic demand the TMs declined from 0.9% to 0.4% (table 14). This correlates with the increasing share of trucking in the TMs, as discussed above. A possible explanation for this trend is the growth in the efficiency of the freight transportation industries, possibly stemming from deregulation, as well as the advent of more efficient supply chain management practices, such as just-in-time production and distribution. Table 16 shows the relative change in freight rail and trucking prices, gross output (or revenues), and total factor productivity relative to the economy from 1981 to 1996. However, these price, output, and productivity figures cannot explain the relative growth in trucking compared to rail, as rail has exhibited both higher productivity gains and larger declines in prices.

Together these three trends explain the relative stability of transportation as a share of GDP, with the increasing trends in trade, particularly the surplus in equipment, compensating for the decreasing trends in investment and the TMs. A look at table 14 again

TABLE 16 Price, Output, and Productivity Changes in Rail and Trucking Relative to the Economy: 1981–1996
Index, 1981 = 100

	1981	1986	1991	1996
Prices				
Freight rail[1]	100.0	95.9	79.5	68.1
Trucking	100.0	96.8	82.9	74.9
Output				
Freight rail	100.0	88.1	80.4	77.2
Trucking	100.0	107.9	112.9	158.6
Productivity				
Freight rail	100.0	104.7	128.4	146.8
Trucking	100.0	102.3	120.8	134.3

[1] Freight rail refers to the two largest rail lines in Canada: Canadian Pacific and Canadian National.

Source: Transport Canada, internal statistics.

shows that because trade is excluded from domestic demand, transportation as a share of domestic demand is much lower (14.7% in 1996) than it is as a share of GDP (20.7%), with a similar pattern of volatility (e.g., a low of 14.4% in 1991 and a high of 16.3% in 1981). A possible declining trend can also be observed with the two transportation shares recorded in the 1990s (1991, 14.4%; 1996, 14.7%) representing the lowest shares of domestic demand of all years. This may reflect competition from new consumer durables, such as personal computers.

Overlying these three trends as the major determinant of transportation as a share of GDP is the volatility of fuel as a share of GDP, which can be particularly associated with the market strategies of the OPEC oil cartel. The peak in fuel shares in 1981 (e.g., 2.0% of GDP in consumption and 1.0% in indirect domestic demand), correlates with the peaks in transportation share of GDP and domestic demand (1981 was the most atypical of all years surveyed), along with the restructuring recession of 1991. The trend in fuel shares has followed a triangular pattern over the years assessed, with low and similar fuel shares of GDP in 1971 and 1996, and a peak in 1981; the total transportation share of GDP in 1971 and 1996 was also similar and shows a peak in 1981.

One of the advantages of using a commodity-based classification of transportation demand, as defined in the IO tables, is that it allows for different levels of macroeconomic analysis. Table 17 illustrates an aggregation of the demand for transportation by the different types of commodities, divided in a standard manner into goods, services, and indirect taxes. As can be noted, transportation is fairly evenly split between transportation goods and services, with a slight predominance of transportation goods over services up until 1986 and then again in 1996, corresponding to the growth in equipment exports. The largest transportation commodity is equipment,

TABLE 17 Transportation Demand by Categorization as Goods and Services
Millions of Canadian dollars

	1971		1976		1981		1986		1991		1996	
	Dollars	%	Dollars	%	Dollars	%	Dollars	%	Dollars	%	Dollars	%
Transportation goods												
Transportation equipment	4,294	26.4	8,017	25.4	12,975	21.5	23,311	28.0	26,047	25.8	49,439	35.1
Transportation fuel	1,151	7.1	3,736	11.9	10,128	16.8	7,081	8.5	8,964	8.9	10,491	7.5
Transportation construction	2,107	13.0	3,019	9.6	6,118	10.1	5,888	7.1	8,101	8.0	7,269	5.2
Inventory and scrap	–47	–0.3	170	0.5	992	1.6	1,247	1.5	–1,019	–1.0	–1,238	–0.9
Total	**7,505**	**46.2**	**14,942**	**47.4**	**30,213**	**50.0**	**37,528**	**45.0**	**42,093**	**41.7**	**65,962**	**46.9**
Transportation services												
Transportation industries	2,183	13.4	3,822	12.1	8,395	13.9	11,964	14.4	15,820	15.7	20,820	14.8
Transportation margins	2,666	16.4	5,094	16.2	9,098	15.1	12,319	14.8	12,712	12.6	16,808	11.9
Other transportation services	2,568	15.8	4,979	15.8	8,654	14.3	13,974	16.8	19,130	19.0	23,450	16.7
Total	**7,416**	**45.7**	**13,895**	**44.1**	**26,147**	**43.3**	**38,257**	**45.9**	**47,661**	**47.2**	**61,078**	**43.4**
Indirect taxes	**1,313**	**8.1**	**2,678**	**8.5**	**4,041**	**6.7**	**7,583**	**9.1**	**11,150**	**11.1**	**13,655**	**9.7**
Total, transportation	**16,234**		**31,515**		**60,402**		**83,368**		**100,905**		**140,695**	

accounting for a high of 35.1% of total transportation demand in 1996, with a low corresponding to the 1981 fuel price peak (21.5%). The second largest category (other transportation services) is also primarily associated with equipment, notably trade margins and equipment repairs in transportation consumption. Three of the trends discussed above are also highlighted in looking at the demand for transportation by commodity class: the impact and volatility of fuel prices, the declining share of transportation construction, and the steadily declining share of the TMs.

CONCLUSION

This paper has used the IO tables maintained in the CNA to assess the share of transportation-related demand, both direct and indirect, in Canadian GDP from 1971 to 1996. The industry and commodity classification used in this paper is from the standard industrial classification system that was specific to the Canadian national accounts. This classification system has now been replaced by the NAICS, which will be common to the Canadian, U.S., and Mexican national accounts. NAICS should allow for future work involving a similar methodology to compare trends in the relative share of transportation in the national economies of Canada, Mexico, and the United States. Another future development related to IO data that would enable a refining of these estimates, while using a similar methodology, is the integration of one or more transport satellite accounts, such as private trucking, within the IO tables.

REFERENCES

Han, X. and B. Fang. 1998. Measuring Transportation in the U.S. Economy. *Journal of Transportation and Statistics* 1(1).

Fang, B., X. Han, A. Lawson, and S. Okubo. 2000. U.S. Transportation Satellite Accounts for 1996. *Survey of Current Business,* May.

Statistics Canada. 1989. *A User's Guide to the System of National Accounts.*

ACKNOWLEDGMENT

The author would like to thank the two anonymous referees for their helpful comments, as well as Dr. Xiaoli Han.

Author's address: Jeff Harris, Senior Advisor, Economic Analysis, Place de Ville, Tower C, Ottawa, Ontario, Canada, K1A 0N5. Email: harrijb@tc.gc.ca.

Effects of Alcohol and Highway Speed Policies on Motor Vehicle Crashes Involving Older Drivers

PATRICK McCARTHY
Georgia Institute of Technology

ABSTRACT

This paper explores aspects of highway safety with a focus on crashes involving older drivers. As the "baby boomers" age and move into retirement, a larger proportion of older drivers will be using the nation's roads. The analysis here develops and estimates econometric models using a panel dataset that includes each county in California and spans an 18-year period, 1981–1998. The models are estimated using feasible generalized least squares techniques that account for cross-section heterogeneity, adjust for county-specific first-order serial correlation, and correct for nonconstant variances due to the large differences in county sizes across the state. The results indicate that the set of explanatory variables for crashes involving older drivers is not identical to the set for crashes involving younger drivers. Among the factors that have large effects on older driver crashes are risk exposure, energy and alcohol prices, alcohol availability, and increased speed limits on higher speed roads.

INTRODUCTION

As the U.S. population ages, the proportion of older drivers on our nation's highways increases. Between 1986 and 1996, the total number of older drivers

KEYWORDS: older driver, highway safety, public health, speed limits, alcohol policy.

grew by 45% in comparison with a 13% increase in the total number of licensed drivers. Although we would expect a higher proportion of older drivers to affect the distribution of crashes, existing literature does not address this subject in great detail. It is well known, for example, that older drivers have less exposure and, accordingly, fewer crashes. At the same time, physical disabilities, a greater consumption of legal medication, and slower reaction times have all led to higher crash rates among the elderly relative to their younger driver counterparts (McCloskey et al. 1994; Hu et al. 1998; Lundberg et al. 1998). In order to continue developing public policy that is relevant to current and future characteristics of highway users, policymakers can benefit by gaining a better understanding of the causes of crashes among older drivers and how changes in policies may affect the distribution of crashes in this group.

The broad objective of this research is to study the relationship between highway safety public policy and highway crashes among older drivers. More specifically, this study focuses on the policies related to increased speed limits and those affecting the monetary and time (i.e., resource) costs associated with driving under the influence of alcohol.

A number of papers have analyzed the deterrent effects of legislation that have increased the penalties associated with drinking and driving or the enforcement of drinking and driving laws. Among the more recent analyses, Chaloupka et al. (1993) found that restrictive administrative per se (APS) laws significantly reduced alcohol-related crashes. Under APS laws, a driver's license is revoked or suspended if his or her blood alcohol concentration (BAC) level equals or exceeds the state legal limits, a sanction that is independent of other penalties if the driver is convicted. Other per se sanctions (mandatory jail sentences and community service laws) were ineffective. Legge and Park (1994) concluded that APS and other per se laws have the largest impact on single-vehicle nighttime crashes, whereas laws requiring some jail time for a first conviction or a fine for a first offender had no effect. Zador et al. (1989) also found that APS and other per se laws reduced fatal crashes. Peck (1991) found California's license suspension program has been more effective than alcohol rehabilitation programs in reducing crash risk for driving under the influence (DUI). Rogers (1995) concluded that reductions in the BAC level were effective. In contrast to most of this literature, Evans et al. (1991) found no evidence that sanctions contained in "punitive legislation" were effective deterrents of traffic fatalities.

There is also a significant literature on the highway safety effects of raised speed limits on higher speed roads (e.g., U.S. Interstate highways and the German autobahn), much of which is summarized in two Transportation Research Board reports (1984, 1998). As a result of the 1970s energy crisis, maximum speed limits on U.S. Interstates were federally set in 1975 at 55 mph. Complementing the savings in energy, this policy reduced fatal crashes, and the 55 mph speed limit continued long after the energy crisis was over. In 1987, Congress passed legislation that permitted states to raise the speed limit to 65 mph on rural Interstate highways. U.S. speed limits were again relaxed in 1995, when Congress ceded to states the authority to set speed limits on higher speed roads. Existing literature indicates that higher speed limits increase average speeds, but the speed increase is typically less than the increase in limits (e.g., the 10 mph increase in 1987 speed limits generally resulted in a 4 mph increase in average speeds). The extensive literature on the 1987 increase and the available, but more minimal, literature for the 1995 relaxation indicate that higher limits produce more fatal crashes and fatalities on the affected roads (McCarthy 2000). However, there is much less certainty about the systemwide effects of raising speed limits on higher speed roads due to "tainting" and "diversion" effects.[1]

Notwithstanding the extensive literature on the effects of alcohol-related and speed limit policies, with some focus on the younger driver, these studies

[1] Speed limit tainting effects occur when changes in speed limits on one set of roads (e.g., Interstate highways) not only induce increased actual speeds on the affected roads but also lead to speed increases on non-affected roads (e.g., arterials), that is, roads where speed limits did not change. Speed limit diversion effects occur when changes in speed limits on one road divert traffic to other roads, including the affected road. For example, a higher speed limit on rural Interstates is expected to divert some traffic from slower roads where speed limits have not changed to the Interstate highway's increased speed limit.

TABLE 1 Definitions of Variables

Dependent variables	
Ttlge60	Total crashes involving drivers ≥ 60 years of age
Fatge60	Fatal crashes involving drivers ≥ 60 years of age
Injrge60	Injury crashes involving drivers ≥ 60 years of age
Pdoge60	Property damage only crashes involving drivers ≥ 60 years of age
Ttllt60	Total crashes involving drivers < 60 years of age
Fatlt60	Fatal crashes involving drivers < 60 years of age
Injrlt60	Injury crashes involving drivers < 60 years of age
Pdolt60	Property damage only crashes involving drivers < 60 years of age
Explanatory variables	
Rpcinc	County real per capita income: 1982–84 US $
Cpigas	Regional consumer price index for gasoline: 1982–84 = 100
Cpialc	Regional consumer price index for alcohol: 1982–84 = 100
Uerate	County unemployment rate
Vmtge60	County vehicle-miles traveled (vmt): drivers ≥ 60 years old
Vmtlt60	County vmt: drivers < 60 years old
Totvmt	Total county vmt
Popden	Persons per square mile in the county
Pctge60	Share of the county population ≥ 60 years old
Alclic	Number of retail alcohol licenses in the county
AB541	Dummy variable for an omnibus DUI prevention law (effective Jan. 1, 1982), which equals 0 for 1981 and 1 for 1982–98
APS	Dummy variable for an administrative per se license suspension law (effective July 1, 1990), which equals 0 for 1981–89 and 1 for 1990–98
Slmt_65	Dummy variable for 65 mph speed limits on rural Interstate highways (effective May 1987), which equals 0 for 1981–86 and 1 for 1987–98
Slmt_70	Dummy variable for 70 mph speed limits on portions of the state's Interstate highways (enacted in January 1997), which equals 0 for 1981–96 and 1 for 1997–98
Pc_dui	Per capita DUI arrests in the county
Pcmcyc	Per capita motorcycle registrations in the county

say little about the impact of highway safety policy on older drivers.

DATA

This analysis uses the state of California as a case study for evaluating highway crashes among older drivers. For this analysis, an older driver is defined as a driver who is 60 years of age or older. The cross-section unit of observation in the analysis is a California county and the time period is one year. In total, the analysis includes 58 counties over an 18-year period (1981 through 1998), resulting in a total of 1,024 observations. Among the sources of data for this study were the California Highway Patrol (crash data), the California Bureau of Criminal Statistics (arrest data), the California Department of Alcohol Beverage Control (alcohol license data), the California Department of Finance (demographic data, price indexes, per capita county income), and the California Department of Labor (unemployment rates).

Descriptive Statistics

Table 1 provides definitions of the variables used in this study. Table 2 gives descriptive statistics for the dependent variables.[2] Each of these variables is subdivided into two broad age groups—those involving

[2] The dependent variable is defined in levels (i.e., number of crashes) in order to estimate the sensitivity of older and younger driver-involved crashes to changes in vehicle-miles traveled (vmt) exposure, which is not possible when the dependent variable is defined as a rate (e.g., crashes per 100 million vmt). Tables 7 and 8 below show that crash sensitivity to vmt varies by age group as well as by crash type, which is discussed in more detail later in this paper.

TABLE 2 Descriptive Statistics: Dependent Variables

Variable		Mean	Standard deviation
Ttlge60	overall	1195.32	2527.11
	across-counties		2532.24
	across-years		280.26
Fatge60	overall	10.61	19.28
	across-counties		19.12
	across-years		3.48
Injrge60	overall	526.19	1257.33
	across-counties		1261.96
	across-years		119.57
Pdoge60	overall	658.52	1273.48
	across-counties		1269.68
	across-years		189.60
Ttllt60	overall	7492.10	17902.68
	across-counties		17949.75
	across-years		1887.81
Fatlt60	overall	60.70	122.32
	across-counties		120.56
	across-years		25.77
Injrlt60	overall	3103.00	7890.45
	across-counties		7885.63
	across-years		1043.86
Pdolt60	overall	4328.40	9988.92
	across-counties		10016.22
	across-years		1043.56

Key: Overall standard deviation: calculated across all groups and time periods. Across-counties standard deviation: obtained by first calculating the 18-year average for each county and then calculating the standard deviation across counties. Across-years standard deviation: obtained by first calculating, for each year, the average over all counties and then calculating the standard deviation for the 18-year period.

drivers 60 years of age or older (*older* drivers) and those involving drivers less than 60 years of age (*younger* drivers).

Included in table 2 is the overall mean of the variable in the sample and three measures of variation. The overall measure is the standard deviation across all groups and time periods. The "across-counties" measure is the standard deviation across the 58 counties (i.e., averaging over the 18-year period for each county and then calculating the measure of variation), and the "across-years" measure is the variation across the 18-year period (i.e., averaging all counties for each year and then calculating the measure of variation over the period).

Looking at the means, we see that younger drivers were involved in over six times as many crashes per county per year as older drivers, 7,492 versus 1,195. (The younger population is also about 6 times the size of the older population, but younger drivers account for 7.6 times as many vehicle-miles traveled (vmt) as older drivers.) This ratio was also true for fatal crashes. On average, there were 61 fatal crashes per county-year involving younger drivers compared with 11 fatal crashes involving older drivers. Nonfatal injury and property damage only (PDO) crashes show similar patterns between the two groups of drivers. Also not surprisingly, the overall and across-counties measures of variation are very similar, whereas the across-years variation is an order of magnitude smaller. An across-counties measure that virtually mimics the overall measure is not surprising, because it reflects the large variation in population and vmt across counties, as does the overall measure.

In contrast, the across-years measure reflects the average variation across the time period. We would expect that for any given county, the year-to-year variation in crashes would be relatively small, which is consistent with the results reported in table 2.

Table 3 provides summary statistics for the explanatory variables used in the econometric analysis.[3] For the entire panel, average per capita income was $13,736 and the annual countywide unemployment rate was relatively high at 10%. The mean price index for gasoline was close to its base index value of 100, whereas the mean price index for alcohol was 28% above the base level.

Looking at the exposure-related variables, annual estimated per-county travel for older drivers, who made up 16.8% of the population, averaged 501 million vmt, whereas estimated per-county average vmt for younger drivers was substantially higher at

[3] A correlation analysis indicates that the explanatory variables are not in general highly correlated. With two exceptions, correlation coefficients are less than 0.50. Vmt measures, not surprisingly, are highly correlated. Also, the number of retail alcohol licenses is highly correlated with exposure measures, which reflects the relationship between population and alcohol licenses.

TABLE 3 Descriptive Statistics: Explanatory Variables

		Mean	Standard deviation
Economic variables			
Rpcinc	overall	13736.4	3444.1
	across-counties		3327.3
	across-years		985.5
Cpigas	overall	98.93	9.89
	across-counties		1.01
	across-years		9.84
Cpialc	overall	128.73	24.30
	across-counties		3.42
	across-years		24.07
Uerate	overall	0.101	0.047
	across-counties		0.041
	across-years		0.023
Exposure variables			
Vmtge60	overall	5.01E+08	1.01E+09
	across-counties		1.00E+09
	across-years		1.53E+08
Vmtlt60	overall	3.80E+09	8.45E+09
	across-counties		8.42E+09
	across-years		1.27E+09
Pctge60	overall	0.168	0.042
	across-counties		0.041
	across-years		0.009
Popden	overall	560.27	2088.10
	across-counties		2103.30
	across-years		90.93
Alcohol variables			
Alclic	overall	14272.9	29331.9
	across-counties		29496.8
	across-years		2116.3
Pc_dui	overall	0.013	0.007
	across-counties		0.005
	across-years		0.005
Regulatory variables			
AB541	overall	0.944	0.229
	across-years		0.229
APS	overall	0.500	0.500
	across-years		0.500
Slmt_65	overall	0.667	0.472
	across-years		0.472
Slmt_70	overall	0.167	0.373
	across-years		0.373

Key:
Overall standard deviation: calculated across all groups and time periods.
Across-counties standard deviation: obtained by first calculating the 18-year average for each county and then calculating the standard deviation across counties.
Across-years standard deviation: obtained by first calculating, for each year, the average over all counties and then calculating the standard deviation for the 18-year period.

3.8 billion vmt. Population density averaged 560 persons per square mile.[4]

In addition to the price index for alcohol, the analysis includes two alcohol-related variables—the number of retail establishments selling alcohol and the number of misdemeanor and felony arrests for DUIs. Overall, there are over 14,000 licenses to sell alcohol per county, but the overall and across-counties standard deviations are quite high. This reflects the large variation in population, and, therefore, alcohol consumption, across counties. On the other hand, if we normalize for population, as is done for per capita DUI arrests, we see that, based on an overall mean of 0.013 arrests per capita, the across-counties and across-years variations are equal—0.005 in each case.

The model also includes four regulatory variables that are specific to driving under the influence of alcohol and to highway speeds. These variables do not have an across-counties measure, which reflects the fact that these are statewide rather than county laws so there is only temporal and no cross-section variation in these variables.

Assembly Bill 541 (AB541) was an omnibus bill that increased various penalties for driving under the influence of alcohol. Senate Bill 1623 (APS) enacted an administrative per se BAC level at 0.08 under which a driver's license is immediately revoked upon arrest for driving with a BAC equal to

[4] Because countywide vmt data do not exist for California, a methodology was developed to estimate age-gender vmt data by county from available aggregate annual vmt data for California. The procedure included the following steps: 1) using aggregate annual data, "vmt per driver" was regressed on a constant term, "% of statewide drivers ≥ 60," "% of statewide drivers ≤ 24," and "#persons per driver" (R^2 = 0.97 from this regression); 2) using countywide values for each of these explanatory variables, the regression model coefficients were used to predict the countywide vmt per driver; 3) countywide vmt was obtained by multiplying the number of drivers in each county by the estimated countywide vmt per driver. In order to allocate the estimated countywide vmt to various age-gender specific groups, data on the proportion of annual vmt per driver, disaggregated by alternative age-gender categories, were obtained from the Nationwide Personal Transportation Survey (NPTS) for 1983, 1990, and 1995. These data were then used to scale the estimated countywide vmt data. Specifically, NPTS data for 1983 (1990; 1995) were used to develop the weights for observations in this study during the period 1981–1986 (1987–1993; 1994–1998) to obtain vmt in each county for different age-gender groups.

or greater than 0.08 regardless of whether the person demonstrates any behavioral signs of alcohol impairment.[5] Respectively, AB541 and APS were active for 94% and 50% of the time period covered in this analysis.

In addition to these laws, California relaxed speed limits on its high-speed roads. In 1987, California increased speed limits on rural Interstates from 55 mph to 65 mph, and, on selected roads, it raised speed limits from 65 mph to 70 mph in 1995. These two laws were active 66% and 16.6% of the time during the 18-year period.

ECONOMETRIC METHODOLOGY

The econometric formulation for this study assumes that crashes among older drivers, y_{it}, are a function of a set of explanatory variables, x_{it}, and can be expressed as

$$\begin{aligned} y_{it} &= \alpha + \beta' x_{it} + u_i + e_{it} \\ i &= 1,\ldots,58;\ t = 1981,\ldots,1998 \end{aligned} \quad (1)$$

where i indexes the county and t indexes the year. α is a scalar parameter and β is a parameter vector, both of which are to be estimated. e_{it} is an error term assumed to have a mean of zero and a constant variance. For a given county, i, the term u_i is constant over time but is assumed to vary by county. Depending on the distribution assumption associated with u_i, the resulting statistical model will either be a fixed-effects or a random-effects model. If u_i ($i = 1,\ldots,I$) is assumed to be a fixed parameter, then, in addition to α and β, the model estimates each effect u_i (normalizing on one of the cross sections). The estimator for this model is also referred to as the across-years estimator because it is equivalent to estimating the

$$(y_{it} - \bar{y}_i) = (x_{it} - \bar{x}_{it})\beta + (e_{it} - \bar{e}_i),$$
$$(i = 1,\ldots, I;\ t = 1,\ldots, T).$$

Alternatively, if u_i is assumed to vary randomly (e.g., the cross-section units are a sample from a larger set of cross-section units), then the model's error term becomes ($u_i + e_{it}$), which is assumed to satisfy the standard assumptions of a zero mean, constant variance, and zero correlation with the explanatory variables. The choice between fixed effects and random effects specification generally revolves around the correlation between u_i and x_{it} in the random effects specification. In particular, if the random effect u_i is correlated with the vector of explanatory variables x_{it} (i.e., corr(u_i, x_{it}) ≠ 0), then the parameter estimates are unbiased, but the standard errors are biased and we have no confidence in our t-statistics.

How can we determine whether the independence assumption between u_i and x_{it} is reasonable? Theoretically, the issue hinges on whether there are unobserved time-invariant effects that are correlated with a subset of the included variables. In a fixed effects specification, u_i is a fixed parameter and the correlation between u_i and x_{it} does not affect the model's properties. However, in a random effects specification, u_i reflects time-invariant effects that become part of the error structure. If these effects are correlated with the set of explanatory variables, biases in the standard errors arise. As an example, consider a county's topography, which is time invariant. Counties with less mountainous terrains will likely have greater vmt. Thus, a random effects specification will not be able to determine the extent to which higher crashes are due to more traveling and how much are due to unobserved topography.

In general, u_i captures county heterogeneity, which means that it reflects the net effect of unobserved variables (e.g., topography in the above example or the presence of "through routes") on the dependent variable. Thus, in u_i's picking up the influence of a variable that is correlated with an explanatory variable in the model, the independence assumption is violated and a random effects specification is not valid.

In addition to a fixed or random effects specification, panel data may suffer from serially correlated errors if the time span is sufficiently long and from heteroskedastic errors if the cross-section units have different scales. It is likely that the dataset for this analysis includes both problems, given the 18-year time horizon for these data and the fact that some counties in California are heavily populated (e.g., Los Angeles, San Diego, and San Francisco), whereas other counties have considerably smaller populations (e.g., Alpine and Tulare counties) and, accordingly, many fewer highway crashes. To

[5] Under California's law, a 30-day temporary license is issued to allow for due process, which provides drivers with time to challenge the suspension.

account simultaneously for serial correlation and heteroskedasticity, a feasible generalized least squares (FGLS) approach provides an alternative estimator. The model for this estimator is

$$y_{it} = \alpha_i + \beta' x_{it} + e_{it} \qquad i = 1,..., \mathrm{I};\ t = 1,...,T$$
$$e_{it} = \rho_i e_{i,t-1} + \varepsilon_{it} \qquad i = 1,..., \mathrm{I};\ t = 1,...,T$$
$$\mathrm{var}(\varepsilon_{it}) = \sigma_i^2 \qquad i = 1,..., \mathrm{I};\ t = 1,...,T \qquad (2)$$

which accommodates fixed effects, first order serial correlation, and cross-section heteroskedasticity.[6]

For this study, fixed effects, random effects, and FGLS models were estimated. Since, theoretically and empirically, an FGLS specification provided the best fits to the data for older drivers, only these results will be reported.[7] For each of the FGLS models, the estimation results and marginal effects of selected variables will be discussed. In addition to reporting results for older driver involved crashes, similar results will be reported for highway crashes involving drivers less than 60 years of age.

Tables 4 and 5 report the estimation results for older drivers and tables 6 and 7 report comparable results for younger drivers.

Estimation Results for Older Drivers

For crashes involving older drivers, tables 4 and 5 report the FGLS estimation results. In addition to estimating separate autocorrelation coefficients for each county and county-specific variances, these models include a full set of fixed effects for 57 of the 58 counties in California, normalizing on Yuba County. With regard to serial correlation, note that average estimated ρ varies across crash types and there is significant variation across counties, as reflected in the range of estimates for each crash type. For example, the estimated range of correlation coefficients varied from –0.46 to 0.47 for fatal crashes (with an average of –0.02), whereas the estimated range was 0.18 to 0.95 for PDO crashes (with an average of 0.71). This indicates there is likely to be significant bias in the variance estimates if the serial correlation is ignored.

Overall, the results reported in table 4 column (a) are consistent with expectations. For total crashes, an increase in real per capita income (Rpcinc) or unemployment rate (Uerate) has relatively little effect on crashes involving older drivers. But to the extent that an effect is present, a weaker economy is seen to increase the frequency of crashes. On the other hand, the price indexes for alcohol (Cpialc) and gasoline (Cpigas) have strong negative effects on total crashes. If all else remains constant, an

TABLE 4 FGLS Regression: Involved Drivers in Total and Fatal Crashes, ≥ 60[1]

	Total crashes (a)		Fatal crashes (b)	
Variable	Coeff	*z*-stat	Coeff	*z*-stat
Rpcinc	–3.60E–04	–0.67	–1.00E–04	–1.45
Cpigas	–0.262	–4.35	–0.012	–1.48
Cpialc	–0.118	–1.98	–0.010	–1.28
Uerate	15.901	0.99	2.044	1.08
Vmtge60	3.91E–07	3.47	—	—
Vmtlt60	3.31E–08	1.71	—	—
Totvmt	—	—	2.83E–10	1.78
Popden	–0.169	–3.15	–2.33E–04	–0.18
Pctge60	147.575	3.14	5.054	0.78
Alclic	0.016	4.06	4.18E–04	4.53
AB541	–2.901	–1.80	0.033	0.14
APS	0.022	0.01	–0.071	–0.23
Slmt_65	2.062	1.24	0.713	3.66
Slmt_70	4.108	2.67	0.084	0.48
Pc_dui[2]	98.784	2.51	–19.779	–0.14
Constant	138.594	5.11	3.993	2.17
	ρ range: (–0.131, 0.929)		ρ range: (–0.465, 0.470)	
	mean ρ: 0.660		ρ mean : –0.022	
	Wald $\chi^2(71)$ = 31837.1		Wald $\chi^2(71)$ = 2416.6	
	Prob > $\chi^2(71)$ = 0.0		Prob > $\chi^2(71)$ = 0.0	

[1] Estimates of the constant term and fixed effects are not reported for 57 counties, normalizing on the 58th county, Yuba.

[2] Predicted Pc_dui was used in the fatal crashes equation.

[6] FGLS preserves the first time series observation in each cross section by applying the Prais-Winston transformation. See Baltagi (1995, p. 83).

[7] A commonly used alternative methodology is a negative binomial model that accounts for the overdispersion property typically seen in highway crash data (and the data for this analysis are no exception) by relaxing the mean-variance equivalence property of the Poisson model. Fixed effects negative binomial models generally account for cross-section heterogeneity and overdispersion but not heteroskedasticity and serial correlation. FGLS, on the other hand, accounts for cross-section heterogeneity, heteroskedasticity, and serial correlation but does not address the overdispersion problem. As will be seen later, the estimated serial correlation coefficients varied across crash types and varied significantly across counties, which indicates that in models that fail to adjust for serial correlation, the coefficient estimates are inefficient and the estimated variances of the coefficients will be biased.

TABLE 5 FGLS Regression: Involved Drivers in Injury and PDO Crashes, ≥ 60[1]

Variable	Injury crashes (a) Coeff	z-stat	PDO crashes (b) Coeff	z-stat
Rpcinc	–6.00E–04	–1.39	4.59E–04	1.39
Cpigas	–0.128	–2.99	–0.124	–3.39
Cpialc	–0.190	–4.72	0.015	0.41
Uerate	–9.786	–0.88	9.707	1.07
Vmtge60	1.89E–07	3.95	—	—
Vmtlt60	2.55E–08	3.30	—	—
Totvmt	—	—	4.15E–08	6.53
Popden	–8.07E–03	–0.26	–0.189	–3.79
Pctge60	67.756	2.00	72.219	2.55
Alclic	9.14E–03	5.51	9.88E–03	3.36
AB541	0.369	0.33	–2.147	–2.17
APS	6.513	4.19	–3.329	–2.27
Slmt_65	1.616	1.49	1.439	1.39
Slmt_70	1.388	1.38	1.317	1.37
Pc_dui	10.984	0.42	45.035	1.93
Constant	57.206	4.63	76.007	4.74
	ρ range: (–0.261, 0.863)		ρ range: (–0.179, 0.953)	
	mean ρ: 0.471		mean ρ: 0.710	
	Wald $\chi^2(71)$ = 47027.8		Wald $\chi^2(71)$ = 12211.8	
	Prob > $\chi^2(71)$ = 0.0		Prob > $\chi^2(71)$ = 0.0	

[1] Estimates of the constant term and fixed effects are not reported for 57 counties, normalizing on the 58th county, Yuba.

TABLE 6 FGLS Regression: Involved Drivers in Total and Fatal Crashes, < 60[1]

Variable	Total crashes (a) Coeff	z-stat	Fatal crashes (b) Coeff	z-stat
Rpcinc	2.67E–02	4.65	–9.87E–04	–6.11
Cpigas	–6.842	–8.90	–0.069	–2.89
Cpialc	–7.910	–8.12	–0.116	–5.30
Uerate	945.280	5.98	–9.812	–1.60
Vmtlt60	1.37E–07	2.73	—	—
log(Totvmt)	—	—	2.757	6.56
Popden	0.852	2.36	–3.70E–02	–5.56
Pctge60	–1878.053	–10.54	–23.971	–4.70
Alclic	0.279	23.77	3.95E–03	17.07
AB541	–52.419	–4.83	–2.261	–3.47
APS	250.065	8.58	1.965	2.32
Slmt_65	72.421	6.22	0.627	1.09
Slmt_70	–48.308	–3.66	–0.444	–0.82
Pc_dui[2]	–32947.750	–6.72	85.661	3.57
Pcmcyc	–2922.915	–4.92	–136.231	–5.41
Const	2060.190	9.66	–6.947	–0.78
	ρ range: (–0.035, 0.963)		range: (–0.133, 0.898)	
	mean ρ: 0.670		mean ρ: 0.381	
	Wald $\chi^2(37)$ = 13330.9		Wald $\chi^2(37)$ = 5764.4	
	Prob > $\chi^2(37)$ = 0.0		Prob > $\chi^2(37)$ = 0.0	

[1] Not reported are estimates of the constant term and fixed effects for 24 consolidated metropolitan statistical area counties—Alameda, Butte, Contra Costa, El Dorado, Kern, Los Angeles, Marin, Napa, Orange, Placer, Riverside, Sacramento, San Benito, San Bernardino, San Diego, San Francisco, San Luis Obispo, San Mateo, Santa Clara, Santa Cruz, Solano, Sonoma, Sutter, and Ventura.

[2] Predicted pc_dui was used in the total crashes equation.

increase in the consumption price of gasoline or alcohol reduces crashes involving older drivers.

As expected, risk exposure produces more crashes. From the results, we see that a 100 million vmt increase by older drivers (Vmtge60) leads to 39 more crashes involving older drivers. Further, we see that younger driver exposure has a similar effect on older driver crashes. Specifically, a 1 billion vmt increase in younger driver exposure (Vmtlt60) produces 33 additional total crashes involving older drivers. Also, the greater the share of the population that is older than 60 (Pctge60), the greater the number of crashes. Given that vmt is constant, this result is consistent with older drivers having reduced driving skills relative to their younger driver counterparts. However, as reported below, the share of the population that is older than 60 has a negative and significant impact on younger driver crashes. The positive sign for older (negative for younger) involved crashes suggests that this variable may be capturing an aspect of vmt exposure that is not reflected in the included aggregate vmt variable.[8] Increases in population density are also associated with fewer older driver crashes. A 100-person increase per square mile (Popden) leads to 17 fewer crashes.

The last set of variables in table 4 column (a) relates to statewide alcohol and highway speed policies. First, and consistent with other research, alcohol availability (Alclic) is detrimental to highway safety, producing 1.6 additional crashes per 100 increase in the number of licenses. Second, the model includes two major pieces of alcohol-related

[8] I would like to thank the editors and anonymous referees for suggesting this as a possible explanation for the systematically different results.

TABLE 7 FGLS Regression: Involved Drivers in Injury and PDO Crashes, < 60[1]

Variable	Injury crashes (a) Coeff	Injury crashes (a) z-stat	PDO crashes (b) Coeff	PDO crashes (b) z-stat
Rpcinc	–3.80E–03	–2.85	1.95E–02	5.07
Cpigas	–0.829	–4.59	–4.0889	–8.27
Cpialc	–0.537	–3.20	–4.643	–7.08
Uerate	–8.545	–0.22	637.308	6.21
Vmtlt60	2.89E–08	1.59	6.16E–08	1.72
Popden	0.650	4.96	0.590	2.00
Pctge60	–303.521	–5.69	–954.108	–7.59
Alclic	0.126	29.37	0.160	19.09
AB541	–28.288	–5.75	–23.961	–3.59
APS	29.428	4.80	146.169	7.67
Slmt_65	11.545	2.57	42.969	5.92
Slmt_70	–2.853	–0.66	–23.468	–2.77
Pc_dui[2]	–204.864	–0.93	–19481.330	–6.02
Pcmcyc	–61.949	–0.53	–1919.889	–4.55
Const	242.387	7.19	1131.578	7.90
	ρ range: (–0.196, 0.968)		ρ range: (0.058, 0.947)	
	mean ρ: 0.552		mean ρ: 0.742	
	Wald $\chi^2(37)$ = 23414.5		Wald $\chi^2(37)$ = 9247.1	
	Prob > $\chi^2(37)$ = 0.0		Prob > $\chi^2(37)$ = 0.0	

[1] Not reported are estimates of the constant term and fixed effects for 24 consolidated metropolitan statistical area counties—Alameda, Butte, Contra Costa, El Dorado, Kern, Los Angeles, Marin, Napa, Orange, Placer, Riverside, Sacramento, San Benito, San Bernardino, San Diego, San Francisco, San Luis Obispo, San Mateo, Santa Clara, Santa Cruz, Solano, Sonoma, Sutter, and Ventura.

[2] Predicted Pc_dui was used in the total crashes equation.

legislation passed in California. AB541 was an omnibus bill that raised the cost of driving under the influence of alcohol in various ways and was California's first major effort at reducing drinking-and-driving crashes. The net effect on crashes was in the desired direction, leading to an average of 2.9 fewer crashes per county-year. On the other hand, APS, which implemented a 0.08 administrative per se law, has had no identifiable effect on total crashes.

DUI enforcement, defined as per capita arrests (Pc_dui), is significant but has an unexpected positive sign. There are three possible explanations for this. First, per capita arrests are the product of the probability of being stopped and the probability of being arrested given that one is stopped. Thus, a positive sign could result if an increase in arrests per stopped driver (reflecting, for example, police targeting DUI drivers) is associated with a contemporaneous decline in overall traffic enforcement (which reduces the likelihood of being stopped). Second, the effect on total crashes may reflect a net distribution effect on the different types of crashes. By inducing drinking drivers at the margin to behave less recklessly (e.g., consume fewer drinks, drive slower), more stringent enforcement, for example, could result in fewer fatal crashes but an increasing number of nonfatal crashes. Third, there may be an endogeneity problem in that an increase in crashes today leads to higher enforcement. Using a Hausman specification test statistic to test for endogeneity, the null hypothesis that Pc_dui was exogenous could not be rejected.[9]

The last two variables in table 4 column (a) identify the estimated effect of higher speed limits on crashes involving older drivers. The 65 mph speed limit (Slmt_65) had a positive but statistically insignificant effect on total crashes, whereas the more recent relaxation of speed limits (Slmt_70) produced a strong effect, generating four additional older driver crashes.

Column (b) in table 4 reports the FGLS estimation results for fatal crashes involving older drivers. Although the results are similar, there are some interesting differences. First, the state of the economy, whether it is measured by per capita income (Rpcinc) or the unemployment rate (Uerate), has a similar and somewhat stronger effect on fatal crashes than total crashes. Moreover, the effects are symmetric. Gains in per capita income and reductions in the unemployment rate both reduce fatal crashes. In that increases in income allow drivers to

[9] For a discussion of the Hausman specification test, see Pindyck and Rubinfeld (1991, p. 303–304). Instrumental variables for per capita arrests included other explanatory variables in the model. As noted by the editors, in the present framework, endogeneity is a problem if an increase in the incidence of alcohol-related crashes leads to more enforcement, However, if the level of enforcement is determined independently by the extent of drinking and driving rather than crashes, then enforcement is an instrumental variable for drinking and driving and the positive sign is consistent with expectations. It is likely that both effects are operating to some degree, which is consistent with the mixed results obtained from the Hausman specification test.

spend more on safer vehicles, better tires, and so forth to increase their level of safety (assuming that safety is a normal good) but drive faster and hence less safely (reflecting an increase in the value of time), the negative sign on Rpcinc implies that the safety effect dominates. And increases in gasoline (Cpigas) and alcohol prices (Cpialc) lead to fewer fatal crashes involving older drivers.

As expected, greater risk exposure produces more fatal crashes, where the results indicate that an increase of 10 billion vmt in total vmt produces an additional 2.8 fatal crashes among older drivers.[10] But in contrast to the results in table 4 column (a), neither the share of older drivers nor the population density has significant effects on fatal crashes.

Turning to the regulatory variables, alcohol availability significantly increases the number of older driver-involved fatal crashes, but neither AB541 nor APS had an identifiable effect. At the same time, and in contrast to the results in column (a), increases in per capita DUI arrests reduce fatal crashes but the result was not statistically significant.[11]

Also in contrast to the results for total crashes, the effect of higher speed limits has opposite implications for fatal crashes. Although the 1987 speed limit relaxation had a marginal effect on total crashes, its effect on fatal crashes involving older drivers was stronger and significant, leading to a 0.71 additional fatal crash involving older drivers per county-year (or 41 annual fatal crashes statewide). The 1997 law, however, which allowed states to raise limits above 65 mph, although significantly increasing total crashes involving older drivers, had no effect on fatal crashes.

Table 5 columns (a) and (b) summarize the estimation results for nonfatal injury and PDO crashes. The results reported in column (a) for nonfatal injury crashes are similar to the results in table 4 column (a) for total crashes with a few exceptions. First, per capita DUI arrests has no effect on nonfatal injury crashes. Second, APS has a positive sign and is statistically significant, indicating that passage of the 0.08 administrative per se law actually increased nonfatal injury crashes involving older drivers. Theoretically, APS is expected to reduce crashes. The law mandates that police revoke the license of a driver arrested for a DUI offence and issue a 30-day temporary driving permit. Drivers pay minor fees to get their licenses reinstated. Also, other legislation passed in early 1990 reduced the BAC level from 0.10 to 0.08 as per se evidence of impaired driving and increased various sanctions for a DUI offense. By raising the expected cost of a DUI event, we would expect these laws to reduce the incidence of crashes, all else being constant. Although APS had little impact on total crashes, the positive and significant effect identified in table 5 column (a) is inconsistent with higher expected costs and may reflect the laws' distribution effects in the post-1990 driving environment.

Finally, table 5 column (b) reports the results for older driver PDO crashes. These crashes are most sensitive to the price index for gasoline, population density, the share of drivers 60 years of age or older, and the number of alcohol licenses. In addition, the best fit included total vmt, which found that a 100 million increase in vmt would lead to four additional PDO crashes involving older drivers. In this case, both AB541 and APS reduced the number of PDO crashes involving older drivers. Recall that APS was not significant for total crashes but was positive and significant for nonfatal injury crashes. In contrast to this latter result, in table 5 column (b), APS significantly reduces PDO crashes, which is consistent with expectations.[12]

Estimation Results for Younger Drivers

For purposes of comparison, FGLS models were estimated for county crashes involving younger

[10] In preliminary models, various vmt specifications were employed, including total vmt, older driver vmt, younger driver vmt, and logarithmic transformations. The results reported represent the best model fits. In contrast to the total crashes results reported in table 4 column (a) often the inclusion of both younger and older driver vmt led to robustness problems due to the high level of collinearity between these exposure measures.

[11] For the fatal crashes equation, the hypothesis that Pc_dui was exogenous could be rejected at the 0.05 level. In this case, Pc_dui was regressed on a set of explanatory variables and the predicted value included in the fatal crashes equation.

[12] As will be seen below for younger drivers, APS has strong positive effects across all crash types, which is consistent with the results for injury crashes among older drivers but inconsistent with expectations. Further research is needed to isolate the mechanisms through which APS affects crash frequency and severity.

drivers. One difference in these models from those for the older drivers is that including a full set of fixed effects led to poorer model fits. In order to control for major sources of cross-section heterogeneity without including a full set of fixed effects parameters, the results in tables 6 and 7 include fixed effects for those counties that are part of a consolidated metropolitan statistical area (CMSA). In total, the models include 24 fixed effects parameters in addition to autocorrelation and variance parameters associated with each of the 58 counties.[13] Similar to tables 4 and 5, the results in tables 6 and 7 reflect large variation in the average as well as the range of estimated serial correlation coefficients.

Table 6 columns (a) and (b) give the results for total and fatal crashes involving younger drivers. From a statistical perspective, the results are stronger in that the *z*-statistics for the significant variables reject the null hypothesis at higher levels of significance. There are a number of differences in total crash results between older and younger involved drivers. Some differences are:

- Higher population densities are detrimental to highway safety in that population density significantly increased the number of nonfatal injury and PDO crashes among younger involved drivers (but high population density reduced fatal accidents for younger drivers). For older involved drivers, there was no effect on nonfatal injury crashes and a negative effect on PDO crashes.
- A higher county share of older persons increased total crashes among older involved drivers but reduced total crashes among younger involved drivers.
- APS significantly increased the total number of crashes involving younger drivers.
- For younger involved crashes, the 1987 relaxed speed limit (Slmt_65) significantly increased total crashes whereas the 1995 relaxation significantly reduced total crashes.
- Per capita DUI arrests had a positive effect on older involved crashes; for younger involved crashes, this variable had a negative and significant effect on total crashes but a positive effect on fatal crashes. The total number of younger driver-involved crashes decreased with the number of per capita motorcycles registrations in a county.

A similar comparison of fatal crash determinants in table 4 column (b) for older involved drivers with that of younger involved drivers in table 6 column (b) also reveals a number of differences. In particular:

- An increase in the unemployment rate reduced fatal crashes among the younger driver-involved crashes, whereas this variable had no significant effect for the older drivers involved in crashes.
- Increases in population density and the share of older drivers in the population had strong beneficial effects on safety, whereas no effect was found for fatal crashes involving older drivers.
- AB541 significantly reduced fatal crashes among the younger group, while APS increased fatal crashes, effects that were absent in the older driver group.
- Neither of the two increases in speed limits affected fatal crashes involving younger drivers.
- Per capita DUI arrests significantly increased fatal crashes, as it did for total crashes, and per capita motorcycles reduced fatal crashes involving younger drivers.

In the fatal crash equation, the two results that were unexpected but yet were robust to alternative specifications were the positive sign on APS and the positive sign on per capita DUI arrests. As noted earlier, the expectation was that APS would have a negative effect. Yet for total and fatal crashes for younger involved drivers, the sign and significance were robust.

For per capita DUI arrests, there is the possibility of endogeneity problems. This was explored in some detail for per capita arrests. Neither alternative variable specifications nor using "predicted per capita arrests" (based on a panel regression of per capita DUI arrests on the set of explanatory vari-

[13] In preliminary runs of the model, a full set of fixed effects for all but the normalized county led to convergence problems. Because many differences among counties are expected to reflect factors related to urbanization, a separate fixed effect was included for each CMSA county in order to capture this heterogeneity. For the reported models, the interpretation of the CMSA fixed effect parameters is the impact on crashes relative to all rural counties in California.

ables in the model) altered the finding in table 6 column (b). Further, fatal crashes were not found to be a significant determinant of per capita arrests in a panel regression analysis of arrests. This suggests that endogeneity is not a large problem. In contrast to expectations that per capita arrests decrease the most serious crashes at the cost of less serious crashes, there is some evidence that the reverse may have occurred. As seen in Table 7 column (b), per capita arrests was negative and statistically significant, which, when combined with the fatal crash results, suggest a redistribution from less serious to more serious crashes. A possible explanation is that the police are targeting the wrong group in their DUI enforcement efforts.

Table 7 columns (a) and (b) summarize the estimation results for nonfatal injury and PDO crashes among younger drivers involved in crashes. For each of these models, the results generally have similar implications with respect to price, exposure, population density, the share of older persons in the population, alcohol availability, and regulatory variables. But there are qualitative differences regarding the effects of real per capita income (reducing injury but increasing PDO crashes) and the unemployment rate (which has no effect on injury crashes).[14]

Elasticity Estimates

Tables 8 and 9 report sensitivity measures associated with model variables in the older and younger driver groups, respectively. For crashes involving older drivers, table 8 provides several interesting insights. First, fatal crashes are an order of magnitude more sensitive to changes in gasoline and alcohol price indexes than are less serious crashes. A 10% increase in energy prices, for example, reduces fatal crashes 1.3% in comparison with an approximate 0.2% decrease in injury and PDO crashes. Second, older drivers are most likely to be involved in PDO crashes and least likely to be involved in a fatal crash as their risk exposure increases.[15] Third, and consistent with the higher alcohol price sensitivity, fatal crashes among older drivers are more than twice as sensitive to alcohol availability in compari-

TABLE 8 Sensitivity Measures:[1] Involved Drivers, ≥ 60

Variable	Total	Fatal	Injury	PDO
Rpcinc	†–0.004	–0.130	–0.016	0.010
Cpigas	–0.022	–0.130	–0.024	–0.019
Cpialc	–0.013	–0.109	–0.047	†0.003
Uerate	†0.001	†0.019	†–0.002	†0.001
Vmt_lt60	0.106	—	0.184	—
Vmt_ge60	0.167	—	0.184	—
Totvmt	—	0.114	—	0.272
Popden	–0.080	†–0.012	†–0.009	–0.162
Pctge60	0.021	†0.080	0.022	0.019
Alclic	0.189	0.562	0.250	0.216
Pc_dui	0.001	†–0.027	†0.000	0.001
Ab541*	–2.901	†0.033	†0.369	–2.147
Sb1623*	0.022	–0.071	6.513	–3.329
Slmt_65*	†2.06	0.713	1.616	1.439
Slmt_70*	4.108	†0.084	1.388	1.317

[1] Elasticities for continuous variables; marginal effects for dummy variables (marked *).

Key: † = coefficient estimate was not significant at ≤ 0.20.

TABLE 9 Sensitivity Measures:[1] Involved Drivers, < 60

Variable	Total	Fatal	Injury	PDO
Rpcinc	0.052	–0.232	–0.018	0.066
Cpigas	–0.096	–0.117	–0.028	–0.099
Cpialc	–0.144	–0.256	–0.024	–0.147
Uerate	0.013	–0.017	†0.000	0.016
Vmt_lt60	0.073	—	0.037	0.057
Logvmtlt60	—	0.047	—	—
Popden	0.067	–0.355	0.124	0.081
Pctge60	–0.045	–0.069	–0.017	–0.039
Alclic	0.562	0.966	0.614	0.564
Pc_dui	–0.067	0.019	†0.000	–0.063
Ab541*	–52.420	–2.261	–28.290	–23.960
Sb1623*	250.070	1.965	29.430	146.17
Slmt_65*	72.420	†0.627	11.540	42.970
Slmt_70*	–48.310	†–0.444	†–2.853	–23.470
Pcmcyc	–0.011	–0.065	†–0.002	–0.013

[1] Elasticities for continuous variables; marginal effects for dummy variables (marked *).

Key: † = coefficient estimate was not significant at ≤ 0.20.

[14] Similar to the total and fatal crash equations involving younger drivers, the effect of motorcycle registrations was consistently negative and significant in the PDO equations. Given the high fatality rate among motorcyclists (e.g., the number of fatalities per vmt was 26 times higher than passenger car occupants (USDOT 2002)), this result was unexpected and may be capturing part of the smaller total exposure among the motorcycle population in comparison with that of the motor vehicle population.

[15] Since the best model fits for fatal and PDO crashes include total rather than older driver vmt, this assumes that older driver vmt increases in proportion to total vmt.

son with less serious crashes. A 4% increase in the number of retail licenses to sell alcohol results in an approximate 1% increase in injury and PDO crashes but a 2% increase in fatal crashes among older drivers. Fourth, the higher speed limits enacted in 1987 were detrimental to older drivers in that all types of crashes increased. However, the 1995 relaxation has had, to date, no impact on fatal crashes involving older drivers.

Table 9 reports the sensitivity measures for younger drivers involved in crashes. In this table, there are some interesting similarities and contrasts to those reported in table 8. First, crashes among younger drivers show sensitivities to gasoline and alcohol that are similar to those observed in table 8. Fatal crashes were most sensitive and an order of magnitude more sensitive than less serious crashes.

Second, although PDO crashes among younger drivers were most sensitive to increased exposure, there was a much smaller difference in fatal, nonfatal injury, and PDO crash sensitivity than was seen for older drivers.[16] A further contrast with table 8 is that, in table 9, vmt elasticity for each crash type is almost an order of magnitude smaller. Crashes involving younger drivers were much less sensitive to risk exposure than those among older drivers.

Third, table 9 shows that an increased share of older persons in the population reduced all crashes and the largest effect is associated with fatal crashes. Combined with the results in tables 4 and 5, an increased share of older persons in the population redistributed crashes away from younger involved and toward older involved drivers.

Fourth, highway safety among younger drivers was sensitive to the availability of alcohol, as in the older driver group. For younger drivers, the pattern was similar to that for older drivers in that fatal crashes were most sensitive to alcohol availability. The difference in the level of sensitivity for younger involved drivers was approximately double that for older involved drivers.

Fifth, and also in contrast to the older driver results, AB541 was beneficial to younger driver highway safety regardless of severity level. Last, neither the 1987 nor the 1995 higher speed limits increased fatal crashes involving younger drivers, although the 1987 relaxation did increase the incidence of nonfatal injury crashes.

SUMMARY AND CONCLUSIONS

The objective of this analysis was to obtain some insights on economic and regulatory factors that are important determinants of highway crashes involving older drivers. With the "graying of America," older drivers will make up a larger proportion of the traveling population, and it is important to understand the impact this is likely to have on highway crashes, particularly those involving fatalities or injuries. Further, the study analyzed younger drivers in order to identify differential impacts that public policy may have on older versus younger driver-involved crashes. The major conclusions that are evident from this study are as follows.

- The set of explanatory variables for crashes involving older drivers differ from those variables for crashes involving younger drivers.
- Among the subset of explanatory variables for crashes among older and younger drivers, crash sensitivity to these factors show large quantitative differences and, in some cases, sign differences.
- Fatal crashes involving older drivers were sensitive to risk exposure. This has serious implications for highway safety as the population ages and risk exposure among this group of drivers increases.
- Increases in gasoline and alcohol taxes, which can be justified on economic grounds for a number of reasons, can provide safety benefits in reducing fatal crashes involving older drivers.
- Reducing the number of retail establishments that sell alcohol will reduce the number of crashes involving older drivers.
- DUI enforcement, measured by per capita DUI arrests and after correcting for endogeneity bias, had little effect on fatal and noninjury crashes involving older drivers. For crashes involving younger drivers, a significant increase in crashes was found, arguing for further research to better understand the relationship between measures of DUI enforcement and highway safety.
- For older drivers, raising speed limits to 65 mph on Interstates increased fatal crashes in this group.

[16] For fatal crashes, this implies that younger driver vmt increases in proportion to total vmt.

- The omnibus deterrent law passed in 1981 (AB541) had a significant effect on all crashes involving younger drivers but no impact on fatal and nonfatal injury crashes involving older drivers. This identifies a differential policy effect for the two groups. However, this result is tentative because the absence of an effect for crashes involving older drivers may reflect the fact that there is only one year of data for the pre-law period.
- The results for California's 1990 administrative per se law produced mixed results, with increasing nonfatal injury crashes but decreasing PDO crashes involving older drivers. For crashes involving younger drivers, and contrary to expectations, the effect was consistently positive and statistically significant, calling for more research in order to understand the structural relationship between APS and changes in highway safety.

Two modeling implications flow from this analysis. First, relative to fixed effects negative binomial methods, FGLS was used to estimate the models reported in this analysis. FGLS estimation comes at the cost of not directly accounting for overdispersion that is typical in crash data. For the models reported, normality tests were conducted on the errors and these tests uniformly rejected the null hypothesis, which likely reflects remaining problems with overdispersion.[17] At the same time, negative binomial models that do not adequately account for heteroskedasticity and serial correlation are inefficient and will produce incorrect standard errors that invalidate standard hypothesis tests. Additional research needs to be done to better understand the tradeoffs and empirical importance in using alternative panel data estimation techniques.

Second, all of the models reported in this analysis focused on the impact that alternative explanatory variables had on the number of highway crashes. Alternatively, the focus could be on crash rates rather than levels. Each of the models reported in tables 4 through 7 were re-estimated using crash rates rather than levels.[18] Qualitatively, the estimation results were broadly, but not uniformly, consistent with the results reported in these tables. Alcohol and gasoline price indexes had similar effects to that identified in this analysis. With respect to the alcohol policy variables, AB541 was uniformly negative in the rate equation, whereas APS had similar effects on crash rates as it did on levels, but these were not consistent. For crashes involving older drivers, the effects of the 1987 and 1995 speed limit increases were generally consistent with those for crash levels. This was not true for the younger group.[19] Further, the effect on crash rates of an increase in the number of alcohol retail licenses was uniformly negative and always significant.

Further study is needed to establish the relationships between the effects that determining factors have on crash levels vis-à-vis crash rates. For example, consistent with expectations, increasing the number of alcohol licenses is expected to increase the number of crashes (all else remaining constant), because a larger number of licenses is expected to reflect greater alcohol consumption. However, an increase in the number of licenses could increase or decrease crashes per vmt, because crash rates are not controlling for vmt. In order to set appropriate highway safety policy for a growing older population, it is important to understand the effects that policies have on alternative measures of highway safety.

REFERENCES

Baltagi, B.H. 1995. *Econometric Analysis of Panel Data.* New York, NY: John Wiley and Sons.

Chaloupka, F., H. Saffer, and M. Grossman. 1993. Alcohol Control Policies and Motor Vehicle Fatalities. *Journal of Legal Studies* 22:161–186.

Evans, W.N., D. Neville, and J.D. Graham. 1991. General Deterrence of Drunk Driving: Evaluation of Recent American Policies. *Risk Analysis* 11(2):279–289.

[17] Normality tests were based on the Shapiro-Wilk test statistic (Shapiro and Wilk 1965).

[18] In these models, vmt was omitted from the estimating equations.

[19] The results were robust for fatal crashes. However, in the level model for total crashes, the 1987 law coefficient was positive and significant (for the rate model it was negative and insignificant) and the 1995 law coefficient was negative and significant (positive and significant for the rate model). For injury crashes, the 1987 and 1995 laws carried opposite signs in the rate relative to the level models and were significant at a 0.10 level. For PDO crashes, the 1987 and 1995 laws in the rate model also had opposite signs from those in the level models, but only in the 1997 case was the coefficient statistically significant.

Hu, P.S., D.A. Trumble, D.J. Foley, J.W. Eberhard, and R.B. Wallace. 1998. Crash Risks of Older Drivers: A Panel Data Analysis. *Accident Analysis and Prevention* 30(5):569–581.

Legge, J.S. and J. Park. 1994. Policies To Reduce Alcohol-Impaired Driving: Evaluating Elements of Deterrence. *Social Science Quarterly* 75:594–606.

Lundberg, C., L. Hakamies-Blomquist, D. Almkvist, and K. Johansson. 1998. Impairments of Some Cognitive Functions are Common in Crash-Involved Older Drivers. *Accident Analysis and Prevention* 30(3):371–377.

McCarthy, P. 2000. Effect of Speed Limits on Speed Distributions and Highway Safety: A Survey of Recent Literature. *Transport Reviews* 21:31–50.

McCloskey, L.W., T.D. Koepsell, M.E. Wolf, and D.M. Bucher. 1994. Motor Vehicle Collision Injuries and Sensory Impairments of Older Drivers. *Aging* 23:267–273.

Peck, R.C. 1991. The General and Specific Deterrent Effects of DUI Sanctions: A Review of California's Experiences. *Alcohol, Drugs, and Driving* 7(1):13–42.

Pindyck, R. and D. Rubinfeld. 1991. *Econometric Models & Economic Forecasts (Third Edition)*. New York, NY: McGraw-Hill, Inc.

Rogers, P. 1995. *The General Deterrent Impact of California's 0.08% Blood Alcohol Concentration Limit and Administrative per se License Suspension Laws: Volume 1,* Report No. 158. Sacramento, CA: California Department of Motor Vehicles.

Transportation Research Board. 1984. 55: *A Decade of Experience,* Special Report 204. Washington, DC: National Research Council.

_____. 1998. *Managing Speed: Review of Current Practice for Setting and Enforcing Speed Limits,* Special Report 254. Washington, DC: National Research Council.

Shapiro, S. and M. Wilk. 1965. An Analysis of Variance Test for Normality (Complete Samples). *Biometrika* 52:591–611

U.S. Department of Transportation (USDOT), National Highway Traffic Safety Administration, National Center for Statistics and Analysis. *Traffic Safety Facts 2001: A Compilation of Motor Vehicle Crash Data from the Fatality Analysis Reporting System and the General Estimates System.* Washington, DC. December.

Zador, P.L., A.K. Lund, M. Fields, and K. Weinberg. 1989. *Fatal Crash Involvement and Laws Against Alcohol-Impaired Driving.* Arlington, VA: Insurance Institute for Highway Safety.

ACKNOWLEDGMENTS

The research reported in this paper was sponsored by the National Institute on Alcohol Abuse and Alcoholism, National Institutes of Health, Grant #5R03AA012148-03. A previous version of this paper was presented at the 1st International Congress on Transportation Research in Athens, Greece, in February 2002. I would like to thank three anonymous referees and the Journal's Editors for their constructive comments and suggestions. I would also like to thank Aruna Srinivasan, a graduate research assistant in the School of Economics, who estimated the models reported in this paper.

Author address: Patrick McCarthy, School of Economics, Ivan Allan College, Georgia Institute of Technology, Atlanta, GA 30318. Email: mccarthy@econ.gatech.edu.

Estimation and Accuracy of Origin-Destination Highway Freight Weight and Value Flows

PAUL METAXATOS
University of Illinois at Chicago

ABSTRACT

This paper proposes a spatial interaction modeling framework and implements a maximum likelihood estimation of highway freight weight and value flows using the gravity model. The computation of the standard error of the flow estimates provides the basis for measuring the level of accuracy of the estimates. The results provide evidence of the suitability of gravity models for freight forecasting given the excellent fit and the small variances.

INTRODUCTION

The measurement of freight movements requires tracking freight flows across geographic and political boundaries. This is a particularly challenging task given the current capabilities for state and regional data acquisition. Various mathematical approaches have been implemented (Memmot 1983; USDOT 1996; Cambridge Systematics 1997) to circumvent this problem, but none, to the best of this author's knowledge, proposes a measure to assess the accuracy of the computed flows.

This paper proposes to fill this gap using developments in spatial interaction modeling that have not been demonstrated on a large scale to date. The methodology computes maximum likelihood flow

KEYWORDS: freight origin-destination flow estimation, covariance of estimates, gravity model.

estimates and obtains their covariance matrices that, in turn, may be used to obtain confidence intervals and carry out certain tests of hypotheses. The approach can accommodate the large number of origins and destinations typically encountered in freight (and passenger) travel forecasting.

The methodology was applied to highway freight weight and value flows of international trade traffic between seaports or border ports and destination states (see Metaxatos (2002) for details). The variances of the flow estimates computed were remarkably small. The demand for freight transportation flows can then be estimated within a desired confidence level. Moreover, the empirical analysis undertaken provides evidence that the theoretical framework proposed in this paper is rich enough for freight demand forecasting applications.

THEORETICAL FRAMEWORK

Commodity shipments in this paper are thought to be realized patterns of spatial interactions that typically result from many independent decisions by individual firms, each constituting a relevant subsystem within the economy as a whole. Hence, if the travel behavior of each firm is modeled as a very small interaction process, the resultant interaction process can be taken to be the superposition of all these processes. It may be argued that for large collections of small frequency processes, the resulting superimposed process is approximately Poisson and, therefore, completely characterized by its associated mean interaction frequencies (Sen and Smith 1995).

In this light, assuming that the observations N_{ij} of shipment weight and value between origin seaports/border ports of entry i and destination states j can be described by the gravity model, then

$$\begin{aligned} N_{ij} &= T_{ij} + \varepsilon_{ij} \\ T_{ij} &= E(N_{ij}) = A_i B_j F_{ij} \quad \forall i, j \end{aligned} \tag{1}$$

In this paper, T_{ij}'s (the stochastic term) are interpreted as the expected international trade traffic flow (in terms of weight and value) carried by highway from external station i to state j. The A_i's are factors related to the origin zone i and the B_j's are destination-related factors. The F_{ij}'s are factors that reflect the separation between i and j. A common form that is general enough for most applications is

$$F_{ij} = \exp\left[\sum_k \theta_k c_{ij}^{(k)}\right] \quad \forall i, j \tag{2}$$

This form is called an exponential form and $c_{ij}^{(k)}$ are different measures of separation, while θ_k's are parameters to be estimated. Potential measures of separation include travel time, distance, generalized costs, etc.

In the gravity model, observable quantities $N_{i*} = \Sigma_j N_{ij}$, $N_{*j} = \Sigma_i N_{ij}$, N_{ij} and their expected values T_{i*}, T_{*j}, and T_{ij} are described by means of an underlying structure consisting of unobservable quantities A_i, B_j, and F_{ij}. Similar situations abound in statistics. In moving average models, for example, observations are described by means of unobservable parameters. A like situation exists in analysis of variance models.

Although the origin and destination factors are unobservable, they do have physical interpretations. For example, if for some origin i, there are two destinations j and j' such that $F_{ij} = F_{ij'}$, then $T_{ij}/T_{ij'} = B_j/B_{j'}$. Thus other factors being equal, T_{ij} is proportional to B_j (but, in general, not proportional to T_{*j}), and is called the *attractiveness* of j. Similarly, the origin factor A_i may also be called the *emissiveness* of i.

Clearly, B_jF_{ij} is the effect of the destination factor B_j at i, or the *accessibility* of j as perceived from i. This is a spatial analogy of the temporal concept of present value in economics, where a dollar earned in n years in the future is worth only $(1 + \sigma)^{-n}$ now, where σ is the interest rate. Similarly, A_iF_{ij} is the effect of the origin factor A_i at j. The sum $\alpha_i = \sum_j B_j F_{ij}$ may be called the total accessibility of all destinations at i, and the sum $\beta_j = \sum_i A_i F_{ij}$ may be called the total accessibility of all origins at j. If, for example, T_{i*} is kept fixed as α_i increases, the push A_i decreases. Thus, as the competition α_i from the destinations increases, the push at i decreases. From the point of view of someone at i, α_i measures accessibility; from the viewpoint at j, it measures competition. Similar statements can be made about B_j.

Maximum Likelihood Estimation

The model (1) will be estimated using maximum likelihood (ML). Maximum likelihood estimates have desirable asymptotic properties (consistency, efficiency, and asymptotic normality) and are robust to distributional assumptions for realistic departures from the Poisson assumption (note that the multinomial distribution leads to identical estimates with the Poisson distribution). Furthermore, they are essentially unbiased even for a very small sample of flows (Sen and Smith 1995).

Under some mild conditions (Sen and Smith 1995), the ML estimate of $\boldsymbol{\zeta} = (A(1),\ldots,A(I), B(1),\ldots,B(J),\theta_1,\ldots,\theta_Q)'$ exists and that of $\boldsymbol{\theta}$ is unique. The estimates of $\boldsymbol{A}$ and $\boldsymbol{B}$ are not unique; however, if one $A(i)$ or $B(j)$ is chosen to be an arbitrary positive number, the remaining $A(i)$'s and $B(j)$'s are unique under the previous mild conditions. It can be proved that the ML estimate of $\boldsymbol{\zeta}$ results in the solution to the following system of equations

$$T_{i*} = N_{i*} \quad \forall i \in I \tag{3}$$

$$T_{*j} = N_{*j} \quad \forall j \in J \tag{4}$$

$$\sum_{ij} c_{ij}^{(q)} T_{ij} = \sum_{ij} c_{ij}^{(q)} N_{ij} \quad \forall q \in Q \tag{5}$$

where the operator * indicates summation with respect to the subscript it replaces (e.g., $T_{i*} = \sum_j T_{ij}$, $T_{*j} = \sum_i T_{ij}$). A number of standard numerical methods or more specialized procedures can be used to solve equations (3) through (5). The three procedures adapted for the paper are the Deming-Stefan-Furness (DSF) procedure, the linearized DSF procedure, and the Modified Scoring procedure. An account of the development of the three procedures is given in Sen and Smith (1995). For completeness of the presentation, some of the details for each procedure follow.

The DSF Procedure for Parameters A_i and B_j

The DSF procedure gives values of T_{ij} for any choice of ($\boldsymbol{\theta}'$), and N_{i*}, N_{*j}, with $\sum_i N_{i*} = \sum_j N_{*j}$. Generally speaking, this procedure adjusts the rows (columns) of a two-dimensional table in each even (odd) iteration. After choosing an initial value for the column balancing coefficient, $B_j^0 = 1$, say, the DSF procedure iterates as follows (the index r denotes the iteration number):

$$A_i^{(2r-1)} = O_i \Big/ \sum_{j=j+1}^{J} B_j^{(2r-2)} F_{ij} \qquad \forall i \tag{6}$$

$$B_j^{(2r)} = D_j \Big/ \sum_{i=1}^{I} A_i^{(2r-1)} F_{ij} \qquad \forall j \tag{7}$$

where $O_i = T_{i*}$, $D_j = T_{*j}$, and F_{ij} is a function of the separation measures $c_{ij}^{(q)}$. Upon convergence (see Sen and Smith (1995) for a proof of convergence), the values of T_{ij} are given by

$$T_{ij}^{(2r)} = A(i)^{(2r-1)} B(j)^{(2r)} F_{ij}.$$

In this paper, we chose a fairly stringent criterion for convergence as follows

$$\sum_{i=1}^{I} |O_i - T_{i*}^{(2r)}| + \sum_{j=1}^{J} |D_j - T_{*j}^{(2r)}| < \delta \tag{8}$$

where $\delta = 10^{-12}$. The algorithm attained this criterion in less than 100 iterations.

The DSF procedure essentially expresses T_{ij} as a function of $\boldsymbol{\theta}$. These values of T_{ij} could then be used in equation (5) to solve for an updated value of $\boldsymbol{\theta}$. In general, however, as $\boldsymbol{\theta}$ changes, equations (3) and (4) could be violated, unless these changes are very small. This is achieved by using a linearized version of the DSF procedure, called the LDSF procedure (Weber and Sen 1985), which is computationally very attractive.

The LDSF Procedure for T_{ij}

Let us assume that we have run the DSF procedure and obtained a good set of T_{ij} that solves equations (3) and (4) for any given θ. This means that $O_i = T_{i*} = N_{i*}$ and $D_j = T_{*j} = N_{*j}$. Define by $\Delta\boldsymbol{O} = (\Delta O_1,\ldots,\Delta O_I)'$ and $\Delta\boldsymbol{D} = (\Delta D_1,\ldots,\Delta D_J)'$ to be small changes in the values of $\boldsymbol{O} = (O_1,\ldots,O_I)'$ and $\boldsymbol{D} = (D_1,\ldots,D_J)'$, respectively. Also, let ΔF_{ij}, be a small change in $F_{ij} = \exp[\boldsymbol{\theta}'\boldsymbol{c}_{ij}]$. It can be proved (Sen and Smith 1995) that the corresponding small change, ΔT_{ij}, in each T_{ij}, so that $\Delta T_{i*} = \Delta O_i$, and $\Delta T_{*j} = \Delta D_j$ can be obtained by the LDSF procedure, which iterates as follows

$$\Delta T_{ij}^{(2r-1)} = \Delta T_{ij}^{(2r-2)} + (T_{ij}/O_i)(\Delta O_i - \Delta T_{i*}^{(2r-2)}) \tag{9}$$

$$\Delta T_{ij}^{(2r)} = \Delta T_{ij}^{(2r-1)} + (T_{ij}/D_j)(\Delta D_j - \Delta T_{*j}^{(2r-1)}) \tag{10}$$

for $i \in I$, $j \in J$ and $r = 0, 1, 2,...$, with initial T_{ij} values given by

$$\Delta T_{ij}^{(0)} = (T_{ij}/F_{ij})\Delta F_{ij} \tag{11}$$

A proof for the convergence of the procedure is given in Weber and Sen (1985).

Changes in T_{ij} as a Function of a Change in θ

For a small change $\Delta\boldsymbol{\theta}$ in $\boldsymbol{\theta}$ and a small change 0 so that $\Delta \boldsymbol{O} = \Delta \boldsymbol{D} = \mathbf{0}$, it can be proved (Yun and Sen 1994) that an approximation for the corresponding small change ΔT_{ij} for each T_{ij} for all $i \in I$ and $j \in J$ is given by equation 12.

$$\Delta T_{ij} \approx \Delta\boldsymbol{\theta}\left\{ \boldsymbol{c}_{ij} T_{ij} - \sum_j \boldsymbol{c}_{ij} T_{ij}\left(\frac{T_{ij}}{O_i}\right) - \sum_i \boldsymbol{c}_{ij} T_{ij}\left(\frac{T_{ij}}{D_j}\right) + \sum_i \left[\sum_j \boldsymbol{c}_{ij} T_{ij}\left(\frac{T_{ij}}{O_i}\right)\right]\left(\frac{T_{ij}}{D_j}\right) \right\} = \boldsymbol{S}_{ij}\Delta\boldsymbol{\theta} \tag{12}$$

The $S_{ij}^{(k)}$'s are constants, and $O_i = T_{i*}$, $D_j = T_{*j}$. Therefore, if T_{ij}'s are known, the only unknown in equation (12) is the $\Delta\boldsymbol{\theta}$. The solution for the $\Delta\boldsymbol{\theta}$ will be the topic of the next section.

Estimation of θ Using the Modified Scoring (MS) Procedure

So far, for an initial value for $\boldsymbol{\theta}$, we obtained $T_{ij}(\boldsymbol{\theta})$ by using the DSF procedure to solve equations (3) and (4). We then changed $\boldsymbol{\theta}$ to $\boldsymbol{\theta} + \Delta\boldsymbol{\theta}$ and computed, using the LDSF procedure, with $\Delta \boldsymbol{O} = \Delta \boldsymbol{D} = \mathbf{0}$, the corresponding change, $\Delta T_{ij}(\boldsymbol{\theta}, \Delta\boldsymbol{\theta})$ in $T_{ij}(\boldsymbol{\theta})$ as a function of $\Delta\boldsymbol{\theta}$.

We are ready now to insert the $[T_{ij}(\boldsymbol{\theta}) + \Delta T_{ij}(\boldsymbol{\theta}, \Delta\boldsymbol{\theta})]$'s into the left side of equation (5) and solve the resultant equation for $\Delta\boldsymbol{\theta}$. Inserting $T_{ij} + \Delta T_{ij}$ in place of T_{ij} in (5) and using (12), equations (3) and (4) would remain approximately satisfied, while obtaining the following system of Q linear equations with Q unknowns (the $\Delta\theta_q$'s):

$$\begin{aligned} \sum_{ij} c_{ij}^{(1)}(\Delta T_{ij}) &= \sum_{ij} c_{ij}^{(1)}(N_{ij} - T_{ij}) \\ &\vdots \\ \sum_{ij} c_{ij}^{(Q)}(\Delta T_{ij}) &= \sum_{ij} c_{ij}^{(Q)}(N_{ij} - T_{ij}) \end{aligned} \tag{13}$$

This system of equations can be solved by any standard solution method such as Gaussian elimination.

The current solution for $\boldsymbol{\theta}$ at iteration r is updated next using the formula

$$\boldsymbol{\theta}^r = \boldsymbol{\theta}^{r-1} + \Delta\boldsymbol{\theta}^{r-1} \tag{14}$$

If the corrections $\Delta\boldsymbol{\theta}^{r-1}$ have become negligible, the values of $\boldsymbol{\theta}$ have been stabilized and the MS procedure terminates. Otherwise, new T_{ij}'s are obtained from the DSF procedure and the MS procedure continues. There is no guarantee that the MS procedure always converges (Sen and Smith 1995), although our computational experience is positive.

Goodness of Fit

Under the previous assumption that observations N_{ij} are independently Poisson distributed, the (Pearson) X^2 statistic,

$$X^2 = \sum_{ij} \frac{(N_{ij} - \hat{T}_{ij})^2}{\hat{T}_{ij}} \tag{15}$$

where $\hat{T}_{ij}$ is an estimate of T_{ij}, is an appropriate measure of the overall fit of a model. Moreover, when $\hat{T}_{ij}$ is obtained using maximum likelihood, equation (15) has a X^2 distribution with df = $IJ - I - J - K + 1$ degrees of freedom (Bishop et al. 1975; Rao 1973).

If $\hat{T}_{ij} \approx T_{ij}$, then $X^2 = Z^2$, where

$$Z^2 = \sum_{ij} \frac{(N_{ij} - T_{ij})^2}{T_{ij}} \tag{16}$$

Since $E(N_{ij}) = T_{ij}$ and because N_{ij} have the Poisson distribution,

$$\mathrm{var}(N_{ij}) = E(N_{ij} - T_{ij})^2 = T_{ij} \tag{17}$$

Therefore, $E(Z^2) = IJ$, where I is the number of origin zones i, and J the number of destinations j. Equivalently, $E(Z^2/IJ) = 1$. Thus, the so-called "X^2-ratio," X^2/df, has an expectation that is asymptotically 1. It can be shown (Sen and Smith 1995) that the variance of the X^2-ratio is

$$\mathrm{var}[Z^2/(IJ)] \approx \sum_{ij} [(T_{ij} I^2 J^2)^{-1} + 2(IJ)^{-2}] \tag{18}$$

Hence, if T_{ij}'s are bounded away from zero (which is the case in exponential gravity models with finite parameters $\boldsymbol{\theta}$), the variance of $Z^2/IJ \to 0$, as

$IJ \to \infty$. It follows that when $\hat{T}_{ij} \to T_{ij}$ and T_{ij}'s are bounded away from zero, the variance of $X^2/df \to 0$.

In practical applications, since the Poisson assumption seldom holds perfectly (as is the case here, where every pound or dollar value of shipment does not travel independently of each other pound or dollar value), an X^2 ratio less than 2 is a good indication that the gravity model fits the data well (Sen and Smith 1995).

Covariance of Maximum Likelihood Estimates

Covariance of $\hat{\theta}_q$'s

Let small case letters stand for the logarithms of corresponding capital letters (e.g., $t_{ij} = log[T_{ij}]$, $a(i) = log[A(i)]$, $b(j) = log[B(j)]$. The model (1) may be written as

$$t_{ij} = a(i) + b(j) + \sum_q \theta_q c_{ij}^{(q)} \quad \forall i \in I, j \in J, q \in Q \quad (19)$$

Let M denote the coefficient matrix of the right side of the system of equations (19). The matrix M is not of full rank (Sen and Smith 1995). However, the matrix $M_{(2)}$ obtained by deleting one of the first $I + J$ columns of M is of full rank and has dimension $IJ \times (I + J + Q - 1)$. Let diag($\cdot$) stand for a diagonal matrix, the diagonal elements of which are given within the parentheses. Then compute the matrix $M'_{(2)} \cdot \text{diag}(\boldsymbol{T}) \cdot M_{(2)}$ from the equation

$$M'_{(2)} \cdot \text{diag}(\boldsymbol{T}) \cdot M_{(2)} = \begin{pmatrix} U_1 & U_2 \\ U'_2 & U_3 \end{pmatrix} \quad (20)$$

where

$$U_1 = \begin{pmatrix} V_1 & V_2 \\ V'_2 & V_3 \end{pmatrix} \quad (21)$$

$$U_2 = \begin{pmatrix} W_1 \\ W_2 \end{pmatrix} \quad (22)$$

$U_3 = ((u_{pq}))$ with $u_{pq} = \Sigma_{ij} c_{ij}^{(p)} c_{ij}^{(q)} T_{ij}, W_1 = ((w_{iq}^{(1)}))$ with $w_{iq}^{(1)} = \Sigma_j c_{ij}^{(q)} T_{ij}$, $W_2 = ((w_{jq}^{(2)}))$ with $w_{jq}^{(2)} = \Sigma_i c_{ij}^{(q)} T_{ij}$, $V_1 = \text{diag}(T_{1*},\ldots,T_{I*})$, $V_2 = ((T_{ij}))$ and $V_3 = \text{diag}(T_{*1},\ldots,T_{*J-1})$. Notice that the subscript j in each of the matrices above goes only up to $J - 1$.

Matrix $M'_{(2)} \cdot \text{diag}(\boldsymbol{T}) \cdot M_{(2)}$, a square matrix of dimension $(I + J + Q - 1)$, is the covariance matrix of $M'_{(2)} \cdot \text{diag}(\boldsymbol{N})$. This is because the N_{ij}'s have independent Poisson distributions and the covariance matrix Cov($\boldsymbol{N}$) of $\boldsymbol{N}$ is diag($\boldsymbol{T}$). It can be shown (Sen and Smith 1995) that the covariance matrix of $(\hat{A}(1),\ldots,\hat{A}(I), \hat{B}(1),\ldots,\hat{B}(J-1), \hat{\theta}_1,\ldots,\hat{\theta}_Q)'$ is

$$\Phi^{-1} \cdot M'_{(2)} \cdot \text{diag}(\boldsymbol{T}) \cdot M_{(2)} \cdot (\Phi^{-1})' \quad (23)$$

where

$$\Phi = M'_{(2)} \cdot \text{diag}(T) \cdot M_{(2)} \cdot \text{diag}(1/A(1),\ldots,1/A(I),$$
$$1/B(1),\ldots,1/B(J-1), 1,\ldots,1) \quad (24)$$

and

$$\Phi^{-1} = \text{diag}(1/A(1),\ldots,1/A(I), 1/B(1),\ldots,1/B(J-1),$$
$$1,\ldots,1) \cdot (M'_{(2)} \cdot \text{diag}(T) \cdot M_{(2)})^{-1} \quad (25)$$

Notice that using equations (23) through (25), another expression for the covariance matrix of $(\hat{A}(1),\ldots,(\hat{A}(I), \hat{B}(1),\ldots,\hat{B}(J)-1), \hat{\theta}_1,\ldots,\hat{\theta}_Q)'$ can be written by as follows

$$\text{diag}(1/A(1),\ldots,1/A(I), 1/B(1),\ldots,1/B(J-1),1,\ldots1) \cdot$$
$$(M'_{(2)} \cdot \text{diag}T \cdot M_{(2)})^{-1} \cdot \text{diag}(1/A(1),\ldots,1/A(I),$$
$$1/B(1),\ldots,1/B(J-1), 1,\ldots,1) \quad (26)$$

The bottom right $Q \times Q$ submatrix of matrix (26) is the estimated covariance matrix of $\hat{\boldsymbol{\theta}}$. The bottom right $Q \times Q$ submatrix of the inverse of (20) is (Rao 1973)

$$\left(U_3 - U'_2 U_1^{-1} U_2\right)^{-1} \quad (27)$$

From equation (26), it follows that equation (27) is the covariance matrix of $\hat{\boldsymbol{\theta}}$.

Covariance of $\hat{T}_{ij}$'s

Having obtained Cov($\boldsymbol{A},\boldsymbol{B},\boldsymbol{\theta}$) from equation (23) or using equation (26), and since $B(J)$ is set equal to a constant, and its variance and covariances involving it are zeros, it can readily be seen that the covariance matrix of $\hat{T}$, denoted by the symbol Cov($\hat{T}$), is

$$\mathrm{Cov}(\hat{T}) = \Psi_1 \cdot \mathrm{Cov}(A, B, \theta) \cdot \Psi_1' \tag{28}$$

with

$$\Psi_1 = \mathrm{diag}(T) \cdot M \cdot \mathrm{diag}(1/A(1),...,1/A(I), 1/B(1),...,1/B(J), 1,...,1) \tag{29}$$

and M the coefficient matrix of the right side of the system of equations (19).

Short-Term Forecasting

The shipment of goods is affected by a multitude of factors (USDOT 1996, table 2.1). The weight and value of commodities shipped, characteristics of immediate concern in this study in particular, may be affected by the economy as a whole, globalization of business, international trade agreements, just-in-time inventory practices (weight only), packaging materials (weight only), economic regulation/deregulation, publicly provided infrastructure (weight only), user charges and other taxes, changes in truck size and weight limits (weight only), and technological advances. The discussion below is based on the assumption that, in the short term (e.g., three to five years) the compound effect of these factors on the size of the weight and value characteristics of commodity shipments is consistent with Poisson randomness.

Let the random variable $N_{ij}^{(f)}$ be a future observation[1] of the flow from i to j. In the short run, under the assumption that the separation configuration will not change during the forecast period, we may conjecture that $N_{ij}^{(f)}$ and N_{ij} will be highly (serially) correlated. Thus, we may then argue that the variance of the difference of future and present observations will be smaller than the variance of future observations alone. Hence,

$$\mathrm{var}\left(N_{ij}^{(f)} - N_{ij}\right) = \mathrm{var}\left(N_{ij}^{(f)}\right) + \mathrm{var}(N_{ij}) - 2\,\mathrm{Cov}\left(N_{ij}^{(f)} N_{ij}\right) \tag{30}$$

where var($\cdot$) stands for the "variance of" and Cov($\cdot$) for the "covariance of," will be small. In such cases, especially if N_{ij}'s are known, Sen and Smith (1995, chapter 5) suggest that it would be preferable to predict $(N_{ij}^{(f)} - N_{ij})$'s and add the predictions to the N_{ij}'s. A possible set of predictions for $(N_{ij}^{(f)} - N_{ij})$'s are $(\hat{T}_{ij}^{(f)} - \hat{T}_{ij})$'s. Thus, if the future is not too far off,

$$N_{ij} + (\hat{T}_{ij}^{(f)} - \hat{T}_{ij}) \tag{31}$$

may yield a better prediction than $\hat{T}_{ij}$. The computation of $\Delta\hat{T}_{ij} = \hat{T}_{ij}^{(f)} - \hat{T}_{ij}$ can be made easily using the LDSF procedure described earlier.

Other Forecasts

Four separate cases are discussed here. The first case assumes that $c_{ij}^{(q)f}$ would be available exogenously, $\boldsymbol{\theta}^f$ would be the estimate $\hat{\boldsymbol{\theta}}$ from the base period assumed to remain unchanged into the forecast period, and the estimates $A^f(i)$ and $B^f(j)$ could be base period estimates assumed to remain unchanged into the forecast period, or one or both of them could be exogenous. Then, T_{ij} for the forecast period may be estimated by

$$T_{ij}^f = A^f(i)B^f(j)\exp[(\boldsymbol{\theta}^f)'\boldsymbol{c}_{ij}{}^f] \tag{32}$$

The calculation of Cov($\hat{\boldsymbol{T}}^f$) requires the computation of the covariance matrices of the $A^f(i)$'s, $B^f(j)$'s, and $\boldsymbol{\theta}^f$. For those estimates that remain unchanged from the base period, the appropriate covariance matrix is the one for the base period. For estimates obtained exogenously, the covariance matrix needs to be supplied exogenously. Covariances of estimates obtained from different independent data are usually assumed to be zeros.

The second case assumes that $c_{ij}^{(q)f}$ and $\boldsymbol{\theta}^f$ would be available as before and the estimates $B^f(j)$ and T_{*j}^f could be base period estimates assumed to remain unchanged into the forecast period or could be obtained exogenously. Then, T_{ij} for the forecast period may be estimated by

$$T_{ij}^f = \frac{T_{*j}^f B^f(j)\exp[(\boldsymbol{\theta}^f)'\boldsymbol{c}_{ij}{}^f]}{\Sigma_j B^f(j)\exp[(\boldsymbol{\theta}^f)'\boldsymbol{c}_{ij}{}^f]} \tag{33}$$

The second case assumes that $c_{ij}^{(q)f}$ and $\boldsymbol{\theta}^f$ would be available as before and the estimates $A^f(i)$ and T_{i*}^f could be base period estimates assumed to remain unchanged into the forecast period or could

[1] The superscript f denotes that the quantity is an estimate for the forecast period.

be obtained exogenously. Then, T_{ij} for the forecast period may be estimated by

$$T_{ij}^{f} = \frac{T_{i*}^{f} A^{f}(i)\exp[(\boldsymbol{\theta}^{f})'\boldsymbol{c}_{ij}{}^{f}]}{\sum_j A^{f}(i)\exp[(\boldsymbol{\theta}^{f})'\boldsymbol{c}_{ij}{}^{f}]} \tag{34}$$

The computation of the covariance matrices for the second and third cases is similar to the first case once the appropriate Jacobians with respect to $A^f(i)$'s, $B^f(j)$'s and $\boldsymbol{\theta}^f$ have been obtained. Recall that equation (29) is the respective Jacobian for the first case.

A fourth case, arising when $c_{ij}^{(q)f}$, $\boldsymbol{\theta}^f$, T_{i*}^f, and T_{*j}^f are available, requires the use of the DSF procedure to generate the $A^f(\mathrm{i})$ and $B^f(j)$. Then, T_{ij} for the forecast period may be estimated by equation (32). The computation of the covariance matrices for this case is treated in detail by Sen and Smith (1995, p. 440).

It is interesting to observe at this point that origin-destination travel times and/or costs (which depend on routes chosen and traffic congestion on the routes) usually become available only after the gravity model has been applied to forecast flows. This difficulty affects only the forecasting phase (not the estimation phase, when we use base period data for which travel times can be observed or computed using available information) and can be alleviated by combining the distribution and assignment stages of the process into a single stage.

EMPIRICAL ANALYSIS

Data Issues

Freight Shipments

Oak Ridge National Laboratory (ORNL) provided two sets of origin-destination flow data for freight shipment weight and value between origin seaports/ border ports of entry i and destination states j from the following sources:

- **The Transborder Surface Freight Database.** This database (http://www.bts.gov/transborder/prod.html) is distributed by the Bureau of Transportation Statistics and contains freight flow data by commodity type and surface mode of transportation (rail, truck, pipeline, or mail) for U.S. exports to and imports from Canada and Mexico.
- **The Port Import Export Reporting Service (PIERS) Database.** This commercial database (http://www.piers.com/about/default.asp), offered by the *Journal of Commerce*, offers statistics on global cargo movements transiting seaports in the United States, Mexico, and South America to companies around the globe.

Two matrices, N_{ij}, were developed: one for shipment weight and one for shipment value for each of the two databases with dimensions 128 x 50 and 144 x 50 for the Transborder and PIERS databases, respectively.

The limitations of these two data sources are well documented (Meyburg and Mbwana 2002). A caveat for the Transborder Database is that it is a customs, not a transportation, dataset, resulting in inconsistencies between U.S. and Canadian data and issues related to the accuracy of transshipment data.[2] One caveat for using PIERS is that reported origins and destinations may be billing addresses rather than shipment points.

Separation Measures

In identifying specific types of spatial separation that tend to impede or enhance the likelihood of interactions between points of entry and destination sites, the most obvious type involves physical space, as exemplified by travel distance and travel time, which are quantifiable in terms of meaningful units of measurement. ORNL provided two sets with separate impedance matrices, $c_{ic}^{(k)}$, k = 1, 2 from origin point i to destination county c.

The first dataset included impedances between origin seaports and destination counties, the second set between origin border ports and destination counties. Both measures were computed from ORNL's National Highway Network (NHN). The first of the two measures is the route (network) distance in miles; the second computes a function of travel time for different functional classifications of highway segments and adds these time penalties together while tracing the previous network routes. The previous impedance matrices presented two problems: 1) more than 50% of the cells were empty because not all seaports/border ports of entry are connected to each county through the NHN; and 2) occasionally, freight shipments moved

[2] See http://www.bts.gov/ntda/tbscd/desc.html, as of January 2004, for more information.

between origin-destination pairs with a missing connection (separation measure).

In order to deal with both problems, we devised and implemented the following stepwise procedure for the distance matrix (the first separation measure, $c_{ic}^{(1)}$).

1. Centroid coordinates were computed for all counties nationwide.
2. Spherical distances in miles from each county to every other county.
3. For every destination county with a missing connection to an origin point (seaport/border port of entry), the closest county with an existing connection to an origin point was estimated based on the previous county-to-county distances.
4. The missing connection from an origin point to a destination county was finally computed to be the distance between the same point of origin and the closest destination county augmented by 130% the (airline) distance between the two destination counties (as a proxy to the actual road distance). This simplification was made under the assumption that a missing connection would imply that the destination county is off the NHN and thus would require additional time to be accessed from its closest county on the NHN.

The above procedure resulted in the construction of a synthesized separation measure comprising both route distance (the ORNL estimate) and multiples of airline distance (our estimate). It is important to take a closer look at those two components of the (re-estimated) first separation measure $c_{ic}^{(1)}$.

The airline distance between origin point i and destination county c, even augmented by 30%, is only an approximation of the actual miles traveled and represents a surrogate for the complex set of factors that express the difficulty of overcoming separation. On the other hand, the route distance for the same origin-destination pair is the total over-the-road distance on a realistic route. As a measure of separation, the route distance provides better accuracy for the eastern half of the limited access highway system of the United States, which has a higher level of complexity resulting in less circuitous routes than in the western half.

Whereas the distance measure developed above is a measure that depends more or less on the physical characteristics of the road link, the impedance measure, a function of link travel time as estimated by ORNL, depends on the special roadway type or certain conditions encountered on the link. ORNL's estimating procedure could not be replicated in this paper for the missing impedances. A surrogate value was computed based on the average speed obtained by ORNL's distance and time estimates, and the synthesized distances.

The procedure classified ORNL's distance and time estimates into 10 deciles. The stratum average speed was computed by the ratio of the average distance and time for that stratum (table 1). Each

TABLE 1 Average Distance, Time, and Speed by Decile

	PIERS Database			Transborder Database		
Decile	Average distance (miles)	Average time (min.)	Estimated speed (mph)	Average distance (miles)	Average time (min.)	Estimated speed (mph)
1	230.69	241.56	57.30	381.11	400.32	57.12
2	468.65	465.70	60.38	708.18	711.44	59.73
3	642.53	625.89	61.59	929.82	920.37	60.61
4	797.46	769.51	62.17	1,121.06	1,093.26	61.52
5	954.44	915.38	62.56	1,291.37	1,251.02	61.93
6	1,127.45	1,077.43	62.78	1,457.91	1,405.62	62.23
7	1,324.50	1,261.04	63.01	1,632.94	1,569.44	62.42
8	1,596.70	1,512.70	63.33	1,850.64	1,771.00	62.69
9	2,109.35	1,979.77	63.92	2,166.76	2,048.56	63.46
10	2,958.11	2,963.42	59.89	2,757.87	2,641.23	62.65

missing time impedance was then estimated by the ratio of its corresponding distance estimate and the average speed of the stratum to which the distance estimate belongs.

Clearly, this procedure imposes some inconsistency in the imputation of travel times, because the average speed estimate is based on decile groupings of the initial estimates. The consistency of the procedure could be improved if in a second iteration the decile groupings are based on all distances (initial and imputed), but I did not attempt this for this paper.

The previous steps resulted in two re-estimated sets of separation measures: 1) a distance matrix and a travel time matrix between origin seaports and destination counties; and 2) a distance matrix and a travel time matrix between origin points of entry and destination counties. The final step in preparing the separation measures for model estimation was to estimate these measures between origin points and destination states rather than counties. This was made possible simply by computing for each origin point i and all destination counties $c \in C_j$ in state j the average distance and travel time. That is,

$$\forall i \in I,\ c_{ij}^{(1)} = \Sigma_n c_{ic_n}^{(1)}/n,$$

and

$$c_{ij}^{(2)} = \Sigma_n c_{ic_n}^{(2)}/n,\ c_n \in C_j$$

where n is the cardinality of C_j. Note that had we avoided the intermediate step of origin-to-destination county distance estimation, say, by computing distances to state geographical centers, origin-destination pairs with both ends in the same state would have been indistinguishable in terms of their separation.

Other Relevant Variables

As mentioned earlier, the weight and value of commodities shipped may be affected by a nexus of factors that are either origin-specific, destination-specific, or solely dependent on separation between an origin and a destination. Although study resources did not permit the collection of pertinent data on these factors, it could be interesting to observe how such factors can be accommodated in the proposed theoretical framework for future reference.

For origin-specific and destination-specific factors, Sen and Smith (1995) propose the following exploratory analysis: estimate A_i's and B_j's first using the maximum likelihood procedure described earlier, and then use these estimates as the dependent variable values in a model fitting procedure. Thus, the estimates of the A_i's can be associated with the origin-specific factors, while the estimates of the B_j's can be associated with the destination-specific factors.

Factors that can be solely attributed to the separation between an origin and destination can be accommodated in a deterrence function, an example of which is indeed the most general exponential function in equation (2). Openshaw and Connolly (1977) have compiled a list of possible functions and empirically compared several of them.

Results

The only decision necessary to apply the procedures described above (i.e., DSF, LDSF, and MS) is the choice of the flow unit. In the case of passenger transportation, a flow unit of 1 (consistent with a Poisson or multinomial distribution assumption) would be reasonable. In the case of freight shipments of goods, a basic unit of flow would appear to be a trainload (for shipments by rail) or a truckload (for shipments by truck). In the absence of mode-specific information as well as information related to the variation in modal size, we experimented with different values and determined an "optimal" (with regard to providing the best model fit) basic unit of flow of 100,000 pounds or dollars. This is not surprising given that the bulk of flows are long-distance shipments usually performed by large[3] trucks or rail (on a limited scale for the particular data) with an average shipment value of $2.03 per pound for the Transborder data and $1.39 per pound for the PIERS data. Interestingly, our results appeared to be quite insensitive to the choice of the flow unit, primarily because the flows were inordinately large. This is in agreement with previous work (Sen and Pruthi 1983).

The procedures described above were run to a tight convergence. For each iteration of the MS procedure, the DSF procedure attained a 10E–12 convergence as defined by equation (8) in less than 100 iterations. The MS procedure itself attained a

[3] Currently, the gross weight limit for a 6-axle combination truck is 80,000 pounds.

10E–06 convergence as defined by the right-hand side of equation (13) in less than 20 iterations.

Parameter Estimates and Model Fit for PIERS Data

Separate gravity models were estimated for both seaport-to-state (seaport data) and border port-to-state (transborder) weight and value flows. In particular, for the seaport data, several model specifications were tested based on transformations of the distance and time separation measures, $c_{ij}^{(1)}$ and $c_{ij}^{(2)}$, respectively, after a careful examination of residuals that showed the presence of outliers.

In the analysis of residuals, two difficulties were addressed. The first difficulty is the unequal variances of residuals, a consequence of the Poisson distribution. This problem was handled by considering, as Cochran suggests (Rao 1973, p. 393), instead of the plain residuals $N_{ij} - \hat{T}_{ij}$, the components

$$\sqrt{N_{ij}} - \sqrt{\hat{T}_{ij}} \tag{35}$$

The second difficulty stems from the very large number of residuals (the number is $I \times J$, the number of all origin-destination pairs). To tackle this concern, we followed a procedure described in Sen and Smith (1995, p. 456) consisting of drawing normal (rankit) plots of residuals. If the residuals are normal, such a plot would lie approximately on a straight line. Points that deviate sharply near the ends were tagged as outliers. Unusual jumps in the plots or other strongly nonlinear shapes could signal the need for transformations of the $c_{ij}^{(k)}$'s or for additional $c_{ij}^{(k)}$'s. There are clearly other ways of examining residuals. It is important to note here that the effort to conduct residual analysis in the gravity model case should not be seen as less than that for linear models where residual analysis is routinely carried out.

Indeed, the examination of residuals signaled the need for transformation of the $c_{ij}^{(k)}$'s. We found the square root transformation provided adequate fit for both weight flow and value flow estimates (table 2). The model provides an excellent fit for the data, because the Chi-square statistic hovers around its expected value of 1. In addition, the cell-to-cell weight flow estimates correlate very well with the data (Pearson correlation coefficient, r = 0.89, p < 0.01. The cell-to-cell value flow estimates also correlate very well with the data (r = 0.98, p < 0.01). Moreover, the weight and value length distributions in figure 1 (1% corresponds to 2,292,872,822 pounds) and figure 2 (1% corresponds to $3,198,123,011), respectively, seem to corroborate the previous results.

TABLE 2 Parameter Estimates for PIERS Data

	Weight flows		Value flows	
Parameter	Sample mean	Standard deviation	Sample mean	Standard deviation
θ_1	11.1227	1.1646	3.7854	0.3963
θ_2	–16.7447	1.7533	–11.1437	1.1668
X^2 ratio	0.7253	0.0000	1.1760	0.0000

Of interest in table 2 is the appearance of $\hat{\theta}_1$ with a positive value. This is due to collinearity between the used impedances, distance and travel time, a phenomenon that is known to adversely affect the sign of parameter estimates. Dropping one of the impedances would bias the parameter estimates left in the model. Given the robustness of the maximum likelihood procedure above in collinearity situations, all available impedance measures were retained (Sen and Smith 1995, chapter 5). After all, the sign of $\hat{\theta}_1$ would have changed to a negative value had we reparameterized the model and considered instead of travel time as the first impedance measure, the difference between distance and travel time (in appropriate units).

FIGURE 1 PIERS Data: Weight Length Frequency Distribution

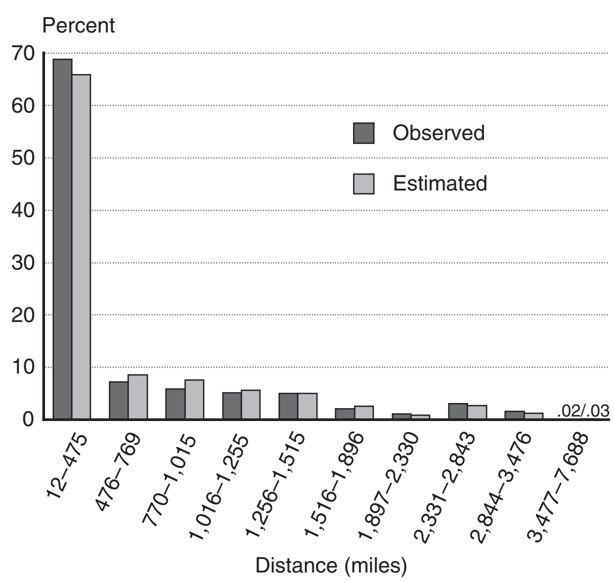
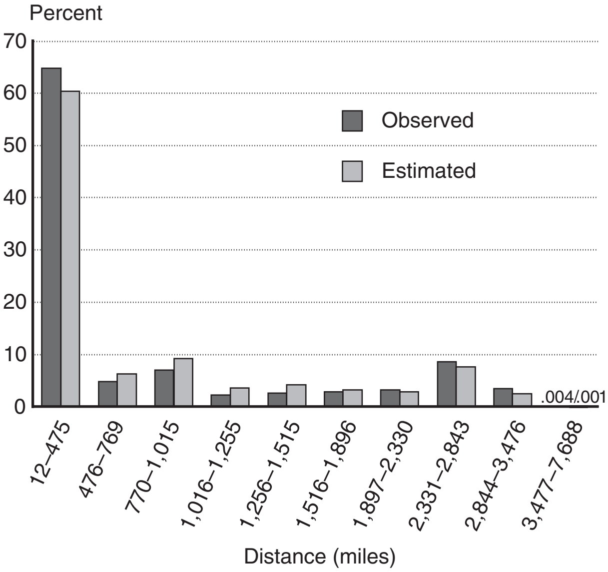


FIGURE 2 PIERS Data: Value Length Frequency Distribution



In the end, the model specifications for the weight and value flows, respectively, are

$$T_{ij} = A_i B_j \exp\left(\theta_1 \sqrt{c_{ij}^{(1)}} + \theta_2 \sqrt{c_{ij}^{(2)}} \right)$$

$$T_{ij} = A_i B_j \exp\left(\theta_1 c_{ij}^{(1)} + \theta_2 \sqrt{c_{ij}^{(2)}} \right) \quad (36)$$

Parameter Estimates and Model Fit for Transborder Data

In the case of Transborder data, we conducted a similar investigation of residuals (as above with the PIERS data) and observed the presence of a few remaining outliers, despite using transformed $c_{ij}^{(k)}$'s. These outliers were mainly states receiving an unusually large share of shipments compared with other states. The discovery was made after squaring the residuals in equation (35) and adding them together for all origin points of entry for each destination state. Destination states with a large share of shipments had a much larger sum of squares of residuals than other destinations. This was understandable, because these large destination states attract more shipments from longer distances and suggest the use of an additional impedance variable

$$c_{ij}^{(3)} = \delta \sqrt{\sqrt{c_{ij}^{(2)}}}$$

for the weight flows, and

$$c_{ij}^{(3)} = \delta \log[c_{ij}^{(2)}]$$

for the value flows, where δ is 1 for destinations with more than a 10% share in weight (4% in value) and 0 otherwise. In effect, we sought a different parameter estimate for the square root/log of distance for destination states with large shares.

The previous use of the indicator (dummy) variable δ is typical in spatial analysis for the treatment of residuals (Sen and Smith 1995) and introduces the effects of spatial structure on flow patterns (Gensler and Meade 1988; Fotheringham and O'Kelly 1989; Lo 1991), which is not the primary intent of this paper. More complex measures of relative location have been developed and tested empirically (Boots and Kanaroglou 1988; Lowe and Sen 1996). In the end, the model specifications for the weight and value flows, respectively, are

$$T_{ij} = A_i B_j \exp\left(\theta_1 \sqrt{c_{ij}^{(1)}} + \theta_2 \sqrt{c_{ij}^{(2)}} + \theta_3 \delta \sqrt{\sqrt{c_{ij}^{(2)}}} \right)$$

$$T_{ij} = A_i B_j \exp(\theta_1 c_{ij}^{(1)} + \theta_2 c_{ij}^{(2)} + \theta_3 \delta \log[c_{ij}^{(2)}]) \quad (37)$$

The previous steps removed all remaining outliers. Parameter estimates and goodness-of-fit statistics are shown in table 3. The previous observation regarding the sign of $\hat{\theta}_1$ applies. The model fits the data very well, as the Chi-square ratios for both weight and value remain under 2. In addition, the cell-to-cell weight flow estimates correlate very well with the data (Pearson correlation coefficient, $r = 0.91$, $p < 0.01$). The cell-to-cell value flow estimates also correlate very well with the data ($r = 0.94$, $p < 0.01$). Moreover, the weight and value length distributions in figure 3 (1% corresponds to 767,519,815 pounds) and figure 4 (1% corresponds to $1,565,313,371), respectively, seem to corroborate the previous results.

Computation of Covariance of Estimates for PIERS Data

The previous point estimates of the $\hat{\boldsymbol{\theta}}$ parameters and the flows $\hat{T}$ were used in the methodology described earlier to obtain the covariance of these estimates. The procedure can accommodate any

TABLE 3 Parameter Estimates for Transborder Data

	Weight flows		Value flows	
Parameter	Sample mean	Standard deviation	Sample mean	Standard deviation
θ_1	14.8743	1.5575	11.1739	1.1700
θ_2	–21.4296	2.2439	–13.6992	1.4344
θ_3	–2.1181	0.2217	–0.5534	0.0579
X^2 ratio	1.9642	0.0000	1.9012	0.0000

FIGURE 3 Transborder Data: Weight Length Frequency Distribution

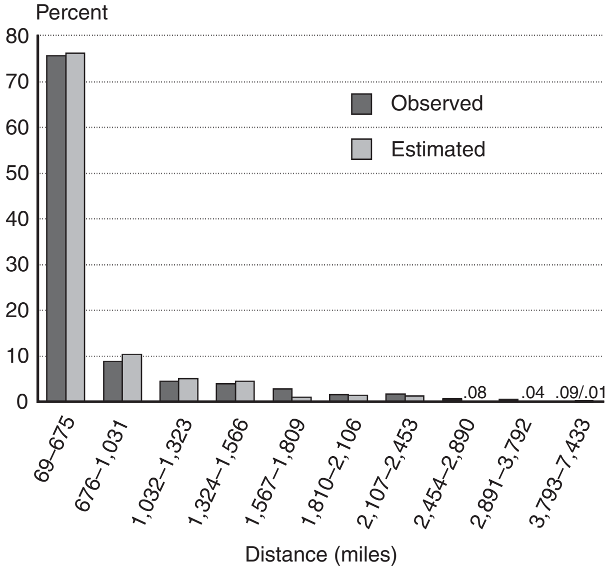
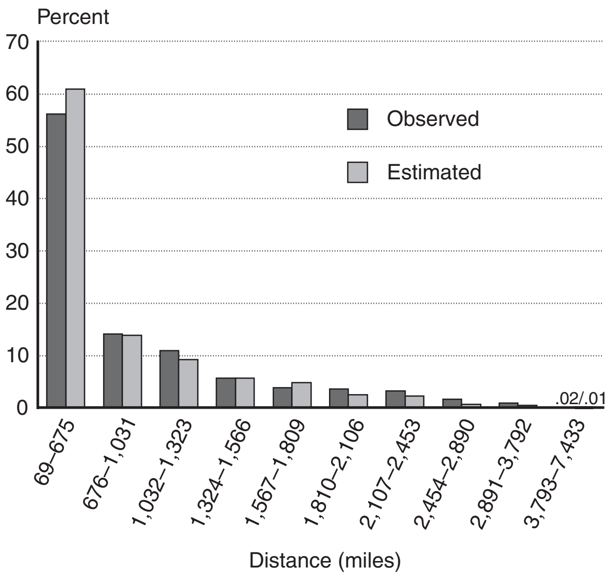


FIGURE 4 Transborder Data: Value Length Frequency Distribution



reasonable number of separation measures, $c_{ij}^{(k)}$, and the large number of origins and destinations typically encountered in practice.

The computational requirements of the procedure are no longer prohibitive. We were able to obtain the covariance matrix of a 144 × 50 flow matrix, a 7200 × 7200 matrix, in a little more than an hour on a Pentium III, 800 MHz, 512MB RAM laptop computer running a FORTRAN 77 compiler. Replacing T_{ij}'s by their estimates in equation (27), the covariance matrix of $\hat{\boldsymbol{\theta}}$ for the weight flows from the PIERS dataset was found to be

$$\left(\begin{array}{ccc} & \hat{\theta}_1 & \hat{\theta}_2 \\ \hat{\theta}_1 & 2.62830473 & -2.81464286 \\ \hat{\theta}_2 & -2.81464292 & 3.02464367 \end{array}\right) \tag{38}$$

The correlation between $\hat{\theta}_1$ (the parameter estimate for the distance measure) and $\hat{\theta}_2$ (the parameter estimate for the travel time measure) is readily apparent from equation (38). The negative sign of the covariances should not be disconcerting. It shows that, if for some small shift in the observations, $\hat{\theta}_1$ were to increase, $\hat{\theta}_2$ would decrease to "compensate."

Similarly, the covariance matrix of $\hat{\boldsymbol{\theta}}$ for the value flows from the PIERS dataset was found to be

$$\left(\begin{array}{ccc} & \hat{\theta}_1 & \hat{\theta}_2 \\ \hat{\theta}_1 & 1.53388134 & -1.66648732 \\ \hat{\theta}_2 & -1.66648735 & 1.81284692 \end{array}\right) \tag{39}$$

The correlation between $\hat{\theta}_1$ and $\hat{\theta}_2$ is similarly apparent from equation (39).

Computation of Covariance of Estimates for Transborder Data

The covariance matrices of $\hat{\boldsymbol{\theta}}$ from the Transborder dataset for the weight and value flows, respectively, were found to be

$$\left(\begin{array}{cccc} & \hat{\theta}_1 & \hat{\theta}_2 & \hat{\theta}_3 \\ \hat{\theta}_1 & 9.26053517 & -9.88856511 & 0.08654230 \\ \hat{\theta}_2 & -9.88856442 & 10.60290046 & -0.15624519 \\ \hat{\theta}_3 & 0.08654113 & -0.15624389 & 0.67484586 \end{array}\right) \tag{40}$$

and

$$\left(\begin{array}{cccc} & \hat{\theta}_1 & \hat{\theta}_2 & \hat{\theta}_3 \\ \hat{\theta}_1 & 2.87729568 & -3.13541449 & 0.01302693 \\ \hat{\theta}_2 & -3.13541410 & 3.42693579 & -0.02059522 \\ \hat{\theta}_3 & 0.01302676 & -0.02059503 & 0.01040623 \end{array}\right) \tag{41}$$

These last two covariance matrices in equations (40) and (41) show relatively high correlation between $\hat{\theta}_1$ (the parameter estimate for the distance

TABLE 4 An Illustration of N_{ij}, $\hat{T}_{ij}$ and 90% Confidence Intervals

N_{ij} (observed flows) (pounds)	$\hat{T}_{ij}$ (estimated flows) (pounds)	$N_{ij}-\hat{T}_{ij}$ % difference	90% confidence interval of $\hat{T}_{ij}$	
			Lower	Upper
4,850,905.00	4,764,382.12	1.78	4,764,179.19	4,764,585.05
87,454.00	88,012.62	0.63	87,934.47	88,090.77
39,922,216.00	38,847,527.51	2.69	38,845,275.15	38,849,779.87
16,002,283,960.00	15,556,162,969.96	2.78	15,555,972,990.43	15,556,352,949.49
216,787,500.00	211,385,682.81	2.49	211,374,117.32	211,397,248.30
4,174,427.00	4,266,148.17	2.19	4,266,017.84	4,266,278.50
3,076,999.00	3,100,768.36	0.77	3,100,346.82	3,101,189.90
2,029,122,669.00	2,108,940,894.95	3.93	2,108,896,284.86	2,108,985,505.04
693,089.00	664,724.67	4.09	664,638.45	664,810.89

measure) and $\hat{\theta}_2$ (the parameter estimate for the travel time measure) and relatively low correlation between $\hat{\theta}_3$ (the parameter estimate for the transformed travel time measure) and $\hat{\theta}_1$ or $\hat{\theta}_2$. The interpretation of the negative signs is the same as above.

For illustration purposes, an example of the usefulness of the computed covariance matrices for $\hat{T}_{ij}$'s in computing confidence intervals is shown in table 4. The table shows a few N_{ij}'s, $\hat{T}_{ij}$'s, percentage difference between N_{ij} and $\hat{T}_{ij}$, and 90% confidence intervals constructed by adding and subtracting 1.65 times the standard error available from the variances.

CONCLUSIONS

The most characteristic and indeed more restrictive feature of the Poisson distribution is the equality between its mean and variance. Many types of spatial interaction phenomena exhibit restricted (variance less than their means) or extra variation (variance greater than their means). Yet, the framework described above is very robust. Its robustness stems from the many types of asymptotic results that establish the Poisson distribution as the unique limiting form for a wide range of interaction processes when the population size increases and the overall influence of each individual interaction decreases.

More specifically, this research demonstrates that, to the extent the behavior of each shipping firm can be thought of as a very small interaction process, the resultant interaction process (i.e., the realized pattern of origin-destination commodity shipments) can be thought of as the superposition of all these processes and thus characterized by gravity models. The limited empirical analysis in this paper confirmed this conjecture and thus establishes an extremely rich framework for future experimentation.

Future applications of the proposed framework could conduct additional exploratory analyses to determine which factors affecting the demand for freight shipments are origin-based, destination-based, or separation-based. Another important issue needing more attention in the future is the requirement that, when the proposed modeling framework is used for forecasting, the travel times and costs used as impedance measures should be consistent with those obtained during the traffic assignment.

In summary, this paper has provided evidence that the movement of freight shipments can now be estimated within a desired confidence level as a result of maximum likelihood estimation of Poisson gravity models. The freight transportation modeler has two procedures available for computing reliable information within a predetermined accuracy: one for computing freight flow estimates and one for computing covariance matrices. These procedures can accommodate any reasonable number of separation measures, $c_{ij}^{(k)}$, and the large number of origins and destinations typically encountered in practice.

ACKNOWLEDGMENTS

This paper reports on research conducted by the author for the Bureau of Transportation Statistics, U.S. Department of Transportation.

REFERENCES

Bishop, Y.M.M., S.E. Fienberg, and P.W. Holland. 1975. *Discreet Multivariate Analysis: Theory and Practice*. Cambridge, MA: MIT Press.

Boots, B. and P. Kanaroglou. 1988. Incorporating the Effects of Spatial Structure in Discrete Choice Models of Migration. *Journal of Regional Science* 28:495–507.

Cambridge Systematics. 1997. A Guidebook for Forecasting Freight Transportation Demand, NCHRP Report 388. Transportation Research Board, National Research Council.

Fotheringham, A. and M. O'Kelly. 1989. Spatial Interaction Models: Formulations and Application. *Studies in Operational Regional Science*. Boston, MA: Kluwer Academic Publishers.

Gensler, W. and M. Meade. 1988. Locational and Population Factors in Health Care-Seeking Behavior in Savannah, Georgia. *Health Services Research* 22:443–462.

Lo, L. 1991. Spatial Structure and Spatial Interaction: A Simulation Approach. *Environment and Planning A* 23:1279–1300.

Lowe, J. and A. Sen. 1996. Gravity Model Applications in Health Planning Analysis of an Urban Hospital Market. *Journal of Regional Science* 36:(3)437–462.

Memmot, F.W. 1983. Application of Statewide Freight Demand Forecasting Techniques, NCHRP Report 260. Transportation Research Board, National Research Council.

Metaxatos, P. 2002. International Trade Traffic Study (Section 5115): Measuring the Ton-Miles and Value-Miles of International Trade Traffic Carried by Highway for Each State: Estimation of Origin-Destination Flows and Confidence Intervals, Final Report, prepared for the U.S. Department of Transportation, Bureau of Transportation Statistics. March.

Meyburg, A.H. and J.R. Mbwana, eds. 2002. Conference Synthesis: Data Needs in the Changing World of Logistics and Freight Transportation. New York Department of Transportation, Albany, NY. Available at http://www.dot.state.ny.us/ttss/conference/synthesis.pdf.

Openshaw, S. and C.J. Connolly. 1977. Empirical Derived Deterrence Functions for Maximum Performance Spatial Interaction Models. *Environment and Planning A* 9:1067–1079.

Rao, C.R. 1973. *Linear Statistical Inference and Its Application*. New York, NY: Wiley.

Sen, A. and T.E. Smith. 1995. *Gravity Models of Spatial Interaction Behavior.* New York, NY: Springer-Verlag

Sen, A. and R.K. Pruthi. 1983. Least Squares Calibration of the Gravity Model When Intrazonal Flows are Unknown. *Environment and Planning A* 15:1545–1550.

U.S. Department of Transportation (USDOT). 1996. *Quick Response Freight Manual,* DOT-T-97-10. Washington, DC. September 1996.

Weber, J.S. and A. Sen. 1985. On the Sensitivity of Gravity Model Forecasts. *Journal of Regional Science* 25:317–336.

Yun S. and A. Sen. 1994. Computation of Maximum Likelihood Estimates of Gravity Model Parameters. *Journal of Regional Science* 34:199–216.

Author address: Paul Metaxatos, University of Illinois at Chicago, 412 South Peoria Street, Suite 340, Chicago, IL 60607. Email: pavlos@uic.edu.

Estimation and Accuracy of Origin-Destination Highway Freight Weight and Value Flows

PAUL METAXATOS
University of Illinois at Chicago

ABSTRACT

This paper proposes a spatial interaction modeling framework and implements a maximum likelihood estimation of highway freight weight and value flows using the gravity model. The computation of the standard error of the flow estimates provides the basis for measuring the level of accuracy of the estimates. The results provide evidence of the suitability of gravity models for freight forecasting given the excellent fit and the small variances.

INTRODUCTION

The measurement of freight movements requires tracking freight flows across geographic and political boundaries. This is a particularly challenging task given the current capabilities for state and regional data acquisition. Various mathematical approaches have been implemented (Memmot 1983; USDOT 1996; Cambridge Systematics 1997) to circumvent this problem, but none, to the best of this author's knowledge, proposes a measure to assess the accuracy of the computed flows.

This paper proposes to fill this gap using developments in spatial interaction modeling that have not been demonstrated on a large scale to date. The methodology computes maximum likelihood flow

KEYWORDS: freight origin-destination flow estimation, covariance of estimates, gravity model.

estimates and obtains their covariance matrices that, in turn, may be used to obtain confidence intervals and carry out certain tests of hypotheses. The approach can accommodate the large number of origins and destinations typically encountered in freight (and passenger) travel forecasting.

The methodology was applied to highway freight weight and value flows of international trade traffic between seaports or border ports and destination states (see Metaxatos (2002) for details). The variances of the flow estimates computed were remarkably small. The demand for freight transportation flows can then be estimated within a desired confidence level. Moreover, the empirical analysis undertaken provides evidence that the theoretical framework proposed in this paper is rich enough for freight demand forecasting applications.

THEORETICAL FRAMEWORK

Commodity shipments in this paper are thought to be realized patterns of spatial interactions that typically result from many independent decisions by individual firms, each constituting a relevant subsystem within the economy as a whole. Hence, if the travel behavior of each firm is modeled as a very small interaction process, the resultant interaction process can be taken to be the superposition of all these processes. It may be argued that for large collections of small frequency processes, the resulting superimposed process is approximately Poisson and, therefore, completely characterized by its associated mean interaction frequencies (Sen and Smith 1995).

In this light, assuming that the observations N_{ij} of shipment weight and value between origin seaports/border ports of entry i and destination states j can be described by the gravity model, then

$$\begin{aligned} N_{ij} &= T_{ij} + \varepsilon_{ij} \\ T_{ij} &= E(N_{ij}) = A_i B_j F_{ij} \quad \forall i, j \end{aligned} \tag{1}$$

In this paper, T_{ij}'s (the stochastic term) are interpreted as the expected international trade traffic flow (in terms of weight and value) carried by highway from external station i to state j. The A_i's are factors related to the origin zone i and the B_j's are destination-related factors. The F_{ij}'s are factors that reflect the separation between i and j. A common form that is general enough for most applications is

$$F_{ij} = \exp\left[\sum_k \theta_k c_{ij}^{(k)}\right] \quad \forall i, j \tag{2}$$

This form is called an exponential form and $c_{ij}^{(k)}$ are different measures of separation, while θ_k's are parameters to be estimated. Potential measures of separation include travel time, distance, generalized costs, etc.

In the gravity model, observable quantities $N_{i*} = \Sigma_j N_{ij}$, $N_{*j} = \Sigma_i N_{ij}$, N_{ij} and their expected values T_{i*}, T_{*j}, and T_{ij} are described by means of an underlying structure consisting of unobservable quantities A_i, B_j, and F_{ij}. Similar situations abound in statistics. In moving average models, for example, observations are described by means of unobservable parameters. A like situation exists in analysis of variance models.

Although the origin and destination factors are unobservable, they do have physical interpretations. For example, if for some origin i, there are two destinations j and j' such that $F_{ij} = F_{ij'}$, then $T_{ij}/T_{ij'} = B_j/B_{j'}$. Thus other factors being equal, T_{ij} is proportional to B_j (but, in general, not proportional to T_{*j}), and is called the *attractiveness* of j. Similarly, the origin factor A_i may also be called the *emissiveness* of i.

Clearly, B_jF_{ij} is the effect of the destination factor B_j at i, or the *accessibility* of j as perceived from i. This is a spatial analogy of the temporal concept of present value in economics, where a dollar earned in n years in the future is worth only $(1 + \sigma)^{-n}$ now, where σ is the interest rate. Similarly, A_iF_{ij} is the effect of the origin factor A_i at j. The sum $\alpha_i = \sum_j B_j F_{ij}$ may be called the total accessibility of all destinations at i, and the sum $\beta_j = \sum_i A_i F_{ij}$ may be called the total accessibility of all origins at j. If, for example, T_{i*} is kept fixed as α_i increases, the push A_i decreases. Thus, as the competition α_i from the destinations increases, the push at i decreases. From the point of view of someone at i, α_i measures accessibility; from the viewpoint at j, it measures competition. Similar statements can be made about B_j.

Maximum Likelihood Estimation

The model (1) will be estimated using maximum likelihood (ML). Maximum likelihood estimates have desirable asymptotic properties (consistency, efficiency, and asymptotic normality) and are robust to distributional assumptions for realistic departures from the Poisson assumption (note that the multinomial distribution leads to identical estimates with the Poisson distribution). Furthermore, they are essentially unbiased even for a very small sample of flows (Sen and Smith 1995).

Under some mild conditions (Sen and Smith 1995), the ML estimate of $\boldsymbol{\zeta} = (A(1),\ldots,A(I), B(1),\ldots,B(J),\theta_1,\ldots,\theta_Q)'$ exists and that of $\boldsymbol{\theta}$ is unique. The estimates of $\boldsymbol{A}$ and $\boldsymbol{B}$ are not unique; however, if one $A(i)$ or $B(j)$ is chosen to be an arbitrary positive number, the remaining $A(i)$'s and $B(j)$'s are unique under the previous mild conditions. It can be proved that the ML estimate of $\boldsymbol{\zeta}$ results in the solution to the following system of equations

$$T_{i*} = N_{i*} \quad \forall i \in I \tag{3}$$

$$T_{*j} = N_{*j} \quad \forall j \in J \tag{4}$$

$$\sum_{ij} c_{ij}^{(q)} T_{ij} = \sum_{ij} c_{ij}^{(q)} N_{ij} \quad \forall q \in Q \tag{5}$$

where the operator * indicates summation with respect to the subscript it replaces (e.g., $T_{i*} = \Sigma_j T_{ij}$, $T_{*j} = \Sigma_i T_{ij}$). A number of standard numerical methods or more specialized procedures can be used to solve equations (3) through (5). The three procedures adapted for the paper are the Deming-Stefan-Furness (DSF) procedure, the linearized DSF procedure, and the Modified Scoring procedure. An account of the development of the three procedures is given in Sen and Smith (1995). For completeness of the presentation, some of the details for each procedure follow.

The DSF Procedure for Parameters A_i and B_j

The DSF procedure gives values of T_{ij} for any choice of ($\boldsymbol{\theta}'$), and N_{i*}, N_{*j}, with $\Sigma_i N_{i*} = \Sigma_j N_{*j}$. Generally speaking, this procedure adjusts the rows (columns) of a two-dimensional table in each even (odd) iteration. After choosing an initial value for the column balancing coefficient, $B_j^0 = 1$, say, the DSF procedure iterates as follows (the index r denotes the iteration number):

$$A_i^{(2r-1)} = O_i \Big/ \sum_{j=j+1}^{J} B_j^{(2r-2)} F_{ij} \qquad \forall i \tag{6}$$

$$B_j^{(2r)} = D_j \Big/ \sum_{i=1}^{I} A_i^{(2r-1)} F_{ij} \qquad \forall j \tag{7}$$

where $O_i = T_{i*}$, $D_j = T_{*j}$, and F_{ij} is a function of the separation measures $c_{ij}^{(q)}$. Upon convergence (see Sen and Smith (1995) for a proof of convergence), the values of T_{ij} are given by

$$T_{ij}^{(2r)} = A(i)^{(2r-1)} B(j)^{(2r)} F_{ij}.$$

In this paper, we chose a fairly stringent criterion for convergence as follows

$$\sum_{i=1}^{I} |O_i - T_{i*}^{(2r)}| + \sum_{j=1}^{J} |D_j - T_{*j}^{(2r)}| < \delta \tag{8}$$

where $\delta = 10^{-12}$. The algorithm attained this criterion in less than 100 iterations.

The DSF procedure essentially expresses T_{ij} as a function of $\boldsymbol{\theta}$. These values of T_{ij} could then be used in equation (5) to solve for an updated value of $\boldsymbol{\theta}$. In general, however, as $\boldsymbol{\theta}$ changes, equations (3) and (4) could be violated, unless these changes are very small. This is achieved by using a linearized version of the DSF procedure, called the LDSF procedure (Weber and Sen 1985), which is computationally very attractive.

The LDSF Procedure for T_{ij}

Let us assume that we have run the DSF procedure and obtained a good set of T_{ij} that solves equations (3) and (4) for any given θ. This means that $O_i = T_{i*} = N_{i*}$ and $D_j = T_{*j} = N_{*j}$. Define by $\Delta\boldsymbol{O} = (\Delta O_1,\ldots,\Delta O_I)'$ and $\Delta\boldsymbol{D} = (\Delta D_1,\ldots,\Delta D_J)'$ to be small changes in the values of $\mathbf{O} = (O_1,\ldots,O_I)'$ and $\boldsymbol{D} = (D_1,\ldots,D_J)'$, respectively. Also, let ΔF_{ij}, be a small change in $F_{ij} = \exp[\boldsymbol{\theta}'\boldsymbol{c}_{ij}]$. It can be proved (Sen and Smith 1995) that the corresponding small change, ΔT_{ij}, in each T_{ij}, so that $\Delta T_{i*} = \Delta O_i$, and $\Delta T_{*j} = \Delta D_j$ can be obtained by the LDSF procedure, which iterates as follows

$$\Delta T_{ij}^{(2r-1)} = \Delta T_{ij}^{(2r-2)} + (T_{ij}/O_i)(\Delta O_i - \Delta T_{i*}^{(2r-2)}) \tag{9}$$

$$\Delta T_{ij}^{(2r)} = \Delta T_{ij}^{(2r-1)} + (T_{ij}/D_j)(\Delta D_j - \Delta T_{*j}^{(2r-1)}) \tag{10}$$

for $i \in I$, $j \in J$ and $r = 0, 1, 2, \ldots$, with initial T_{ij} values given by

$$\Delta T_{ij}^{(0)} = (T_{ij}/F_{ij})\Delta F_{ij} \tag{11}$$

A proof for the convergence of the procedure is given in Weber and Sen (1985).

Changes in T_{ij} as a Function of a Change in θ

For a small change $\Delta\boldsymbol{\theta}$ in $\boldsymbol{\theta}$ and a small change 0 so that $\Delta O = \Delta D = \mathbf{0}$, it can be proved (Yun and Sen 1994) that an approximation for the corresponding small change ΔT_{ij} for each T_{ij} for all $i \in I$ and $j \in J$ is given by equation 12.

$$\Delta T_{ij} \approx \Delta\boldsymbol{\theta}\left\{ \boldsymbol{c}_{ij}T_{ij} - \sum_j \boldsymbol{c}_{ij}T_{ij}\left(\frac{T_{ij}}{O_i}\right) - \sum_i \boldsymbol{c}_{ij}T_{ij}\left(\frac{T_{ij}}{D_j}\right) + \sum_i \left[\sum_j \boldsymbol{c}_{ij}T_{ij}\left(\frac{T_{ij}}{O_i}\right)\right]\left(\frac{T_{ij}}{D_j}\right)\right\} = \boldsymbol{S}_{ij}\Delta\boldsymbol{\theta} \tag{12}$$

The $S_{ij}^{(k)}$'s are constants, and $O_i = T_{i*}$, $D_j = T_{*j}$. Therefore, if T_{ij}'s are known, the only unknown in equation (12) is the $\Delta\boldsymbol{\theta}$. The solution for the $\Delta\boldsymbol{\theta}$ will be the topic of the next section.

Estimation of θ Using the Modified Scoring (MS) Procedure

So far, for an initial value for $\boldsymbol{\theta}$, we obtained $T_{ij}(\boldsymbol{\theta})$ by using the DSF procedure to solve equations (3) and (4). We then changed $\boldsymbol{\theta}$ to $\boldsymbol{\theta} + \Delta\boldsymbol{\theta}$ and computed, using the LDSF procedure, with $\Delta O = \Delta D = \mathbf{0}$, the corresponding change, $\Delta T_{ij}(\boldsymbol{\theta}, \Delta\boldsymbol{\theta})$ in $T_{ij}(\boldsymbol{\theta})$ as a function of $\Delta\boldsymbol{\theta}$.

We are ready now to insert the $[T_{ij}(\boldsymbol{\theta}) + \Delta T_{ij}(\boldsymbol{\theta}, \Delta\boldsymbol{\theta})]$'s into the left side of equation (5) and solve the resultant equation for $\Delta\boldsymbol{\theta}$. Inserting $T_{ij} + \Delta T_{ij}$ in place of T_{ij} in (5) and using (12), equations (3) and (4) would remain approximately satisfied, while obtaining the following system of Q linear equations with Q unknowns (the $\Delta\theta_q$'s):

$$\begin{aligned} \sum_{ij} c_{ij}^{(1)}(\Delta T_{ij}) &= \sum_{ij} c_{ij}^{(1)}(N_{ij} - T_{ij}) \\ &\vdots \\ \sum_{ij} c_{ij}^{(Q)}(\Delta T_{ij}) &= \sum_{ij} c_{ij}^{(Q)}(N_{ij} - T_{ij}) \end{aligned} \tag{13}$$

This system of equations can be solved by any standard solution method such as Gaussian elimination.

The current solution for $\boldsymbol{\theta}$ at iteration r is updated next using the formula

$$\boldsymbol{\theta}^r = \boldsymbol{\theta}^{r-1} + \Delta\boldsymbol{\theta}^{r-1} \tag{14}$$

If the corrections $\Delta\boldsymbol{\theta}^{r-1}$ have become negligible, the values of $\boldsymbol{\theta}$ have been stabilized and the MS procedure terminates. Otherwise, new T_{ij}'s are obtained from the DSF procedure and the MS procedure continues. There is no guarantee that the MS procedure always converges (Sen and Smith 1995), although our computational experience is positive.

Goodness of Fit

Under the previous assumption that observations N_{ij} are independently Poisson distributed, the (Pearson) X^2 statistic,

$$X^2 = \sum_{ij} \frac{(N_{ij} - \hat{T}_{ij})^2}{\hat{T}_{ij}} \tag{15}$$

where $\hat{T}_{ij}$ is an estimate of T_{ij}, is an appropriate measure of the overall fit of a model. Moreover, when $\hat{T}_{ij}$ is obtained using maximum likelihood, equation (15) has a X^2 distribution with df = $IJ - I - J - K + 1$ degrees of freedom (Bishop et al. 1975; Rao 1973).

If $\hat{T}_{ij} \approx T_{ij}$, then $X^2 = Z^2$, where

$$Z^2 = \sum_{ij} \frac{(N_{ij} - T_{ij})^2}{T_{ij}} \tag{16}$$

Since $E(N_{ij}) = T_{ij}$ and because N_{ij} have the Poisson distribution,

$$\text{var}(N_{ij}) = E(N_{ij} - T_{ij})^2 = T_{ij} \tag{17}$$

Therefore, $E(Z^2) = IJ$, where I is the number of origin zones i, and J the number of destinations j. Equivalently, $E(Z^2/IJ) = 1$. Thus, the so-called "X^2-ratio," X^2/df, has an expectation that is asymptotically 1. It can be shown (Sen and Smith 1995) that the variance of the X^2-ratio is

$$\text{var}[Z^2/(IJ)] \approx \sum_{ij}[(T_{ij}I^2J^2)^{-1} + 2(IJ)^{-2}] \tag{18}$$

Hence, if T_{ij}'s are bounded away from zero (which is the case in exponential gravity models with finite parameters $\boldsymbol{\theta}$), the variance of $Z^2/IJ \to 0$, as

$IJ \to \infty$. It follows that when $\hat{T}_{ij} \to T_{ij}$ and T_{ij}'s are bounded away from zero, the variance of $X^2/df \to 0$.

In practical applications, since the Poisson assumption seldom holds perfectly (as is the case here, where every pound or dollar value of shipment does not travel independently of each other pound or dollar value), an X^2 ratio less than 2 is a good indication that the gravity model fits the data well (Sen and Smith 1995).

Covariance of Maximum Likelihood Estimates

Covariance of $\hat{\theta}_q$'s

Let small case letters stand for the logarithms of corresponding capital letters (e.g., $t_{ij} = log[T_{ij}]$, $a(i) = log[A(i)]$, $b(j) = log[B(j)]$. The model (1) may be written as

$$t_{ij} = a(i) + b(j) + \sum_q \theta_q c_{ij}^{(q)} \quad \forall i \in I, j \in J, q \in Q \quad (19)$$

Let M denote the coefficient matrix of the right side of the system of equations (19). The matrix M is not of full rank (Sen and Smith 1995). However, the matrix $M_{(2)}$ obtained by deleting one of the first $I + J$ columns of M is of full rank and has dimension $IJ \times (I + J + Q - 1)$. Let diag$(\cdot)$ stand for a diagonal matrix, the diagonal elements of which are given within the parentheses. Then compute the matrix $M'_{(2)} \cdot \mathrm{diag}(\boldsymbol{T}) \cdot M_{(2)}$ from the equation

$$M'_{(2)} \cdot \mathrm{diag}(\boldsymbol{T}) \cdot M_{(2)} = \begin{pmatrix} U_1 & U_2 \\ U'_2 & U_3 \end{pmatrix} \quad (20)$$

where

$$U_1 = \begin{pmatrix} V_1 & V_2 \\ V'_2 & V_3 \end{pmatrix} \quad (21)$$

$$U_2 = \begin{pmatrix} W_1 \\ W_2 \end{pmatrix} \quad (22)$$

$U_3 = ((u_{pq}))$ with $u_{pq} = \Sigma_{ij} c_{ij}^{(p)} c_{ij}^{(q)} T_{ij}, W_1 = ((w_{iq}^{(1)}))$ with $w_{iq}^{(1)} = \Sigma_j c_{ij}^{(q)} T_{ij}$, $W_2 = ((w_{jq}^{(2)}))$ with $w_{jq}^{(2)} = \Sigma_i c_{ij}^{(q)} T_{ij}$, $V_1 = \mathrm{diag}(T_{1*},...,T_{I*})$, $V_2 = ((T_{ij}))$ and $V_3 = \mathrm{diag}(T_{*1},...,T_{*J-1})$. Notice that the subscript j in each of the matrices above goes only up to $J - 1$.

Matrix $M'_{(2)} \cdot \mathrm{diag}(\boldsymbol{T}) \cdot M_{(2)}$, a square matrix of dimension $(I + J + Q - 1)$, is the covariance matrix of $M'_{(2)} \cdot \mathrm{diag}(\boldsymbol{N})$. This is because the N_{ij}'s have independent Poisson distributions and the covariance matrix Cov($\boldsymbol{N}$) of $\boldsymbol{N}$ is diag($\boldsymbol{T}$). It can be shown (Sen and Smith 1995) that the covariance matrix of $(\hat{A}(1),...,\hat{A}(I), \hat{B}(1),...,\hat{B}(J-1), \hat{\theta}_1,...,\hat{\theta}_Q)'$ is

$$\Phi^{-1} \cdot M'_{(2)} \cdot \mathrm{diag}(\boldsymbol{T}) \cdot M_{(2)} \cdot (\Phi^{-1})' \quad (23)$$

where

$$\Phi = M'_{(2)} \cdot \mathrm{diag}(T) \cdot M_{(2)} \cdot \mathrm{diag}(1/A(1),...,1/A(I), 1/B(1),...,1/B(J-1), 1,...,1) \quad (24)$$

and

$$\Phi^{-1} = \mathrm{diag}(1/A(1),...,1/A(I), 1/B(1),...,1/B(J-1), 1,...,1) \cdot (M'_{(2)} \cdot \mathrm{diag}(T) \cdot M_{(2)})^{-1} \quad (25)$$

Notice that using equations (23) through (25), another expression for the covariance matrix of $(\hat{A}(1),...,(\hat{A}(I), \hat{B}(1),...,\hat{B}(J)-1), \hat{\theta}_1,...,\hat{\theta}_Q)'$ can be written by as follows

$$\mathrm{diag}(1/A(1),...,1/A(I), 1/B(1),...,1/B(J-1),1,...1) \cdot (M'_{(2)} \cdot \mathrm{diag}T \cdot M_{(2)})^{-1} \cdot \mathrm{diag}(1/A(1),...,1/A(I), 1/B(1),...,1/B(J-1), 1,...,1) \quad (26)$$

The bottom right $Q \times Q$ submatrix of matrix (26) is the estimated covariance matrix of $\hat{\boldsymbol{\theta}}$. The bottom right $Q \times Q$ submatrix of the inverse of (20) is (Rao 1973)

$$\left(U_3 - U'_2 U_1^{-1} U_2\right)^{-1} \quad (27)$$

From equation (26), it follows that equation (27) is the covariance matrix of $\hat{\boldsymbol{\theta}}$.

Covariance of $\hat{T}_{ij}$'s

Having obtained Cov($\boldsymbol{A}$,$\boldsymbol{B}$,$\boldsymbol{\theta}$) from equation (23) or using equation (26), and since $B(J)$ is set equal to a constant, and its variance and covariances involving it are zeros, it can readily be seen that the covariance matrix of $\hat{T}$, denoted by the symbol $\mathrm{Cov}(\hat{T})$, is

$$\mathrm{Cov}(\hat{T}) = \Psi_1 \cdot \mathrm{Cov}(A, B, \boldsymbol{\theta}) \cdot \Psi'_1 \tag{28}$$

with

$$\Psi_1 = \mathrm{diag}(T) \cdot M \cdot \mathrm{diag}(1/A(1),...,1/A(I),$$

$$1/B(1),...,1/B(J), 1,...,1) \tag{29}$$

and M the coefficient matrix of the right side of the system of equations (19).

Short-Term Forecasting

The shipment of goods is affected by a multitude of factors (USDOT 1996, table 2.1). The weight and value of commodities shipped, characteristics of immediate concern in this study in particular, may be affected by the economy as a whole, globalization of business, international trade agreements, just-in-time inventory practices (weight only), packaging materials (weight only), economic regulation/deregulation, publicly provided infrastructure (weight only), user charges and other taxes, changes in truck size and weight limits (weight only), and technological advances. The discussion below is based on the assumption that, in the short term (e.g., three to five years) the compound effect of these factors on the size of the weight and value characteristics of commodity shipments is consistent with Poisson randomness.

Let the random variable $N_{ij}^{(f)}$ be a future observation[1] of the flow from i to j. In the short run, under the assumption that the separation configuration will not change during the forecast period, we may conjecture that $N_{ij}^{(f)}$ and N_{ij} will be highly (serially) correlated. Thus, we may then argue that the variance of the difference of future and present observations will be smaller than the variance of future observations alone. Hence,

$$\mathrm{var}\left(N_{ij}^{(f)} - N_{ij}\right) = \mathrm{var}\left(N_{ij}^{(f)}\right) + \mathrm{var}(N_{ij}) - 2\,\mathrm{Cov}\left(N_{ij}^{(f)} N_{ij}\right) \tag{30}$$

where $\mathrm{var}(\cdot)$ stands for the "variance of" and $\mathrm{Cov}(\cdot)$ for the "covariance of," will be small. In such cases, especially if N_{ij}'s are known, Sen and Smith (1995, chapter 5) suggest that it would be preferable to predict $(N_{ij}^{(f)} - N_{ij})$'s and add the predictions to the N_{ij}'s. A possible set of predictions for $(N_{ij}^{(f)} - N_{ij})$'s are $(\hat{T}_{ij}^{(f)} - \hat{T}_{ij})$'s. Thus, if the future is not too far off,

$$N_{ij} + (\hat{T}_{ij}^{(f)} - \hat{T}_{ij}) \tag{31}$$

may yield a better prediction than $\hat{T}_{ij}$. The computation of $\Delta\hat{T}_{ij} = \hat{T}_{ij}^{(f)} - \hat{T}_{ij}$ can be made easily using the LDSF procedure described earlier.

[1] The superscript f denotes that the quantity is an estimate for the forecast period.

Other Forecasts

Four separate cases are discussed here. The first case assumes that $c_{ij}^{(q)f}$ would be available exogenously, $\boldsymbol{\theta}^f$ would be the estimate $\hat{\boldsymbol{\theta}}$ from the base period assumed to remain unchanged into the forecast period, and the estimates $A^f(i)$ and $B^f(j)$ could be base period estimates assumed to remain unchanged into the forecast period, or one or both of them could be exogenous. Then, T_{ij} for the forecast period may be estimated by

$$T_{ij}^f = A^f(i)B^f(j)\exp[(\boldsymbol{\theta}^f)'\boldsymbol{c}_{ij}{}^f] \tag{32}$$

The calculation of $\mathrm{Cov}(\hat{\boldsymbol{T}}^f)$ requires the computation of the covariance matrices of the $A^f(i)$'s, $B^f(j)$'s, and $\boldsymbol{\theta}^f$. For those estimates that remain unchanged from the base period, the appropriate covariance matrix is the one for the base period. For estimates obtained exogenously, the covariance matrix needs to be supplied exogenously. Covariances of estimates obtained from different independent data are usually assumed to be zeros.

The second case assumes that $c_{ij}^{(q)f}$ and $\boldsymbol{\theta}^f$ would be available as before and the estimates $B^f(j)$ and T_{*j}^f could be base period estimates assumed to remain unchanged into the forecast period or could be obtained exogenously. Then, T_{ij} for the forecast period may be estimated by

$$T_{ij}^f = \frac{T_{*j}^f B^f(j)\exp[(\boldsymbol{\theta}^f)'\boldsymbol{c}_{ij}{}^f]}{\sum_j B^f(j)\exp[(\boldsymbol{\theta}^f)'\boldsymbol{c}_{ij}{}^f]} \tag{33}$$

The second case assumes that $c_{ij}^{(q)f}$ and $\boldsymbol{\theta}^f$ would be available as before and the estimates $A^f(i)$ and T_{i*}^f could be base period estimates assumed to remain unchanged into the forecast period or could be obtained exogenously. Then, T_{ij} for the forecast period may be estimated by

$$T_{ij}^f = \frac{T_{i*}^f A^f(i)\exp[(\boldsymbol{\theta}^f)'\boldsymbol{c}_{ij}{}^f]}{\sum_j A^f(i)\exp[(\boldsymbol{\theta}^f)'\boldsymbol{c}_{ij}{}^f]} \tag{34}$$

The computation of the covariance matrices for the second and third cases is similar to the first case once the appropriate Jacobians with respect to $A^f(i)$'s, $B^f(j)$'s and $\boldsymbol{\theta}^f$ have been obtained. Recall that equation (29) is the respective Jacobian for the first case.

A fourth case, arising when $c_{ij}^{(q)f}$, $\boldsymbol{\theta}^f$, T_{i*}^f, and T_{*j}^f are available, requires the use of the DSF procedure to generate the $A^f(i)$ and $B^f(j)$. Then, T_{ij} for the forecast period may be estimated by equation (32). The computation of the covariance matrices for this case is treated in detail by Sen and Smith (1995, p. 440).

It is interesting to observe at this point that origin-destination travel times and/or costs (which depend on routes chosen and traffic congestion on the routes) usually become available only after the gravity model has been applied to forecast flows. This difficulty affects only the forecasting phase (not the estimation phase, when we use base period data for which travel times can be observed or computed using available information) and can be alleviated by combining the distribution and assignment stages of the process into a single stage.

EMPIRICAL ANALYSIS

Data Issues

Freight Shipments

Oak Ridge National Laboratory (ORNL) provided two sets of origin-destination flow data for freight shipment weight and value between origin seaports/border ports of entry i and destination states j from the following sources:

- **The Transborder Surface Freight Database.** This database (http://www.bts.gov/transborder/prod.html) is distributed by the Bureau of Transportation Statistics and contains freight flow data by commodity type and surface mode of transportation (rail, truck, pipeline, or mail) for U.S. exports to and imports from Canada and Mexico.
- **The Port Import Export Reporting Service (PIERS) Database.** This commercial database (http://www.piers.com/about/default.asp), offered by the *Journal of Commerce*, offers statistics on global cargo movements transiting seaports in the United States, Mexico, and South America to companies around the globe.

Two matrices, N_{ij}, were developed: one for shipment weight and one for shipment value for each of the two databases with dimensions 128 x 50 and 144 x 50 for the Transborder and PIERS databases, respectively.

The limitations of these two data sources are well documented (Meyburg and Mbwana 2002). A caveat for the Transborder Database is that it is a customs, not a transportation, dataset, resulting in inconsistencies between U.S. and Canadian data and issues related to the accuracy of transshipment data.[2] One caveat for using PIERS is that reported origins and destinations may be billing addresses rather than shipment points.

Separation Measures

In identifying specific types of spatial separation that tend to impede or enhance the likelihood of interactions between points of entry and destination sites, the most obvious type involves physical space, as exemplified by travel distance and travel time, which are quantifiable in terms of meaningful units of measurement. ORNL provided two sets with separate impedance matrices, $c_{ic}^{(k)}$, k = 1, 2 from origin point i to destination county c.

The first dataset included impedances between origin seaports and destination counties, the second set between origin border ports and destination counties. Both measures were computed from ORNL's National Highway Network (NHN). The first of the two measures is the route (network) distance in miles; the second computes a function of travel time for different functional classifications of highway segments and adds these time penalties together while tracing the previous network routes. The previous impedance matrices presented two problems: 1) more than 50% of the cells were empty because not all seaports/border ports of entry are connected to each county through the NHN; and 2) occasionally, freight shipments moved

[2] See http://www.bts.gov/ntda/tbscd/desc.html, as of January 2004, for more information.

between origin-destination pairs with a missing connection (separation measure).

In order to deal with both problems, we devised and implemented the following stepwise procedure for the distance matrix (the first separation measure, $c_{ic}^{(1)}$).

1. Centroid coordinates were computed for all counties nationwide.
2. Spherical distances in miles from each county to every other county.
3. For every destination county with a missing connection to an origin point (seaport/border port of entry), the closest county with an existing connection to an origin point was estimated based on the previous county-to-county distances.
4. The missing connection from an origin point to a destination county was finally computed to be the distance between the same point of origin and the closest destination county augmented by 130% the (airline) distance between the two destination counties (as a proxy to the actual road distance). This simplification was made under the assumption that a missing connection would imply that the destination county is off the NHN and thus would require additional time to be accessed from its closest county on the NHN.

The above procedure resulted in the construction of a synthesized separation measure comprising both route distance (the ORNL estimate) and multiples of airline distance (our estimate). It is important to take a closer look at those two components of the (re-estimated) first separation measure $c_{ic}^{(1)}$.

The airline distance between origin point i and destination county c, even augmented by 30%, is only an approximation of the actual miles traveled and represents a surrogate for the complex set of factors that express the difficulty of overcoming separation. On the other hand, the route distance for the same origin-destination pair is the total over-the-road distance on a realistic route. As a measure of separation, the route distance provides better accuracy for the eastern half of the limited access highway system of the United States, which has a higher level of complexity resulting in less circuitous routes than in the western half.

Whereas the distance measure developed above is a measure that depends more or less on the physical characteristics of the road link, the impedance measure, a function of link travel time as estimated by ORNL, depends on the special roadway type or certain conditions encountered on the link. ORNL's estimating procedure could not be replicated in this paper for the missing impedances. A surrogate value was computed based on the average speed obtained by ORNL's distance and time estimates, and the synthesized distances.

The procedure classified ORNL's distance and time estimates into 10 deciles. The stratum average speed was computed by the ratio of the average distance and time for that stratum (table 1). Each

TABLE 1 Average Distance, Time, and Speed by Decile

	PIERS Database			Transborder Database		
Decile	Average distance (miles)	Average time (min.)	Estimated speed (mph)	Average distance (miles)	Average time (min.)	Estimated speed (mph)
1	230.69	241.56	57.30	381.11	400.32	57.12
2	468.65	465.70	60.38	708.18	711.44	59.73
3	642.53	625.89	61.59	929.82	920.37	60.61
4	797.46	769.51	62.17	1,121.06	1,093.26	61.52
5	954.44	915.38	62.56	1,291.37	1,251.02	61.93
6	1,127.45	1,077.43	62.78	1,457.91	1,405.62	62.23
7	1,324.50	1,261.04	63.01	1,632.94	1,569.44	62.42
8	1,596.70	1,512.70	63.33	1,850.64	1,771.00	62.69
9	2,109.35	1,979.77	63.92	2,166.76	2,048.56	63.46
10	2,958.11	2,963.42	59.89	2,757.87	2,641.23	62.65

missing time impedance was then estimated by the ratio of its corresponding distance estimate and the average speed of the stratum to which the distance estimate belongs.

Clearly, this procedure imposes some inconsistency in the imputation of travel times, because the average speed estimate is based on decile groupings of the initial estimates. The consistency of the procedure could be improved if in a second iteration the decile groupings are based on all distances (initial and imputed), but I did not attempt this for this paper.

The previous steps resulted in two re-estimated sets of separation measures: 1) a distance matrix and a travel time matrix between origin seaports and destination counties; and 2) a distance matrix and a travel time matrix between origin points of entry and destination counties. The final step in preparing the separation measures for model estimation was to estimate these measures between origin points and destination states rather than counties. This was made possible simply by computing for each origin point i and all destination counties $c \in C_j$ in state j the average distance and travel time. That is,

$$\forall i \in I,\ c_{ij}^{(1)} = \textstyle\sum_n c_{ic_n}^{(1)}/n,$$

and

$$c_{ij}^{(2)} = \textstyle\sum_n c_{ic_n}^{(2)}/n,\ c_n \in C_j$$

where n is the cardinality of C_j. Note that had we avoided the intermediate step of origin-to-destination county distance estimation, say, by computing distances to state geographical centers, origin-destination pairs with both ends in the same state would have been indistinguishable in terms of their separation.

Other Relevant Variables

As mentioned earlier, the weight and value of commodities shipped may be affected by a nexus of factors that are either origin-specific, destination-specific, or solely dependent on separation between an origin and a destination. Although study resources did not permit the collection of pertinent data on these factors, it could be interesting to observe how such factors can be accommodated in the proposed theoretical framework for future reference.

For origin-specific and destination-specific factors, Sen and Smith (1995) propose the following exploratory analysis: estimate A_i's and B_j's first using the maximum likelihood procedure described earlier, and then use these estimates as the dependent variable values in a model fitting procedure. Thus, the estimates of the A_i's can be associated with the origin-specific factors, while the estimates of the B_j's can be associated with the destination-specific factors.

Factors that can be solely attributed to the separation between an origin and destination can be accommodated in a deterrence function, an example of which is indeed the most general exponential function in equation (2). Openshaw and Connolly (1977) have compiled a list of possible functions and empirically compared several of them.

Results

The only decision necessary to apply the procedures described above (i.e., DSF, LDSF, and MS) is the choice of the flow unit. In the case of passenger transportation, a flow unit of 1 (consistent with a Poisson or multinomial distribution assumption) would be reasonable. In the case of freight shipments of goods, a basic unit of flow would appear to be a trainload (for shipments by rail) or a truckload (for shipments by truck). In the absence of mode-specific information as well as information related to the variation in modal size, we experimented with different values and determined an "optimal" (with regard to providing the best model fit) basic unit of flow of 100,000 pounds or dollars. This is not surprising given that the bulk of flows are long-distance shipments usually performed by large[3] trucks or rail (on a limited scale for the particular data) with an average shipment value of \$2.03 per pound for the Transborder data and \$1.39 per pound for the PIERS data. Interestingly, our results appeared to be quite insensitive to the choice of the flow unit, primarily because the flows were inordinately large. This is in agreement with previous work (Sen and Pruthi 1983).

The procedures described above were run to a tight convergence. For each iteration of the MS procedure, the DSF procedure attained a 10E–12 convergence as defined by equation (8) in less than 100 iterations. The MS procedure itself attained a

[3] Currently, the gross weight limit for a 6-axle combination truck is 80,000 pounds.

$10E$–06 convergence as defined by the right-hand side of equation (13) in less than 20 iterations.

Parameter Estimates and Model Fit for PIERS Data

Separate gravity models were estimated for both seaport-to-state (seaport data) and border port-to-state (transborder) weight and value flows. In particular, for the seaport data, several model specifications were tested based on transformations of the distance and time separation measures, $c_{ij}^{(1)}$ and $c_{ij}^{(2)}$, respectively, after a careful examination of residuals that showed the presence of outliers.

In the analysis of residuals, two difficulties were addressed. The first difficulty is the unequal variances of residuals, a consequence of the Poisson distribution. This problem was handled by considering, as Cochran suggests (Rao 1973, p. 393), instead of the plain residuals $N_{ij} - \hat{T}_{ij}$, the components

$$\sqrt{N_{ij}} - \sqrt{\hat{T}_{ij}} \tag{35}$$

The second difficulty stems from the very large number of residuals (the number is $I \times J$, the number of all origin-destination pairs). To tackle this concern, we followed a procedure described in Sen and Smith (1995, p. 456) consisting of drawing normal (rankit) plots of residuals. If the residuals are normal, such a plot would lie approximately on a straight line. Points that deviate sharply near the ends were tagged as outliers. Unusual jumps in the plots or other strongly nonlinear shapes could signal the need for transformations of the $c_{ij}^{(k)}$'s or for additional $c_{ij}^{(k)}$'s. There are clearly other ways of examining residuals. It is important to note here that the effort to conduct residual analysis in the gravity model case should not be seen as less than that for linear models where residual analysis is routinely carried out.

Indeed, the examination of residuals signaled the need for transformation of the $c_{ij}^{(k)}$'s. We found the square root transformation provided adequate fit for both weight flow and value flow estimates (table 2). The model provides an excellent fit for the data, because the Chi-square statistic hovers around its expected value of 1. In addition, the cell-to-cell weight flow estimates correlate very well with the data (Pearson correlation coefficient, $r = 0.89$, $p < 0.01$. The cell-to-cell value flow estimates also correlate very well with the data ($r = 0.98$, $p < 0.01$). Moreover, the weight and value length distributions in figure 1 (1% corresponds to 2,292,872,822 pounds) and figure 2 (1% corresponds to \$3,198,123,011), respectively, seem to corroborate the previous results.

TABLE 2 Parameter Estimates for PIERS Data

Parameter	Weight flows: Sample mean	Weight flows: Standard deviation	Value flows: Sample mean	Value flows: Standard deviation
θ_1	11.1227	1.1646	3.7854	0.3963
θ_2	–16.7447	1.7533	–11.1437	1.1668
X^2 ratio	0.7253	0.0000	1.1760	0.0000

Of interest in table 2 is the appearance of $\hat{\theta}_1$ with a positive value. This is due to collinearity between the used impedances, distance and travel time, a phenomenon that is known to adversely affect the sign of parameter estimates. Dropping one of the impedances would bias the parameter estimates left in the model. Given the robustness of the maximum likelihood procedure above in collinearity situations, all available impedance measures were retained (Sen and Smith 1995, chapter 5). After all, the sign of $\hat{\theta}_1$ would have changed to a negative value had we reparameterized the model and considered instead of travel time as the first impedance measure, the difference between distance and travel time (in appropriate units).

FIGURE 1 PIERS Data: Weight Length Frequency Distribution

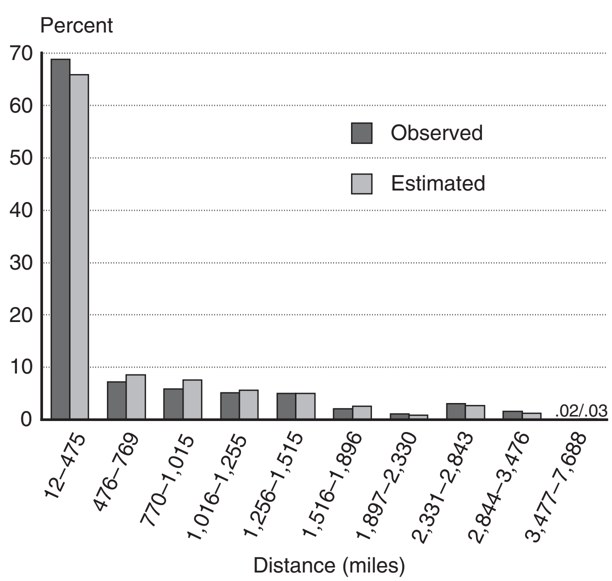
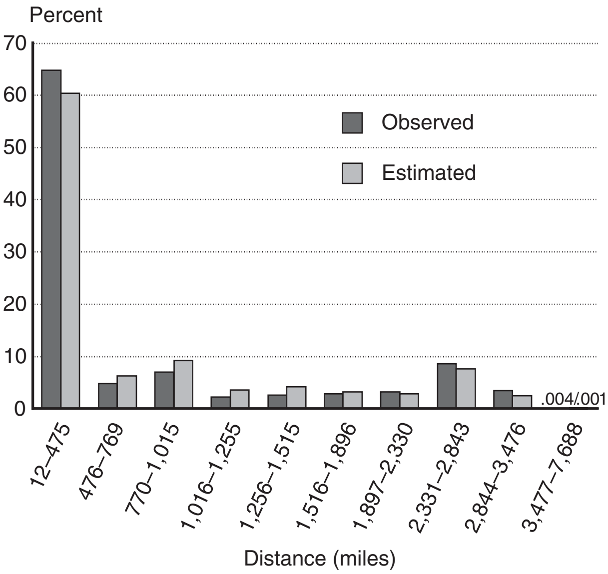


FIGURE 2 PIERS Data: Value Length Frequency Distribution



In the end, the model specifications for the weight and value flows, respectively, are

$$T_{ij} = A_i B_j \exp\left(\theta_1 \sqrt{c_{ij}^{(1)}} + \theta_2 \sqrt{c_{ij}^{(2)}} \right)$$

$$T_{ij} = A_i B_j \exp\left(\theta_1 c_{ij}^{(1)} + \theta_2 \sqrt{c_{ij}^{(2)}} \right) \tag{36}$$

Parameter Estimates and Model Fit for Transborder Data

In the case of Transborder data, we conducted a similar investigation of residuals (as above with the PIERS data) and observed the presence of a few remaining outliers, despite using transformed $c_{ij}^{(k)}$'s. These outliers were mainly states receiving an unusually large share of shipments compared with other states. The discovery was made after squaring the residuals in equation (35) and adding them together for all origin points of entry for each destination state. Destination states with a large share of shipments had a much larger sum of squares of residuals than other destinations. This was understandable, because these large destination states attract more shipments from longer distances and suggest the use of an additional impedance variable

$$c_{ij}^{(3)} = \delta \sqrt{\sqrt{c_{ij}^{(2)}}}$$

for the weight flows, and

$$c_{ij}^{(3)} = \delta \log[c_{ij}^{(2)}]$$

for the value flows, where δ is 1 for destinations with more than a 10% share in weight (4% in value) and 0 otherwise. In effect, we sought a different parameter estimate for the square root/log of distance for destination states with large shares.

The previous use of the indicator (dummy) variable δ is typical in spatial analysis for the treatment of residuals (Sen and Smith 1995) and introduces the effects of spatial structure on flow patterns (Gensler and Meade 1988; Fotheringham and O'Kelly 1989; Lo 1991), which is not the primary intent of this paper. More complex measures of relative location have been developed and tested empirically (Boots and Kanaroglou 1988; Lowe and Sen 1996). In the end, the model specifications for the weight and value flows, respectively, are

$$T_{ij} = A_i B_j \exp\left(\theta_1 \sqrt{c_{ij}^{(1)}} + \theta_2 \sqrt{c_{ij}^{(2)}} + \theta_3 \delta \sqrt{\sqrt{c_{ij}^{(2)}}} \right)$$

$$T_{ij} = A_i B_j \exp(\theta_1 c_{ij}^{(1)} + \theta_2 c_{ij}^{(2)} + \theta_3 \delta \log[c_{ij}^{(2)}]) \tag{37}$$

The previous steps removed all remaining outliers. Parameter estimates and goodness-of-fit statistics are shown in table 3. The previous observation regarding the sign of $\hat{\theta}_1$ applies. The model fits the data very well, as the Chi-square ratios for both weight and value remain under 2. In addition, the cell-to-cell weight flow estimates correlate very well with the data (Pearson correlation coefficient, $r = 0.91$, $p < 0.01$). The cell-to-cell value flow estimates also correlate very well with the data ($r = 0.94$, $p < 0.01$). Moreover, the weight and value length distributions in figure 3 (1% corresponds to 767,519,815 pounds) and figure 4 (1% corresponds to $1,565,313,371), respectively, seem to corroborate the previous results.

Computation of Covariance of Estimates for PIERS Data

The previous point estimates of the $\hat{\boldsymbol{\theta}}$ parameters and the flows $\hat{T}$ were used in the methodology described earlier to obtain the covariance of these estimates. The procedure can accommodate any

TABLE 3 Parameter Estimates for Transborder Data

	Weight flows		Value flows	
Parameter	Sample mean	Standard deviation	Sample mean	Standard deviation
θ_1	14.8743	1.5575	11.1739	1.1700
θ_2	–21.4296	2.2439	–13.6992	1.4344
θ_3	–2.1181	0.2217	–0.5534	0.0579
X^2 ratio	1.9642	0.0000	1.9012	0.0000

FIGURE 3 Transborder Data: Weight Length Frequency Distribution

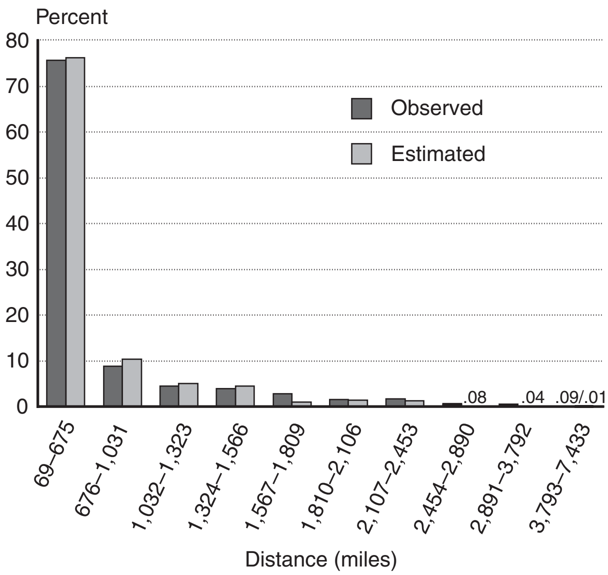
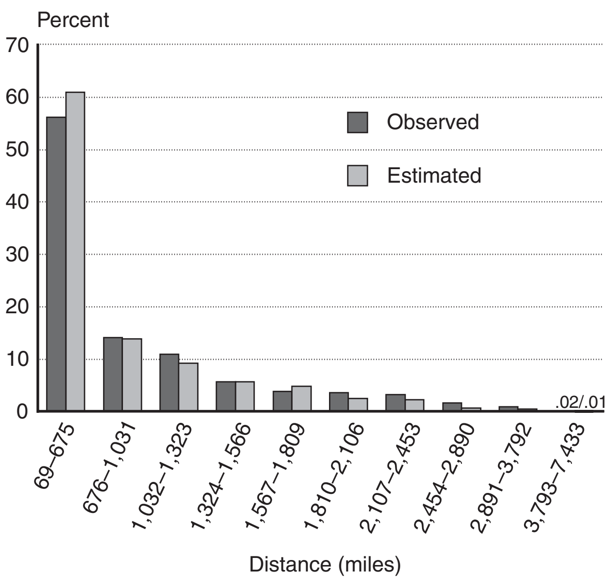


FIGURE 4 Transborder Data: Value Length Frequency Distribution



reasonable number of separation measures, $c_{ij}^{(k)}$, and the large number of origins and destinations typically encountered in practice.

The computational requirements of the procedure are no longer prohibitive. We were able to obtain the covariance matrix of a 144 × 50 flow matrix, a 7200 × 7200 matrix, in a little more than an hour on a Pentium III, 800 MHz, 512MB RAM laptop computer running a FORTRAN 77 compiler. Replacing T_{ij}'s by their estimates in equation (27), the covariance matrix of $\hat{\boldsymbol{\theta}}$ for the weight flows from the PIERS dataset was found to be

$$\left(\begin{array}{ccc} & \hat{\theta}_1 & \hat{\theta}_2 \\ \hat{\theta}_1 & 2.62830473 & -2.81464286 \\ \hat{\theta}_2 & -2.81464292 & 3.02464367 \end{array}\right) \tag{38}$$

The correlation between $\hat{\theta}_1$ (the parameter estimate for the distance measure) and $\hat{\theta}_2$ (the parameter estimate for the travel time measure) is readily apparent from equation (38). The negative sign of the covariances should not be disconcerting. It shows that, if for some small shift in the observations, $\hat{\theta}_1$ were to increase, $\hat{\theta}_2$ would decrease to "compensate."

Similarly, the covariance matrix of $\hat{\boldsymbol{\theta}}$ for the value flows from the PIERS dataset was found to be

$$\left(\begin{array}{ccc} & \hat{\theta}_1 & \hat{\theta}_2 \\ \hat{\theta}_1 & 1.53388134 & -1.66648732 \\ \hat{\theta}_2 & -1.66648735 & 1.81284692 \end{array}\right) \tag{39}$$

The correlation between $\hat{\theta}_1$ and $\hat{\theta}_2$ is similarly apparent from equation (39).

Computation of Covariance of Estimates for Transborder Data

The covariance matrices of $\hat{\boldsymbol{\theta}}$ from the Transborder dataset for the weight and value flows, respectively, were found to be

$$\left(\begin{array}{cccc} & \hat{\theta}_1 & \hat{\theta}_2 & \hat{\theta}_3 \\ \hat{\theta}_1 & 9.26053517 & -9.88856511 & 0.08654230 \\ \hat{\theta}_2 & -9.88856442 & 10.60290046 & -0.15624519 \\ \hat{\theta}_3 & 0.08654113 & -0.15624389 & 0.67484586 \end{array}\right) \tag{40}$$

and

$$\left(\begin{array}{cccc} & \hat{\theta}_1 & \hat{\theta}_2 & \hat{\theta}_3 \\ \hat{\theta}_1 & 2.87729568 & -3.13541449 & 0.01302693 \\ \hat{\theta}_2 & -3.13541410 & 3.42693579 & -0.02059522 \\ \hat{\theta}_3 & 0.01302676 & -0.02059503 & 0.01040623 \end{array}\right) \tag{41}$$

These last two covariance matrices in equations (40) and (41) show relatively high correlation between $\hat{\theta}_1$ (the parameter estimate for the distance

TABLE 4 An Illustration of N_{ij}, $\hat{T}_{ij}$ and 90% Confidence Intervals

N_{ij} (observed flows) (pounds)	$\hat{T}_{ij}$ (estimated flows) (pounds)	$N_{ij}-\hat{T}_{ij}$ % difference	90% confidence interval of $\hat{T}_{ij}$ Lower	Upper
4,850,905.00	4,764,382.12	1.78	4,764,179.19	4,764,585.05
87,454.00	88,012.62	0.63	87,934.47	88,090.77
39,922,216.00	38,847,527.51	2.69	38,845,275.15	38,849,779.87
16,002,283,960.00	15,556,162,969.96	2.78	15,555,972,990.43	15,556,352,949.49
216,787,500.00	211,385,682.81	2.49	211,374,117.32	211,397,248.30
4,174,427.00	4,266,148.17	2.19	4,266,017.84	4,266,278.50
3,076,999.00	3,100,768.36	0.77	3,100,346.82	3,101,189.90
2,029,122,669.00	2,108,940,894.95	3.93	2,108,896,284.86	2,108,985,505.04
693,089.00	664,724.67	4.09	664,638.45	664,810.89

measure) and $\hat{\theta}_2$ (the parameter estimate for the travel time measure) and relatively low correlation between $\hat{\theta}_3$ (the parameter estimate for the transformed travel time measure) and $\hat{\theta}_1$ or $\hat{\theta}_2$. The interpretation of the negative signs is the same as above.

For illustration purposes, an example of the usefulness of the computed covariance matrices for $\hat{T}_{ij}$'s in computing confidence intervals is shown in table 4. The table shows a few N_{ij}'s, $\hat{T}_{ij}$'s, percentage difference between N_{ij} and $\hat{T}_{ij}$, and 90% confidence intervals constructed by adding and subtracting 1.65 times the standard error available from the variances.

CONCLUSIONS

The most characteristic and indeed more restrictive feature of the Poisson distribution is the equality between its mean and variance. Many types of spatial interaction phenomena exhibit restricted (variance less than their means) or extra variation (variance greater than their means). Yet, the framework described above is very robust. Its robustness stems from the many types of asymptotic results that establish the Poisson distribution as the unique limiting form for a wide range of interaction processes when the population size increases and the overall influence of each individual interaction decreases.

More specifically, this research demonstrates that, to the extent the behavior of each shipping firm can be thought of as a very small interaction process, the resultant interaction process (i.e., the realized pattern of origin-destination commodity shipments) can be thought of as the superposition of all these processes and thus characterized by gravity models. The limited empirical analysis in this paper confirmed this conjecture and thus establishes an extremely rich framework for future experimentation.

Future applications of the proposed framework could conduct additional exploratory analyses to determine which factors affecting the demand for freight shipments are origin-based, destination-based, or separation-based. Another important issue needing more attention in the future is the requirement that, when the proposed modeling framework is used for forecasting, the travel times and costs used as impedance measures should be consistent with those obtained during the traffic assignment.

In summary, this paper has provided evidence that the movement of freight shipments can now be estimated within a desired confidence level as a result of maximum likelihood estimation of Poisson gravity models. The freight transportation modeler has two procedures available for computing reliable information within a predetermined accuracy: one for computing freight flow estimates and one for computing covariance matrices. These procedures can accommodate any reasonable number of separation measures, $c_{ij}^{(k)}$, and the large number of origins and destinations typically encountered in practice.

ACKNOWLEDGMENTS

This paper reports on research conducted by the author for the Bureau of Transportation Statistics, U.S. Department of Transportation.

REFERENCES

Bishop, Y.M.M., S.E. Fienberg, and P.W. Holland. 1975. *Discreet Multivariate Analysis: Theory and Practice*. Cambridge, MA: MIT Press.

Boots, B. and P. Kanaroglou. 1988. Incorporating the Effects of Spatial Structure in Discrete Choice Models of Migration. *Journal of Regional Science* 28:495–507.

Cambridge Systematics. 1997. A Guidebook for Forecasting Freight Transportation Demand, NCHRP Report 388. Transportation Research Board, National Research Council.

Fotheringham, A. and M. O'Kelly. 1989. Spatial Interaction Models: Formulations and Application. *Studies in Operational Regional Science*. Boston, MA: Kluwer Academic Publishers.

Gensler, W. and M. Meade. 1988. Locational and Population Factors in Health Care-Seeking Behavior in Savannah, Georgia. *Health Services Research* 22:443–462.

Lo, L. 1991. Spatial Structure and Spatial Interaction: A Simulation Approach. *Environment and Planning A* 23:1279–1300.

Lowe, J. and A. Sen. 1996. Gravity Model Applications in Health Planning Analysis of an Urban Hospital Market. *Journal of Regional Science* 36:(3)437–462.

Memmot, F.W. 1983. Application of Statewide Freight Demand Forecasting Techniques, NCHRP Report 260. Transportation Research Board, National Research Council.

Metaxatos, P. 2002. International Trade Traffic Study (Section 5115): Measuring the Ton-Miles and Value-Miles of International Trade Traffic Carried by Highway for Each State: Estimation of Origin-Destination Flows and Confidence Intervals, Final Report, prepared for the U.S. Department of Transportation, Bureau of Transportation Statistics. March.

Meyburg, A.H. and J.R. Mbwana, eds. 2002. Conference Synthesis: Data Needs in the Changing World of Logistics and Freight Transportation. New York Department of Transportation, Albany, NY. Available at http://www.dot.state.ny.us/ttss/conference/synthesis.pdf.

Openshaw, S. and C.J. Connolly. 1977. Empirical Derived Deterrence Functions for Maximum Performance Spatial Interaction Models. *Environment and Planning A* 9:1067–1079.

Rao, C.R. 1973. *Linear Statistical Inference and Its Application.* New York, NY: Wiley.

Sen, A. and T.E. Smith. 1995. *Gravity Models of Spatial Interaction Behavior.* New York, NY: Springer-Verlag

Sen, A. and R.K. Pruthi. 1983. Least Squares Calibration of the Gravity Model When Intrazonal Flows are Unknown. *Environment and Planning A* 15:1545–1550.

U.S. Department of Transportation (USDOT). 1996. *Quick Response Freight Manual,* DOT-T-97-10. Washington, DC. September 1996.

Weber, J.S. and A. Sen. 1985. On the Sensitivity of Gravity Model Forecasts. *Journal of Regional Science* 25:317–336.

Yun S. and A. Sen. 1994. Computation of Maximum Likelihood Estimates of Gravity Model Parameters. *Journal of Regional Science* 34:199–216.

Author address: Paul Metaxatos, University of Illinois at Chicago, 412 South Peoria Street, Suite 340, Chicago, IL 60607. Email: pavlos@uic.edu.

Predicting the Construction of New Highway Links

DAVID LEVINSON
University of Minnesota

RAMACHANDRA KARAMALAPUTI
Capital One

ABSTRACT

This paper examines new highway construction based on the status of the network, traffic demand, project costs, and budget constraints. The data span two decades and consist of descriptions of physical attributes of the network, the construction and expansion history, and average annual daily traffic values on each of the links. An algorithm is developed to designate adjacent and parallel links in a large network. A nonlinear cost model for new construction and highway expansion is developed for the Minneapolis-St. Paul metropolitan area. Results show that new links providing greater potential access are more likely to be constructed and that more links will be constructed when the budget is larger, which supports the underlying economic theory. The models developed here have important implications for planning and forecasting, allowing us to predict how networks might be altered in the future in response to changing conditions.

INTRODUCTION

The 240 km of paved road in the United States in 1900 increased to about 6.4 million km in 2000, providing virtually 100% of the U.S. population with almost immediate access to paved roadways

KEYWORDS: highway construction, cost model, transportation forecasting, network growth, mixed logit.

(USDOT 2002). The growth or decline of transportation networks obviously affect a region's social and economic activities, yet the dynamics of how such changes occur is one of the least understood areas in transportation research and regional science. This lack of understanding is often revealed in the long-range planning efforts of metropolitan planning organizations (MPOs), where transportation network change is treated exclusively as the result of top-down decisionmaking. In fact, changes to the transportation network are the result of numerous small decisions (and some large ones) by semi-autonomous entities (firms, developers, towns, cities, counties, state department of transportation districts, MPOs, and states) in response to market conditions and policy initiatives. Understanding how markets and policies translate into facilities on the ground is essential for scientific understanding and for improving forecasting, planning, policymaking, and evaluation.

The study of network growth has been limited. Taaffe et al. (1963) explored the economic, political, and social forces behind infrastructure expansion in underdeveloped countries and found that roads are initially developed to connect regions of economic activity and feeder roads later connect to these initial investments. Garrison and Marble (1965) observed that connections to the nearest large neighbor explained the order of rail network construction in Ireland. Grübler (1990) found that the growth of infrastructure follows a logistic curve and road infrastructure in developed countries has reached saturation levels. Yamins et al. (2003) developed a simulation that grows urban roads using simple connectivity rules proportional to the activity at locations. Yerra and Levinson (2003) developed a simulation model to capture the expansion of existing links. The results also show that hierarchical arrangements of roads (i.e., specific routes with continuous attributes), are emergent properties of transportation networks.[1] Several studies have examined specific networks, for example, the London Underground (Barker and Robbins 1975), but no general theoretical framework has been given for incremental network growth at the microscopic level.

Our study focuses on understanding the conditions under which new links are constructed (as opposed to existing links being improved) on a highway network. The construction of new links can be modeled in several ways, assuming we have the location of possible and existing nodes. We could assume that all (or a very large number of) nodes are connected, but at some very slow speed, and then use a network investment model to improve selected links while allowing others to whither, much as a neural network learns. In contrast to this process, we could assume that, for every node, there is a set of possible nodes it can connect with (neighbors within a certain radius to which it is not already connected). The connections made depend on underlying conditions.

It is the second approach that we investigate in this paper. Specifically, we want to understand the effects of travel demand, cost of construction, budget, and the surrounding conditions on the generation of new links. A highway network is thought to be expanded or constructed due to congested traffic conditions or in anticipation of regional economic development. Limited budgets and existing land uses constrain the number of new links constructed in a given period. The traffic level on parallel links is expected to be a highly significant factor in new construction. Also, the number of potential trips on the new link is thought to be an important factor in its initiation.

Theory and statistical techniques used for this study are explained in the next section of this article. The following section provides a description of the data used and its assembly. In that section, adjacent and parallel links are designated and the model used to estimate the cost of construction is described. The next section presents the model we used and poses the specific hypotheses. Results are then presented, followed by conclusions.

THEORY AND STATISTICAL MODELING

Construction of a new link may alleviate traffic congestion or open a new area to development by

[1] Specific routes with continuous attributes imply that connecting links do not necessarily have the same attributes. For instance, two four-lane links connecting at an intersection with two two-lane links imply that the four-lane links are part of a continuous route and the two-lane links are part of a different route. Because these routes are differentiated, some must be more important than others, which produces a hierarchy of roads.

increasing accessibility. Such a link could lead to the availability of additional routes and cause traffic patterns to change. Each of the variables considered affect either the supply curve or the demand curve and can shift the equilibrium.

A higher transportation budget (B) increases the ability to expand or construct highways resulting in an outward shift in the demand curve. Increasing the cost of construction decreases its likelihood. Previous studies empirically show that, in mature networks, the capacity added to the system decreases over time (Nakicenovic 1988; Grübler 1990). A marginal increase in capacity decreases average travel demand per lane (demand by consumers) as existing capacity increases.

Expanding a link means additional trips on that link due to re-routing and rescheduling of trips and also due to induced demand (Parthasarathi et al. 2003; Levinson and Kanchi 2002; Fulton et al. 2000; Noland 2001). In light of induced demand, the effect of roadway expansion in reducing traffic congestion is not fully understood. Furthermore, although consumers' surplus increases after the expansion, travelers are inconvenienced during its expansion.

Long links take more time to complete and diverting traffic during that period is difficult. The possibility of constructing a new link increases in such scenarios and hence the condition of traffic in the surrounding links is a crucial factor for new link construction. Networks, because of land scarcity, tend to grow more in the peripheries once they reach saturation levels near downtown.

Statistical Theory

Due to the discrete nature of the dependent variable (a new link is constructed or not), we considered discrete choice modeling to be appropriate. Initial modeling was done using a logit model, although there are certain limitations (Haynes et al. 1988; Haynes and Fotheringham 1991). The logit probabilities are derived under the assumption that the unobserved portion of the utility (combined with the error term) is distributed in accordance with the extreme value distribution. The probability of a decisionmaker choosing alternative a is given by

$$P_a = \frac{e^{\beta' X_a}}{\sum_a e^{\beta' X_a}}$$

Logit models assume that tastes are invariant across the population and consequently estimate fixed coefficients for the variables. In general, individual tastes vary across the population and this variation in tastes (random effects) should be included in modeling. In mixed (random parameter) logit models, the unobserved heterogeneity (individual-specific effects) is taken care of (McFadden and Train 2000; Train and Brownstone 1999; Hensher 2001). The likelihood function is similar to logit models, but the coefficient of some variables is not fixed across the population. Because our models did not suggest varying coefficients for variables, only the constant taste variable is assumed to vary across the population. In this model, the utility of an alternative a is given by

$$U_a = \beta' X_a + (\eta_a + \varepsilon_a)$$

where
U_a = utility of alternative a,
X = vector of variables,
η_a = random term with zero mean (any distribution),
ε_a = random term with extreme value distribution.

Given the value of random term η from its distribution, the choice probability is again logit. Because we do not know the value of the random term, we integrate the logit probability over all values of the random term using its density function. The choice probability of an alternative for an individual is then given by

$$\int \frac{\exp(\beta' x_a)}{\sum_a \exp(\beta' x_a)} f(\eta / \Omega) d\eta$$

where Ω s are the distributional parameters of the random term in the likelihood function. This is called *mixed logit,* because the terms are split into a mixture of distributions. The integral above does not have a closed form in general. To overcome this problem, for each individual we need to average over a range of simulated values of likelihood. Taking this value as the probability, the log-likelihood function of multiplication of simulated probabilities of all individuals is maximized to obtain the coefficients of utility functions of each alternative.

The log-likelihood of the simulated probability is a biased estimator of the true probability. This bias decreases as the number of draws increases. In the simulations, values from a uniform random generator are converted to the distribution of the random variable. When using a random number generator, it is sometimes possible that large sections of the distribution are not generated. The uniform random number generator does not guarantee uniform coverage in a given simulation (even coverage is guaranteed only on infinite draws). Halton draws[2] have been suggested for their specific advantage of even coverage. It has been found that 125 Halton draws are as efficient as 2,000 random draws in simulations of this kind (Bhat 2001). To reduce the computational time and increase the efficiency, Halton draws were used for this study.

DATA

The dataset for this study is built using data from three different sources. The Metropolitan Council of the Twin Cities of Minneapolis and St. Paul, Minnesota, provided network data for 1995 with length and location of each link. Each link is identified by its start and end nodes. Data on average annual daily traffic were obtained from the Transportation Information Systems Division of the Minnesota Department of Transportation. Data on construction of new links and expansion of the existing links were obtained from the local Transportation Improvement Program and the Hennepin County Capital Budget for 1978 to 1998. Data on new county highways are available only for Hennepin County. Hennepin is the largest of the seven counties and contains the city of Minneapolis. Using the investment data, a network for each of the years is built with 1995 used as the base network. The remaining dataset was integrated using ArcView geographic information systems software and custom computer programs.

[2] Halton draws are generated using a prime number, p, as a seed. The interval (0,1) is divided into n equal intervals and those form the first numbers of the sequence. Each of the n intervals is again divided equally into n sub-intervals. The new sets of numbers are arranged in a particular fashion to continue the sequence. For a detailed discussion, see Bhat (2001).

While links to be expanded are chosen from the existing network, when a new link is going to be constructed, it is selected from a set of possible links between nodes. In the case of the Twin Cities network, creation of new nodes because of new construction was not observed. The possible set of new links is, therefore, based only on existing nodes. Theoretically, a node can be connected to any of the remaining nodes.[3]

The mean length of newly constructed links was 0.68 km and the maximum was 4.54 km. Because of the large number of possible connections and high redundancy levels within the radius of 4.54 km, a shorter range of possible lengths was considered. In the new scenario, only links between 200 meters and 3.2 km in length were considered. These lengths were arrived at by removing new construction in the five percentile regions on both ends of the dataset. We observed that new nodes are seldom created by new construction in the Twin Cities network. This indicates that a possible set of new links should be such that they do not cross any of the existing links that are of higher-level hierarchy than the link being constructed, because doing so would create a new node. However, new links can cross lower level roads without technically intersecting them (via overpasses).

With the above restrictions, each node was found to have on average a set of 10 possible connections, with 29,804 possible new links. We found, however, that only 69 bi-directional new links (all highways) were actually constructed in the past two decades. There were, of course, many lower-level roads built and other higher-level roads widened, but those are not addressed here.

Adjacent and Parallel Links in a Network

We needed to compute the potential amount of traffic a newly constructed link might serve based on the traffic on the nodes it connects. The links that would be connected, "adjacent links," were divided into two categories: supplier links and consumer

[3] Freeway interchanges were treated as a single node for this purpose. With the current network, this possible connection can be made with any of the nodes at interchanges. To overcome this feature in the dataset, all the nodes within 50 meters of each other were given the same node number. A computer program was written to accomplish this task and the resulting node set was used to investigate new construction.

links. Supplier links would supply traffic to the new link, while consumer links are the links the traffic would move to after traversing the new link. A link (*ij*) that is a supplier to another link (*jk*) may be a consumer link in the other direction (*ji* receives traffic from *kj*). A computer program was written to enumerate adjacent links for each of the possible links.

The parallel link can be thought of as the link that would bear most of the diverted traffic if the link in consideration were closed. This definition is extended to new links by assuming the link is constructed and then finding the parallel link in the existing network. It is necessary to identify parallel links, because they are the links currently serving the traffic of that area. Because of the large number of possible new links, parallel links were not identified using traffic assignment. Rather, parallel links were assigned to each of the possible links using fuzzy theory (Zadeh 1992; Kosko 1993). Fuzzy theory assumes a continuous truth-value rather than the deterministic Boolean values used conventionally. The sum composition method combined with appropriate weights was found suitable for our purposes.

In general, a parallel link is in the proximity of link *L*, approximately parallel to it in orientation and of comparable length. Four attributes are defined to satisfy the above requirements. The first attribute is based on the angular difference between the orientations of the two links, which should be as small as possible. The second attribute is the perpendicular distance from mid-point of link *L* to the other link divided by length of link *L*. The third attribute is the sum of the distance between the start and end nodes of the two links being compared. The final attribute takes the ratio of lengths of the two links into consideration. Mathematically, the four attributes are defined as follows:

1. *Para* = 1 – *(angular difference) / 45*
2. *Perp* = 1 – *a*(perpendicular distance) / length of link L*
3. *Dist* = 1 – *b*(sum of node distances) / length of link L*
4. *Comp* = 1 – *c*(lratio – 1)*

where *perpendicular distance* is from the center of link *L* to the other link, *node distances* are distances between the corresponding start and end nodes, *lratio* is the ratio of length of the probable parallel link to the *length of link L* or the inverse of it, whichever is greater.[4]

In sum composition, computing the truth-value of each attribute and then summing these values gives the fuzzy output. Here, we modified this method by weighing the truth-values of the attributes based on the importance of each attribute in relation to others. The assumed parameters of *a, b,* and *c* and the assumed weights of the attributes are given in table 1. These values were calibrated to match our expectations of what should be the most parallel link using a few sample links. One parallel link was selected for each link.

TABLE 1 Assumed Values of Weights and Parameters of Parallel Link Attributes

Attribute	Weight	Parameter
Para	0.5	—
Perp	0.5	$a = 0.40$
Dist	1.0	$b = 0.25$
Comp	0.5	$c = 0.50$

Cost Model

A cost function is needed to estimate the cost of possible new construction. Investment data obtained from the Metropolitan Council's Transportation Improvement Program and the Hennepin County Capital Improvement Program were used to estimate a modified Cobb-Douglas (log-log) model. This model was used to account for the non-linear behavior of some of the explanatory variables.

$$ln(E_{ij}) = a + b_1 ln(L_{ij}{*}\Delta C_{ij}) + b_2 N + b_3 H_I + b_4 H_S + b_5 Y + b_6 ln(P) + b_7 X$$

where

E_{ij} = cost to construct or expand the link (in nominal thousands of dollars),

$L_{ij}{*}\Delta C_{ij}$ = lane kilometers of construction (Length * Increase in number of lanes),

[4] *Dist* and *Perp* differ in that *Perp* considers the perpendicular (or shortest) distance between the links, while *Dist* looks at the distances between the beginnings and ends of the links, which may not be perpendicular. In a perfect grid network, the two variables would measure the same thing, but most networks are not perfect grids (see Levinson and Karamalaputi 2003).

N = dummy variable 1 if new construction or 0 if expansion,

H_I, H_S = dummy variables for Interstate highways (H_I) and state highways (H_S), (default = county highways)

Y = year of completion (1979),

P = period (duration) of construction (in years),

X = distance of the link from the nearest downtown (Minneapolis or St. Paul) (in km).

The data consist of both expansions and new construction projects totaling 76 observations (more than 1 link can be expanded in a single project). Results of the model are shown in table 2. The coefficient of lane kilometers of construction ($L_{ij}*\Delta C_{ij}$) is less than one, indicating economies of scale in construction. As can be expected, the cost of a new construction project (N) is higher than expanding an existing link. The cost of construction increases with the hierarchy of the road (H) (H = 0 represents a county road). The year variable (Y) controls for inflation and the improving quality of the road construction. Longer duration projects (P) cost more and construction becomes costlier over time. The distance from the nearest downtown, entered as a linear variable (X), shows that the project cost decreases as it moves away from downtown areas. Downtown areas have higher traffic flows and land costs and hence restrict the construction flexibility, generating the extra cost.

TABLE 2 Regression Coefficients for Cost Model

Description of the variable	Variable	Coef.	$P > \mid t \mid$
Cost of construction	(E_{ij})	—	—
Lane-kilometers of construction	$ln(L_{ij}*\Delta C_{ij})$	0.50	0.00*
Dummy for new construction	N	0.39	0.04*
Dummy for interstate roads	H_I	1.97	0.00*
Dummy for state roads	H_S	0.56	0.02*
Log of (year–1979)	$ln(Y)$	0.75	0.00*
Log of period of construction	$ln(P)$	0.16	0.06*
Distance from nearest downtown	X	–0.03	0.04*
Constant		5.79	0.00*

Number of observations: 76

Adj. R^2: 0.77

* Significant at 90% confidence interval.

MODEL

Due to the few new links built over the last two decades, construction was assumed to occur in five-year intervals and the dataset was built accordingly. The budget over these five years was summed to act as a budget constraint. Nodes connecting only local roads (below county highways) were not considered in modeling, because we did not have data on new construction of such roads. New construction in the next time interval can be modeled as

$$N_{ijt+1} = f(L_{ij}, C_p, L_p, Q_p/C_p, A, E_{ij}, B, T, X, D)$$

where

N_{ijt+1} = dummy for new construction of link ij in period (t + 1),

L_{ij} = length of link ij along the road,

C_p = capacity of the parallel link,

L_p = length of the parallel link,

Q_p = flow on the parallel link,

Q_p / C_p = congestion measure on the parallel link,

A = product of total supplier link flows and total consumer link flows (access),

E_{ij} = cost of constructing the new link,

B = transportation department's budget constraint,

T = time period of construction,

X = distance from the nearest downtown,

D = number of nodes within the interval of 200 meters and 3.2 km.

Volumes on the links are directional.

Variable A can be considered an accessibility measure of the new link. It represents the effect of supplier link flows and consumer link flows on the probability of new construction. The effect of surrounding conditions was expected to be prominent in the construction of a new link compared with a link expansion. Based on that theory, the hypotheses are as follows:

- High congestion on the parallel link (Q_p/C_p) favors construction of the link to relieve traffic on the parallel link.

- Higher capacity of the parallel link (C_p) decreases the likelihood of new construction, as capacity is already available. However, high capacity links are less likely to be expanded.
- Longer links (L_{ij}) are less likely to be expanded because of the longer duration of construction.
- A longer length of a parallel link (L_p) favors new construction, because longer links tend not to be expanded as often due to the duration inherent in such an expansion.
- High expected costs of construction (E_{ij}) on the new link decreases the probability that the link will be built, while a higher transportation budget (B) increases that probability.
- A higher access score (A) for a link increases the chances of construction.
- As was observed in the literature, road construction has declined over time and was expected to be reflected in a negative sign on the year (T) variable.
- New links have a higher probability of being constructed far from downtown (X), as land acquisition is easier there.
- A large node density (D) in the surrounding area results in fewer new links being constructed, because the number of links is high in these areas.

Binomial logit and mixed logit modeling were used to analyze the dataset. The results are shown in the following section.

RESULTS

A binomial logit model was used to estimate the construction of a new link between existing nodes. Results of the regression models are given in table 3. Variables C_p, E_{ij}, and X are negative and significant while the variables L_p, A, T, and B are positive and significant.

As has been noted earlier, the construction of a new link depends heavily on its surrounding conditions and alternate route conditions. The longer the parallel link (L_p), the higher the probability of a new link. This might be interpreted as reflecting the cost involved in the expansion of the longer parallel link and also as a result of the traffic diversion problems on the parallel link if it were expanded.

The capacity of the parallel link (C_p) is negative and significant, supporting this hypothesis. High capacity links already serve high volumes of traffic in an area (generated in or passing through that area) and hence reduce the need for a new link.

A high access measure (A) between two nodes tends to increase the probability of new construction connecting those nodes. Access is directly proportional to the total time savings due to new construction and hence it is logical that high demand between two nodes has this effect.

A higher cost of constructing a new link (E_{ij}) reduces its probability of expansion, as expected. Also, more new construction is possible when the budget (B) is higher.

Contrary to our hypothesis, distance to the nearest downtown (X) variable is negative and significant, indicating that new links are more likely to be built nearer to downtown than in the suburbs. This probably reflects the completion of the Interstate Highway System in the Twin Cities, which saw the urban links finished last (in the past 20 years), while suburban links were completed as long as 40 years ago.

More new links are being constructed with the passage of time (T), refuting the hypothesis. Earlier studies showed decreasing expansion rates for existing links. This may reflect a policy shift from expansion to new construction. Expanding a road leads to traffic inconvenience during construction, a problem that can be avoided by new construction, which may explain the reasoning behind more new construction.

A mixed logit model was estimated to allow for the taste variances of individual links (i.e., of decisionmakers). Table 3 gives the results of the model. The log likelihood value was improved by 3%, indicating a better model. As mentioned earlier, changes in traffic demand were not considered due to the low number of new links. Considering changes in demand would require dropping one period of observation. The random term was assumed to have a triangular distribution and its estimated standard deviation is given in the table. Models with other possible distributions for the random term did not improve significance. The significant variance in the constant term reflects the variance in the links due to the effects of these

TABLE 3 **Estimated Models for New Construction**

Variable	Hypothesis	Logit Coefficient	Logit z	Mixed logit Coefficient	Mixed logit z
Length of the link (L_{ij})	–S	–5.91E-01	–1.35	1.84E–01	0.27
Capacity of the parallel link(C_p)	–S	–3.31E–01*	–1.86	–4.87E–01*	–1.67
Length of the parallel link(L_p)	+S	4.93E–01*	1.58	1.42E+00*	1.97
Congestion on parallel link (Q_p/C_p)	+S	–1.96E–05	–1.25	–5.23E–05*	–1.87
Access (A)	+S	4.24E–05*	4.70	1.92E–04*	2.71
Time period of construction (T)	–S	7.31E–01*	2.24	3.00E+00*	3.27
Cost of construction (E_{ij})	–S	–2.82E–01*	–2.48	–8.97E–01*	–3.36
Budget (B)	+S	5.68E–06*	3.06	7.78E–06*	3.96
Distance from downtown (X)	+S	–1.21E–01*	–5.69	–3.08E–01*	–6.62
Node density (D)	–S	–6.32E–04	–0.18	–1.27E–02*	–1.73
Constant		–6.09E+00*	–6.02	–1.24E+02*	–2.92
Triangular deviation					
Constant		—	—	4.72E+01*	2.36
Number of observations		89,031		89,031	
Log likelihood		–473.19		–459.68	

* Significant at 90% confidence interval.

omitted variables and the inherent taste variance (of decisionmakers). More data are needed to model new construction with other influencing variables.

However, a mixed logit model can to some extent encompass the effect of these variables. Omitted variables in a model increase the standard error of the estimated variables and thus cloud the significance of some variables. For instance, the number of nodes in the surrounding area (D) is significant in the new model, supporting the hypothesis. The coefficients of variables changed significantly when the unobserved variance was accounted for. The z-values of the mixed logit model are higher than those of the logit model indicating increased reliability of the estimated coefficients. In the case of the mixed logit model, the congestion on the parallel link is negative and significant, refuting our hypothesis (it was insignificant in the logit model).

Out of a network of 29,804 possible new links, there were 69 new links in the time period considered. Of the 69 most likely new construction links as predicted by the models, the logit model identified 17 links that were actually built. The mixed logit model performed better predicting the same 17 links and an additional 5 new links correctly. In view of these results, mixed logit models perform better than conventional discrete choice models.

CONCLUSIONS

This paper developed a model to predict the location of new highway construction based on the surrounding conditions of the new link, the estimated cost of construction, and a budget constraint. A new process for identifying potential construction projects is developed. The methodology used here reduces the number of possible newly constructed links drastically and paves the way for feasible modeling. This paper provides a practical solution to the problem of identifying adjacent and parallel links in a large network. Using the investment data, a model to estimate the cost of potential new construction is developed here.

Results indicate significant dependence on parallel link attributes and potential access to traffic due to the new link. A newly constructed link provides an additional route; hence, its construction depends on the attributes of the links that presently serve the region. A high capacity route is sufficient to cater to the traffic generated in or going through the region and usually does not require a new construction project. New construction projects are less likely to be undertaken if they are costly and are limited by the available budget. New links are unnecessary when the region is well connected, as reflected by the node density variable. Two different types of discrete choice models were estimated to compare their performances. It was found that mixed logit models perform better than logit models and account for unobserved taste variance.

Although politics factor into these decisions, it should be noted that they are constrained by the decisions made in the past and by the present conditions of the network. The models suggest a number of significant factors that lead to new highway construction. The models estimated here also can be used to monitor the growth of the network given projected traffic demand for the existing links and values of model variables at present conditions. This would improve transportation planning by enabling modelers to predict pressures for additional links. Forecasting future demands on the transportation network requires a forecast of the network structure itself. Only with models of new link construction and link expansion can these forecasts be made.

REFERENCES

Barker, T.C. and M. Robbins. 1975. *A History of London Transport*, vols. 1 and 2. London, England: Allen and Unwin.

Bhat, C.R. 2001. Quasi-Random Maximum Simulated Likelihood Estimation of the Mixed Multinomial Logit Model. *Transportation Research B* 35:677–693.

Fulton, L.M., R.B. Noland, D.J. Meszler, and J.V. Thomas. 2000. A Statistical Analysis of the Induced Travel Effects in the U.S. Mid-Atlantic Region. *Journal of Transportation and Statistics* 3(1):1–14.

Garrison, W.L. and D.F. Marble. 1965. A Prolegomenon to the Forecasting of Transportation Development, U.S. Army Aviation Material Labs Technical Report. Office of Technical Services, U.S. Department of Commerce.

Grübler, A. 1990. *The Rise and Fall of Infrastructures: Dynamics of Evolution and Technological Change in Transport*. Heidelberg, Germany: Physica-Verlag.

Haynes, K. and S. Fotheringham. 1991. The Impact of Space on the Application of Discrete Choice Models. *The Review of Regional Studies* 20(2):39–49.

Haynes, K., D. Good, and T. Dignan. 1988. Discrete Spatial Choice Modeling and the Axiom of Independence from Irrelevant Alternatives. *Socio-Economic Planning Sciences* 22(6):241–251.

Hensher, D. 2001. The Valuation of Commuter Travel Time Savings for Car Drivers: Evaluating Alternative Model Specifications. *Transportation* 28(2):101–118.

Kosko, B. 1993. *Fuzzy Thinking: The New Science of Fuzzy Logic*. New York, NY: Hyperion.

Levinson, D. and S. Kanchi. 2002. Road Capacity and the Allocation of Time. *Journal of Transportation and Statistics* 5(1):25–45.

Levinson, D. and R. Karamalaputi. 2003. Induced Supply: A Model of Highway Network Expansion at the Microscopic Level. *Journal of Transport Economics and Policy* 37(3):297–318, September.

McFadden, D. and K. Train. 2000. Mixed MNL Models for Discrete Response. *Journal of Applied Econometrics* 15:447–470.

Nakicenovic, N. 1988. Dynamics and Replacement of U.S. Transport Infrastructures. *Cities and Their Vital Systems: Infrastructure—Past, Present and Future*. Edited by J.H. Ausubel and R. Herman. Washington, DC: National Academy Press.

Noland, R.B. 2001. Relationships Between Highway Capacity and Induced Vehicle Travel. *Transportation Research A* 35(1):47–72, January.

Parthasarathi, P., D. Levinson, and R. Karamalaputi. 2003. Induced Demand: A Microscopic Perspective. *Urban Studies* 40(7):1335–1351.

Taaffe, E.J., R.L. Morrill, and P.R. Gould. 1963. Transport Expansion in Underdeveloped Countries. *Geographical Review* 53:503–529.

Train, K. and D. Brownstone. 1999. Forecasting New Product Penetration with Flexible Substitution Patterns. *Journal of Econometrics* 89:109–129.

U.S. Department of Transportation (USDOT), Bureau of Transportation Statistics. 2002. *Transportation Indicators*. Available at http://www.bts.gov/transtu/indicators/, as of January 2004.

Yamins, D., S. Rasmussen, and D. Fogel. 2003. Growing Urban Roads. *Networks and Spatial Economics* 3:69–85.

Yerra, B. and D. Levinson. 2003. The Emergence of Hierarchy in Transportation Networks, presented at the Western Regional Science Association Meeting, RioRico, AZ, February.

Zadeh, A.L. 1992. *Fuzzy Logic for the Management of the Uncertainty*. New York, NY: Wiley Publications.

Authors' addresses: David Levinson, Assistant Professor, 500 Pillsbury Dr. SE, Dept. of Civil Engineering, University of Minnesota, Minneapolis, MN 55455. E-mail: levin031@tc.umn.edu.

Ramachandra Karamalaputi, Analyst, Capital One, 3470 Kilburn Circle, # 1028, Richmond, VA 23233. E-mail: ramachandra.karamalaputi@capitalone.com.

Demand Elasticity on Tolled Motorways

ANNA MATAS
JOSÉ-LUIS RAYMOND
Universitat Autònoma de Barcelona

ABSTRACT

This paper analyzes the elasticities of demand on tolled motorways in Spain. We use a panel dataset covering an 18-year period, where the cross-section observations correspond to various Spanish tolled motorway sections. A dynamic model is estimated, which allows us to identify short-term and long-term responses to changes in the independent variables. The results show that demand is elastic with respect to the level of economic activity, whereas average elasticity with respect to gasoline price is in line with that estimated in previous studies. For the main variable of interest, the results indicate that demand is relatively sensitive to toll changes, although a wide variation is observed across motorway sections. A statistical analysis reveals that the main factors explaining such differences are related to variables that reflect the quality of alternative and free roads.

INTRODUCTION

In recent years, there has been renewed interest in using tolls to finance road investment, in order to avoid public budget constraints and at the same time to involve the private sector in the provision of infrastructure. In this new context, it is vital to have accu-

KEYWORDS: demand elasticities, toll motorways, intercity road traffic.

rate knowledge of demand behavior for forecasting and evaluation purposes. More precisely, it is necessary to know the elasticity of demand with respect to price, quality, or income, in order to obtain traffic and revenue forecasts or to evaluate potential negative effects such as the misallocation of traffic between tolled roads and parallel untolled roads.

Empirical evidence on demand elasticity on tolled motorways is limited due to the relative scarcity of tolled roads in the world.[1] Furthermore, most of the studies provide average elasticities for specific short road sections, tunnels, or bridges, which are highly dependent on site-specific factors such as the degree of congestion or the availability of alternatives. Because of this, it is difficult to transfer the results to other contexts.

This paper aims to provide new and robust evidence on demand elasticity on tolled motorways with respect to its main determinants, placing special emphasis on toll elasticities. We address this issue by estimating a dynamic demand function using a panel dataset consisting of observations of the Spanish tolled motorway network over the period 1981 to 1998. The results show that the sensitivity of demand to price depends both on the characteristics of the tolled motorways and on those of the alternative free road.

The next section provides a review of demand elasticity with respect to tolls followed by a brief summary of the toll policy in Spain. The model specification and certain relevant econometric issues are discussed next. In the following section, we present the data we used, and then we turn to the model estimation and results. We next carry out a statistical analysis in order to identify the factors that explain the differences in toll elasticities across motorway sections. Finally, the main conclusions of the paper are summarized.

REVIEW OF THE LITERATURE

There is a general consensus that, on average, transportation demand is fairly inelastic with respect to price. The empirical evidence gathered on toll elasticities (table 1) seems to confirm this. The most frequent values fall around –0.2 and –0.3 with a range of –0.03 to –0.50. These values correspond to average demand elasticities. Unfortunately, the potential sources of variation are not taken into account in a formal manner. Nevertheless, some authors do identify the characteristics that will have an impact on the elasticity value.

The lowest values of toll elasticities are usually observed for bridges in highly congested metropolitan areas in the United States. This result can be explained by the low level of the toll fee compared with other components of private car cost, such as parking fees (Harvey 1994). Wuestefeld and Regan (1981) found that response varies according to the purpose of the trips, trip frequency, the existence of a toll-free alternative, and journey length. Hirschman et al. (1995) state that demand is more sensitive in the case of those infrastructures with a good alternative untolled road.

Finally, some authors argue that traffic is sensitive to time-varying pricing schemes. Gifford and Talkington (1996) found evidence that day-of-the-week cross-elasticities are complementary; that is, an increase in toll rates on one day results in a reduction of traffic not only on that day but on other days of the week as well. Burris et al. (2001) showed that travelers responded to the off-peak toll discount implemented on two county bridges in Florida. They also showed that demand elasticities calculated across different off-peak periods varied considerably. These results suggest that the implementation of time-varying pricing schemes can encourage a more efficient use of motorways compared with a uniform toll throughout the day.

TOLL POLICY IN SPAIN

In the early 1970s, tolls were introduced on the road network in Spain mainly to raise revenue to finance construction, operation, and maintenance.[2] According to this objective, each motorway had to cover its own costs, and cross-subsidies between different motorways were not allowed. Initial toll rates were specific to each motorway concession and subsequently increased according to a cost

[1] According to the World Bank (2004), most countries have no toll roads and, where there are such roads, the tolled network typically comprises less than 5% of the entire road network.

[2] Appendix A provides a brief summary of the development of toll roads in Spain.

TABLE 1 Elasticity of Traffic Level with Respect to Tolls

Authors	Results	Context
Wuestefeld and Regan (1981)	Roads between −0.03 and −0.31 Bridges between −0.15 and −0.31 Average value = −0.21	16 tolled infrastructures in the U.S. (roads, bridges, and tunnels)
White (1984), quoted in Oum et al. (1992)	Peak-hours between −0.21 and −0.36 Off-peak hours between −0.14 and −0.29	Bridge in Southampton, UK
Goodwin (1988), quoted in May (1992)	Average value = −0.45	Literature review of a number of previous studies
Ribas, Raymond, and Matas (1988)	Between −0.15 and −0.48	Three intercity motorways in Spain
Jones and Hervik (1992)	Oslo −0.22 Alesund −0.45	Toll ring schemes in Norway
Harvey (1994)	Bridges between −0.05 and −0.15 Roads −0.10	Golden Gate Bridge, San Francisco Bay Bridge, and Everett Turnpike in New Hampshire (U.S.).
Hirschman, McNight, Paaswell, Pucher, and Berechman (1995)	Between −0.09 and −0.50 Average value −0.25 (only significant values quoted)	Six bridges and two tunnels in the New York City area, U.S.
Mauchan and Bonsall (1995)	Whole motorway network −0.40 Intercity motorways −0.25	Simulation model of motorway charging in West Yorkshire, UK
Gifford and Talkington (1996)	Own-elasticity of Friday–Saturday traffic −0.18 Cross-elasticity of Monday–Thursday traffic with respect to Friday toll −0.09	Golden Gate Bridge, San Francisco, U.S.
INRETS (1997), quoted in TRACE (1998)	Between −0.22 and −0.35	French motorways for trips longer than 100 kilometers
Lawley Publications (2000)	−0.20	New Jersey Turnpike, U.S.
Burris, Cain, and Pendyala (2001)	Off-peak period elasticity with respect to off-peak toll discount between −0.03 and −0.36	Lee County, Florida, U.S.

index based on the rate of inflation for the three main factors of production: labor, fuel, and steel. This resulted in a substantial variation in toll rates across the country, with higher tolls per kilometer on those motorways with larger construction costs or lower traffic volume.

However, for various reasons, the toll policy was modified over time and toll rates were not increased as planned. First, the severe economic crisis that the Spanish economy suffered from 1974 to 1984 revealed that certain concessionaires could not break even at initial toll rates. Three of the concessionaires with financial difficulties were taken over by the government, while others were merged with stronger companies. In both cases, the terms of the concession agreements were modified, leading to an increase in the initial toll rates in real terms. Moreover, explicit financial support from the government was allowed for a small share in the motorway network and cross-subsidies appeared among the merged concessionaires.[3] Furthermore, the formula approved for toll revisions was not systematically applied to all the motorways, and, as a consequence, toll rates for different motorways varied over time. Thus, in the 1980s, tolls increased in real terms on 8 of the 10 motorways, although at varying rates, whereas a decrease was observed in the other 2.

[3] It is important to note that the Spanish government assumed a very high level of risk as a consequence of the foreign rate assurance and the loan guarantees offered. The exchange rate losses over the period 1976–1996 were up to 65% of the total investment (Bel 1999).

In 1990, when most of the motorways had already been constructed, the toll revision formula was changed. Since then the new formula has been linked to the consumer price index (CPI) and allows for toll increases equal to 95% of annual CPI growth. This new approach to price regulation should have resulted in a slight decrease in toll rates in real terms over time for all the motorways. However, in practice, this formula was only systematically applied for a short period of time (1990 to 1996) and even then not to all motorways. The reasons again are manifold.

First, in the 1990s, there was renewed interest in the construction of tolled motorways from both central and some regional governments. In the first case, some concession agreements were renegotiated and existing toll rates reduced (even halved) to compensate for the introduction of tolls on upgraded toll-free motorways. Second, toll rates on regional motorways increased well above the average. Finally, the growing political pressure against tolls resulted in a renegotiation of most of the agreements with a reduction of toll rates in exchange for compensating the concessionaires with an extension of the concession period.

Since 1997, those motorways with higher tolls per kilometer have progressively reduced their rates; in some cases, the tolls decreased as much as 40% nominally in one year. Additionally, the rate of value-added tax was lowered from 16% to 7% on all the motorways.

The criterion used to set initial toll rates and changes in the toll policy during the 1990s have resulted in a wide range of variation of rates across the country and over time, which greatly facilitates the econometric estimation of toll elasticities.

MODEL SPECIFICATION

The Demand Equation

The methodology used to estimate the demand function was the panel data approach, where the cross-section observations correspond to motorway sections. This approach has two types of advantages. First, the temporal dimension allows the modeling of the dynamic adjustment of demand resulting from changes in transportation policy and the socioeconomic environment over time. Moreover, the cross-section observations provide more variation in the data, because toll rates vary more between motorway sections than they do over time. It thus solves the problem of insufficient variation in tolls per kilometer that appears in pure time series studies.[4] As a result, the estimated elasticity value captures the rich variation of prices across the different sections, as well as its temporal variation in a given section. Furthermore, by using a panel dataset, the number of observations is increased, which improves the precision of the estimated parameters.

We assumed that the volume of traffic on a motorway section is a function of the monetary and time costs of using the motorway, the monetary and time costs of using the alternative parallel free road or modes, the level of economic activity, and the generation and attraction factors at the origins and destinations. Monetary cost is defined as the sum of three components: toll, gasoline cost, and other vehicle operating costs. All the monetary variables were deflated by the CPI. The level of economic activity was measured as real gross domestic product (GDP); given that trips on motorways are undertaken for both leisure and business purposes, we used real GDP rather than disposable income in order to better capture the level of economic activity. Finally, the amount of traffic on a motorway section depends on the size of the potential market for each of them, which was determined by generation capacity and attraction of the origins and destinations, such as population and employment.

The model can therefore be expressed as follows:

$$Y_{it} = f\left(GDP_t, GP_t, MT_{it}^m, OC_{it}^m, TC_{it}^m, OC_{it}^o, TC_{it}^o, O_i, D_i, u_{it}\right) \quad (1)$$

where

m = motorway,

o = alternative (other) routes or modes,

Y_{it} = traffic volume on motorway section i in period t,

[4] When using a panel dataset, the total variability of a measure has two components: the within component (variability of the sample through time within each section) and the between component (variability of the sample across motorway sections). The panel dataset takes advantage of both sources of variability.

GDP_t = real national GDP in period t,

GP_t = gasoline price in period t deflated by the CPI,

MT_{it}^{m} = motorway toll on section i in period t deflated by the CPI,

OC_{it}^{j} = other vehicle operating costs (i.e., other than tolls and gasoline), $j = m, o$,

TC_{it}^{j} = time costs on section i in period t, $j = m, o$,

O_i = generation factors on section i,

D_i = attraction factors on section i,

u_{it} = error term, normally distributed with mean 0 and variance σ^2.

However, this is an ideal model. The empirical specification we finally estimated was limited by some data issues. Unfortunately, no data were available on other vehicle operating costs or time costs for the whole sample period. An analysis of the transportation costs in Spain allowed us to assume that vehicle operating costs and time costs have remained approximately constant over time on most of the motorway sections although this hypothesis did not hold for some of them. This was the case for seven sections, located around urban areas where both an increase in congestion and changes in the road network have affected the quality of service. These observations were excluded from the sample. The rest of the sections corresponded to interurban motorways where congestion was not a problem on most days. Hence, it can be assumed that time costs have remained relatively constant over time.

In order to take into account the most significant changes in the road network, a set of dummy variables was introduced. For example, the improvement of a parallel free road was captured by a dummy variable that takes the unit value since the opening year. Finally, the generation and attraction factors showed that the difference in traffic volume across motorway sections related mainly to population and the level of economic activity. Given that the dependent variable was observed for very short sections of the motorway and given also the difficulty in identifying how these factors should be measured, we assumed that these factors were captured by the specific fixed effects.[5]

[5] See Voith (1991) for a similar assumption.

Hence, under the assumption that $OC_{it}^{j} = OC_{i}^{j}$ and $TC_{it}^{j} = TC_{i}^{j}$ for $j = m$ and o, the equation can be rewritten as

$$Y_{it} = f\left\{\left(OC_i^m, TC_i^m, OC_i^o, TC_i^o, O_i, D_i\right), GDP_t, GP_t, MT_{it}^{m}, Z_{it}, u_{it}\right\} \tag{2}$$

where Z_{it} is the vector of dummy variables accounting for major changes in the network. These variables are defined in table 2.

One of the advantages of using a panel dataset is that this methodology allows us to explain the differences between cross-section observations not captured by the variables included in the model through the individual fixed effects, α_i. These individual fixed effects are represented by specific intercepts for each motorway section in the sample, and they capture the effect of factors not included in the equation that can be considered fixed over time but vary among motorway sections.

Thus, the demand equation can be rewritten as

$$Y_{it} = f(\alpha_i, GDP_t, GP_t, MT_{it}^{m}, Z_{it}, u_{it}) \tag{3}$$

where α_i captures the variables in parentheses in equation (2).

From a statistical point of view we have validated the assumed hypothesis that certain factors remain relatively constant over time by the application of recursive least squares.[6] This methodology allows us to prove the constancy of the estimated coefficients over time, so the null hypothesis of the structural constancy of the model is not rejected by the data.

Given the low number of temporal observations available for some of the motorway sections, the demand elasticities with respect to GDP and gasoline price are assumed to be the same across all motorway sections. According to the statistical test

[6] The recursive least squares technique estimates the model by adding new temporal observations in a progressive way, thus making it possible to test the stability of the coefficient vector. If the coefficient displays significant variation as more data are added to the estimated equation, it is an indication of instability. In our case, using the standard approach, the calculation of confidence intervals for the estimated recursive coefficients verifies the structural constancy hypothesis.

TABLE 2 Definition of the Dummy Variables Included in the Estimated Demand Equation

Dummy variables	Period	Comment	Expected sign
Z1–Z4	1994–1998	They reflect the negative impact on traffic on the 4 A(2) motorway sections, derived from capacity and quality improvements on the alternative free road.	–
Z5–Z7	1992	They account for the positive impact on the 3 A(4) motorway sections, derived from the Seville World Exhibition in 1992.	+
Z8–Z11	1995–1998	They reflect the negative impact on traffic on 4 A(7) motorway sections as a consequence of the extension of an alternative tollway.	–
Z12, Z14, and Z16	1993–1998	They reflect the negative impact on traffic on 3 A(7) motorway sections, derived from the opening of an alternative free motorway.	–
Z13, Z15, and Z17–Z24	1990–1998	They account for the positive impact on traffic on 10 A(7) motorway sections, due to the extension of this motorway.	+
Z25	1996–1998	It reflects improvements in the free alternative motorway network for the first of the A(19) motorway sections.	–
Z26, Z27, and Z28	1994–1998	They account for the positive impact on traffic on the 3 A(66) sections as a consequence of the improvement and extension of the motorway.	+

Notes: In Spain, the motorways (autopistas) are identified by the letter "A" followed by a number in parentheses.
These variables take the value 1 in the reported period; otherwise they are 0.

applied, these constraints were not rejected by the data.[7] The advantage of estimating a constrained model is that it allows efficiency gains in the estimation of the main parameter of interest, which in our case is toll elasticity. The coefficients of the toll variable, and hence the toll elasticities, are specific for each motorway section. We will, therefore, estimate different toll coefficients for each cross-section unit, which will depend on the characteristics of the motorway and the alternative routes.

To sum up, the traffic volume on motorway section i in period t depends on the individual fixed effects, the level of economic activity, the price of gasoline, and the level of toll. The individual fixed effects capture the effects of factors not included in the equation that remain relatively constant over time but vary among the different motorway sections. As previously mentioned, even for this more parsimonious version of the model, the use of recursive estimation techniques does not reject the temporal stability of the coefficients.

[7] The calculated F statistic for the hypothesis of equal elasticity with respect to gasoline price in all the sections of the motorway included in the sample is 0.944; for the hypothesis of equal elasticity with respect to GDP it is 1.082, while the critical value at a significance level of 5% is 1.22.

Some Econometric Issues

The next step in the model specification process is to decide on the functional form for the demand equation. The first issue we considered is whether the series are stationary or integrated[8] and, in the case of integrated series, whether they are cointegrated or not. The available econometric literature does not offer a clear guide on how to deal with this issue when panel data are used.

In this study, in spite of the short time span for the series (a maximum of 18 years), the traditional unit root tests (Augmented Dickey Fuller and Phillips Perron) were used to test whether the variables were stationary or integrated. The tests were applied to each motorway section. The null hypothesis of unit root was always nonrejected at the usual significance levels of the tests. However, the same tests showed the stationarity of the variables in first differences.

[8] Using integrated series (unit root series) is not a problem if the considered variables are cointegrated, given that in this case the cointegration property guarantees that the estimates are both consistent and efficient. However, if the variables are not cointegrated, this may give rise to the spurious regression problem. For a standard reference to unit root and cointegration tests, see Hamilton (1994). All results from the applied test are available on request.

The following step is an analysis of the series' cointegration. We carried this out using the Engle-Granger and the Cointegration Equation Durbin-Watson tests.[9] In this case, in almost all regressions estimated in levels, the null hypothesis of no cointegration was also nonrejected. Based on this evidence, and following standard econometric practice, all the estimations were carried out using first differences of the variables.

Second, in order to allow for dynamic effects, the starting specification included lags of the dependent and explanatory variables. The search for the final specification followed a general-to-specific process. After simplifying the model with restrictions that were not rejected by the data, a partial adjustment equation was selected. Therefore, both exogenous and lagged dependent variables appear as explanatory variables in the final model.

Finally, given that there are no theoretical arguments that can contribute to the choice of the functional form for the demand equation, we proceeded to select the most appropriate one on the basis of the goodness of fit of the models. We considered three alternatives—the linear model, the semi-log model, and the log-linear model—which are three of the most widely used in estimating aggregate demand models. The criterion used to select among these alternative specifications is based on the comparison of the values of the log of the likelihood function from the three competing models.[10] According to this criterion the log-linear specification was preferred as it showed the highest value for the log of the likelihood function.[11]

According to the three issues previously discussed, the equation to be estimated corresponds to a partial adjustment model where the variables are in first differences of the logarithms. The equation can be written as follows:[12]

$$\begin{aligned}\Delta\ln(Y_{it}) &= \beta_1\Delta\ln(GDP_t) + \beta_2\Delta\ln(GP_t) \\ &\quad + \beta_{3i}\Delta\ln(MT_{it}^{m}) + \varphi\Delta\ln(Y_{it-1}) \\ &\quad + \gamma'\Delta Z_{it} + \varepsilon_{it} \qquad (4)\end{aligned}$$

It must be stressed that in equation (4) using first differences of the variables eliminates the fixed effects from the estimated equation. In other words, the section-specific intercepts appearing in the model expressed in levels vanish from the finally estimated equation.

The presence of the lagged dependent variable as a regressor implies a dynamic structure for the response of the dependent variable to changes in the explanatory variable. That is, individuals do not adjust their behavior in one period, but with a delay. The underlying hypothesis for this specification is that present behavior is also determined by the values of the explanatory variables in the past. Therefore, the estimation of a dynamic model makes it possible to distinguish between short-term and long term effects. In our study, short term refers to the effect on demand occurring within one year of a change in the relevant variable, whereas long-term measures the total response to a change in the independent variable over time.

According to equation (4), the coefficients of the independent variables (β) should be interpreted as short-term elasticities. The long-term elasticities are

[9] The Johansen test was not applied to test cointegration, because this test assumes the existence of feedback between all the variables. In our case, variables such as gasoline price, toll, and GDP must be considered as weakly exogenous in a model trying to explain motorway traffic volume.

[10] The log of the likelihood functions for the linear, semi-log, and log-linear specifications are, respectively:

$$L_0 = C - \frac{T}{2}\ln(SSR_0), L_1 = C - \frac{T}{2}\ln(SSR_1) - \sum_1^T \ln(Y_t),$$

$$L_2 = C - \frac{T}{2}\ln(SSR_2) - \sum_1^T \ln(Y_t)$$

where the constant C is the same for each specification, SSR is the residual sum of squares, Y is the dependent variable and T is the sample size (see Davidson and MacKinnon 1993). The calculated values for these functions are, respectively, –9,768.5, –9,201.8, and –9,183.6

[11] The log-linear functional form is one of the most widely used functional forms in aggregate demand models. In spite of its simplicity, the log-linear specification offers an adequate approximation to the demand curve, at least in the neighborhood of the actual data. This is the usual procedure for selecting among alternative functional forms when estimating aggregate transportation demand equations. For an application of similar procedures see, for example, Oum (1989) and Dargay and Hanley (2002).

[12] This is a standard specification for aggregate demand functions. See, for instance, Dargay and Goodwin (1995), Dargay and Hanley (2002).

$\frac{\beta}{1-\varphi}$, where $1-\varphi$ is the adjustment factor measuring the speed of adjustment. The greater the value of φ the slower the speed of adjustment and the greater the difference between short-term and long-term elasticities.

The concept of mean lag is useful to characterize the dynamic structure of the model. The mean lag is defined as a weighted average of the lag structure of the model, where the weighting coefficient for period *j* is the ratio between the coefficient with lag *j* and the long-term coefficient. The mean lag can be calculated as $\frac{\varphi}{1-\varphi}$.

THE DATA

The data cover all Spanish tolled motorways sections for 18 observation years between 1981 and 1998 (see Ministerio de Fomento Annual). The cross-section observations correspond to the shortest motorway section allowed by data-collection processes, with an average length of 14.7 kilometers. The use of these short sections guarantees that the observed traffic mix is homogeneous.

We eliminated 11 sections: those that experienced significant changes in congestion (either on the motorway or on the alternative routes), those that partially admit toll-free traffic, and those that have open tolls. Not all the motorways sections were observed for all the years in the sample. Only sections for which data were available for at least eight periods were used. Furthermore, those sections belonging to motorways not completely constructed during the observation period were also eliminated to avoid changes in traffic volume that may be due to the progressive extension of the motorway. The final sample was a panel dataset of 72 road sections for 1981 through 1998, although this temporal span was not available for all cross-section units. The total number of observations was 1,135.[13]

The dependent variable is the annual average daily traffic volume in each section, defined as the number of vehicle-kilometers run per year, divided by section length and number of days.[14] The explanatory variables are real GDP, gasoline price, and toll per kilometer, the last two deflated by the CPI. Working with short sections of the motorway made it possible to calculate in a fairly precise way the toll paid per kilometer. GDP and gasoline prices are defined at the national level and take the same value for all sections in the sample, but, as we are working with a panel dataset, these have different values for each year of the sample. Finally, a set of 28 dummy variables captures the most important changes in the road network. These variables, defined in table 2 (page 6), take the value 1 in the reported period and 0 otherwise. The main descriptive statistics for the explanatory variables are defined in table 3.

Before estimating the demand equation, we present two of the main features of the relevant variables in the study: traffic volume and toll paid per kilometer. First, as figure 1 shows, there seems to be a clear relationship between the level of economic activity—measured as GDP—and traffic volume over time. Using aggregate data for all the motorway sections for 1981 through 1998, figure 1 shows the synchronism between the rates of growth of GDP and traffic volume with a correlation coefficient equal to 0.796.[15] This preliminary result is in line with previous studies showing that automobile use is elastic with respect to income.[16] It is also interesting to compare the cycles of GDP and traffic volume.[17] As can be seen in figure 2, the traffic cycle clearly overreacts to the GDP cycle. Therefore, in periods of economic expansion, the cyclical components of

[13] Given that the equation is estimated in first differences of the logarithms and includes the lagged dependent variable, the final number of observations is reduced to 990.

[14] It should be noted that the dependent variable is an aggregate of different types of traffic, of different length and purpose. Therefore, estimated elasticity for each section must be understood as an average value.

[15] The *t* statistic is equal to 5.27 and for 18 observations the null hypothesis of independence will be rejected at a *P*-value of 0.0001. This confirms the narrow relationship that exists between both variables.

[16] For a recent review on this subject, see Graham and Glaister (2002).

[17] Cycles for both variables were obtained through the application of the Hodrick-Prescott filter to the log of the series and by calculating the difference between the observed and trend values. This filter is a standard technique that allowed us to smooth the series in order to obtain an adaptive long-term trend for the variable. It is usual to consider that the difference between the observed series and smoothed series approximates the cycle.

TABLE 3 Descriptive Statistics

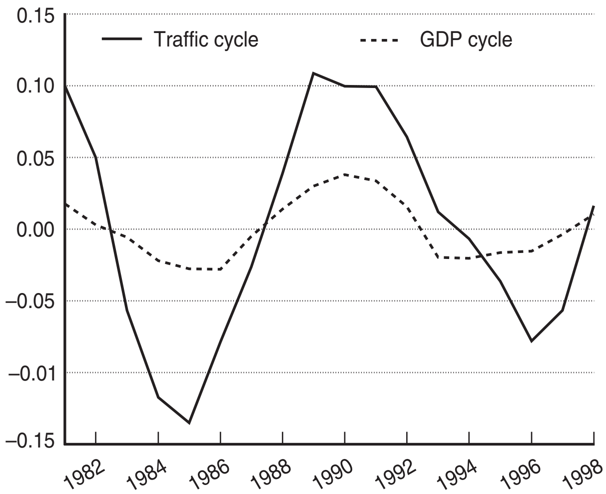
Variables	Mean	Median	Maximum	Minimum	Std. dev.	Observations
Daily traffic volume	11,490	9,460	63,741	1,689	8,821	1,135
Toll (euros per km)[1]	0.091	0.086	0.224	0.037	0.035	1,135
Gasoline price (euros per liter)[1]	0.619	0.533	0.867	0.486	0.139	18
GDP (millions of euros)[1]	219,311	230,277	275,869	174,149	33,619	18
Section length (kms)	14.7	13.9	43.0	2.0	8.2	72

[1] The base year for variables expressed in monetary units is 1992.

traffic volume exceed the corresponding components of GDP, while the opposite occurs during recession.

Second, at the cross-section level, a substantial difference is observed in traffic volume among the different motorways as well as among sections of the same motorway. The daily average traffic volume ranges from 1,689 automobiles per day in the section and year having the lowest volume to 63,741 automobiles per day in the section and year with the highest. Finally, as we explained earlier, we found an extensive price range for initial toll rates. For the whole period, at 1992 prices, the lowest price paid per kilometer was about 0.037 euros, whereas the highest was about 0.22 euros.

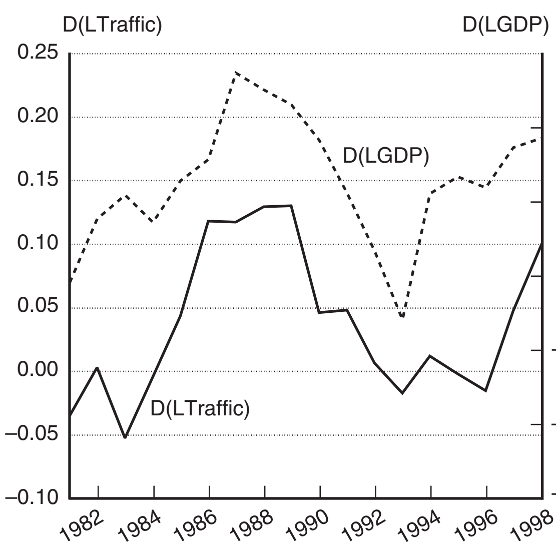
FIGURE 2 Traffic Cycle and GDP Cycle



RESULTS

The results of the estimated model—equation (4)—show that all the estimated coefficients have the expected signs and most of these were estimated with a high degree of precision, as measured by the standard error of the coefficients (see appendix B). Given that heteroscedasticity was observed in the variance of the random term between sections, the model was estimated using weighted least squares (WLS). Comparing ordinary least squares with WLS, the latter procedure results in similar estimates while the standard errors decrease. In relation to the toll coefficients, a significant variation across motorway sections was observed.

FIGURE 1 Rate of Growth of GDP and Traffic Volume



A Chi-square test allowed us to clearly reject the null hypothesis of equality of toll coefficients across all sections.[18] On the other hand, the differences in the value of the toll coefficient (which, given the model specification, correspond to short-term elasticity) could be explained by certain motorway characteristics. First, adjacent sections in the same motorway present very similar elasticities. Second,

[18] The calculated Chi-square statistic was 113.12, while the critical value for 71 degrees of freedom (d.f.) at a significance level of 5% is 52.0.

the more inelastic sections are located on corridors with a high volume of traffic—mainly the motorways along the Mediterranean coast. Third, demand is seen to be more elastic where a good alternative free road exists.

The observed results suggest the possibility of reestimating the model introducing the hypothesis of equality of toll elasticities across those motorway sections that showed similar coefficients in the initial general model. The introduction of equality constraints among coefficients, not rejected by the data (see below), makes it possible to obtain more precise estimates of the coefficients by reducing both the number of coefficients to be estimated and the multicollinearity. In fact, we followed standard econometric methodology that recommends going from the general to more parsimonious model. These constraints were introduced by classifying the motorway sections into the following groups according to the toll coefficient estimated in the general model:

1. Low short-term elasticity: sections with toll coefficients between 0 and −0.3.
2. Middle-low short-term elasticity: sections with toll coefficients between −0.3 and −0.4.
3. Middle-high short-term elasticity: sections with toll coefficients between −0.4 and −0.6.
4. High short-term elasticity: sections with toll coefficients larger than −0.6 in absolute value.

Thus, four different coefficients for toll elasticities are now estimated. Detailed results of this final model are reported in table 4 and correspond to the WLS estimation. The application of a Chi-square test did not reject the hypothesis of classifying the motorway sections into four groups according to their estimated toll elasticity.[19] The model fits the data well and all estimated coefficients are highly significant. The level of economic activity has a positive effect on traffic volume, whereas gasoline price and the toll have a negative influence. With respect to toll coefficients, significant differences among them can be observed, according to the grouping by sections mentioned above. All dummy variables take the expected signs described earlier in table 2. It should be noted that the inclusion of these variables increased the statistical significance of toll variable estimates without modifying their value in any noticeable way.

The short-term and long-term elasticities are summarized in table 5. As can be seen from the *t*-statistics, all the estimated coefficients are significant at *P*-values clearly lower than the conventional 0.05 or 0.01 levels. A lag parameter equal to 0.366 implies a long-term effect of about 1.58 times the short-term effect. This result reflects a wider range of opportunities and available options open to individuals over a longer time span. However, the period of adjustment is relatively short, with 87% of total adjustment taking place within the first two years.

Traffic on tolled motorways is shown to be elastic in relation to GDP, with elasticity values equal to 0.89 and 1.41 for the short- and long-term, respectively. This result confirms what is intuitively obvious in figures 1 and 2. Elasticity with respect to gasoline price equals –0.34 in the short term and –0.53 in the long term. Our results are consistent with those reported in the literature,[20] although they are closer to the maximum reported values. In the context of this paper, relatively high value for gasoline price elasticity can be expected, compared with other estimates carried out for freeways, given that when the gasoline price is increased tolled motorway users can shift to a parallel free road.

The estimated coefficients on the toll variables provide evidence that demand is sensitive to toll variations. This conclusion is supported by the high precision, measured by the standard error, with which the elasticities have been estimated. Nonetheless, significant differences were observed among groups of motorways. For the first group, demand is shown to be rather inelastic. Short-term and long-term elasticities are equal to –0.21 and –0.33, respectively. However, for the remaining groups, demand becomes more price elastic. For those

[19] The calculated Chi-square statistic was 13.27, while the critical value for 68 d.f. at a significance level of 5% was 49.0.

[20] For a literature review of such findings, see Goodwin (1992), Oum et al. (1992), Johansson and Schipper (1997), Espey (1998), and de Jong and Gunn (2001). Graham and Glaister (2002) provide an extensive international review of demand elasticity with respect to fuel price.

TABLE 4 Estimated Demand Equation
Dependent variable: D(LTRAFFIC)
Estimation method: weighted least squares
Total system (unbalanced) observations: 990

Variable	Coefficient	Std. error	*t*-statistic	*P*-value
D(LGDP)	0.8901	0.0409	21.7605	0.0000
D(LPGAS)	–0.3367	0.0153	–22.0050	0.0000
D(LTRAFFIC(–1))	0.3659	0.0158	23.1470	0.0000
D(LTOLL1)	–0.2092	0.0177	–11.813	0.0000
D(LTOLL2)	–0.3707	0.0147	–25.248	0.0000
D(LTOLL3)	–0.4449	0.0225	–19.801	0.0000
D(LTOLL4)	–0.8286	0.0844	–9.8179	0.0000
D(Z1)	–0.0517	0.0260	–1.9919	0.0467
D(Z2)	–0.0689	0.0239	–2.8782	0.0041
D(Z3)	–0.0718	0.0246	–2.9179	0.0036
D(Z4)	–0.0519	0.0263	–1.9745	0.0486
D(Z5)	0.1549	0.0396	3.9136	0.0001
D(Z6)	0.1690	0.0364	4.6466	0.0000
D(Z7)	0.1196	0.0583	2.0507	0.0406
D(Z8)	–0.0679	0.0219	–3.1035	0.0020
D(Z9)	–0.0623	0.0201	–3.1029	0.0020
D(Z10)	–0.0656	0.0286	–2.2911	0.0222
D(Z11)	–0.0425	0.0227	–1.8689	0.0619
D(Z12)	–0.0550	0.0250	–2.2015	0.0279
D(Z13)	0.0746	0.0251	2.9698	0.0031
D(Z14)	–0.0337	0.0201	–1.6798	0.0933
D(Z15)	0.0626	0.0202	3.0952	0.0020
D(Z16)	–0.0360	0.0188	–1.9175	0.0555
D(Z17)	0.0498	0.0188	2.6367	0.0085
D(Z18)	0.0445	0.0192	2.3154	0.0208
D(Z19)	0.0404	0.0153	2.6397	0.0084
D(Z20)	0.0529	0.0134	3.9563	0.0001
D(Z21)	0.1698	0.0433	3.9163	0.0001
D(Z22)	0.0812	0.0163	4.9712	0.0000
D(Z23)	0.0822	0.0207	3.9686	0.0001
D(Z24)	0.1379	0.0187	7.3903	0.0000
D(Z25)	–0.1366	0.0488	–2.7992	0.0052
D(Z26)	0.0864	0.0177	4.8746	0.0000
D(Z27)	0.0751	0.0175	4.2902	0.0000
D(Z28)	0.0451	0.0206	2.1897	0.0288
R^2 (average for the motorway sections)	0.74			
First order autocorrelation coefficient (pooling estimation for the motorway sections)	0.019			

Note: All variables are defined in first differences (D) of the logarithm (L). GDP = gross domestic product; PGAS = gasoline price; TRAFFIC = average daily traffic volume; TOLL1 = low toll elasticity group; TOLL2 = low-medium toll elasticity group, TOLL3 = medium-high toll elasticity group; TOLL4 = high toll elasticity group; D(Z1) to D(Z28) = first differences of the dummy variables to account for changes in the road network.

TABLE 5 Estimated Short-Term and Long-Term Elasticities[1]

Variable	Short-term elasticity	*t*-statistic	Long-term elasticity	*t*-statistic
GDP elasticity	0.890	21.76	1.405	27.85
Gasoline price elasticity	–0.337	–22.01	–0.531	–18.50
Toll elasticity group 1	–0.209	–11.81	–0.330	–11.42
Toll elasticity group 2	–0.371	–25.25	–0.585	–21.71
Toll elasticity group 3	–0.445	–19.80	–0.702	–17.66
Toll elasticity group 4	–0.828	–9.82	–1.307	–9.81

[1] Group 1 includes 21 sections; group 2, 25 sections; group 3, 21 sections; and group 4, 5 sections.

motorway sections classified in group 4, elasticities are over –0.8 in the short term and well above unity in the long term. These differences prove that the demand response to toll variations depends on the particular characteristics of the motorway and alternative routes. In the next section, we provide some evidence of these characteristics.

VARIATION OF TOLL ELASTICITIES ACROSS MOTORWAYS

Once it has been proved that toll elasticities vary across motorway sections, it is interesting to consider which are the main variables that explain such differences. With this purpose in mind, we estimated an ordered probit model[21] where the dependent variable is the category in which the tolled section falls, ranging from 1 to 4 (low, middle-low, middle-high, and high categories of toll elasticities).

The set of explanatory variables is limited by data availability. First, we were able to gather information on average speed and the percentage of heavy vehicles with respect to total traffic on the parallel free road; these variables reflect the quality of the alternative road. Second, two characteristics of the motorway have been included: section length and a dummy for sections in tourist areas. There are a priori reasons to expect that traffic in tourist areas will be less sensitive to price. It might well be that foreign visitors, due to a lack of information given that they are occasional users, have more inelastic demands than frequent motorway users. Moreover, congestion in these areas on the free alternative roads is rather high during summer months due to their low capacity and the high volume of short-distance traffic for which tolled motorways are not a feasible option. This increased congestion can further reduce demand elasticity.

The number of observations in this model falls from 72 to 52, as we could not gather all the required information for all sections. Table 6 shows the results of the estimated equation. Because the interpretation of the coefficients of the model was not straightforward, we calculated the change in the estimated frequencies (probabilities) after a change in the explanatory variable. Baseline frequencies were calculated for the mean value of the variables in the sample, and the tourism dummy takes value 1. In order to simulate the change in probabilities a 10% increase in each variable was assumed. Results are presented in table 7.

The estimated frequencies show that demand is more sensitive to price when the alternative free road is of better quality. That is, the higher the speed on the alternative road the more elastic demand is with respect to tolls. On the contrary, when the percentage of heavy vehicles on the alternative road increases, the roadway segment shifts into a more inelastic demand category. Additionally, demand is slightly more elastic on longer motorway sections. This can be explained by the fact that demand is more sensitive to price when the total amount to be paid is larger. Finally, as could be expected, motorway demand in tourist areas is more inelastic.

CONCLUSIONS

The estimation of a dynamic demand function on tolled motorways has made it possible to identify the behavioral responses of users to changes in the

[21] Alternatively, a logit ordered model was estimated with very similar results.

TABLE 6 Estimation Results of the Ordered Probit Model
Dependent variable: category of toll elasticity (from 1 to 4)
Robust *t*-statistics

Variable	Coefficient	Std. error	*t*-statistic	*P*-value
Speed on the alternative road	0.032	0.010	3.298	0.001
Percentage of heavy vehicles on the alternative road	–0.053	0.017	–4.233	0.000
Motorway section length	0.024	0.010	4.268	0.000
Tourist dummy	–1.227	0.358	–3.340	0.001
Limit_1	0.919	0.862	1.214	0.226
Limit_2	2.393	0.920	3.014	0.003
Limit_3	3.666	0.962	3.814	0.000
Observations	52			
Likelihood ratio-statistic	25.60 (critical value at 5% = 9.49)			

Notes: The limit points are the estimates of the threshold coefficients of the distribution function. That is, if $F(X'\beta)$ is the distribution function of the unobserved continuous latent variable, the ordered probit model implies that:

If $F(X'\beta) \leq$ Limit_1, then the dependent variable falls into category 1 (low elasticity).

If Limit_1< $F(X'\beta) \leq$ Limit_2, then the dependent variable falls into category 2 (middle-low elasticity).

If Limit_2 < $F(X'\beta) \leq$ Limit_3, then the dependent variable falls into category 3 (middle-high elasticity).

If $F(X'\beta) >$ Limit_3, then the dependent variable falls into category 4 (high elasticity).

TABLE 7 Estimated Probabilities

Motorway group elasticity	Baseline	10% increase in speed on alternative road	10% increase in heavy vehicles on alternative road	10% increase in section length	Tourism dummy = 0
Low	0.522	0.410	0.574	0.500	0.121
Middle-low	0.415	0.484	0.377	0.430	0.498
Middle-high	0.060	0.100	0.047	0.067	0.323
High	0.003	0.006	0.002	0.003	0.058

Note: The baseline values taken by the explanatory variables are: speed = 88.9 km/hr; percentage of heavy vehicles = 24.9%; section length = 23.4 km; and tourism dummy = 1.

explanatory variables. First, traffic on the tolled motorways is shown to be strongly correlated with the level of economic activity in such a way that, during periods of growth, traffic increases clearly exceed GDP growth, with the opposite occurring during recessions.

Travel demand is shown to be less sensitive to gasoline prices and tolls than it is to GDP. Elasticity with respect to gasoline price is about – 0.3, whereas a wide range of variation appears in toll elasticities across motorway sections. The model results prove that an average aggregate toll elasticity cannot be used for forecasting or evaluation purposes. According to individual estimates, the sections were classified into four categories for which short-term elasticity ranged from – 0.21 in the most inelastic sections to – 0.83 in the most elastic. This range of variation can be explained by those variables related to the quality of the alternative roads, the length of the motorway section, and the location of the motorway in a tourist area. The more congested the alternative roads are, the higher the time benefits of using the tolled motorway will be, with demand consequently being more inelastic.

The finding that the sensitivity of demand to tolls can be higher than the average values found in the literature confirms that tolling motorways can have a significant impact on traffic. Setting a toll on a motorway can result in a misallocation of traffic between the tolled motorway and the parallel free road. There are several examples in Spain of underused motorway sections while the alternative road is severely congested, with a consequent increase in maintenance and environmental costs. In such cases, decreasing the toll may improve traffic allocation and, hence, reduce the total costs of using the infrastructure. Moreover, it should be noted that investment in alternative roads or transportation modes would imply a more elastic demand for motorway users, because they can take advantage of a wider range of choices in traveling to their destinations. Thus, decisions about toll levels on the motorways are not independent of investment policy for transportation infrastructure.

ACKNOWLEDGMENTS

This study has had the inestimable dedication of our research assistant, Alvaro Angeriz. We also wish to thank Javier Asensio for his helpful suggestions and comments. Additionally, we are grateful to Mark Burris for providing certain references to toll elasticities in the United States.

Financial support for this study was provided by the Spanish Ministerio de Fomento and the CICYT program (project SEC2003-7928). The authors gratefully acknowledge the support received from the Ministerio de Fomento, from the Asociación de Sociedades Españolas Concesionarias de Autopistas, and from the concessionary companies.

REFERENCES

Bel, G. 1999. Financiación de Infraestructuras Viarias: La Economía Política de los Peajes. *Papeles de Economía Española* 82:123–139.

Burris, M.W., A. Cain, and R.M. Pendyala. 2001. Impact of Variable Pricing on Temporal Distribution of Travel Demand. *Transportation Research Record* 1747:36 – 43.

Dargay, J.M. and P.B. Goodwin. 1995. Evaluation of Consumer Surplus with Dynamic Demand. *Journal of Transport Economics and Policy* 29:179–193.

Dargay, J.M. and M. Hanly. 2002. The Demand for Local Bus Services in England. *Journal of Transport Economics and Policy* 36:73–91.

Davidson, R. and J.G. MacKinnon. 1993. *Estimation and Inference in Econometrics*. New York, NY: Oxford University Press.

De Jong, G. and H. Gunn. 2001. Recent Evidence on Car Cost and Time Elasticities of Travel Demand in Europe. *Journal of Transport Economics and Policy* 35:137–160.

Espey, M. 1998. Gasoline Demand Revisited: An International Meta-Analysis of Elasticities. *Energy Economics* 20:273–295.

Gifford, J.L. and S.W. Talkington. 1996. Demand Elasticity Under Time-Varying Prices: Case Study of Day-of-Week Varying Tolls on the Golden Gate Bridge. *Transportation Research Record* 1558:55–59.

Gómez-Ibañez, J.A. and J.R. Meyer. 1993. *Going Private: The International Experience with Transport Privatization.* Washington, DC: The Brookings Institution.

Goodwin, P.B. 1992. A Review of New Demand Elasticities with Special Reference to Short and Long Run Effects of Price Changes. *Journal of Transport Economics and Policy* 26:155–169.

Graham, D.J. and S. Glaister. 2002. The Demand for Automobile Fuel: A Survey of Elasticities. *Journal of Transport Economics and Policy* 36:1–26.

Hamilton, J.D. 1994. *Time Series Analysis.* Princeton, NJ: Princeton University Press.

Harvey, G. 1994. Transportation Pricing Behavior. *Curbing Gridlock: Peak-Period Fees To Relieve Traffic Congestion,* 2, Transportation Research Board Special Report 242. Washington, DC: National Academy Press. pp. 89–114.

Hirschman, I., C. McNight, R.E. Paaswell, J. Pucher, and J. Berechman. 1995. Bridge and Tunnel Toll Elasticities in New York: Some Recent Evidence. *Transportation* 22:97–113.

Johansson, O. and L. Schipper. 1997. Measuring the Long-Run Fuel Demand of Cars. *Journal of Transport Economics and Policy* 31:272–292.

Jones, P. and A. Hervik. 1992. Restraining Car Traffic in European Cities: An Emerging Role for Road Pricing. *Transportation Research A* 26:133–145.

Lawley Publications. 2000. Traffic Responses to Toll Increases Remain Inelastic. *The Urban Transportation Monitor* 14(10).

Mauchan, A. and P. Bonsall. 1995. Model Predictions of the Effects of Motorway Charging in West Yorkshire. *Traffic, Engineering and Control* 36:206–212.

May, A.D. 1992. Road Pricing: An International Perspective. *Transportation* 19:313–333.

Ministerio de Fomento. 2003. Memoria. Delegación del Gobierno en las Sociedades Concesionarias de Autopistas de Peaje, Secretaria de Estado de Infraestructuras y Transportes, Ministerio de Fomento, Madrid, Spain.

______. Annual. *El Tráfico en las Autopistas de Peaje*. Madrid, Spain: Dirección General de Carreteras, Ministerio de Fomento.

Oum, T.H. 1989. Alternative Demand Models and Their Elasticity Estimates. *Journal of Transport Economics and Policy* 13:163–187.

Oum, T.H., W.G. Waters, and J.S. Yong. 1992. Concepts of Price Elasticities of Transport Demand and Recent Empirical Estimates. *Journal of Transport Economics and Policy* 26:139–154.

Ribas, E., J.L. Raymond, and A. Matas. 1988. Estudi Sobre la Elasticitat Preu de la Demanda de Tràfic per Autopista. Departament de Política Territorial i Obres Públiques, Generalitat de Catalunya, Barcelona, Spain.

TRACE Consortium. 1998. *Deliverable 1: Outcomes of Review on Elasticities and Values of Time*. The Hague, Netherlands.

Voith, R. 1991. The Long-Run Elasticity of Demand for Commuter Rail Transportation. *Journal of Urban Economics* 30:360–372.

World Bank. 2004. *Toll Roads and Concessions*. Available at http://www.worldbank.org/transport/roads/toll_rds.htm, as of Feb. 9, 2004.

Wuestefeld, N.H. and E.J. Regan. 1981. Impact of Rate Increases on Toll Facilities. *Traffic Quarterly* 34:639–55.

Author addresses: Anna Matas, Departament d'Economía Aplicada, Facultat Ciències Econòmiques, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain. Email: anna.matas@uab.es.

José-Luis Raymond, Departament d'Economía i Història Econòmica, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain. Email: josep.raymond@uab.es.

APPENDIX A

In Spain, the tolled motorway construction policy of the 1960s granted concessions to private companies both for their construction and operation. As a result of this policy, 1,800 kilometers of tolled motorways, called *autopistas*, were completed by the end of the 1970s, serving demand along two main traffic corridors.

Once the main traffic corridors had been concessioned and, simultaneously, with the Spanish economy suffering the effects of the energy crisis, private capital was no longer interested in the construction of autopistas. In the mid-1970s, the concession of planned motorways was increasingly difficult, and some were postponed. By the end of that decade, the policy was abandoned.

In the 1980s, the need for significant expansion of the road network was evident, and the government decided to finance this with national tax revenue. Approximately 5,500 kilometers of untolled motorways were constructed by 1998, covering most of the national network. The main exceptions were the concessions granted by the regional government of Catalonia to construct and operate some tolled motorways. By 1998, the Spanish national highway network consisted of 9,637 kilometers, of which 2,072 kilometers were tolled motorways, 6,185 kilometers were untolled motorways, and 1,380 kilometers were two-lane freeways.

More recently, and due to severe public budget constraints, a new program of private tolled motorways was initiated. Nevertheless, the scope of private tolled roads in Spain is currently limited. For a review of the development of Spain's motorways, see Gómez-Ibañez and Meyer (1993), and for an analysis of the present situation, see Ministerio de Fomento (2003).

APPENDIX B

TABLE B1 Estimated General Model
Dependent variable: D(LTRAFFIC)
Estimation method: weighted least squares
Total system (unbalanced) observations: 990

	Coefficient	Std. error	*t*-statistic	*P*-value
D(LGDP)	0.8942	0.0410	21.8117	0.0000
D(LPGAS)	–0.3369	0.0157	–21.4261	0.0000
D(LTRAFFIC(–1))	0.3642	0.0160	22.6942	0.0000
D(LTOLL1)	–0.0371	0.2279	–0.1626	0.8709
D(LTOLL2)	–0.2754	0.2630	–1.0471	0.2954
D(LTOLL3)	–0.2338	0.2631	–0.8886	0.3745
D(LTOLL4)	–0.3237	0.2548	–1.2704	0.2043
D(LTOLL5)	–0.1799	0.2101	–0.8565	0.3920
D(LTOLL6)	–0.4766	0.1451	–3.2853	0.0011
D(LTOLL7)	–0.4758	0.1274	–3.7354	0.0002
D(LTOLL8)	–0.4630	0.1362	–3.3982	0.0007
D(LTOLL9)	–0.2811	0.1296	–2.1689	0.0304
D(LTOLL10)	–0.4903	0.1258	–3.8977	0.0001
D(LTOLL11)	–0.4512	0.1157	–3.9003	0.0001
D(LTOLL12)	–0.5876	0.1893	–3.1042	0.0020
D(LTOLL13)	–0.3588	0.1705	–2.1048	0.0356
D(LTOLL14)	–0.3555	0.2371	–1.4994	0.1341
D(LTOLL15)	–0.1118	0.1281	–0.8725	0.3832
D(LTOLL16)	–0.3144	0.1504	–2.0901	0.0369
D(LTOLL17)	–0.0580	0.1998	–0.2902	0.7717
D(LTOLL18)	–0.1894	0.1477	–1.2828	0.1999
D(LTOLL19)	–0.3475	0.2299	–1.5118	0.1309
D(LTOLL20)	–0.3417	0.1964	–1.7398	0.0822
D(LTOLL21)	–0.4309	0.2585	–1.6672	0.0958
D(LTOLL22)	–0.3609	0.4125	–0.8748	0.3819
D(LTOLL23)	–0.1517	0.0422	–3.5972	0.0003
D(LTOLL24)	–0.1716	0.0390	–4.4061	0.0000
D(LTOLL25)	–0.1815	0.0456	–3.9764	0.0001
D(LTOLL26)	–0.2204	0.0629	–3.5052	0.0005
D(LTOLL27)	–0.2831	0.0827	–3.4209	0.0007
D(LTOLL28)	–0.2653	0.0608	–4.3657	0.0000
D(LTOLL29)	–0.3136	0.0413	–7.5963	0.0000
D(LTOLL30)	–0.4182	0.0646	–6.4717	0.0000
D(LTOLL31)	–0.4387	0.0534	–8.2101	0.0000
D(LTOLL32)	–0.4062	0.0465	–8.7398	0.0000
D(LTOLL33)	–0.3620	0.0537	–6.7361	0.0000
D(LTOLL34)	–0.3915	0.0414	–9.4648	0.0000
D(LTOLL35)	–0.3447	0.0342	–10.0812	0.0000
D(LTOLL36)	–0.3582	0.1184	–3.0254	0.0026

TABLE B1 Estimated General Model (continued)
Dependent variable: D(LTRAFFIC)
Estimation method: weighted least squares
Total system (unbalanced) observations: 990

	Coefficient	Std. error	*t*-statistic	*P*-value
D(LTOLL37)	–0.3701	0.0445	–8.3134	0.0000
D(LTOLL38)	–0.3904	0.0512	–7.6203	0.0000
D(LTOLL39)	–0.3992	0.0489	–8.1651	0.0000
D(LTOLL40)	–0.5231	0.2504	–2.0892	0.0370
D(LTOLL41)	–0.4556	0.2466	–1.8476	0.0650
D(LTOLL42)	–0.4489	0.1657	–2.7082	0.0069
D(LTOLL43)	–0.4662	0.2451	–1.9018	0.0575
D(LTOLL44)	–0.3729	0.1515	–2.4614	0.0140
D(LTOLL45)	–0.4115	0.1619	–2.5420	0.0112
D(LTOLL46)	–0.4045	0.2312	–1.7495	0.0806
D(LTOLL47)	–0.5029	0.1554	–3.2366	0.0013
D(LTOLL48)	–0.0931	0.2265	–0.4108	0.6813
D(LTOLL49)	–0.2227	0.1657	–1.3439	0.1793
D(LTOLL50)	–0.1935	0.0913	–2.1199	0.0343
D(LTOLL51)	–0.3617	0.0650	–5.5628	0.0000
D(LTOLL52)	–0.4411	0.0647	–6.8186	0.0000
D(LTOLL53)	–0.8415	0.1494	–5.6340	0.0000
D(LTOLL54)	–0.8140	0.1435	–5.6730	0.0000
D(LTOLL55)	–0.8301	0.1714	–4.8438	0.0000
D(LTOLL56)	–0.3729	0.1065	–3.5004	0.0005
D(LTOLL57)	–0.3294	0.0839	–3.9255	0.0001
D(LTOLL58)	–0.3569	0.1728	–2.0648	0.0392
D(LTOLL59)	–0.3863	0.0936	–4.1281	0.0000
D(LTOLL60)	–0.5015	0.0830	–6.0381	0.0000
D(LTOLL61)	–0.5248	0.1946	–2.6970	0.0071
D(LTOLL62)	–0.4816	0.1138	–4.2314	0.0000
D(LTOLL63)	–0.3233	0.1335	–2.4213	0.0157
D(LTOLL64)	–0.3922	0.0990	–3.9625	0.0001
D(LTOLL65)	–0.4431	0.1168	–3.7933	0.0002
D(LTOLL66)	–0.3706	0.1427	–2.5963	0.0096
D(LTOLL67)	–0.3451	0.0534	–6.4635	0.0000
D(LTOLL68)	–0.3692	0.0562	–6.5701	0.0000
D(LTOLL69)	–0.4417	0.1117	–3.9532	0.0001
D(LTOLL70)	–0.8108	0.2983	–2.7182	0.0067
D(LTOLL71)	–0.8798	0.3854	–2.2825	0.0227
D(LTOLL72)	–0.2516	0.3066	–0.8208	0.4120
D(Z1)	–0.0517	0.0259	–1.9918	0.0467
D(Z2)	–0.0688	0.0240	–2.8680	0.0042
D(Z3)	–0.0720	0.0246	–2.9222	0.0036
D(Z4)	–0.0518	0.0259	–1.9971	0.0461

(*continues on next page*)

TABLE B1 Estimated General Model (continued)
Dependent variable: D(LTRAFFIC)
Estimation method: weighted least squares
Total system (unbalanced) observations: 990

	Coefficient	Std. error	*t*-statistic	*P*-value
D(Z5)	0.1548	0.0395	3.9192	0.0001
D(Z6)	0.1689	0.0363	4.6490	0.0000
D(Z7)	0.1197	0.0577	2.0755	0.0382
D(Z8)	–0.0687	0.0214	–3.2092	0.0014
D(Z9)	–0.0621	0.0198	–3.1309	0.0018
D(Z10)	–0.0689	0.0283	–2.4343	0.0151
D(Z11)	–0.0428	0.0228	–1.8769	0.0609
D(Z12)	–0.0554	0.0242	–2.2876	0.0224
D(Z13)	0.0726	0.0248	2.9326	0.0034
D(Z14)	–0.0337	0.0198	–1.7058	0.0884
D(Z15)	0.0622	0.0202	3.0823	0.0021
D(Z16)	–0.0365	0.0172	–2.1197	0.0343
D(Z17)	0.0474	0.0175	2.7040	0.0070
D(Z18)	0.0439	0.0194	2.2611	0.0240
D(Z19)	0.0419	0.0155	2.7086	0.0069
D(Z20)	0.0515	0.0130	3.9495	0.0001
D(Z21)	0.1690	0.0440	3.8452	0.0001
D(Z22)	0.0812	0.0168	4.8386	0.0000
D(Z23)	0.0836	0.0209	4.0077	0.0001
D(Z24)	0.1399	0.0187	7.4758	0.0000
D(Z25)	–0.1377	0.0492	–2.7967	0.0053
D(Z26)	0.0863	0.0178	4.8569	0.0000
D(Z27)	0.0749	0.0175	4.2714	0.0000
D(Z28)	0.0449	0.0206	2.1811	0.0294

Notes: All the variables are defined in first differences (D) of the logarithm (L). GDP = gross domestic product; PGAS = gasoline price; TRAFFIC = average daily traffic volume; TOLL = toll paid per km for the 72 motorway sections; D(Z1)–D(Z28) = the first differences of dummy variables to account for changes in the road network.

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

REVIEWERS for 2003

A

Montasir Abbas	Texas A&M University System, College Station, TX
Saurabh Amin	Texas A&M University System, College Station, TX
William Anderson	Boston University, Boston, MA
Anthony Apostolides	Bureau of Transportation Statistics, USDOT, Washington, DC
Kay Axhausen	Institute for Transport Planning and Systems, Zürich, Switzerland

B

Dipasis Bhadra	MITRE, Bedford, MA
Charles Bowman	Bureau of Labor Statistics, Department of Labor, Washington, DC
Kurt Brannas	University of Umea, Umea, Sweden
Michael Bronzini	George Mason University, Fairfax, VA
Jeremy Broughton	TRL, Berkshire, UK

C

David Chien	Bureau of Transportation Statistics, USDOT, Washington, DC
Stephen Clark	Morley Town Council, Leeds, UK
Timothy Coburn	Abilene Christine University, Abilene, TX

D

Kay Drucker	Bureau of Transportation Statistics, USDOT, Washington, DC
Gary A. Davis	University of Minnesota, Minneapolis, MN

E

David Ellis	Texas A&M University System, College Station, TX

F

Patrick Flanagan	Bureau of Transportation Statistics, USDOT, Washington, DC

G

Marie Gantz	University of Kentucky, Lexington, KY
Richard Golaszewski	Federal Aviation Administration, USDOT, Washington, DC
Tony Gomez-Ibañez	Harvard University, Cambridge, MA
Pilar Gonzales	Universidad del País Vasco, Bilbao, Spain
Jason Goodfriend	Bureau of Transportation Statistics, USDOT, Washington, DC
Michael Greenwald	University of Wisconsin, Milwaukee, WI
Richard Gruberg	Federal Motor Carrier Safety Administration, USDOT, Washington, DC

H

Shauna Hallmark	Iowa State University, Ames, IA
Xiaoli Han	Macrosys Research and Technology, Washington, DC
Susan Handy	University of California at Davis, Davis, CA
David Hensher	University of Sydney, Sydney, Australia
David Hartgen	University of North Carolina at Charlotte, Charlotte, NC
Patricia Hu	Oak Ridge National Laboratory, Oak Ridge, TN

K

Bill Knee	Oakridge National Laboratory, Oakridge, TN
Siem Jan Koopman	Free University of Amsterdam, Amsterdam, Netherlands

L

Kajal Lahiri	University at Albany, Albany, NY
Richard Lawton	University of the West of England, Bristol, England
Janice Lent	Bureau of Transportation Statistics, USDOT, Washington, DC
James P. LeSage	University of Toledo, Toledo, OH
David Levinson	University of Minnesota, Minneapolis, MN

M

Anna Matas	University of Barcelona, Barcelona, Spain
Michael McNerney	DMJM Aviation, Tampa, FL
Nigel Meade	Imperial College of London, London, England
Jeffrey Memmott	Bureau of Transportation Statistics, USDOT, Washington, DC
Paz Moral	Universidad del País Vasco, Bilbao, Spain

N

Nagaraj Neerchal	University of Maryland, Baltimore County, Baltimore, MD
Debbie Neimeier	University of California at Davis, Davis, CA

O

Keith Ord	Georgetown University, Washington, DC

P

Byungkyu (Brian) Park	University of Virginia, Charlottesville, VA
Wayne Perry	George Mason University, Fairfax, VA

R

Robert Raeside	Napier University, Edinburgh, Scotland
Aura Reggiani	University of Bologna, Bologna, Italy

S

Mitsuru Saito	Brigham Young University, Provo, UT
Terry Shelton	Federal Motor Carrier Safety Administration, USDOT, Washington, DC
Caesar Singh	Bureau of Transportation Statistics, USDOT, Washington, DC
Kumares Sinha	Purdue University, West Lafayette, IN
Clifford Spiegelman	Texas A&M University, College Station, TX
Frank Southworth	Oakridge National Laboratory, Oakridge, TN

T

Piyushimita (Vonu) Thakuriah	University of Illinois at Chicago, Chicago, IL
Paul Thomas	Continental Air, Houston, TX
Dawn Tucker	Bureau of Transportation Statistics, USDOT, Washington, DC

W

Othmar Winkler	Georgetown University, Washington, DC

Y

Melinda Yang	Bureau of Automotive Repair, Sacramento, CA
Peg Young	Bureau of Transportation Statistics, USDOT, Washington, DC
Theodore Younglove	University of California at Riverside, Riverside, CA

INDEX FOR VOLUME 6*

* A complete index of all volumes of the journal is available online at www.bts.gov.

BUREAU OF TRANSPORTATION STATISTICS UNITED STATES DEPARTMENT OF TRANSPORTATION

JOURNAL OF TRANSPORTATION AND STATISTICS

Volume 6 Numbers 2/3, 2003
ISSN 1094-8848

CONTENTS