



# JOURNAL OF TRANSPORTATION AND STATISTICS

U.S. Department of Transportation  
Research and Innovative Technology Administration  
Bureau of Transportation Statistics

# JOURNAL OF TRANSPORTATION AND STATISTICS

PEGYOUNG Editor-in-Chief  
CAESAR SINGH Associate Editor  
JEFFERY MEMMOTT Associate Editor  
KAY DRUCKER Associate Editor  
WILLIAM MOORE Managing Editor  
MARSHA FENN Managing Editor through 3/17/06  
JENNIFER BRADY Data Review Editor  
VINCENTYAO Book Review Editor  
DORINDA EDMONDSON Desktop Publisher  
ALPHA GLASS WINGFIELD Editorial Assistant

## EDITORIAL BOARD

DAVID BANKS Duke University  
KENNETH BUTTON George Mason University  
TIMOTHY COBURN Abilene Christian University  
STEPHEN FIENBERG Carnegie Mellon University  
GENEVIEVE GIULIANO University of Southern California  
JOSE GOMEZ-IBANEZ Harvard University  
DAVID GREENE Oak Ridge National Laboratory  
MARK HANSEN University of California at Berkeley  
KINGSLEY HAYNES George Mason University  
DAVID HENSHER University of Sydney  
PATRICIA HU Oak Ridge National Laboratory  
T.R. LAKSHMANAN Boston University  
TIMOTHY LOMAX Texas Transportation Institute  
PETER NIJKAMP Free University  
KEITH ORD Georgetown University  
ALAN PISARSKI Consultant  
ROBERT RAESIDE Napier University  
JEROME SACKS National Institute of Statistical Sciences  
TERRY SHELTON U.S. Department of Transportation  
KUMARES SINHA Purdue University  
CLIFFORD SPIEGELMAN Texas A&M University  
PETER STOPHER University of Sydney  
PIYUSHIMITA (VONU)THAKURIAH University of Illinois at Chicago  
DARRENTIMOTHY U.S. Department of Transportation  
MARTIN WACHS RAND Corporation  
C. MICHAEL WALTON The University of Texas at Austin  
SIMON WASHINGTON Arizona State University  
JOHN V. WELLS U.S. Department of Transportation

The views presented in the articles in this journal are those of the authors and not necessarily the views of the Bureau of Transportation Statistics. All material contained in this journal is in the public domain and may be used and reprinted without special permission; citation as to source is required.

**A PEER-REVIEWED JOURNAL**

**JOURNAL OF TRANSPORTATION  
AND STATISTICS**

**Volume 9 Number 1, 2006  
ISSN 1094-8848**

U.S. Department of Transportation  
Research and Innovative Technology Administration  
Bureau of Transportation Statistics

**U.S. Department of  
Transportation**

NORMAN Y. MINETA  
Secretary

MARIA CINO  
Deputy Secretary

**Research and Innovative  
Technology Administration**

ASHOK G. KAVEESHWAR  
Administrator

ERIC C. PETERSON  
Deputy Administrator

**Bureau of Transportation  
Statistics**

WILLIAM M. BANNISTER  
Acting Deputy Director

**The *Journal of Transportation and Statistics* is  
published by the**

Bureau of Transportation Statistics  
Research and Innovative Technology Administration  
U.S. Department of Transportation  
400 7th Street SW, Room 4117  
Washington, DC 20590  
USA  
journal@bts.gov

**Subscription information**

The *Journal of Transportation and Statistics* is on hiatus and is currently out of production; to receive copies of available back issues:

mail Product Orders  
Bureau of Transportation Statistics  
Research and Innovative Technology Administration  
U.S. Department of Transportation  
400 7th Street SW, Room 4117  
Washington, DC 20590  
USA

phone 202.366.DATA  
internet www.bts.dot.gov

**Information Service**

email answers@bts.gov  
phone 800.853.1351

Cover design Alpha E. G. Wingfield  
Cover photo Michael C. Fries

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.

**Contents**

**Letter from the Editor-in-Chief** ..... v

**Papers in This Issue**

Impacts of Productivity Changes in Air Transportation on Profits, Prices, and Labor Compensation: 1990–2001  
*Anthony D. Apostolides* ..... 1

Speed as a Risk Factor in Serious Run-off-Road Crashes: Bayesian Case-Control Analysis with  
Case Speed Uncertainty  
*Gary A. Davis, Sujay Davuluri, and Jianping Pei* ..... 17

Speed Estimation for Air Quality Analysis  
*Huafeng Gong, Mei Chen, Jesse Mayes, and Rob Bostrom* ..... 29

Measurement Errors in Poisson Regressions: A Simulation Study Based on Travel Frequency Data  
*Erling Häggström Lundevaller* ..... 39

Measuring Variability in Urban Traffic Flow by Use of Principal Component Analysis  
*Theodore Tsekeris and Antony Stathopoulos* ..... 49

Frequency and Severity of Belgian Road Traffic Accidents Studied by State-Space Methods  
*Elke Hermans, Geert Wets, and Filip Van Den Bossche* ..... 63

**Reviewers for 2006** ..... 77

**Index for Volume 9** ..... 79



# Letter from the Editor-in-Chief

*Our revels now are ended.*

William Shakespeare, THE TEMPEST

Dear Readers,

To those of you who have not heard the news, this issue marks the start of a hiatus for the *Journal of Transportation and Statistics*. What this means is that, because of budget cuts, we do not expect to receive funding to publish the journal during the next few years. The value of JT&S was never the question. Senior management of the Research and Innovative Technology Administration understood the contribution made by all involved in this work; however, difficult decisions had to be made and funding was only available for core projects.

I would like to thank everyone who participated in this venture, particularly the authors from around the world who, through these nine years, provided excellent research in the field of statistical and economic analysis in transportation. JT&S was able to showcase new and innovative techniques to measure aspects of transportation and present data and other information not available elsewhere. The reviewers of the articles strengthened the quality of work published. Our distinguished Editorial Board provided support, insight, and valuable advice that led to the continual improvement of the journal. Our dedicated Associate Editors (David Chien, Caesar Singh, Jeffery Memmott, and Kay Drucker) worked hard to ensure the quality of the published articles. I would also like to thank Jennifer Brady, our Data Review Editor, for highlighting new data from the Bureau of Transportation Statistics, and Vincent Yao, our Book Review Editor.

The publishing staff, which over the years included Marsha Fenn, Dorinda Edmondson, William Moore, Alpha Glass, Martha Courtney, Lorisa Smith, Susan Hoffmeyer, Deborah Moore, and Darcy Herman, produced a professional publication that rivaled its counterparts in the academic environment. Special thanks go to Marsha Fenn, the Managing Editor, who was an integral part of the journal from its inception. And, finally, the quality standards established by the previous Editors-in-Chief, David L. Greene, David Banks, and John V. Wells, made my work a simpler task.

Finally, let me thank you, the readers, for without you the journal would have been pointless. I hope you found the articles to be informative, engaging, and interesting. Thank you for your support, and I hope our paths will cross in the future.

A handwritten signature in black ink that reads "Peg Young". The signature is written in a cursive, flowing style with a long horizontal flourish extending to the right.

**PEG YOUNG, Ph.D.**

Editor-in-Chief, JT&S



# Impacts of Productivity Changes in Air Transportation on Profits, Prices, and Labor Compensation: 1990–2001

ANTHONY D. APOSTOLIDES

---

Bureau of Transportation Statistics  
Research and Innovative Technology Administration  
U.S. Department of Transportation  
Washington, DC 20590  
anthony.apostolides@dot.gov

---

## ABSTRACT

The objective of the paper is to assess the impacts of productivity changes in air transportation since 1990 in three basic economic areas: 1) industry profits, 2) consumers/users of air transportation services, and 3) industry employees. In this regard, an analysis is initially carried out between productivity measures and industry profits. Comparisons are also made between the general price level of the economy and several price indexes of the air transportation subsector. Also, an evaluation is conducted of labor compensation in air transportation, the U.S. economy, and other transportation industries. The analysis results in several findings. First, there is a marked association between productivity changes in air transportation and industry profits. Second, the benefit of productivity increases in air transportation does not seem to have transferred to consumers of air passenger services in the form of lower prices. On the other hand, users of scheduled cargo services did seem to benefit from lower prices. Finally, a portion of the benefit of productivity increases went to industry labor in the form of relatively high levels of labor compensation.

---

KEYWORDS: Air transportation, productivity, profits, prices, labor compensation.

## INTRODUCTION

The objective of this paper is to measure productivity and assess the impact of productivity increases in air transportation in three economic areas: 1) industry profits, 2) consumers, and 3) airline employees.

According to economic theory, changes in productivity in an industry (firm) can affect profits, prices, and labor compensation. Increases in productivity are expected to result in higher profits for the industry. Subsequently, there can be positive impacts on consumers and on the employees of the industry.

In elaborating on the above theoretical framework, the basic benefit of increased productivity is that more output can be produced with the same quantity of inputs (some inputs can be of improved quality). Alternatively, the same output can be produced with fewer resources. Other things being equal, this results in a bigger difference between total revenues and total costs, and thus higher profits for the industry. The existence of higher profits can subsequently be followed by three effects:

1. the firms in the industry can keep a portion of the increased profits for internal use;
2. the firms can decrease prices for their service to the consumers, or—perhaps more likely—they may increase prices less than they would in the absence of productivity increases; and
3. the firms can provide higher compensation to their employees (in the form of higher wages and/or fringe benefits).

The assessment of this paper applies this theoretical framework to the air transportation subsector. Greater profits benefit the air carriers directly. With regard to users, a decrease in prices for passengers increases their real incomes. For shippers of air freight, a decrease in prices reduces their (distribution) costs. In addition, higher profits resulting in increased labor compensation for airline employees raises real incomes. Such increases in real incomes, to consumers and labor, are the important contributions of greater productivity. Increases in real incomes lead to more consumption, which contributes to the economy's growth.

High increases of productivity imply a higher likelihood that the above effects would occur. A decline in productivity could reverse the positive

effects of a productivity increase, resulting in declines in labor wages and, in extreme cases, bankruptcies of companies, accompanied by job losses.

## Data and Period of Analysis

The paper uses a consistent set of the most recent data available—for the 1990 to 2001 period. These data refer to the main variables needed for the industry analysis: productivity (labor and multifactor), profits, prices (various types), and labor compensation. Additional data used relate to the U.S. business sectors and the U.S. economy.

Industry data used in this paper are classified under the North America Industry Classification System (NAICS). Labor productivity is examined for three transportation industries/subsectors: air transportation data refer to NAICS industry number 481, line-haul railroads refer to NAICS 48211, and general freight trucking long-distance refer to NAICS 48412. Comparisons are also made with the U.S. business sector. The words “industry” and “subsector” are used interchangeably in the paper.

## LABOR AND MULTIFACTOR PRODUCTIVITY

This section examines changes in labor and multifactor productivity in the U.S. air transportation subsector during the 1990 to 2001 period. It also examines data on productivity of the U.S. economy and the two other transportation subsectors—railroads and trucking.

Labor productivity is defined as output per unit of labor and is calculated by dividing output by a measure of labor input used in the production of the output. For air transportation, output is measured in terms of passenger-miles and ton-miles; and for rail and trucking, output is measured in terms of ton-miles. Labor productivity can be affected by factors that include improved labor skills and training as well as by physical capital per worker.

Multifactor productivity relates to the productivity of all the inputs used in the production process. These include labor, capital (with land), and intermediate inputs. Multifactor productivity is a more comprehensive measure of productivity than labor productivity. It indicates the overall production efficiency of an industry as it relates to increases in

industry output that are not accounted for by increases in the factor inputs. The analysis of the specific impacts, or potential benefits, of productivity increases in an industry is the basic objective of this study.

To evaluate labor productivity in air transportation, data on levels of labor productivity in that industry, over time, are plotted in figure 1. These data indicate that labor productivity increased from 1990 until 1997, when it reached its peak. In 1998, labor productivity declined and stayed at this lower level until 2000. In 2001, it declined again, quite significantly. This was affected by the drop in output/demand as a result of the catastrophic events of September 11, 2001 (9/11), and by a recession in that year.

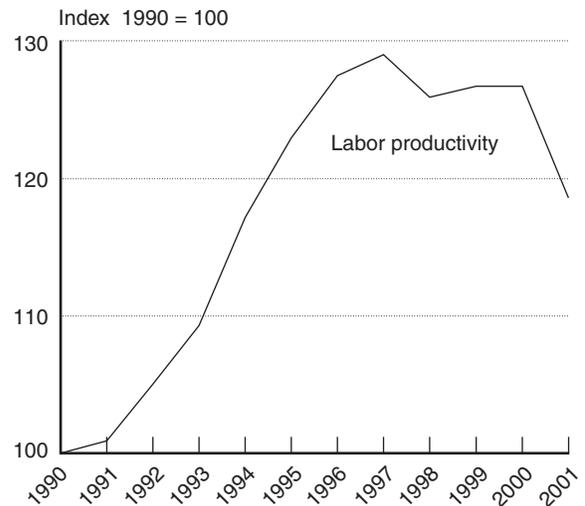
To compare labor productivity in air transportation with the other two transportation industries and the U.S. business sector, relevant data are plotted in figure 2. There, one can observe that between 1990 and 2000 (and with the exception of 1991 to 1993), labor productivity in air transportation increased faster than in long-distance trucking and the U.S. business sector. In 2001, however, labor productivity in air transportation declined while that of the U.S. business sector increased.

Rail transportation was the one subsector in which labor productivity increased faster than labor productivity in air transportation. Rail transportation had continual increases in labor productivity over time. In fact, labor productivity in this subsector continued to increase in 2001 even as it declined in air transportation and trucking.

In order to make comparisons from another perspective, growth rates of labor productivity are presented in table 1. These growth rates show that, over the 1990 to 2000 period, labor productivity in air transportation grew at a higher annual rate (2.4%) than it did in the U.S. business sector (2.0%) and in trucking (1.7%).

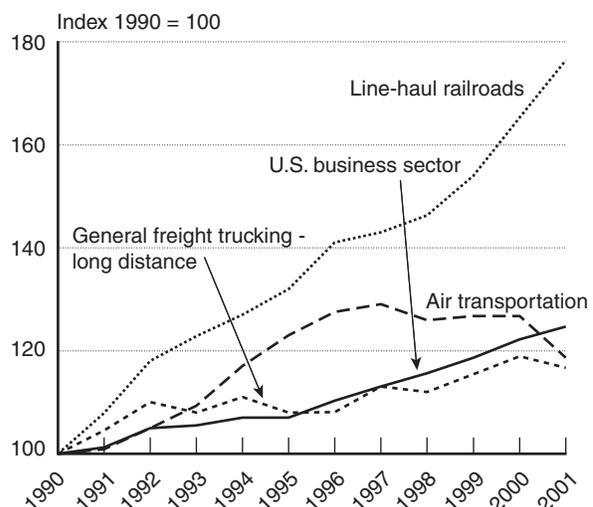
Between 1990 and 2001, however, the growth rate for air transportation was lower (1.6%) than that of the U.S. business sector (2%) and just above trucking (1.4%). These data also indicate a significant drop in the annual growth rate of labor productivity in air transportation between 1990 and 2000 (2.4%) and 1990 and 2001 (1.6%). This sud-

**FIGURE 1 Labor Productivity in Air Transportation**



Source: The data on which this chart is based were obtained from Bureau of Labor Statistics internet site, section on Productivity, subsection on Productivity and Costs.

**FIGURE 2 Labor Productivity in Transportation and the U.S. Business Sector**



Source: The data on which this chart is based were obtained from Bureau of Labor Statistics internet site, section on Productivity, subsection on Productivity and Costs.

den drop when 2001 data are included reflects the significant impact of 9/11 on this subsector. After that date, output of air transportation dropped immediately and significantly while the labor force in air transportation also declined, but with a time lag. In both time periods, rail transportation experienced the highest growth rate of labor productivity. Air transportation productivity, in 2001, was affected more adversely than productivity in the U.S. economy and in the trucking industry, and significantly more adversely than in the railroad indus-

**TABLE 1 Growth Rates of Labor Productivity in Transportation**  
(Growth rates—average annual percentage rate)

	1990– 2000	1990– 2001	1990– 1995	1995– 2000	1995– 2001
Air transportation	2.4	1.6	4.2	0.6	–0.6
Line-haul railroads	5.2	5.3	5.7	4.6	5.0
General freight trucking— long distance	1.7	1.4	1.5	2.0	1.3
U.S. business sector	2.0	2.0	1.4	2.6	2.5

Source: The data on which these growth rates are based were obtained from the Bureau of Labor Statistics Internet site, section on Productivity, subsection on Productivity and Costs

try. There was a recession in 2001 that affected the economy and output of industries; air transportation would seem to have been particularly affected by the events of 9/11. Nevertheless, labor productivity in air transportation increased significantly over the analysis period.

### Multifactor Productivity

With regard to multifactor productivity (MFP), the plots presented in figure 3 show that MFP in air transportation was at higher levels than that of the U.S. business sector over the period of analysis, indicating higher growth rates. Over 1990 to 2000, multifactor productivity in air transportation grew at an annual rate of 1.9% while in the U.S. private business sector it grew at an annual rate of 0.9% (appendix table 1).

These data indicate that, over 1990 to 2000, both labor and multifactor productivity in air transportation generally increased. The same observation applies to the 1990 to 2001 period, with the qualifications noted. The paper proceeds to assess the impacts of this productivity increase in the three areas mentioned previously— profits of air carriers, prices paid by users, and labor compensation of airline employees.

### PRODUCTIVITY AND PROFITABILITY

The basic equation illustrating the calculation of profit of an enterprise is:

$$\text{Profit} = \text{total revenues} - \text{total costs.}$$

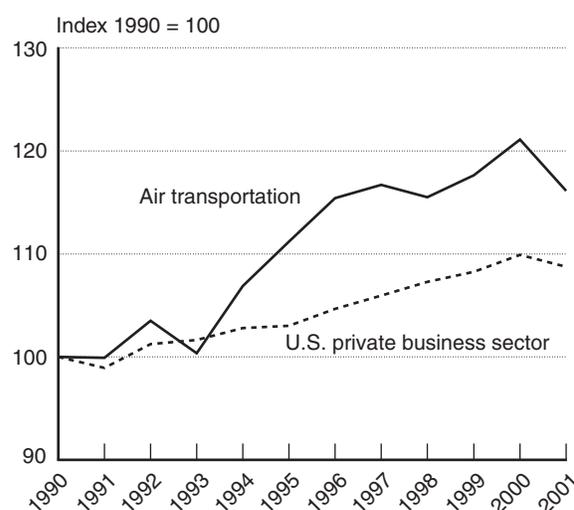
Total revenues consist of the quantity of items sold multiplied by the price per item. In air transportation, the items would relate to tickets for passengers or tons-miles of freight. Total costs are composed of fixed and variable costs. For air carriers, fixed costs

would include the periodic payments made for the purchase of an airplane, while variable costs would include fuel and labor costs.

The basic source for data on profits in air transportation (net income after taxes) is the Bureau of Transportation Statistics (BTS), Office of Airline Information (OAI). These data can be obtained from TranStats, a database on the BTS Internet site that provides data on net income for various sizes of airlines (Majors, National, Regionals, and Small). Also, the *Airline Quarterly Financial Review*, by the Office of the Secretary of Transportation, presents profit data for Major air carriers.

Table 2 presents annual data on productivity and profits in air transportation for the analysis period. These data indicate that, particularly since 1995, operations in the air transportation industry resulted in profits that were maintained over time, up to year

**FIGURE 3 Multifactor Productivity in Air Transportation and U.S. Business Sector**



Source: Data on which this chart is based were obtained from Bureau of Labor Statistics Internet site, section on Productivity, subsection on Multifactor Productivity.

**TABLE 2 Productivity and Profits in Air Transportation**

Year	Labor productivity	Multifactor productivity	Net income (\$millions, all carriers)	Net income (\$millions, Majors)
1992	105.0			
1993	109.3	100.4	\$272	
1994	117.2	106.9	(\$344)	(\$578)
1995	123.0	111.2	\$2,340	\$2,235
1996	127.5	115.4	\$2,804	\$2,779
1997	129.0	116.7	\$5,168	\$5,488
1998	125.9	115.5	\$4,531	\$4,577
1999	126.7	117.6	\$5,357	\$5,075
2000	126.7	121.1	\$2,533	\$2,599
2001	118.6	116.1	(\$8,171)	(\$7,139)

Note: Numbers in parentheses (in columns 3 and 4) indicate losses.

Sources: For data in columns 1 and 2, same as chart 1 and 2. For data in column 3, from BTS, TranStats (on the Internet); data refer to all air carriers. For data in column 4, from DOT, Airline Quarterly Financial Review; data refer to Majors.

2000. In 2000, industry profits declined although they were still positive. In 2001, industry profits became negative; they were affected significantly by the events of 9/11, which suddenly reduced demand for air travel, and the industry was not able to reduce costs proportionately.

The data in this table indicate that there is an association between increases in the productivity measures and profits of the air transportation subsector. From 1995 to 1997, industry productivity (labor and MFP) increased, and industry profits increased. One also notes that 1997 was the year in which air transportation earned the highest amount of profits, and in that year the industry experienced the highest level of labor productivity. During 1998, productivity (labor and MFP) decreased and profits decreased; and during 1999, productivity increased and profits increased. Finally, during 2001, productivity decreased and profits decreased, affected by the events of 9/11. On the other hand, MFP increased in 1994 and 2000, but industry profit declined during these years. Overall, these data are consistent with economic theory predicting a relationship between productivity and profits.

In order to quantify the association between profits and productivity, Spearman rank correlation coefficients were calculated, and they are presented in appendix table 2. An asterisk next to the coefficient indicates significance at the 95% level, given the number of observations (Kvanli 1988, chapter 4). All four coefficients relating to labor productiv-

ity and profits indicate a positive and significant association between the two variables at the 95% level. These coefficients range from 0.77 to 0.83, which shows a substantial association between the two variables.

The rank correlation coefficient is lower between multifactor productivity and profits. One of the four coefficients calculated between these variables is significant at the 95% level (0.74). The calculation of this coefficient does not include data for 2001. The rank correlation between MFP and profits of Majors (OST data) for 1994 to 2001 is 0.45; when data for 2001 are dropped, the coefficient increases to 0.61.

Therefore, visual observation and correlation coefficients indicate a rather marked association between productivity and profits in air transportation.<sup>1</sup> This substantiates and is consistent with economic theory, which predicts the basic benefit of

<sup>1</sup> In addition, regression analysis was used to estimate the relationship between profits (dependent variable) and productivity (independent variable). Better results were obtained when data for year 2001 were dropped. The estimated equation using profits from TranStats and labor productivity is:

$$\begin{aligned} \text{Profits (OAI)} &= f(\text{Labor Productivity}) \quad n = 8 \\ \text{Profits} &= -28,845.11 + 257.20 \text{ Labor Productivity} \\ &\quad (9,425.23) \quad (76.43) \\ t\text{-statistic} &\quad -3.06 \quad 3.37 \\ \text{Adjusted R-squared} &= 0.60 \\ \text{Durbin-Watson} &= 1.93 \end{aligned}$$

productivity increases (in an industry) is a rise in profits.

## **PRODUCTIVITY AND PRICES**

Rising profits of an industry can impact industry prices. If the price of air transportation were to decrease, or increase by slower rates, as a result of greater productivity, then the users of air transportation services would benefit. Lower prices for consumers/passengers increase the purchasing power of consumers' incomes, that is, increase their real income and thus their standard of living. Lower prices for producers of goods that use air transportation services for freight shipments would contribute to lower costs for these producers, and thus higher profits. The occurrence and extent of lower prices—or slower growth of prices—as a result of productivity increases are more likely in industries characterized by a relatively high degree of competition.

In order to evaluate the relationship between productivity and prices in air transportation, an initial comparison is made of price changes in that industry, over time, with price changes in the general economy. The objective is to assess whether greater productivity in air transportation was accompanied by relatively small price increases, or price declines, compared with prices in the general economy. If that occurred, there would be indications that a portion of the benefit from productivity increases (higher profits) went to consumers/users of air transportation services.

Table 3 presents relevant price data for the economy and air transportation. Prices in the general economy are measured by the Consumer Price Index (CPI), while the prices of air transportation services are measured by several price indexes to cover the various segments of the industry. These segments are consumers/passengers and entities using air cargo services. One price index is the CPI for air transportation (CPI-AT), which measures the prices that consumers pay for air transportation services (column 5). This index includes domestic and international air travel. Data are also presented for three other price indexes: the Producer Price Index (PPI) for scheduled passenger service—domestic and international (in column 7 of the table); the PPI for scheduled passenger service—domestic (column 9);

and the PPI for scheduled air cargo (column 11). Growth rates of prices are computed in the columns next to the indexes.

The Bureau of Labor Statistics (BLS) publishes the CPI for the economy and both a CPI and a PPI for airfares. The CPI for commercial air travel is based on prices listed by the airlines in the SABRE system, a reservation system used by many travel agencies. This index measures changes in the prices paid by consumers for airline trips, including taxes and any distribution costs not received by the air carriers, such as travel agents' fees. The PPI-Air Travel measures changes in revenues received by producers of airline trips.

The CPI-Air Travel includes trips purchased from foreign carriers while the PPI-Air Travel excludes these. Monthly prices for the two programs are gathered from different data sources: CPI prices come from the SABRE system, while PPI prices are gathered directly from airline pricing departments.

The data in table 3 show that, since 1990, prices of air transportation for scheduled passenger service increased significantly faster than the CPI of the economy. Moreover, prices of domestic passenger service increased substantially faster than prices of domestic and international services, combined. The data indicate that while the CPI rose by 36% over the 1990 to 2001 period, prices of passenger service increased by 61% based on the CPI-airline fare, which includes domestic and international air travel; by 81% for PPI-domestic and international; and by 101% for PPI-domestic service.

On the other hand, prices of scheduled air cargo increased by a substantially lower percentage than the general price level. These prices rose by less than 10% over the period of analysis, compared with CPI growth of 36%. Consequently, prices of air cargo also increased by a significantly lower percentage than prices in the passenger segment of the air transportation industry.

These data indicate that although productivity and profits went up in air transportation, prices for passenger service also tended to increase at relatively high rates. In this segment of the industry, the providers of transportation services appear to have kept that part of the benefit of productivity increases.

**TABLE 3 Productivity and Prices in Air Transportation**

Year	Labor productivity (1)	Multifactor productivity (2)	Consumer Price Index—U.S. (1982-84=100) (3)	Growth rate—CPI (4)	Consumer Price Index—airline fare (1982-84=100) (5)	Growth rate—airline fare (6)	PPI-air transportation, scheduled passenger-domestic, international (Dec.1989=100) (7)	Growth rate—PPI, scheduled passenger-domestic, international (8)	PPI-air transportation, scheduled passenger-domestic (Dec.1989=100) (9)	Growth rate—PPI, scheduled passenger-domestic (10)	Scheduled air cargo (Dec.1987=100) (11)	Growth rate—PPI, scheduled air cargo (12)
1990	100.0	100.0	130.7		148.4		110.6		111.3		102.0	
1991	100.9	99.9	136.2	4.2	155.2	4.6	122.4	10.7	125.1	12.4	105.2	3.1
1992	105.0	103.5	140.3	3.0	155.2	0.0	114.8	-6.2	115.4	-7.8	107.0	1.7
1993	109.3	100.4	144.5	3.0	178.7	15.1	126.8	10.5	131.2	13.7	112.1	4.8
1994	117.2	106.9	148.2	2.6	185.5	3.8	130.6	3.0	136.4	4.0	109.5	-2.3
1995	123.0	111.2	152.4	2.8	189.7	2.3	137.8	5.5	144.8	6.2	111.2	1.6
1996	127.5	115.4	156.9	3.0	192.5	1.5	148.1	7.5	160.1	10.6	108.7	-2.2
1997	129.0	116.7	160.5	2.3	199.2	3.5	153.9	3.9	167.9	4.9	107.3	-1.3
1998	125.9	115.5	163.0	1.6	205.3	3.1	152.6	-0.8	165.2	-1.6	104.8	-2.3
1999	126.7	117.6	166.6	2.2	218.8	6.6	161.2	5.6	174.2	5.4	106.9	2.0
2000	126.7	121.1	172.2	3.4	239.4	9.4	186.5	15.7	208.1	19.5	110.4	3.3
2001	118.6	116.1	177.1	2.8	239.4	0.0	200.6	7.6	223.8	7.5	112.0	1.5
Percent increase												
1990-2001	18.6	16.1	35.5		61.3		81.4		101.1		9.8	

Key: CPI = Consumer Price Index, PPI = Producer Price Index.

Sources: Data for columns 1 and 2, same as Figures 1 and 2. Data for columns 3 and 5, from BLS website, Consumer Price Index. Data for columns 7, 9, and 11, from BLS website, Producer Price Indexes. Column 5 includes domestic and international travel.

Note: Data for CPI refer to "All Urban Consumers". Data in column 11 are based on Standard Industrial Classification.

On the other hand, it appears that the providers of scheduled air cargo services returned a portion of the benefit of rising productivity to users of these services in the form of relatively lower price hikes. Prices in this segment of the airline industry increased significantly less than the CPI of the economy or the industry passenger segment. In fact, several times during the period of analysis, there occurred price decreases in scheduled air cargo services.

In explaining price changes in the passenger and cargo segments of the air transportation subsector over time, one notes that in the case of passenger service it is the consumers/passengers (typically individually) who are dealing with the providers of air services (air carriers). The individual consumers do not possess much market power with which to negotiate prices for the services they buy—although in recent periods, the Internet has provided more information on ticket prices.

The air transportation industry would be characterized as an oligopoly in the national market or regional markets. Also, a number of mergers and acquisitions in the industry in the 1980s and 1990s resulted in a substantially smaller number of domestic air carriers. According to economic theory, the fewer the number of sellers in an industry, the lower the degree of competition to affect restraints in price increases. This seems to apply to the passenger segment of air transportation.

On the other hand, the purchasers of scheduled air cargo services tend to be business enterprises, often of substantial size, that typically have good information on the available prices for these services. They also tend to provide substantial and repeat business to the providers of air cargo services. Therefore, these enterprises can have substantial market power to use in obtaining advantageous prices for freight transportation services.

Recently, BTS began calculating its own Air Travel Price Index (ATPI).<sup>2</sup> This index measures prices actually paid by passengers rather than prices published in airline price schedules. Data are presented in appendix table 3 to enable comparisons between the ATPI and other price indexes from BLS for 1995 to 2001. These calculations show two results:

1. The ATPI increased significantly less than the CPI, the CPI-airfare, or the PPI-Air Transportation. A recent article that compared the U.S.-Origin ATPI with the BLS Air Travel Index found a significant difference between their increases. The authors stated that this was probably due mainly to: 1) the different methodologies/formulas used in the creation of the indexes, and 2) the ATPI's inclusion of special discount fares (Lent and Dorfman 2005).

These price changes shown by the ATPI indicate that the benefits of productivity increases also accrued to the consumers of air transportation. This is different than the results based on BLS price indexes. This may be a topic for future research.

2. Within the ATPI, the ATPI U.S.-Origin increased substantially more than the ATPI-Foreign, which actually declined. Both sets of price indexes—the CPI-Air and PPIs from BLS, and the ATPI—indicate that prices of domestic air transportation increased faster than prices of international air travel. In attempting to explain such differences, one notes that typically increasing prices would be affected by increasing production costs or by the degree of competition in the industry. The production costs of domestic and international travel would not be expected to diverge significantly over time. The other factor is the degree of competition. Available information indicates that the degree of competition in domestic air transportation has decreased in the domestic market over the period of analysis.

In this regard, a study by U.S. Department of Transportation (DOT 1999), which covered 1992 to 1997, stated a number of findings indicating such a situation. These findings include:

---

<sup>2</sup> There are three primary ATPI series. The U.S. Origin ATPI measures changes in the cost of itineraries originating in the United States, whether the destinations are domestic or international. The Foreign-Origin ATPI measures changes in the cost of itineraries within a foreign origin and a U.S. destination. The Full-Scope ATPI combines the domestic and foreign-origin itineraries. The CPI-Air Fare and ATPI both cover domestic and international travel. However, the CPI is U.S.-Origin only; thus, it is more limited in scope than the Full-Scope ATPI.

1. In short-distance markets without low-fare competition, inflation adjusted fares are substantially higher (26%), not lower, than prederegulation fares. These markets account for about one-fourth of total domestic passengers.
2. There was a reversal of growth in low-fare competition in the last year of the period—1997. Markets with low-fare competition have significantly lower fares on average—often less than one-half—than similar markets without such competition.
3. One observes high fares in short distance markets at hub airports where one major network airline has a dominant market share. Average fares at some of these airports can be 50% to 60% higher when compared with more competitive markets.
4. New entrants in the airline industry experienced difficulties that can contribute to a decrease in low-fare competition. A number of factors make it difficult for new airlines to enter a hub market. These factors include: higher frequency service made available by hub-and-spoke systems; frequent flyer programs; travel agent commissions bonuses (overrides); and lack of gates and ticket counter or takeoff and landing slots for new competitors at certain airports.

The DOT study “concluded that unfair exclusionary practices have been a key reason that competition from new low-fare carriers has not been able to penetrate concentrated hubs...” (USDOT 1999, p.8). In addition, another study assessed predatory pricing in air transportation (Oster and Strong 2001). This study found that the early years of airline deregulation were characterized by periods of significant competition among the major established airlines as well as by competition from new-entrant carriers and from carriers formerly confined to intrastate markets. However, in the mid-to late-1980s, considerable industry consolidation occurred as a result of a wave of mergers. A number of these mergers involved the acquisition of larger carriers such as Frontier, Republic, Eastern, Ozark, Western, and Piedmont.

Following these mergers, the source of deregulation’s benefits began to change. The benefits gradu-

ally became more attributable to the actions of a small number of low-fare carriers rather than to the actions of major network airlines. By the late 1990s, the domestic route networks of major airlines had become fairly stable and were built around hub airports, typically dominated by a single carrier. These hub-based networks established geographic areas in which each major network airline has substantial presence and market power, especially in short-haul smaller markets.

Some of the responses of the incumbent network carriers to entry by low-fare carriers resulted in concerns, by government and others, about the use of predatory pricing or unfair methods of competition. In one example (described in more detail in appendix 4), after a new, low-fare airline entered a particular market, the major network carrier responded by adding more flights on the entrant’s network, by offering bonus miles, by offering special agent commission overrides,<sup>3</sup> and by matching the fares of the entrant in that particular market. As a result, within one month after the entrant began service, losses forced it to reduce its service to one flight a day, and soon thereafter, it exited the market altogether.

The study also examined 12 cases during the 1994 to 1997 period that involved short- to medium-haul flights and entailed a major network carrier hub, at one or both ends, and a new entrant (Oster and Strong 2001, p.10). The main features of the cases are described in appendix 5. The results include the lowering of average fares by the major carriers after the new entry, the exiting of the entrants, and the subsequent increase in fares by the major carriers. The authors of the study point out that predatory practices may be a rational strategy in the airline industry because short-run revenue losses may be recouped in the longer term. Such aggressive responses by major network incumbents to new entry can drive entrants from specific routes. Moreover, they provide a signal to other prospective

---

<sup>3</sup> Travel Agent Commission Overrides (TACOS) are special bonus commissions paid by an airline to travel agents as a reward for booking a targeted proportion or number of passengers on that airline. Such overrides, of which travelers are typically not aware, provide incentives to travel agents to steer some travelers from one airline to another. These overrides can also serve as a barrier to entry.

entrants that despite high fares being charged in a number of markets, any new entry will be met with a response that renders unprofitable the entrant's operation. This results in barriers to entry that can contribute to higher prices.

From the perspective of the consumers, there have been complaints by the Consumer Federation of America (CFA) with regard to competition and prices in air transportation. In testimony to Congress, Mark N. Cooper, the Director of Research of CFA, pointed out that 25 states filed comments in support of the DOT's antipredation rule that identified 15 airports at which the dominant firm had a market share in excess of 70%. Another half dozen airports had a dominant carrier, with 50% to 70% market share (Cooper 2001).

Mr. Cooper noted that airline markets are generally highly concentrated, and most routes have fewer than four carriers. He pointed to one study which found that, measured at the airports, the Hirschman-Herfindahl Index (HHI) was just under 3,300; this is equivalent to three airlines per airport. However, when measured by city pairs, the HHI was over 5,000—the equivalent of 2 airlines per route.<sup>4</sup> He noted that because there is a high level of concentration, one should not be surprised to find that anticompetitive behavior and changes in market structure have a significant impact on fares. Exercising market power is easy in such highly concentrated markets.

With regard to competition in the international market, a DOT study found that as transatlantic deregulation unfolds, competition intensifies and provides price benefits to consumers. This was apparently affected by open skies bilateral agreements that have provided carriers the operating flexibility necessary to improve and expand services. This new flexibility for carriers to respond to marketplace demands has led to downward pressure on

---

<sup>4</sup> Ibid. Cooper 2001. The HHI is calculated by expressing the market share of each firm in the industry as a percentage, squaring these figures, and adding them. For example, if in an industry, two firms control 50% of the market each, the index would be  $50^2 + 50^2 = 2500 + 2500 = 5000$ . For an industry in which each of four firms controls 25% percent of the market, the HHI would be:  $25^2 + 25^2 + 25^2 + 25^2 = 625 + 625 + 625 + 625 = 2500$ . Source: Case and Fair, 1994, pp. 378.

prices, due both to increased supply and increased competition (USDOT OST 2000, p.2).

Data for 1996 to 1999 show decreases in price fares in international air travel. During this time period, average fares (not adjusted for inflation) to open-sky countries declined by 20% (compared with 1996). Moreover, they decreased by various percentages that approached 15% in connecting markets beyond European gateways (USDOT OST 2000, p.3).

In summary, we can see that in using BLS price data, prices of passenger service rose higher than the CPI. Thus, it would appear that the air carriers kept that part of the benefit of the productivity increase. On the other hand, prices of air cargo services increased relatively slowly. Thus, the users of these services were able to benefit from greater productivity in the industry.

Within the passenger segment of the air transportation industry, price data indicate that prices for domestic air transportation services rose faster than for international air transportation. This seems to be consistent with studies that indicated a trend toward decreased competition in the domestic market segment, resulting from increased concentration in the industry and predatory pricing behavior of network carriers toward low-cost entrants. In the international segment, a government study showed prices to have declined during several years in the decade of the 1990s.

## **PRODUCTIVITY AND LABOR COMPENSATION**

The other potential effect of increasing productivity in an industry (or firm) is for a portion of the benefit to go to the employees in the form of higher labor compensation (wages and fringe benefits). In order to evaluate this possibility for air transportation, data are presented in table 4 on compensation per worker for that industry and for the U.S. economy (average labor compensation for all civilian workers), as well as for line-haul railroads and general freight trucking. These data are in current and constant dollars. In current dollars, they indicate that labor compensation in air transportation grew relatively faster, over time, than in the overall economy and in the two other transportation industries.

**TABLE 4 Labor Compensation per Employee**  
(In current and constant dollars)

Year	In current dollars				CPI (1982–84=100)	In constant dollars			
	Air transportation	Line-haul railroads	General freight trucking, long distance	U.S. economy		Air transportation	Line-haul railroads	General freight trucking, long distance	U.S. economy
1990	\$47,815	\$50,236	\$30,092		130.7	\$36,584	\$38,436	\$23,024	
1991	\$49,799	\$51,947	\$30,605	\$34,216	136.2	\$36,564	\$38,140	\$22,471	\$25,122
1992	\$52,084	\$52,461	\$32,482	\$35,922	140.3	\$37,123	\$37,392	\$23,151	\$25,603
1993	\$53,844	\$54,322	\$31,746	\$37,190	144.5	\$37,263	\$37,593	\$21,970	\$25,737
1994	\$56,286	\$56,256	\$32,738	\$38,064	148.2	\$37,980	\$37,959	\$22,090	\$25,684
1995	\$58,485	\$58,439	\$32,961	\$37,877	152.4	\$38,376	\$38,346	\$21,628	\$24,854
1996	\$59,419	\$60,701	\$32,358	\$38,854	156.9	\$37,871	\$38,688	\$20,623	\$24,764
1997	\$60,742	\$62,259	\$34,882	\$39,978	160.5	\$37,846	\$38,791	\$21,733	\$24,908
1998	\$61,350	\$64,328	\$35,842	\$41,101	163.0	\$37,638	\$39,465	\$21,989	\$25,215
1999	\$62,771	\$64,700	\$37,196	\$42,203	166.6	\$37,678	\$38,835	\$22,326	\$25,332
2000	\$64,736	\$66,782	\$38,746	\$44,013	172.2	\$37,593	\$38,782	\$22,501	\$25,559
2001	\$68,350	\$69,351	\$39,147	\$46,072	177.1	\$38,594	\$39,159	\$22,105	\$26,015
<b>Percentage change over time</b>									
1990–2001	42.9	38.1	30.1		35.5	5.5	1.9	–4.0	
1991–2001	37.3	33.5	27.9	34.7	30.0	5.6	2.7	–1.6	3.6
1992–2001	31.2	32.2	20.5	28.3	26.2	4.0	4.7	–4.5	1.6
1990–2000	35.4	32.9	28.8		31.8	2.8	0.9	–2.3	
1990–1999	31.3	28.8	23.6		27.5	3.0	1.0	–3.0	
1990–1998	28.3	28.1	19.1		24.7	2.9	2.7	–4.5	

Key: CPI = Consumer Price Index.

Sources: Industry data were obtained from BLS staff (personal communication).

Note: Data for the U.S. economy (all civilian workers) were computed as follows: compensation per hour (BLS website) x 2,080 (hours per year).

Labor compensation in air transportation was significantly higher than the U.S. average over the period of analysis. Over 1991 to 2001, labor compensation in air transportation increased by 37%, while for the U.S. economy, it increased by 35%. Moreover, during 1990 to 2001, labor compensation in air transportation increased by 43% in nominal dollars, while compensation in rail increased by 38%, and in trucking it increased by 30%.

In real terms, one observes a similar phenomenon. Real labor compensation in air transportation outpaced inflation, and it increased faster than the mean compensation for the economy, and in the two other transportation industries.

The air transportation subsector is characterized by volatility, with booms and busts, and labor compensation to some extent can be affected by those cyclical movements. In order to check the robustness of the results, percentage rates of change were calculated with different starting and ending years. The results are shown in the bottom part of table 4.

It can be observed in the table that in every case, except one (for the 1992 to 2001 period compared to railroads), labor compensation in air transportation has the highest percentage increase compared to the economy as well as rail and trucking.

Thus, productivity increases in air transportation were accompanied by relatively rapid rises in labor compensation compared with the U.S. economy and the two transportation industries. Labor compensation increases in air transportation would have been affected by a more productive industry. Labor compensation could also have been affected by the existence of labor unions that would attempt to maximize income of their members. This factor is examined below.

The air transportation labor force is characterized by well-entrenched unions in various segments of the industry. All the major airlines have union representation in at least part of their labor force (USGAO 2003). The various labor groups that unions typically represent include pilots, flight

attendants, mechanics, and dispatchers. Sometimes unions represent customer-service agents and clerical workers, aircraft and baggage handling personnel, and flight instructors. Different unions may represent a given employee craft or class at different airlines. The existence of strong labor unions has been described in a recent study related to bankruptcy proceedings of a major airline (United Airlines 2002).

A study by the General Accounting Office points out that although the Railway Labor Act is designed to bring about settlement without unions resorting to strikes, negotiations between the airlines and their unions have sometimes been contentious, and strikes have occurred. Since 1990, negotiations have been marked by nonstrike work actions on the part of unions, such as sickouts and work slowdowns. These actions are designed to place economic pressure on airlines (USGAO 2003, p.1).

In the years since deregulation, the frequency of strikes has declined, but the number of nonstrike work actions has increased. Seventy-five percent of *strikes* occurred prior to 1990. By comparison, all identified *nonstrike work actions*—such as sickouts or refusals to work overtime—and all (six) *presidential interventions* occurred after 1990 (USGAO 2003, p.9). Moreover, the length of time to negotiate airline contracts has increased since deregulation, and particularly since 1990. From 1978 to 1989, the median contract negotiation was 9 months while the median negotiation length from 1990 to 2002 increased to 15 months (USGAO 2003, p.10). Consequently, the activities of strong labor unions in air transportation would have exerted a significant influence in the relatively rapid growth of labor compensation in that industry.

## CONCLUSIONS

The paper assesses the benefits of productivity increases in air transportation during the period 1990 to 2001. The choice of this time period is based on the availability of productivity data that are central to the analysis. The benefits of productivity are shown through subsequent impacts on profits, prices, and labor compensation. The evaluation of these three impacts is dependent on produc-

tivity data; therefore, the data for assessing those impacts are for the same time period.

The results show that labor and multifactor productivity in the air transportation subsector generally increased since 1990 and up to 2000. Productivity increases are expected to result in higher industry profits. Subsequently, a portion of this benefit may be passed on to consumers/users of the industry's services, in the form of lower prices, and/or to industry employees, in the form of higher labor compensation.

There is an association between productivity and profits in the industry. This applies particularly with regard to labor productivity. The increases in labor and multifactor productivity over the period of analysis tended to be accompanied by increased industry profits, which can subsequently impact prices and labor compensation.

With respect to productivity and prices, it appears that consumers of scheduled passenger services did not obtain that part of the benefit of productivity increases. Prices for consumers/passengers continued to increase (rather than decrease) relatively rapidly over time—while noting the different conclusion provided by ATPI data. On the other hand, commercial users of scheduled air cargo services obtained a portion of the benefit from productivity increases as prices for those services increased relatively slowly or declined.

In explaining why consumers of scheduled passenger services did not benefit from productivity increases while commercial users of freight services did, one may note that in the case of passenger services, it is the consumers (individually) who are dealing with the providers of the service. Prices are affected by the relative bargaining power of the buyer and seller and the degree of competition in the industry. The industry is an oligopoly, which implies a relatively low level of competition. Moreover, mergers/acquisitions and bankruptcies reduced the number of air carriers over time, further lessening the degree of industry competition. On the other hand, businesses that purchase scheduled air cargo services tend to have substantial bargaining power (including repeat business) and a good knowledge of prices. These factors can be used to obtain advantageous prices for freight transportation.

Another finding with respect to price is that prices for domestic air travel are shown to have increased considerably faster, over the period of analysis, than prices of international air travel. The analysis indicates that this was affected by a trend toward decreased competition in the domestic market, a result of mergers and thus fewer larger firms who, according to various studies, put up aggressive responses to the entry of low-cost air carriers. With this situation in the domestic market, and with other things constant in the international market, one could explain the evolution of domestic and international prices. In addition, in open-sky countries, prices for international air travel declined over a period of years during the 1990s.

With regard to productivity and compensation, the analysis indicates that a part of the benefit of the productivity increase in air transportation went to the employees of air carriers, in the form of higher labor compensation. This can be observed in terms of levels and changes over time in compensation. In terms of levels, labor compensation in air transportation was significantly higher than the average for the U.S. economy. In addition, labor compensation—in nominal and real terms—in air transportation increased at relatively high rates during the period of analysis. It increased faster than the U.S. average and in the other transportation subsectors—railroads and trucking. One also notes that the relatively strong degree of unionization in air transportation would have been instrumental in labor obtaining a substantial portion of the benefit of increased productivity.

## ACKNOWLEDGMENT

The views expressed in this paper are those of the author and do not necessarily reflect those of the Bureau of Transportation Statistics.

The author would like to thank Jack Wells for reviewing the manuscript and providing helpful comments.

## REFERENCES AND BIBLIOGRAPHY

Adrangi, B., G. Chow, and K. Raffiee. 1997. Airline Deregulation, Safety, and Profitability in the U.S. *Transportation* 44–52.

Air Transport Association (ATA). No date. State of the U.S. Airline Industry, 2002–2003. Available at [www.airlines.org](http://www.airlines.org).

American Bar Association (ABA). Report of the Aviation Committee, 2003 Spring Council Meeting, Section on Public Utility, Communications and Transportation Law.

Baltagi, B.H., J.M. Griffin, and D.P. Rich. 1995. Airline Deregulation: The Cost Pieces of the Puzzle. *International Economic Review* 36(1):245–260.

Case, K.E., and R.C. Fair. 1994. *Principles of Economics*, 3rd ed. Englewood Cliffs, NJ: Prentice Hall, pp. 377–379.

Cooper, M.N., Director of Research, Consumer Federation of America, 2001. Testimony before the House Committee on Energy and Commerce, Subcommittee on Commerce, Trade, and Consumer Protection, March 21.

Goodman, W.C. 2000. Transportation by Air: Job Growth Moderates from Stellar Rates. *Monthly Labor Review* 123:34–47.

Hirsch, B.T. and D.A. Macpherson. 2000. Earnings, Rents and Competition in the Airline Labor Market. *Journal of Labor Economics* 18 (1):125–155.

Kvanli, A.H. 1988. *Statistics: A Computer Integrated Approach*. St. Paul, MN: West Publishing Co.

Lent, J. and A.H. Dorfman. 2005. A Transaction Price Index for Air Travel. *Monthly Labor Review* 128:16–31.

Oster C.V. and J.S. Strong. 2001. Predatory Practices in the U.S. Airline Industry. January 15. Found on the internet.

Oum, T.H., and C. Yu. 1996. A Productivity Comparison of the World's Major Airlines. *Journal of Air Transport Management* 2(3–4):181–195.

United Airlines Informational Brief. 2002. Chapter 11, In the U.S. Bankruptcy Court for the Northern District of Illinois. December 9.

U.S. Department of Labor (USDOL), Bureau of Labor Statistics. May 2005. Price indices. Available at <http://data.bls.gov>.

\_\_\_\_\_. May 2005. Productivity, subsection on Productivity and Costs (for data on labor productivity) and subsection on Multifactor Productivity. Available at <http://data.bls.gov>.

\_\_\_\_\_. 1983. *Trends in Multifactor Productivity*, Bulletin 2178. Washington, DC: U.S. Government Printing Office, September.

U.S. Department of Transportation (USDOT). 1999. Competition in the U.S. Domestic Airline Industry: The Need for a Policy to Prevent Unfair Practices. May. Available at <http://www.dot.gov>.

U.S. Department of Transportation (USDOT), Bureau of Transportation Statistics (BTS). 2002. *National Transportation Statistics*. Washington, DC.

\_\_\_\_\_. TranStats. Available at <http://www.bts.gov>.

U.S. Department of Transportation (USDOT), Office of Aviation Analysis (OAA), Economic and Financial Analysis Division. Various. *Airline Quarterly Financial Review*.

U.S. Department of Transportation (USDOT), Office of the Secretary of Transportation (OST). 2000. *International Aviation*

- Developments, Second Report: Transatlantic Deregulation, the Alliance Effect.* October. Available at <http://www.dot.gov>.
- U.S. General Accounting Office (USGAO). 2003. *Airline Labor Relations: Information on Trends and Impacts of Labor Actions.* Washington, DC. June.
- \_\_\_\_\_. 1999. Aviation Competition: Effects on Consumers from Domestic Airline Alliances Vary. Washington, DC. January.
- \_\_\_\_\_. 1995. International Aviation Airline Alliances Produce Benefits, but Effects on Competition is Uncertain. April.
- Von Nordenflycht, A., and T.A. Kochan. 2003. Labor Contract Negotiations in the Airline Industry. *Monthly Labor Review* 126.
- Yuskavage, R.E. 2001. Issues in the Measurement of Transportation Output: The Perspective of the BEA Industry Accounts, paper presented at the Brookings Workshop on Transportation Output and Productivity, May 4, 2001.

**APPENDIX TABLE 1 Growth of Multifactor Productivity in Air Transportation**  
(Annual percentage rates)

	1990– 2000	1990– 2001	1990– 1995	1995– 2000	1995– 2001
Air transportation	1.9	1.4	2.1	1.7	0.7
U.S. private business sector	1.0	0.8	0.6	1.3	0.9

Source: The data on which these growth rates are based were obtained from the Bureau of Labor Statistics Internet site, section on Productivity, subsection on Multifactor Productivity.

**APPENDIX TABLE 2 Spearman Rank Correlation Coefficients**

Years	Labor productivity and profits (OAI)	Labor productivity and profits (OST)	Multifactor productivity and profits (OAI)	Multifactor productivity and profits (OST)
1993–2001	0.79 *		0.53	
1994–2001		0.83 *		0.45
1993–2000	0.77 *		0.74 *	
1994–2000		0.78 *		0.61

Key: OAI = Office of Airline Information; OST = Office of the Secretary of Transportation.

Notes: For columns 1 and 3, profit data were obtained from TranStats. For columns 2 and 4, profit data (for Majors) were obtained from *Airline Quarterly Financial Review* (various issues).

\* Significant at the 95 percent level. There is a 5% chance of concluding that a positive or negative association exists when in fact it does not.

**APPENDIX TABLE 3 Prices In Air Transportation**

Year	CPI—U.S. (1982-84=100) (1)	Growth rate—CPI (1982-84=100) (2)	CPI— airline fare (1982-84=100) (3)	Growth rate—CPI, airline fare (4)	PPI—Air transportation, scheduled passenger— domestic, international (Dec.1989=100) (5)	Growth rate—PPI, scheduled passenger, domestic, international (6)	PPI—Air transportation, scheduled passenger, domestic (Dec.1989=100) (7)	Growth rate—PPI, scheduled passenger, domestic (8)	Scheduled air cargo (Dec.1987=100) (9)	Growth rate—PPI, scheduled air cargo (10)	ATPI, U.S.- origin (11)	ATPI, foreign- origin (12)	ATPI, full scope (13)
1990	130.7		148.4		110.6		111.3		102.0				
1991	136.2	4.2	155.2	4.6	122.4	10.7	125.1	12.4	105.2	3.1			
1992	140.3	3.0	155.2	0.0	114.8	-6.2	115.4	-7.8	107.0	1.7			
1993	144.5	3.0	178.7	15.1	126.8	10.5	131.2	13.7	112.1	4.8			
1994	148.2	2.6	185.5	3.8	130.6	3.0	136.4	4.0	109.5	-2.3			
1995	152.4	2.8	189.7	2.3	137.8	5.5	144.8	6.2	111.2	1.6	100.1	104.5	100.7
1996	156.9	3.0	192.5	1.5	148.1	7.5	160.1	10.6	108.7	-2.2	98.6	98.7	98.4
1997	160.5	2.3	199.2	3.5	153.9	3.9	167.9	4.9	107.3	-1.3	103.9	95.4	102.4
1998	163.0	1.6	205.3	3.1	152.6	-0.8	165.2	-1.6	104.8	-2.3	100.8	84.3	98.1
1999	166.6	2.2	218.8	6.6	161.2	5.6	174.2	5.4	106.9	2.0	101.6	83.5	98.7
2000	172.2	3.4	239.4	9.4	186.5	15.7	208.1	19.5	110.4	3.3	108.7	86.5	105.2
2001	177.1	2.8	239.4	0.0	200.6	7.6	223.8	7.5	112.0	1.5	109.4	84	105.5
Percentage rates of change													
1995- 2001	16.2		26.2		45.6		54.6		0.7		9.3	-19.6	4.8

Sources: Data for columns 1 and 2, same as Figures 1 and 2. Data for columns 3 and 5, from BLS website, Consumer Price Index. Data for columns 7, 9, and 11, from BLS website, Producer Price Indexes. Column 5 includes domestic and international travel.

Note: Data for CPI refer to "All Urban Consumers." Data in column 9 are based on SIC.

Key: CPI = Consumer Price Index, PPI = Producer Price Index.

#### **APPENDIX 4**

The example involved the Reno-Minneapolis market. In this case, the major network carrier had previously served that market but had withdrawn from it. However, after a new airline entered the market, the major network carrier responded in several ways. First, it added new service overlaid on the entrant's network. This included three new daily nonstop flights from the same origin (Reno) to three different destinations; these were markets served by the entrant and not previously served by the network carrier. Moreover, the network carrier announced that it would begin a second daily flight from the same origin to one of the three destinations (Seattle). In addition, the network carrier announced that it would offer bonus frequent flier miles for the residents of the city of origin (Reno) on the routes that it offered from that city. It also stated that it would offer special travel agent commission overrides on flights to and from the city of origin.

Two days after the above actions, the network carrier also announced air fares to match the fares of the low-cost entrant on the Reno to Minneapolis route. It had initially announced lower fares than the fares of the entrant. It also announced that its fares for nonstop flights between several cities would be the same as those of the entrant's connecting service via Reno.

The entrant began service from Reno to Minneapolis service on April 1, as originally intended, but by May 20 losses forced it to reduce its service to

one flight a day. On June 1, 1993, Reno Air exited the Reno to Minneapolis market. The fares of the network carrier between several cities had dropped sharply in response to the entry of the new small airline into the Reno to Minneapolis market. However, following the exit of the new airline from that market, these fares increased quickly and steadily. In two to three quarters, the fares of the network carrier had increased to a level higher than before the entry of the new entrant. (Source: Oster and Strong, 2001, pp. 9-13)

#### **APPENDIX 5**

In 10 of the 12 cases, the new entrant's fare was at least 50 percent lower than the average fare of the incumbent(s) during the quarter preceding entry. In three-fourths of the cases, within two quarters of new entry, the average fare of the incumbent fell by 1/3 or more. The new entrant exited, in half the cases, within eight quarters after entry. In three of the six cases where the entrant exited, average fares then rose to above pre-entry levels; while in the other three markets, average fares increases above the level of the entry period.

With regard to revenue, in five of the six cases in which the new entrant exited from the market, total incumbent revenues were higher eight quarters later, and had increased sufficiently to offset any revenue losses that came from additional low-fare traffic during the period in which the new entrant was in the market. (Source: Cooper, 2001)

# Speed as a Risk Factor in Serious Run-off-Road Crashes: Bayesian Case-Control Analysis with Case Speed Uncertainty

GARY A. DAVIS<sup>1,\*</sup>  
SUJAY DAVULURI<sup>2</sup>  
JIANPING PEI<sup>1</sup>

---

<sup>1</sup> Department of Civil Engineering  
University of Minnesota  
122 CivE  
500 Pillsbury Drive SE  
Minneapolis, MN 55455

<sup>2</sup> Parsons Brinckerhoff  
999 Third Ave, Suite 2200  
Seattle, WA 98104

---

## ABSTRACT

In the United States, the imposition and subsequent repeal of the 55 mph speed limit has led to an energetic debate on the relationship between speed and the risk of being in a (fatal) crash. In addition, research done in the 1960s and 1970s suggested that crash risk is a U-shaped function of speed, with risk increasing as one travels both faster and slower than what is average on a road. This paper describes two case-control analyses of run-off-road crashes, one using data collected in Adelaide, Australia, and the other using data from Minnesota. In both analyses the speeds of the case vehicles were estimated using accident reconstruction techniques while the speeds of the controls were measured for vehicles traveling the crash site under similar conditions. Bayesian relative risk regression was then used to relate speed to crash risk, and uncertainty in the case speeds was accounted for by treating these as additional unknowns with informative priors. Neither dataset supported the existence of a U-shaped relationship, although risk of a serious or fatal run-off-road crash clearly tended to increase as speed increased.

---

Email addresses:  
Corresponding author—drtrips@umn.edu  
S. Davuluri—Davuluri@pbworld.com  
J. Pei—peixx005@umn.edu

---

KEYWORDS: Case-control, speed limits, logit model, Markov Chain Monte Carlo.

## INTRODUCTION

Determining appropriate speed limits is a problem that continues to exercise engineers, elected officials, and interested citizens, and at first glance this issue seems fairly simple. Compared to a slower vehicle, a vehicle traveling at high speed will go farther while the driver is reacting, take longer to stop, be more likely to sideslip for a given steering angle, and need to absorb more kinetic energy to protect its occupants. This suggests that, other things equal, slower speeds are safer, but complicating this issue is a series of observational studies which claim to find that the crash risk for slow moving vehicles is as high or higher than that of speeding vehicles (Solomon 1964; Cirillo 1968; West and Dunn 1971; Harkey et al. 1990). Each of these studies employs what is essentially a case-control design, where estimates of speeds from a sample of crash-involved vehicles (the cases) are compared with speeds from a sample of vehicles not involved in crashes (the controls). These studies have been subjected to a range of methodological criticisms, and there is no consensus on whether the observed associations between low speed and crash risk reflect actual causal processes, or are simply methodological artifacts.

In this paper, we begin by briefly reviewing these studies along with some more recent work. Based on this review we identify three related methodological issues that should be addressed in a case-control study of speed and crash risk. The first issue arises because the causal role of high (and low) speed probably differs for different crash processes, and to understand this causal role we should conduct separate studies of different types of crashes. This leads to the second issue because breaking a dataset down by type of crash often leads to small samples for which statistical methods based on large-sample asymptotics may not be applicable. The third issue stems from the fact that estimating the speeds of the case vehicles most often requires an after-the-fact reconstruction of the crash. The speed estimates produced by a crash reconstruction are to some degree uncertain, and this uncertainty should be allowed for in a case-control analysis. We then illustrate how a Bayesian analysis can address all of these issues by testing for the existence of a U-shaped relationship between speed and relative

crash risk using two small case-control samples of serious and fatal run-off-road crashes.

## Review of Case-Control Studies of Speed and Crash Risk

Summaries and detailed critiques of the studies by Solomon (1964), Cirillo (1968), and West and Dunn (1971) have been given by Shinar (1998) and Kloeden et al. (1997). In Solomon's (1964) study the pre-crash speeds of approximately 10,000 vehicles involved in crashes on some 600 miles (960 km) of two-lane and four-lane highways were obtained from crash records, from reports by the drivers involved, from witness statements, or from estimates provided by the police officers who were called to the scenes post hoc. Speed and traffic volume data were collected on these same highway stretches, and the ratios of the fraction of crash-involved vehicles with speeds in a given range to that of vehicle-miles of travel for that range were computed. These involvement rates were then plotted against the speed ranges' deviations from the average speed, producing a striking U-shaped relationship. For daytime crashes, the involvement rates were lowest for speeds about 10 mph (16 km/h) faster than the average, while speeds 30 mph (48 km/h) lower than average had involvement rates about 300 times greater than the lowest values. Involvement rates also increased for speeds greater than 10 mph (16 km/h) above the average but not as dramatically as for the lower speeds. In a later study Cirillo (1968) found a similar pattern for rear-end, angle, and sideswipe crashes on U.S. Interstate highways, and more recently Harkey et al. (1990) again found a U-shaped relationship between involvement rate and deviation from average speed for a sample of highway sites in North Carolina and Colorado.

These are provocative findings, but their importance depends on whether or not they correctly identify low speed as a cause of crashes. One potential limitation of these studies, which was pointed out in Shinar (1998) and Kloeden et al. (1997), concerns the data collection procedure. In Solomon's original study the speeds of vehicles making turns were included for the case vehicles but not for the control vehicles. In Cirillo's study attention was

restricted to crashes involving multiple vehicles traveling in the same direction, and on freeways these sorts of crashes tend to occur in congested conditions, which are characterized by reduced speeds. It is not clear what traffic conditions were present when the control speed data were collected but if traffic was uncongested then a situation similar to that of Solomon's study could arise. Finally, Harkey et al. (1990) explicitly stated that their control speed data were collected so as to guarantee that drivers were traveling at essentially freeflow speeds. Clearly this sampling procedure will tend to produce a higher fraction of low speeds in the case sample, without necessarily illuminating the role of low speed as a causal factor. For instance, a driver who slows down in order to turn or because of traffic congestion and is then involved in a crash will obviously have a lower speed than will freely moving drivers, but this does not entail that the slow speed actually caused the crash.

In addition, in each of these studies the procedure used to estimate the speeds of the case vehicles differed from that used to estimate speeds for the controls. Because measurement operations are almost always subject to some degree of error, this use of different estimation procedures means that the measurement error effect on the cases may be different than that for the controls. It is not now possible to assess the extent of measurement error in these studies, and so we cannot say with certainty whether or not measurement error effects have biased their results. However, a striking example of the potential effect of measurement error has been given by White and Wilson (1970), who showed that involvement rate curves similar to Solomon's can be produced by making plausible assumptions about speed measurement errors, even if actual crash risk is independent of speed.

We can hope, but probably should not expect, that a single study will provide the conclusive answer to a policy question. A more realistic expectation is that after a study's findings are presented, a process of critical discussion will identify potential weaknesses, and new studies addressing these weaknesses will then be conducted. Over time then our strongest findings should more closely approximate truth. Some limitations of the Solomon and Cirillo

studies were in fact known by 1970, and West and Dunn (1971) described an effort to improve on this earlier work. In West and Dunn's study, investigation teams visited the crash sites and estimated the case vehicle speeds using crash reconstruction methods. An attempt was also made at using data from nearby magnetic loop detectors to determine some case vehicle speeds, but this was apparently successful only in 9 of 36 attempts. What is interesting about West and Dunn's results is that although they still found a U-shaped relation between deviation from average speed and involvement rate, the estimated involvement rates were all of the same order of magnitude. That is, the estimated rates for the slowest vehicles were only about six times larger than for vehicles traveling near the average speed, compared to the increase of several hundred times found in the Solomon, Cirillo, and Harkey et al. studies.

More recently, important advances in the application of case-control methods to study crash risk have been made by the former Road Accident Research Unit (RARU) at the University of Adelaide (Moore et al. 1995; Kloeden et al., 1997, 2001, 2002). The RARU studies explicitly used case-control designs, where the speeds of the case vehicles were estimated using crash reconstruction methods, while the speeds of the controls were obtained by sampling vehicles traversing the crash locations under conditions similar to those present when the crashes occurred. In Kloeden et al. (1997) all crash locations had posted speed limits of 60 km/h (38 mph), all crashes occurred during daylight and dry weather conditions, and vehicles slowing to make turning maneuvers were excluded from the sample. Nonparametric estimates of relative crash risk were then computed for speeds ranging from 35 km/h to 85 km/h (22 mph to 53 mph), and although relative risk tended to increase as speeds increased above 60 km/h (38 mph) there was no evidence of increased risk for lower speeds. In Kloeden et al. (2001) this approach was applied to crashes occurring on rural roads, and in Kloeden et al. (2002) the original data were reanalyzed using parametric logit models. In both later studies, relative crash risk tended to increase as speed increased but no heightened risk for lower speeds was found.

Carrying on the process of critical discussion and improvement, we can identify two ways in which these findings might be further strengthened. The first arises from the fact that in all the studies considered so far, crashes of different types were combined to compute estimates of relative risk. Because one can reasonably expect that the causal effect of speed might differ for different types of crashes, it is possible that estimates computed by aggregating crash types will be influenced by the relative frequency of the different crash types in the sample. Inspection of table 4.4 in Kloeden et al. (1997) suggests that the difference between case and control speeds does vary across crash types with, for example, run-off-road crashes showing a pronounced difference while pedestrian crashes show little difference. One should consider then whether different types of crashes show different relationships between relative risk and speed, but simply disaggregating the RARU's data by crash type leads to relatively small numbers of cases for each type. Standard statistical methods (e.g., Hosmer and Lemeshow 2000), which are based on the large-sample asymptotic properties of estimators, will not necessarily be applicable.

The second avenue for improvement arises from the use of crash reconstruction methods to estimate the case speeds. Although this is a substantial improvement over what was done in earlier studies, in practice the evidence available from a crash investigation is rarely sufficient to determine all quantities needed to compute an estimate of speed. For example, the formula

$$v = \sqrt{2\mu gd} \quad (1)$$

where  $g$  = gravitational acceleration,  $v$  = vehicle speed,  $d$  = measured skid mark length, and  $\mu$  = tire/pavement friction coefficient, can be used to estimate a vehicle's speed, but only if one also has an estimate for  $\mu$ . Knowing the composition of the road and that it was dry can allow one to arrive at a plausible range for  $\mu$  (Fricke 1990), but the actual value characterizing an actual crash will still be to some degree uncertain. This situation becomes even more complicated if we allow that the measured skid mark length is at best an uncertain estimate of the actual braking distance. The estimates produced by a crash reconstruction are

thus subject to uncertainties, and the appropriate way to account for these uncertainties is still something of an open question. In Davis (1999; 2003) we have illustrated how this can be accomplished by treating crash reconstruction as an exercise in Bayesian inference. Here the reconstructionist's expert opinion regarding plausible values for crash variables is captured using prior probability distributions, and a model of the crash can then be combined with measurements to update these prior distributions, via Bayes theorem. The result is a posterior probability distribution over the values of the crash variables.

Both of these issues, accounting for the differential uncertainty in the case speed estimates and analysis of small samples, can be addressed using Bayesian methods. We will illustrate this using two datasets, where the cases were vehicles involved in fatal or severe run-off-road crashes, and where Bayesian crash reconstruction methods were used to compute posterior probabilities for each case vehicle's initial speed. The case vehicle speeds were then combined with speed measurements for vehicles not involved in crashes (controls), leading to a case-control problem. The posterior speed distributions from the crash reconstructions were used as informative prior distributions for the case vehicle speeds, and logistic regression modeling was then used to test whether or not a U-shaped relationship existed between speed and risk in run-off-road crashes.

### **A "Failure Rate" Model for Run-off-Road Crashes**

As noted above, we will use logistic regression to test for the possibility of a U-shaped relation between speed and crash risk, but before proceeding we would like to show how standard assumptions used in crash analysis lead to the logit model. To see this, assume first that run-off-road crashes arise when (a) a driver finds himself or herself in a crash avoiding situation, and (b) the driver's evasive action is not successful. As a driver traverses a section of road, crash avoidance situations are assumed to arise randomly, with density  $\lambda$ . The success of the evasive action is assumed to depend on the vehicle's speed, denoted by  $v$ , in a manner such that the probability of crash given  $v$  is approximately proportional to  $\exp(g(v))$ , for some function  $g(\cdot)$ . This leads to a proportional hazards model with hazard function

$$h(v) = \lambda \exp(g(v)) \quad (2)$$

Now if  $X$  denotes a random variable giving the distance traveled until being in a crash, the probability of being in a crash while traversing a section of road of length  $x$  is simply

$$P[X \leq x | v] = 1 - e^{-(\lambda x) \exp(g(v))} \quad (3)$$

If run-off-road crashes are rare (i.e.,  $0 < \lambda x \ll 1$ ) (3) can be approximated as

$$(\lambda x) \exp(g(v)) e^{-(\lambda x) \exp(g(v))} \quad (4)$$

which is the probability that the value 1 is taken on by a Poisson random variable  $Y$  with expected value

$$E[Y] = (\lambda x) \exp(g(v)) \quad (5)$$

If we then condition on  $Y = 0$  or  $1$  (so that no one can crash more than once), we get

$$P[Y = 1 | v] \approx \left( \frac{\exp(\log(\lambda x) + g(v))}{1 + \exp(\log(\lambda x) + g(v))} \right) \quad (6)$$

a logit model.

### Case-Control Analyses

Two sources provided the data on run-off-road crashes used in this study. The first was the case-control study conducted by the Road Accident Research Unit (RARU) at the University of Adelaide (Kloeden et al., 1997), which, as we noted earlier, reported data for 151 case vehicles involved in serious or fatal crashes on roads with 60 km/h speed limits. For each case vehicle, four control vehicles were selected by randomly sampling vehicles using the crash site at times when conditions were similar to those when the crash occurred. Control speeds were measured using laser speed guns while the case speeds were estimated using crash reconstruction techniques. Of the 151 cases, 14 were single vehicle run-off-road crashes, and of these 8 involved collisions with objects, where it was possible to measure the deformation (crush) suffered by the vehicles. For two others the case vehicles left measurable yaw marks near the points where the drivers lost control of their vehicles.

As noted above, crash reconstructions are subject to nontrivial uncertainties, and the probability

calculus can be used as a logic for reasoning about these. Our general approach to estimating case vehicle speeds for the RARU data was to develop probabilistic versions of the deterministic methods used by the RARU researchers, and this was done by supplementing their measurements with training data, treating the case vehicle speeds as missing values to be estimated. For the fixed-object crashes, the training sample consisted of 19 staged collisions conducted by the National Highway Traffic Safety Administration (NHTSA), reported in Nystrom and Kost (1992). For the yaw-mark crashes, the training sample was 40 measured speeds and yaw radii tabulated in Semon (1995).

First, for the fixed-object crashes, the following variant of Nystrom and Kost's (1992) model was used to relate measured crush to impact speed

$$c = (v - v_0) / (a_0 + a_1 * w) + \varepsilon \quad (7)$$

Where

$c$  = measured crush

$v$  = impact speed

$v_0$  = highest impact speed producing no crush (taken to be 5 mph)

$w$  = vehicle weight

$a_0, a_1$  = coefficients to be estimated

$\varepsilon$  = error.

The error term  $\varepsilon$  allows for differences between measured and predicted crush, and was assumed to be normally distributed with mean equal to 0 and unknown variance  $\sigma^2$ . Because six of the case vehicles left measurable skid-marks prior to collision, it was also necessary to account for speed lost while skidding. Treating the measured skid-mark as an error-prone observation, its expected value was computed using the RARU's formula

$$\bar{s} = \frac{L v_t^2 - v^2}{2\mu g} \quad (8)$$

where

$v_t$  = denotes the vehicles initial speed

$L$  = fraction of kinetic energy retained between the initiation of braking and the point where the skid-mark begins (taken by the RARU to be 0.8),

$\mu$  = coefficient of tire/pavement friction.

Garrot and Guenther (1982) conducted an extensive comparison of measured versus theoretical

skid-marks, and the differences between these showed a coefficient of variation approximately equal to 0.11. Following the approach described in Davis (2003), the measured skid-mark was assumed to have a log normal distribution, with the mean equal to the natural log of the theoretical length given in equation (8), and a normal variance of 0.01. This gives a coefficient of variation for the measurement error equal to approximately 0.1.

In addition to the likelihood functions for the measured crush and skid lengths, Bayesian analysis requires a prior distribution for the unknown quantities. For estimating the speeds of the fixed-object crash vehicles, the following hierarchical prior distribution was used:

$$\begin{aligned} a_0 &\sim \text{Normal}(0, 10^6), \\ a_1 &\sim \text{Normal}(0, 10^6), \\ \sigma^2 &\sim \text{Inverse Gamma}(0.001, 0.001), \\ \nu \text{ and } \nu_t &\sim \text{Normal}(\alpha, \pi), \\ \alpha &\sim \text{Normal}(40 \text{ mph}, 10^6), \\ \pi &\sim \text{Inverse Gamma}(0.001, 0.001), \text{ and} \\ \mu &\sim \text{Uniform}(0.45, 1.0). \end{aligned}$$

With the exception of  $\mu$ , all these are commonly used “uninformative” priors. For  $\mu$ , the lower bound characterizes a dry, travel polished asphalt pavement while the upper bound characterized a dry, new concrete pavement (Fricke 1990). As noted earlier, all crashes in the RARU sample occurred in dry weather.

Compared to the fixed-object crash model, the yaw-mark model was simpler, but still based on the principle of imputing unknown speeds. Treating the radius of the yaw mark as an error-prone measurement caused by the speed, the standard critical speed formula leads to

$$r = \frac{v^2}{\mu g} + \varepsilon \quad (9)$$

where

$$\begin{aligned} r &= \text{measured yaw radius,} \\ \nu &= \text{vehicle's speed,} \\ \mu &= \text{friction coefficient, and} \\ \varepsilon &= \text{measurement error.} \end{aligned}$$

The error term  $\varepsilon$  was assumed to be normally distributed with mean equal to zero, and unknown variance  $\sigma^2$ . As stated earlier, 40 experimental tests having information on observed speed and radius of curvature were used as a training dataset for esti-

imating the value of  $\mu$ . The two RARU cases were then treated as similar to the 40 tests but with missing speeds. The following priors were used:

$$\begin{aligned} \nu &\sim \text{Normal}(\alpha, \pi), \\ \alpha &\sim \text{Normal}(50, 10^6), \\ \pi &\sim \text{Inverse Gamma}(0.001, 0.001), \text{ and} \\ \mu &\sim \text{Uniform}(0.45, 1). \end{aligned}$$

Posterior distributions for the case vehicle speeds were then computed using the Markov Chain Monte Carlo program WinBUGS (Spiegelhalter et al., 2000), and details of the WinBUGS models have been given in Davis and Davuluri (2002). Table 1 summarizes the case and control data for the 10 RARU crashes.

The second dataset was taken from a set of 46 fatal crashes occurring on Minnesota state highways between January 1, 1997 and June 30, 2000. These were all fatal crashes reported during this time period which occurred near a location where the Minnesota Department of Transportation collected automatic vehicle speed data, and for which crash investigation data could be obtained from the Minnesota State Patrol. The automatic speed data were used to produce control speeds by randomly sampling from the speed measurements taken during an hour when conditions were judged to be similar to those present when the crash occurred. Of the 46 crashes, 22 involved loss of control and running off the road, and of these 9 resulted in collisions with other vehicles, 10 resulted in rollover, and 3 resulted in collisions with fixed objects. For 10 of these it was possible to use crash reconstruction methods to estimate initial speeds. For two, initial speeds were estimated from measured yaw marks using the method described above, while for five a tripped rollover model described in Cooperrider et al. (1990) and Martinez and Schlueter (1996) was adapted to estimate initial speeds. This method divides the roll-over into rolling, tripping, and pre-tripping phases, and then works backward from the vehicle's rest position to estimate the speed at the beginning of each phase. For the three remaining Minnesota crashes, straightforward application of either the yaw-mark mark method or the tripped rollover model was not possible, but special features of these crashes still permitted estimates of initial speeds. In one crash, where the case vehicle jumped

**TABLE 1** Posterior Means and Standard Deviations for the Speeds of the Case Vehicles, and the Measured Speeds for the Control Vehicles, for 10 Road Accident Research Unit Run-off Road Crashes. All Speeds are in Miles per Hour.

Crash no.	Case speeds		Control speeds			
	Mean	Standard deviation	1	2	3	4
1	43.4	4.1	34	35	38	38
2	41.5	3.9	27	30	34	38
3	48.3	4.3	30	34	35	39
4	53.2	3.3	32	34	37	42
5	54.8	4.5	37	43	38	30
6	86.9	9.0	37	37	38	38
7	41.5	2.4	33	34	38	43
8	38.7	2.3	31	40	40	42
9	59.2	3.4	39	39	42	43
10	62.5	3.2	34	38	39	43

a ditch, the fall equation (Fricke 1990) was used to estimate speed at the ditch's edge. In another, where the driver was thrown from the vehicle upon its striking a fence, Searle's (1993) throw equation was used to estimate the vehicle's speed at the fence. In the third, where a driver lost control and rolled his vehicle after being rear-ended, we were able to estimate an initial speed for the rear-ending vehicle from skid and yaw marks. Table 2 summarizes the case and control data for the 10 Minnesota run-off-road crashes. Example code illustrating our reconstruction methods is available from the first author on request.

As indicated in the Introduction, one of the unresolved issues in the debate on speed versus crash risk concerns whether or not crash risk is a U-shaped function of speed, with vehicles traveling at atypically low and high speeds having increased crash risk. If we accept that the role of speed may vary for different types of crashes, depending on the operative processes and circumstances, then appropriate tests for the possibility of a U-shaped relationship should be carried out using data disaggregated by crash type. Otherwise, there is the possibility of obscuring the speed effect by combining processes where speed is and is not causal, or of producing an apparent U-shaped relationship by mixing situations where high speed is causal with other situations where low speed is causal. Disaggregating by type of crash reduces sample sizes however, but as argued earlier, a simple proportional hazards model

relating speed, distance traveled, and crash risk leads to a prospective logit model

$$P[\text{crash}|v] = \exp(b_0 + g(v, \mathbf{b})) / (1 + \exp(b_0 + g(v, \mathbf{b}))) \quad (10)$$

The parameter  $b_0$  can be taken as summarizing the effects of those features shared by the cases and controls at a given location, while the function  $g(v, \mathbf{b})$  describes how crash risk varies with speed and a vector of parameters  $\mathbf{b}$ . Assuming first that both the case and the control speeds are known without error, the fact that the cases and controls are matched by location means that a matched case-control approach can be used, which leads to a likelihood contribution from site  $k$  of the form

$$P[y_{k,0} = 1, y_{k,j} = 0, j = 1, \dots, m] = \exp(g(v_{k,0}, \mathbf{b})) / (\sum \exp(g(v_{k,j}, \mathbf{b}))) \quad (11)$$

(e.g., Hosmer and Lemeshow 2000). Here  $v_{k,0}$  denotes the case vehicle speed at site  $k$ , while  $v_{k,j}$   $j > 0$  denotes the corresponding speeds for the control vehicles. The likelihood function obtained as the product of equation (11) over all case-control sets would then provide the basis for either a Bayesian or a classical approach to estimation. If the case speeds are only known up to some probability distribution, they then become additional quantities to be estimated, and by using those distributions as priors Bayesian estimation is in principle straightforward. In all our analyses, the priors for the case vehicle speeds were taken to be normal distribu-

**TABLE 2** Posterior Means and Standard Deviations for the Speeds of the Case Vehicles, and Measured Speeds of the Control Vehicles, for 10 Minnesota Run-off Road Crashes. All Speeds are in Miles per Hour.

Crash no.	Case speeds		Control speeds									
	Mean	Standard deviation	1	2	3	4	5	6	7	8	9	10
1	69.8	5.6	58	58	66	50	56	65	60	70	72	54
2	71.0	5.0	57	53	38	55	65	51	59	61	55	56
3	71.4	4.4	69	75	82	77	75	82	78	70	77	71
4	81.4	2.1	69	75	89	70	75	77	87	65	77	71
5	59.8	3.6	52	60	56	55	56	56	58	53	54	62
6	80.8	2.1	69	78	71	82	80	70	69	70	70	78
7	74.5	5.0	85	78	85	71	83	72	85	71	63	79
8	67.7	4.0	69	68	65	56	72	75	75	69	70	70
9	80.4	1.7	77	81	78	77	72	75	74	80	70	76
10	73.3	4.0	73	73	71	75	80	82	78	70	87	77

tions, with means and standard deviations as given in tables 1 and 2.

### Estimating the Risk Functions

In the simplest case, a test as to whether or not the risk function is U-shaped can be carried out by comparing a quadratic form for the function  $g(\cdot)$

$$g(v, \mathbf{b}) = b_1(v - v_m) + b_2(v - v_m)^2 \quad (12)$$

to a linear form

$$g(v, \mathbf{b}) = b_1(v - v_m). \quad (13)$$

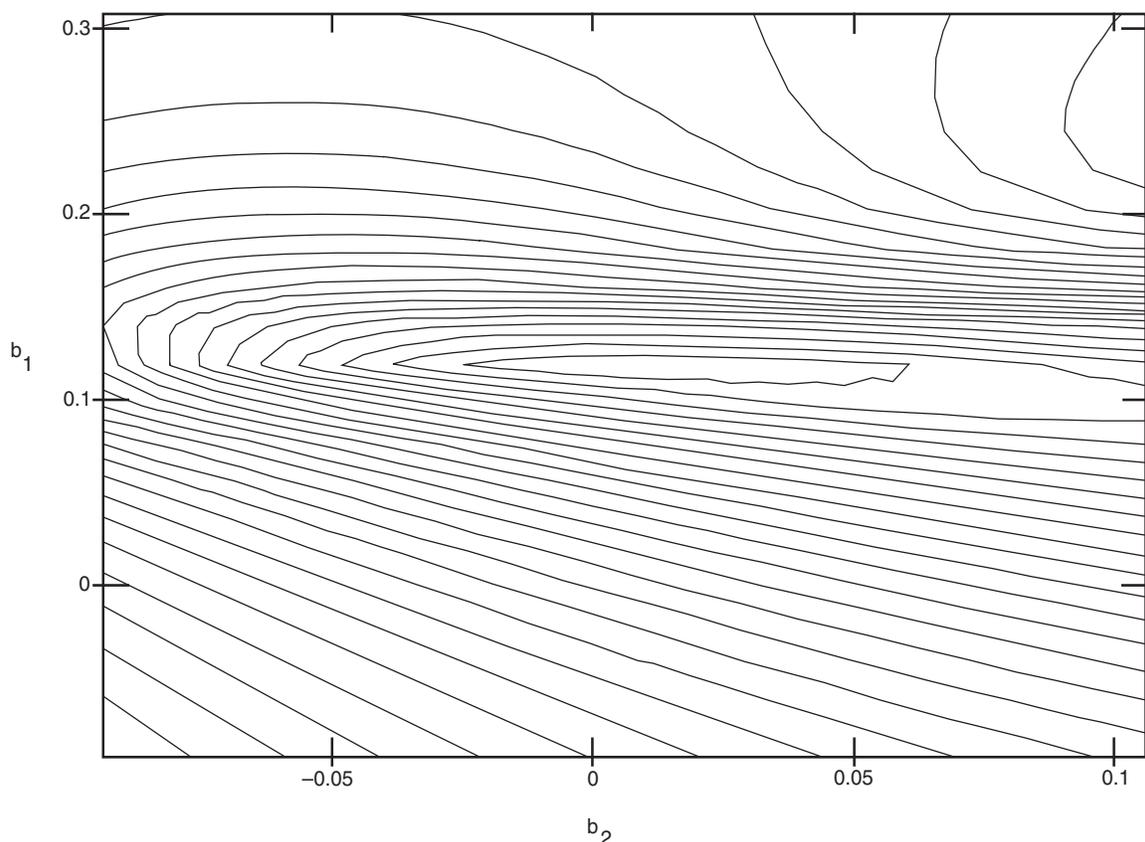
Looking first at the Minnesota crashes, Bayes estimates for the linear model (13) were computed using the Markov Chain Monte Carlo routine WinBUGS with  $v_m$  being fixed, for each case-control set, to the average speed for that set's control population. When we attempted to estimate the quadratic model, however, the MCMC routine was unstable, producing chains with poor mixing properties, and the simulated values for  $b_1$  and  $b_2$  tended to be highly correlated with each other. At least one reason for this can be seen in figure 1, which shows a contour plot of the marginal log-likelihood as a function of  $b_1$  and  $b_2$ . The narrow ridge-shape of this log-likelihood indicates that the data tend to be uninformative about  $b_2$  over a range of values, including zero. For the relative risk function to be U-shaped, however,  $b_2$  must be positive, so additional MCMC runs were conducted with the

prior for  $b_2$  constrained to have support only on the non-negative real numbers. Table 3 displays posterior estimation summaries for the linear and constrained quadratic models as fit to the Minnesota run-off-road data.

The results shown in table 3 indicate that the linear and constrained quadratic models provided roughly equivalent fits to the Minnesota data. The posterior deviances have similar distributions, and the values of the deviance information criteria (DIC) (Spiegelhalter et al. 2002) were approximately equal. Because a convex parabola and a straight line have different implications for the relationship between speed and crash risk, the rough equivalence of these two models may seem contradictory. The contradiction is resolved by looking at the point of minimum risk, which for the quadratic model occurs at a speed equal to  $v_m - b_1 / (2b_2)$ . Substituting the posterior means for  $b_1$  and  $b_2$  into this expression reveals that for the quadratic model minimum risk occurs at about a speed between 10 and 11 mph *below* the average for the controls. Because most (97 out of 100) of the control speeds in table 2 were above these values, what happened was that the quadratic model achieved parity with the linear model simply by being monotonically increasing over the range of the available data. More particularly, the results in table 3 do not appear to support earlier claims that minimum risk tends to occur near the mean or median of the control speeds.

For the RARU data, the Markov Chain Monte Carlo simulations showed poor mixing even when

**FIGURE 1 Contour Plot of Log Likelihood Function of Quadratic Model Fit to the Minnesota Data**



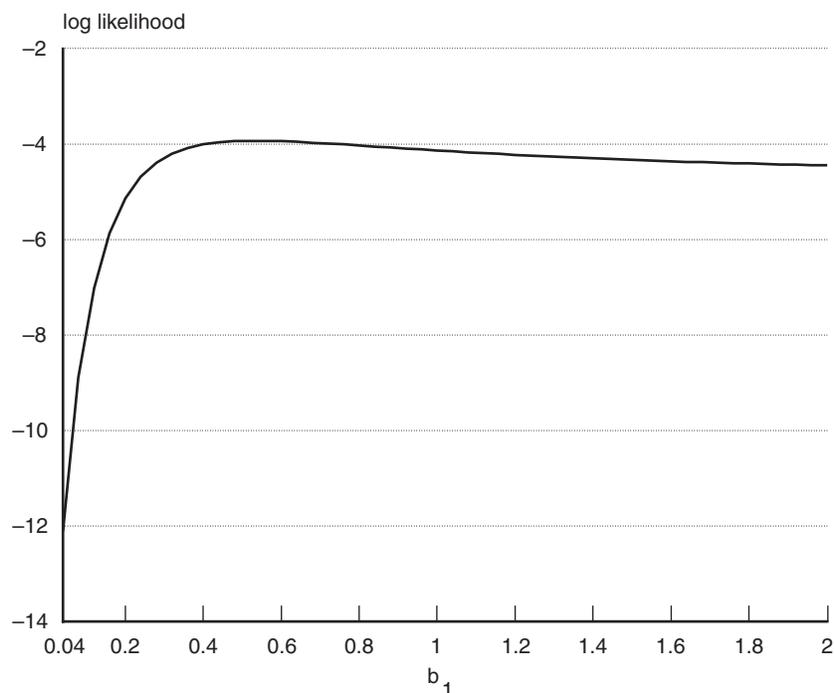
**TABLE 3 Bayesian Parameter Estimates and Goodness of Fit Measures for the Linear and Constrained Quadratic Models Applied to the Minnesota Data**

Parameter/ measure	Linear model			(Constrained) quadratic model		
	Mean	2.5 percentile	97.5 percentile	Mean	2.5 percentile	97.7 percentile
$b_1$	0.19	0.03	0.46	-0.014	-0.013	0.381
$b_2$	—	—	—	0.006	0.0003	0.016
Deviance	40.3	29.5	47.9	40.6	30.8	48.5
DIC			41.1			40.3

attempting to fit the linear model, and again studying a plot of the marginal log-likelihood function was informative. Figure 3 shows this, and the distinctive feature here is how the log-likelihood flattens out for higher values of  $b_1$ , indicating that the matched case-control data contain no information about how large  $b_1$  is. The reason for this is found by inspecting table 1, where it can be seen that most controls speeds are below the posterior mean speed for their corresponding cases, and none are greater than two standard deviations above this mean. To work around this problem, it was decided to supple-

ment the control speeds with those obtained at all other sites in the RARU study. We considered this acceptable because, unlike the Minnesota data, the RARU data were collected under similar road and weather conditions. This produced an unmatched case control study with 10 cases and 604 controls. As with the Minnesota data, WinBUGS was used to compute Bayesian estimates of the parameters for the linear and quadratic models, along with measures of goodness of fit, and these are displayed in table 4. Again, the linear and quadratic models appeared to fit the data about equally well, but the

**FIGURE 2 Matched Case Control Log Likelihood as a Function of  $b_1$  for Linear Model Fit to Road Accident Research Unit Data**



**TABLE 4 Bayesian Parameter Estimates and Goodness of Fit Measures for Linear and Quadratic Models Applied to the RARU Data.**

Parameter/ measure	Linear model			Quadratic model		
	Mean	2.5 percentile	97.5 percentile	Mean	2.5 percentile	97.7 percentile
$b_0$	-7.35	-10.64	-5.29	-7.66	-11.72	-5.36
$b_1$	0.54	0.33	0.84	0.55	0.15	1.11
$b_2$	—	—	—	0.0018	-0.018	0.025
deviance	36.33	26	48	36.3	25.6	48.5
DIC			40			39.6

interpretation is more clear cut. For both the linear and quadratic models the estimates of the intercept and the coefficient for the linear term were essentially equal, while the coefficient for the quadratic term in the quadratic model was essentially centered at zero. The linear and quadratic models thus achieved comparable fits by relying only on the linear terms.

To summarize, for both the Minnesota crashes and the RARU's crashes it appeared that at least over a typical range of speeds, risk of being in a serious or fatal run-off-road crash increases as speed increases. If, in fact, there are situations where low speeds are dangerous, these likely involve processes

or conditions different from those that characterize the crashes in these samples.

## SUMMARY AND CONCLUSION

In the introduction we indicated that a salient issue with regard to the role of speed in road crashes concerns the existence of a U-shaped relationship between speed crash risk. Despite extensive research, a clear resolution of this issue has yet to be achieved. The view we have adopted is that at least some of the current confusion may result from: 1) aggregating crashes that are caused by fundamentally different processes, and 2) failure to account

for uncertainty in an analysis. In this paper, we have showed how pioneering work conducted at the RARU could be combined with recent advances in computation for Bayesian statistical models in order to apply case-control methods to studies with relatively small numbers of cases. Applying the method to two case-control samples, each with 10 serious or fatal run-off-road crashes, we found that these data did not support the existence of a U-shaped relationship between speed and crash risk, although risk did tend to increase as a function of speed.

One implication of this study appears to be that, as common sense tells us, high speed in and of itself is not sufficient to cause a crash. For the 10 Minnesota crashes, other drivers were observed traveling the same road under the same conditions as fast or faster than the crash-involved drivers without being involved in a fatal crash. A reasonable interpretation would be that some type of triggering event, which places the driver in a crash-avoiding situation, is also necessary. This is consistent with Hauer's pyramid (1997, p. 19), which distinguishes normal driving from conflict situations, and from those situations resulting in crashes. Study of the Minnesota crash reports revealed events such as the appearance of a deer in the driver's path, the merging of a slower moving vehicle into the driver's lane, driver distraction leading to a need to avoid a rear-ending collision, and loss of control following the driver's turning to interact with a child in the back seat. The logit model used in this paper assumed that such situations arise randomly with a rate that perhaps differs for different roadways. This parameter was absorbed into the logit model's constant term, and it is well-known that this term cannot be identified from case-control data alone. Because this parameter was not needed in testing for a U-shaped relation between speed and risk, this did not handicap our analysis. (In more traditional crash reconstruction, one in essence conditions on the occurrence of the crash avoidance event, so again it is not necessary to determine how this event arose.) A more complete understanding of how crashes occur however will eventually require determining how crash-avoidance situations arise. At present, though, our ability to model these does not appear

to be as well-developed as our ability to model what happens once a crash sequence has started.

Finally, these results should caution us against using aggregated data to make overly general statements about crash causation. It may be that in other scenarios low speed is a causal factor of crashes. The challenge now is first to identify those scenarios, and then to demonstrate in actual instances how low speed caused these crashes.

## ACKNOWLEDGEMENTS

This research was supported in part by Bureau of Transportation Statistics Contract DTTS-00-G-B004-MN, and in part by the Intelligent Transportation Systems Institute at the University of Minnesota. The authors would like to thank Don Schmaltzbauer of the Minnesota State Patrol and Dan Brannon of the Minnesota Department of Transportation for their assistance in obtaining the Minnesota data.

## REFERENCES

- Cirillo, J. 1968. Interstate System Accident Research: Study II. *Public Roads* 35:71-75.
- Cooperrider, N., T. Thomas, and S. Hammoud. 1990. Testing and Analysis of Vehicle Rollover Behavior, SAE Technical Paper 900366, SAE, Inc, Warrendale, PA.
- Davis, G. 1999. Using Graphical Markov Models and Gibbs Sampling to Reconstruct Vehicle/Pedestrian Accidents. *Proceedings of the Conference on Traffic Safety on Two Continents*. Linkoping, Sweden: Swedish National Road and Transport Institute.
- \_\_\_\_\_. 2003. Bayesian Reconstruction of Traffic Accidents and Causal Effect of Speed in Intersection and Pedestrian Accidents. *Law, Probability and Risk* 2:69-89.
- Davis, G. and S. Davuluri. 2002. Development and Application of Case-Control Methods to Traffic Safety Problems, prepared for the U.S. Department of Transportation, Bureau of Transportation Statistics, Washington, DC.
- Fricke, L. 1990. Traffic Crash Reconstruction. Traffic Institute, Northwestern University, Evanston, IL.
- Garrott, W. and D. Guenther. 1982. Determination of Precrash Parameters from Skid Mark Analysis. *Transportation Research Record* 893:38-36.
- Harkey, D., D. Robertson, and S. Davis. 1990. Assessment of Current Speed Zoning Criteria. *Transportation Research Record* 1281:40-51.
- Hauer, E. 1997. *Observational Before-After Studies in Road Safety*. New York, NY: Elsevier.

- Hosmer, D.W. and S. Lemeshow. 2000. *Applied Logistic Regression*, 2nd ed. New York, NY: Wiley.
- Kloeden, C., A. McLean, V. Moore, and G. Ponte. 1997. Traveling Speed and the Risk of Crash Involvement. NHMRC Road Accident Research Unit, University of Adelaide, Adelaide, Australia.
- Kloeden, C., G. Ponte, and J. McLean. 2001. Traveling Speed and the Risk of Crash Involvement on Rural Roads. Road Accident Research Unit, Adelaide, University of Adelaide, Adelaide, Australia.
- Kloeden, C., J. McLean, and G. Glonek. 2002. Reanalysis of Traveling Speed and the Risk of Crash Involvement in Adelaide South Australia. Road Accident Research Unit, University of Adelaide, Adelaide, Australia.
- Martinez, J., and R. Schlueter. 1996. A Primer on the Reconstruction and Presentation of Rollover Accidents, SAE Technical Paper 960647, SAE Inc., Warrendale, PA.
- Moore, V., J. Dolinis, and A. Woodward. 1995. Vehicle Speed and Risk of a Severe Crash. *Epidemiology* 6:258–262.
- Nystrom, G. and G. Kost. 1992. Application of NHTSA Crash Database to Pole Impact Prediction SAE Technical Paper 920605, SAE Inc., Warrendale, PA.
- Searle, J. 1993. The Physics of Throw Distance in Accident Reconstruction. *Accident Reconstruction: Technology and Animation III*. Warrendale, PA: SAE Inc.
- Semon, M. 1995. Determination of Speed from Yaw Marks. *Forensic Accident Investigation: Motor Vehicles*. Edited by T. Bohan and A. Damask. Charlottesville, VA: Michie Butterworth, 149–198.
- Shinar, D. 1998. Speed and Crashes: A Controversial Topic and an Elusive Relationship. *Managing Speed, Special Report 254*. Washington, DC: Transportation Research Board, 221–276.
- Solomon, D. 1964. *Accidents on Main Rural Highways Related to Speed, Driver and Vehicle*. Washington, DC: Bureau of Public Roads.
- Spiegelhalter, D., N. Best, B. Carlin, and A. van der Linde. 2002. Bayesian Measures of Model Complexity and Fit. *Journal of the Royal Statistical Society B* 64:583–616.
- Spiegelhalter, D., A. Tomas and N. Best. 2000. *WimBUGS Version 1.3 User Manual*. Cambridge, UK: Cambridge University. MRC Biostatistics Unit.
- West, L. and J. Dunn. 1971. Accidents, Speed Deviation and Speed Limits. *Traffic Engineering* 52–55, July.
- White, S. and A. Wilson. 1970. Some Effects of Measurement Errors in Estimating Involvement Rate as a Function of Deviation from Mean Traffic Speed. *Journal of Safety Research* 2:67–72.

# Speed Estimation for Air Quality Analysis

HUAFENG GONG<sup>1</sup>

MEI CHEN<sup>1,\*</sup>

JESSE MAYES<sup>2</sup>

ROB BOSTROM<sup>2</sup>

---

<sup>1</sup> Department of Civil Engineering  
University of Kentucky  
Lexington, KY 40506

<sup>2</sup> Division of Multimodal Programs  
Kentucky Transportation Cabinet  
Frankfort, KY 40622

---

## ABSTRACT

Average speed is an essential input to the air quality analysis MOBILE6 model for the calculation of emissions factors. Traditionally, speed is obtained from travel demand models; however, such models are not usually calibrated to speeds. Furthermore, for rural areas where such models are not available, no reliable method is available for estimating speed. In this study, we developed a procedure based on the model in the Highway Economic Requirement System to estimate average speed using as input various data such as roadway characteristics and traffic conditions. The model was confirmed to be powerful based on the statistical comparisons between the estimated and measured speeds. Various implementation issues including the impact of data quality and potential applications are also discussed.

## BACKGROUND

The increasing use of motor vehicles has resulted in a much degraded air quality in recent decades. The Clean Air Act requires transportation planners to monitor and assess the performance of transportation systems regularly; while the enactment of the Clean Air Act Amendments of 1990 signified the

---

Email addresses:

\*Corresponding author—mchen@engr.uky.edu

H. Gong—hgong0@engr.uky.edu

J. Mayes—jesse.mayes@ky.gov

R. Bostrom—rob.bostrom@ky.gov

---

KEYWORDS: Speed estimation, air quality.

importance of combining travel demand and air pollutant emissions forecasting.

The commonly used air quality analysis model, MOBILE6, provides estimates of current and future emissions from highway motor vehicles. It has been employed by most states in compliance with the U.S. Environmental Protection Agency (EPA) requirement. MOBILE6 is an emissions factor model that employs information such as vehicle classification and age distribution, average operating speed, and vehicle-miles of travel (VMT). The outputs of the model include emissions factors for hydrocarbons (HC), carbon monoxide (CO), nitrogen oxides (NO<sub>x</sub>), carbon dioxide (CO<sub>2</sub>), particulate matter (PM), and toxic pollutants from cars, trucks, and motorcycles under various conditions (Cook and Glover 2002). Even though MOBILE6 has national default values for each category, area-specific inputs on a variety of parameters are preferred (e.g., annual mileage accumulation by vehicle class, average speed distribution by hour and roadway type, distribution of VMT by roadway type, and distribution of VMT by vehicle class).

Among the parameters required by the model, average speed is the most important because emissions rates are highly sensitive to changes in speed. Furthermore, the emissions rates of the three major pollutants, HC, CO, and NO<sub>x</sub>, are also very sensitive to VMT by time of day and average speed (Tang et al. 2003). This calls for an accurate estimate of average operating speed.

Various methodologies can be applied to speed estimation. Dowling et al. (1997) provide a comprehensive review of these methods. Here, we briefly discuss several commonly used methods.

The standard Bureau of Public Roads (BPR) equation was developed in the 1960s. Even though it does not accurately reflect the relationship between volume and speed, it has been widely used as a simple tool to predict mean speed, as shown in the following equation.

$$S = \frac{FFS}{1 + a(v/c)^b}$$

where

$S$  = predicted mean speed  
 $FFS$  = free-flow speed

$v$  = volume  
 $c$  = practical capacity  
 $a$  = 0.15  
 $b$  = 4

The free-flow speed, capacity, and volume can be determined by creating various lookup tables based on area and facility types. The uniform parameter values for  $a$  and  $b$  do not distinguish facilities in different types. This method could result in an estimation error of approximately 40% (Dowling et al. 1997).

Several improvements have been made to enhance the accuracy of the standard BPR equation. Separate curves were fitted for urban interrupted facilities. Data on critical segments of the facility replaced the facility averages. Based on an updated speed-flow relationship, the value of  $a$  was set at 0.05 for signalized facilities and 0.20 for all other facilities, while the value of  $b$  was set at 10. Furthermore, free-flow speed was estimated using an equation instead of the lookup table.

Despite the improved performance of the enhanced BPR technique, BPR-type equations are not capable of addressing the spill-back of physical queues formed at urban interrupted facilities. Therefore, this method should be limited to long-range planning applications that do not usually require high precision (Dowling et al. 1997).

The ARTPLAN technique is a planning procedure developed by the Florida Department of Transportation and is powerful in dealing with urban facilities controlled by signals (Dowling et al. 1997). Subsequently, the model was expanded to cover urban streets with stop sign control and conditions in which demand exceeds capacity. A similar procedure for rural facilities with interrupted flows was also created. Although the ARTPLAN technique outperforms the enhanced BPR technique for mean speed estimation on urban uninterrupted facilities, it still produces large errors. For example, it was observed that for urban arterials the estimation error could be up to 25% or 33% (Dowling et al. 1997).

In Kentucky, travel demand models (TDMs) are the primary tool for obtaining average speed estimates. These models were developed for large urbanized areas such as Louisville, Lexington, and

the Northern Kentucky area. Some smaller urbanized areas also have their own TDMs. The enhanced BPR function is used in the model (Bostrom and Mayes 2003). However, these models do not presently include procedures for calibrating speeds. Furthermore, Kentucky currently has no reliable procedure for estimating speeds in areas without a TDM. Bostrom and Mayes (2003) provide a summary of the air quality attainment issues and highway speed estimation for MOBILE6 in Kentucky.

The objective of this research was to develop a procedure to estimate average speed on different roadway types. This paper evaluates the performance of such a procedure by comparing the estimated speeds with speed data collected in the field. Several issues that arose during the implementation of the model will also be discussed.

## RESEARCH APPROACH

Based on the requirements of the air quality analysis and available data, we developed a procedure based on the internal speed model of the national version (v3.26) of the Highway Economic Requirement System (HERS) (USDOT 2000b), from which the state version (HERS-ST) is derived. HERS is a cost and benefit analysis tool that uses engineering standards and economic criteria to provide decision support on future infrastructure investment levels. HERS consists of a number of internal models that generate intermediate parameters for the cost and benefit analysis. One of the parameters is a speed model that calculates average effective speed (AES) for each segment of a roadway. This information can subsequently be used to calculate the costs of travel time, the external costs, and the total vehicle operating costs.

The HERS speed model requires many data items on facilities and traffic. Such information includes roadway geometric parameters, pavement condition, speed limit, traffic control devices, and traffic composition. Since HERS was designed to run based on the format of the Highway Performance Measurement System (HPMS) sample data, most required data items are available, at least for the sample segments.

The HERS speed model uses an aggregate probabilistic limiting velocity model to determine the free-

flow speed (FFS) on a roadway. The delay due to traffic control devices or the presence of other vehicles on a uninterrupted facility is estimated based on facility type. The average AES is then obtained from the FFS and the delay. Figure 1 shows the general procedure for estimating average effective speed. The complete procedure for the HERS speed model can be found in the HERS Technical Report v3.26 (USDOT 2000b).

Based on the HERS speed model, an Excel macro was programmed to calculate the AES for each roadway segment. The average speeds were then grouped by county and by functional class for the purpose of air quality analysis.

## CASE STUDY AND MODEL VALIDATION

### Input Data

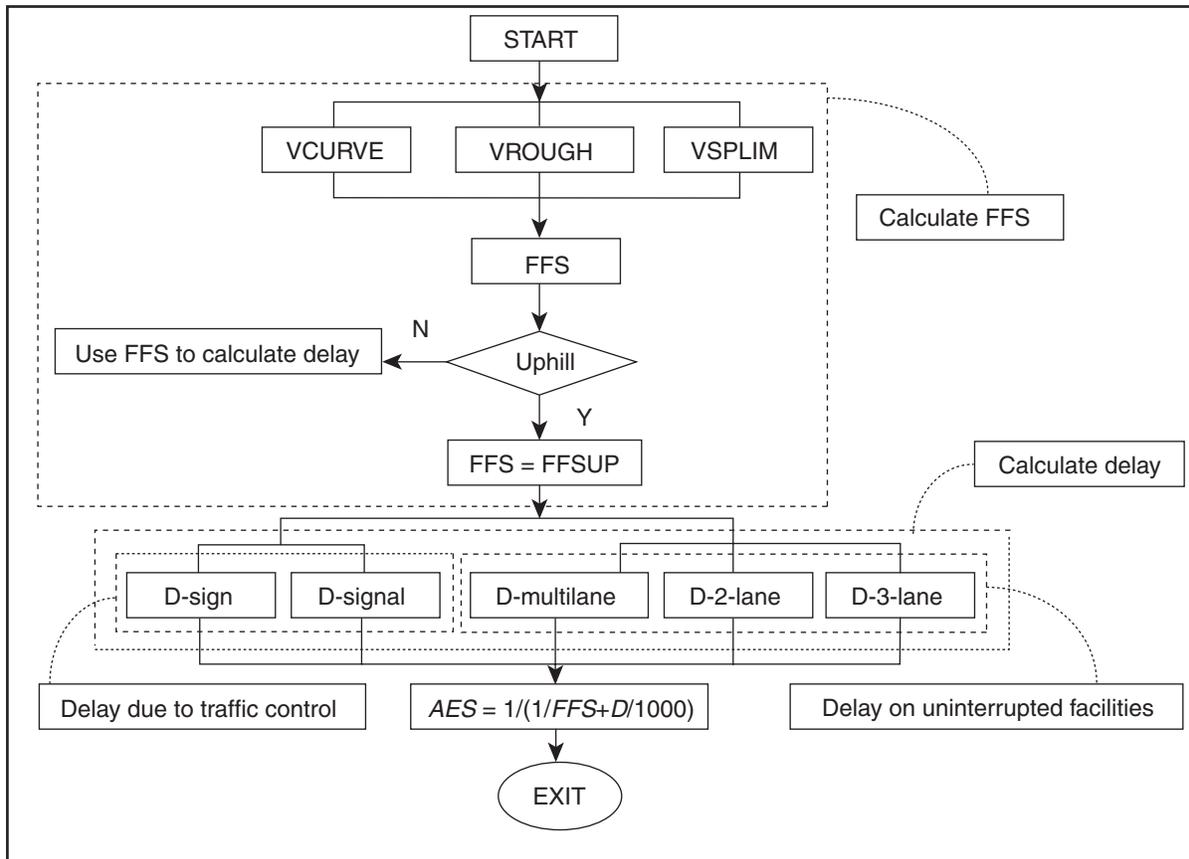
We tested the HERS speed model using the data from the 2002 HPMS extract for Kentucky. This set includes state and locally owned roadways with a total of over 9,000 segments and over 13,500 miles. The mileage breakdown by functional class is shown in table 1. In addition to the data items in the HPMS format, the HERS speed model also needs information on heavy vehicle percentages by vehicle type on the segments (as specified in USDOT 2000b). However, these data are unavailable for most segments. Therefore, a lookup table was created to estimate this information based on the statewide heavy vehicle distribution by functional class and the total heavy vehicle percentages on each segment.

### Model Validation

In order to evaluate the performance of the speed model, we compared the estimates to the field data collected through various efforts. Limited speed data are available in Kentucky, especially after 1995 when the speed limit compliance program was discontinued by the Federal Highway Administration (Bostrom and Mayes 2003).

Two primary sources of speed data exist in Kentucky. One is a study of the impact of speed limit changes on highway safety, in which extensive speed data were collected on various roads in Kentucky (Agent et al. 1997). Another is a recent effort to col-

**FIGURE 1 Estimating Average Effective Speed**



Key:  
 VCURVE = the maximum allowable speed on a curve  
 VROUGH = the maximum allowable ride-severity speed  
 VSPLIM = the maximum speed resulting from the speed limit  
 FFS = free-flow speed  
 FFSUP = free-flow speed on an uphill  
 D = delay  
 AES = average effective speed

lect speed data in Christian County, Kentucky. Although these data were not collected in the same year as the HPMS data extract used for the HERS model, they were chosen to be compared with the model output because they are the most complete sets (in terms of covering various roadway types) of speed data. This time mismatch may introduce some errors to the validation process, especially at the segment level. However, the error at the route (a sequence of segments) level could be less due to the smoothing effect of the aggregation. Additionally, in an attempt to offset the impact of the mismatched time periods, several items (e.g., signal density) in the input data file, to which the speed output may be very sensitive, were updated during the validation process based on field data from Christian County.

This case is also discussed in a later section to illustrate the importance of having accurate input data.

In the 1997 study, speed data were collected on 86 sample routes, covering all highway functional classes except for local roads and rural minor collectors. Based on the beginning and ending mile points for each of these routes, the matching sequence of segments was extracted from the 2002 HPMS data. The average effective speeds for these segments can be obtained via implementing the HERS speed model. The overall average speed for each route containing multiple segments was estimated as total mileage traveled divided by total time spent on the route. Table 2 lists a few sample roadways for which the comparison between measured and estimated speeds was made.

**TABLE 1 Sample Data Summary**

Functional class	Numbers of segments	Mileage
1	115	533
2	822	2,052
6	979	1,633
7	3,138	6,932
8	0	0
9	0	0
11	91	229
12	48	87
14	1,270	661
16	2,009	996
17	535	411
19	0	0
<b>Total</b>	<b>9,007</b>	<b>13,534</b>

For rural Interstates with a 65 mile-per-hour (mph) speed limit, the differences between the estimated and measured speeds ranged from -0.2 mph to 5.9 mph. For urban Interstates and other arterials with a speed limit of 55 mph, such differences ranged from -11.4 mph to 2.2 mph. A paired *t*-test was chosen to test the equality of the underlying population means between the model output and measured samples. Prior to the test, preliminary analyses were conducted to ensure that the data did not violate the assumptions of the test. The first assumption was that the paired differences should be independent of each other, which was satisfied because the speed data came from different roads. Secondly, the paired differences should be normally distributed. The normal probability plot for the

paired differences was constructed in which the close agreement with the straight line was observed. Then, the Lilliefors test for goodness of fit to a normal distribution was conducted. Under the significance level  $\alpha$ , the hypothesis that the paired difference has a normal distribution was accepted.

After the assumptions were confirmed, the paired *t*-test was conducted. With a *p* value of  $5.6 \times 10^{-5}$ , the result recommended that we reject the null hypothesis that the two sets of speeds are from populations with equal means. In other words, the estimated and measure speeds were statistically different. However, the test also showed that the average measured speeds were no more than 1.1 mph higher than that of the estimated speeds when  $\alpha = 0.05$ . The *p* value at this time was 0.08 and the *t* statistic was 1.77 and was lower than the critical *t* statistic (1.99 in this case). The 95% confidence interval for the average difference between the measured and estimated speeds was (0.18, 3.59). This implies that the extent of the differences between estimates and measurements was not very large, although the difference was statistically significant.

To eliminate the potential impact of the speed limit on the sample means, paired *t*-tests were conducted for sample groups with different speed limits. Under the significance level of 0.05, test results showed that for roadways with a 65 mph speed limit (i.e., rural Interstates), the average estimated speed was approximately 1 mph higher than the average measured speed. The 95% confidence inter-

**TABLE 2 Speed Comparison Based on the 1997 Speed Study**

Route	Functional class	Speed limit	Measured speed	Estimated speed
I 24	1,11	65	68.5	71.2
I 64	1,11	65	68.4	69.7
I 65 Jefferson	11	55	59.8	60.7
I 471 Campbell	11,12	55	59.6	61.8
Mountain 9000	2	65	68.3	64.4
Purchase 9003	2,7,12	65	67.0	71.7
W. Kentucky 9001	2,12,14	65	69.2	71.3
US 60, Grayson-Ashland	7	55	54.7	54.9
US 150, Bardstown-Danville	2,6,14	55	59.0	54.3
KY 10, Vanceburg-US 23	2	55	57.6	55.4
KY 15, Whitesburg-Campton	2,7,14	55	58.5	53.9

**TABLE 3 Speed Comparison Based on the 2004 Christian County Survey**

Route	Functional class	Speed limit	Measured speed	Estimated speed
I 24	1	65	72.0	71.8
E 9004	2	65	68.5	72.1
US 41	6	25/35/45/55	53.1	51.1
KY 91	7	55	58.0	58.5
KY 164	7	45	51.2	47.8
KY 1026	8	35	40.5	46.7
KY 1027	8	40	38.4	42.7
CR 1031	9	40	38.9	39.9
CR 1053	9	45	44.8	47.5
I 0024	11	65	74.3	72.0
US 41A	14	25/35/45/55	42.3	35.6
US 68B	14	45	55.1	60.8
KY 115	16	35/45	39.1	39.8
KY 380	16	35	30.8	25.2
KY 911	17	35	36.8	34.9
KY 1007	17	45	32.2	27.3
KY 400	19	35	32.4	37.9

val for the difference between the estimated and measured speeds was (0.69 mph, 2.97 mph) for these roads. For roadways with a speed limit of 55 mph, the HERS speed model underestimated the speed by approximately 2 mph. The 95% confidence interval for the difference between the two was (-3.84 mph, -1.74 mph) for these roads. Although the differences between the estimated and measured speeds were statistically significant, the absolute estimation errors were not substantial as indicated by the mean difference and the confidence intervals.

A larger difference was observed between the estimated and measured speeds on roads with lower functional classes. This was primarily attributable to the model's sensitivity to various factors such as traffic signal density. A detailed discussion on this topic will be presented in the next section.

In 2005, Christian County in Kentucky was designated by EPA as a nonattainment area. It became crucial to obtain accurate speed estimates for different types of roadways in this county in order to establish the future emissions budget. Speed data were collected during a three-month period in summer 2004 on a number of roadways throughout the county. The effort covered approximately 50% of the total mileages (both state and locally maintained) and over 70% of state-maintained facilities

in the county. The sample segments were selected based on the recommendation in FHWA's *Travel Time Collection Handbook* (Turner et al. 1998). Each road was traveled at least twice, once during the peak and once during the offpeak periods.

The HERS model was tested on the same highway segments in Christian County on which the speed survey was conducted. Table 3 shows the comparison between the estimated and measured speeds for several sample roadways in the county. The differences between the two sets of speeds were mostly within 5 mph with few exceptions. However, the paired *t*-test could not be applied in this case because the data violated the assumption that the paired differences between the two sets of speeds should be normally distributed.

Therefore, nonparametric tests need to be used, because they do not usually make distributional assumptions. The most commonly used alternative for the paired *t*-test is the Wilcoxon paired signed rank test. The Wilcoxon signed rank test first sorts the absolute values of the differences (between estimated and measured speeds) from smallest to largest, and then assigns ranks to these absolute values starting with the smallest as rank 1. The sum of the ranks of the positive differences is then calculated. When the null hypothesis (i.e., the median difference in paired data is zero) is true, the sum of the ranks

**TABLE 4 Christian County Speeds Comparison**

Functional class	Sample size	Mileage	Estimated speed (mph)	Measured speed (mph)
1	2	17.3	70.7	71.2
2	6	35.1	56.4	60.3
6	2	7.6	51.1	53.1
7	6	52.4	54.0	54.7
8	30	175.6	47.1	48.5
9	74	188.6	35.7	41.2
11	1	3.3	72.0	74.3
12	—	—	—	—
14	7	15.9	35.7	38.8
16	26	51.6	31.2	30.5
17	7	15.6	29.6	32.5
19	64	68.2	23.0	23.2

of all positive differences is approximately the same as that of the negative differences. The Wilcoxon signed rank test was conducted to compare the estimated and measured speed data. With a  $p$ -value of 0.40 under  $\alpha = 0.05$ , the test did not find enough evidence to reject the null hypothesis that the underlying population speeds had the same median. When the population distribution is symmetric (as was the case for Christian County data), the median is approximately equal to the mean. The test was also conducted for sample differences in each functional class. It subsequently recommended the acceptance of the null hypotheses as well.

Considering the speed variation on highways by various functional classes, the speed samples were grouped according to roadway functional class. The speed sample size and mileage are summarized in table 4 together with the aggregated average speeds from the HERS model and field measurement in Christian County.

### IMPLEMENTATION ISSUES

An Excel-based software tool was developed to implement the HERS speed model on highway data stored in the HPMS format. Additional data items, such as truck percentage breakdowns by truck type, were prepared separately. The tool calculates the average effective speed for each segment and then aggregates them to the county level for each functional class. Specifically, the average travel time on each segment was estimated from the segment length and the average effective speed. The county-

wide average speed was then calculated as total distance traveled (i.e., total length of all segments) in a functional class divided by total travel time on the road segments in that functional class.

### Data Quality

Although the HERS speed model performed very well in estimating the average speed for each roadway segment, its accuracy at the county, regional, or state level is largely dependent on the availability and accuracy of the input data.

### Availability

The HERS model uses highway inventory data in the HPMS sample format to calculate the average speed. However, such data are not available for all highways. Usually, most state-maintained highways are inventoried, but not much information is available for those that are locally maintained unless they are HPMS sample sections. Moreover, many state-maintained roads that are in lower functional classes may not have been inventoried. An accurate estimate of speed would call for adequate samples in each functional class. Intensive effort might be necessary to ensure that enough data are available, particularly for roadways in lower functional classes.

### Accuracy

The accuracy of the input data also affects the performance of the speed model. Like any model, the validity of the speed model output depends on the validity of the input. Some inconsistencies were found in the HPMS extract. For example, the sum

**TABLE 5 Speed Comparison with Changes in Traffic Control Devices**

Route	Function class	Measured speed (mph)	Initial AES (mph)	Updated AES (mph)	Change
US 41A	2	43.7	60.4	41.3	12 signals added
US 68	2	47.9	59.9	46.1	9 signals added, lower speed limits (up to 30 mph reduction)
KY 107	7	47.4	52.3	52.4	2 stop signs added
KY 109	7	46.2	59.6	55.6	11 signals added
KY 164	7	51.2	56.5	48.2	Lower speed limit (up to 10 mph reduction)
KY 380	16	32.0	23.2	25.2	1 signal removed
US 41	16	31.6	39.5	35.2	1 signal added, lane width reduced
KY 911	17	36.8	30.0	34.9	1 signal removed
KY 1007	17	32.0	27.9	32.0	Speed limit increased (up to 10 mph)

Key: AES = average estimated speed.

of the curve (or grade) lengths must equal the segment length (USDOT 2000a); however, a number of segments did not satisfy this requirement. Furthermore, the unavailability of curve (or grade) data on some segments is treated by the speed model as if the segment was all tangent (or leveled), because both scenarios would have a “0” code in the curve (or grade) class fields. However, these segments may not indeed be curve-free (or leveled) as indicated by the horizontal (or vertical) alignment adequacy rating. In other words, the HERS speed model does not distinguish the “no data” scenario from the “no curve (or grade)” scenario. In order to improve the accuracy of the model output, efforts should be made to assure the accuracy of each data item.

Special attention should be paid to the accuracy of data items, such as the density of traffic control devices, because they tend to have a significant impact on the delay estimates. During the model validation process, significant differences between the estimated and measured speeds were observed on several roads. Table 5 lists several roads in Christian County with “Initial AES” estimates significantly different from the observed speeds at the same sites. Further investigation revealed that there are some differences in the density of traffic control devices, speed limit, and lane width between the 2002 data extract and the information collected in 2004. After the input file was updated based on the latest information, the HERS model produced an updated output (also shown in table 5). Significant improvement of estimation accuracy can be seen on many of these roadways.

The HERS model is also quite sensitive to the speed limit, which is one of the parameters used to calculate the free-flow speed. The maximum speed resulting from the speed limit (VSPLIM) is assumed to be at least 6 mph above the speed limit in the HERS speed model.

In Kentucky, the default speed limit for rural highways other than the Interstates and four-lane highways with a median is 55 mph. However, the prevailing speed may be severely restricted by the presence of sharp curves which, as discussed earlier, may not be accurately reported in the HPMS sample data. Nevertheless, the adjustment of the posted speed limit to reflect the prevailing operating speed may not be made on all segments. This is also recognized in the HPMS *Field Manual* (USDOT 2000a) in that it uses the horizontal alignment adequacy rating to describe the curves with design speed less than the prevailing speed limit. Under this circumstance, using the posted speed limits in the data table would yield an unreasonably higher VSPLIM. Combined with the incomplete curve data, which will over-estimate the maximum allowable speed on a curve (VCURVE), higher (and less accurate) estimates of FFS and AES will result.

On the other hand, the item “weighted design speed” in the HPMS data file contains the design speeds weighted by the length of horizontal curves and tangents on a segment. For a number of roadway segments in Kentucky, the weighted design speed could be as low as 40 mph while the posted speed limit is 50 mph. Therefore, in order to reduce the estimation error, in such cases, the effective

speed limit (the lower one between the weighted design speed and posted speed limit) should be used.

Because other factors such as annual average daily travel (AADT) and truck traffic percentage and composition would also affect the average speed, a full-range sensitivity analysis for this model will require an extensive amount of speed data collected in the field. Nevertheless, this study demonstrates the sensitivity of the HERS model as well as the significance of data quality assurance efforts.

### Applications

The HERS speed model was applied to the Kentucky statewide highway inventory data in the HPMS format. Then we grouped average speeds by county and functional class for air quality analysis application. However, a county-level sample size may be too limited to provide reliable speed estimates for each county.

Alternatively, all 120 counties in Kentucky were divided into 3 major groups according to demographic, economic, and topographical characteristics. Although Kentucky is largely a rural state, it contains three major metropolitan areas (Louisville, Northern Kentucky, and Lexington) with typical urban traffic patterns. The eastern Kentucky area is mostly mountainous with many slow-moving coal trucks on the highways; therefore, the statewide speed distribution was obtained for three types of areas—urban, mountainous, and other rural areas (table 6). The areawide speeds by functional class

could then be used to represent the countywide speed distribution. This method preserves the characteristics of each type of area while ensuring a relatively larger sample size to smooth out the impact of stochastic variation, which may result from the limited sample size for one specific county.

The average speed estimates obtained from the HERS model can be used in various applications. In the short term, it can be used as an input to the MOBILE6 model to compute the emissions factor for various automobile-related pollutants. In the long run, if the input data items, such as pavement condition and AADT are updated using their projected values for future years, the HERS speed model will produce the projected average speeds on these roads. Such speeds can be used to estimate the emissions budget for future years.

In addition to air quality-related analysis, speed data can also be used as part of the highway performance measures. The data provide quantitative supplements to the traditional level-of-service indices and serve as the basis for the estimation of travel time, all highly desirable information on highway performance.

### CONCLUSION AND FUTURE RESEARCH

The speed estimation procedure developed in this study is based on the HERS speed model. It uses the HPMS data format to compute speed on each roadway segment. The free-flow speed was first estimated and then adjusted based on delay experienced

**TABLE 6 Statewide Average Speeds by Area Type and Functional Class**

HPMS functional class			Average speed			
			Statewide	Urbanized	Mountainous	Other
1	Rural	Interstate	69.2	70.0	68.5	69.2
2	Rural	Principle arterial	55.4	59.1	52.4	56.6
6	Rural	Minor arterial	45.2	47.0	39.7	46.5
7	Rural	Major collector	44.3	46.8	38.9	46.2
8	Rural	Minor collector	N/A	N/A	N/A	N/A
9	Rural	Local	N/A	N/A	N/A	N/A
11	Urban	Interstate	60.1	58.6	71.6	70.6
12	Urban	Other freeway	62.6	61.0	N/A	65.4
14	Urban	Principle arterial	25.4	21.1	36.2	30.5
16	Urban	Minor arterial	23.1	20.3	26.3	27.9
17	Urban	Collector	31.0	29.4	32.2	33.1
19	Urban	Local	N/A	N/A	N/A	N/A

by each vehicle (on various types of facilities) to obtain the average speed estimate. Although a large number of data items are required as input, these data are available from the annual HPMS submission that is mandatory for all states. However, for those roadways that do not belong to the HPMS sample set (primarily local roads and rural minor collectors), additional data-collection efforts may be necessary.

The model performance was evaluated by two independent speed datasets collected in the field. Various statistical analyses attested to the power of the model for producing accurate speed estimates. Tests also showed that the model was quite sensitive to factors such as the density of traffic control devices. A periodic review and update of such information in the inventory data file may be required to ensure the accuracy of input data to the speed model.

Although default speed distribution by hour is available in MOBILE6, the area-specific hourly speed estimates are needed to increase the prediction accuracy of emissions factors. The next-generation of air quality model, MOVES (Motor Vehicle Emission Simulator), also calls for speed data at a much finer level than the daily average (USEPA 2004). Furthermore, the analysis of hot spots would require delay and queue length by time of day. Currently, an effort is being made to adapt the concept of the HERS speed model to the estimation of hourly speeds. This hourly speed model would provide further detail on the variation of speed, delay, and queue length over time, in addition to accounting for queue spillover during the peak period.

## ACKNOWLEDGMENT

This study is funded by the Kentucky Transportation Cabinet and the U.S. Department of Transpor-

tation, Federal Highway Administration. The views presented in the paper are those of the authors alone.

## REFERENCES

- Agent, K., J.G. Pigman, and J.M. Weber. 1997. *Evaluation of Speed Limits in Kentucky*, Final Report, KYSPR-96-172. Lexington, KY: Kentucky Transportation Center.
- Bostrom, R. and J. Mayes. 2003. Highway Speed Estimation for MODILE6 by Kentucky, available at [transportation.ky.gov/Multimodal/doc/M6\\_speed\\_estimation\\_paper.doc](http://transportation.ky.gov/Multimodal/doc/M6_speed_estimation_paper.doc).
- Cook, R. and E.L. Glover. 2002. *Technical Description of the Toxics Module for MOBILE6.2 and Guidance on Its Use for Emission Inventory Preparation*, EPA420-R-02-029. Washington, DC: U.S. Environmental Protection Agency.
- Dowling, R.G., W. Kittelson, J. Zegeer, and A. Skabardonis. 1997. *National Cooperative Highway Research Report 387: Planning Techniques to Estimate Speeds and Service Volumes for Planning Applications*. Washington, DC: Transportation Research Board.
- Tang, T., M. Roberts, and C. Ho. 2003. *Sensitivity Analysis of MOBILE6: Motor Vehicle Emission Factor Model*, FHWA-RC-Atlanta-03-0007. Washington, DC: U.S. Department of Transportation, Federal Highway Administration.
- Turner, S., W. Eisele, R. Benz, and D. Holdener. 1998. *Travel Time Data Collection Handbook*, FHWA-PL-98-035. Washington, DC: U.S. Department of Transportation, Federal Highway Administration.
- U.S. Department of Transportation (USDOT), Federal Highway Administration. 2000a. *HPMS Field Manual*, OMB No. 21250028. Available at <http://www.fhwa.dot.gov/ohim/hpmsman/hpms.htm>, as of September 2005.
- \_\_\_\_\_. 2006. *Highway Economic Requirements System: Technical Report*, v3.26. Available at [http://www.dot.state.oh.us/gasb34/FHWAAsset\\_Management+GASB\\_34/eei%20team/hers\\_st/documentation/HERS%20Tech%20printready.pdf](http://www.dot.state.oh.us/gasb34/FHWAAsset_Management+GASB_34/eei%20team/hers_st/documentation/HERS%20Tech%20printready.pdf), as of September 2005.
- U.S. Environment Protection Agency (USEPA). 2004. *MOVES2004 Highway Vehicle Population and Activity Data*, EPA420-P-04-020. Washington, DC: December.

# Measurement Errors in Poisson Regressions: A Simulation Study Based on Travel Frequency Data

ERLING HÄGGSTRÖM LUNDEVALLER

---

Department of Statistics  
Umeå University  
SE-901 87, Sweden  
Email: erling.lundevaller@stat.umu.se

---

## ABSTRACT

This paper considers how measurement errors in explanatory variables affect the analysis of a Poisson regression model for frequencies of recreational and shopping trips. Measurement errors can introduce bias into the parameter estimates, and the effects on this particular dataset and model are investigated. The structure of the data, with two observations for each individual, makes it desirable to test for correlation within each individual. It is possible that tests of random effects are sensitive to measurement error. The properties of tests of random individual effects when there are measurement errors are therefore studied in the paper. The results of a simulation study show that classical measurement errors cause severe bias, and Berkson measurement errors produce little bias. The tests for random individual effects work well both with measurement error and negatively correlated responses according to the simulation study.

## INTRODUCTION

An often encountered situation in statistical modeling in various research fields is when the response variable of interest consists of frequencies. A widely used model for frequency data is the Poisson regres-

---

KEYWORDS: Poisson regression, measurement error, travel frequencies, random coefficients.

sion model. This modeling approach makes it possible to relate the response counts to explanatory variables. Although other models for frequency data can also be used, the Poisson model is in many cases the preferred model.

The quality of the data is of central importance for the results of the statistical analysis. One common problem is measurement errors in regressors. It is well known that measurement errors in explanatory variables introduce bias and inconsistency in estimators in linear regression models (e.g., Fuller 1987; Carrol et al. 1995). In a simulation study, Zidek et al. (1996) consider the Poisson regression model and present results indicating additional inconsistency problems under the combination of measurement errors and multicollinearity. Their research suggests potential problems with the validity of the results obtained in studies based on applications of the Poisson regression model.

In this paper, a travel frequency model (e.g., Hausman et al. 1995) assuming a Poisson distribution with potential measurement errors is studied. The data used contain information about the number of recreational and shopping (purchase) trips made by respondents. Recreational trips are made for leisure activities, while purchase trips are made for acquiring goods or services. One topic addressed here is the degree of bias introduced in the estimates of the Poisson regression model when explanatory variables are measured with error. The results presented in this paper add to those of Zidek et al. (1996) in that the effects of measurement errors are studied within a different dataset.

The structure of the data makes it likely that the observations of the same individual are correlated and possibly negatively correlated. Tests of random effects can be applied to confirm correlation between individual's observations. However, le Cessie and van Houwelingen (1995) suggest the use of tests of random effects for model specification testing. A second topic addressed in this paper is how the performance of tests of random effects is effected in models of repeated measurement data under measurement errors in explanatory variables. This problem is interesting because it is likely that tests of random effects are sensitive to different types of misspecifica-

tion (e.g., le Cessie and van Houwelingen 1995). It is, therefore, possible that the tests are sensitive to measurement error as well. The tests considered are two score tests by Jacqmin-Gadda and Commenges (1995) and a test proposed by Häggström Lundevaller and Laitila (2002).

In travel habit surveys, reported travel frequencies can be negatively correlated. One explanation is that a trip of one type makes less time available for a trip of another type. However, the correlation structure from an ordinary random effects specification is positive. Here, the performance of tests of random effects under negative correlation is studied.

## MODEL AND MEASUREMENT ERRORS

Suppose  $T$  measurements are obtained from each of  $n$  respondents. Let  $y_{it}$  denote the  $t$ th measurement from the  $i$ th respondent ( $i = 1, \dots, n; t = 1, \dots, T$ ). Assuming no random effects and that  $y_{it}$  is Poisson distributed with mean  $\mu_{it}$  then the probability density function is

$$\Pr(y_{it}) = e^{-\mu_{it}} \mu_{it}^{y_{it}} / (y_{it}!) \quad y_{it} \in \{0, 1, 2, \dots\}$$

In the generalized linear models context (GLM), the mean  $\mu_{it}$  can be related to a vector of explanatory variables  $x_{it}$  as  $\mu_{it} = \exp(x_{it}'\beta)$ , or equivalently  $\ln \mu_{it} = x_{it}'\beta$  (see McCullagh and Nelder 1989). Estimation of  $\beta$  can be performed by using the maximum likelihood (ML) estimator (see Mad-dala 1983; McCullagh and Nelder 1989).

A frequent problem in regression analysis, linear or nonlinear, is measurement errors in explanatory variables. Fuller (1987) gives an introduction to this subject for linear models and Carrol et al. (1995) treat the problem in the case of nonlinear models. The effect of measurement error in combination with multicollinearity in Poisson regression has been considered by Zidek et al. (1996). They demonstrate that the combination of measurement error and multicollinearity can cause misleading estimation results. The effect of important explanatory variables may be overlooked in an analysis, while the importance of other explanatory variables may be overstated.

In the classical measurement-error model, the explanatory variables with measurement error are assumed to be measured as the true value plus an additive error. This model can be expressed as

$$z_{it}^{obs} = z_{it}^* + u_{it}$$

where

$z_{it}^{obs}$  denotes the observed value,

$z_{it}^*$  denotes the true value, and

$u_{it}$  is a random error term.

The other variables are assumed to be measured without error and are denoted by  $x_{it}$ . In the Poisson regression case, the mean function can be written as

$$E[y|x_{it}, z_{it}^*] = \exp[x_{it}\beta + z_{it}^*\gamma]$$

In the applied model,  $z_{it}^*$  is replaced by  $z_{it}^{obs}$  due to measurement errors and the mean of the applied model equals

$$E[y|x_{it}, z_{it}^{obs}] = \exp[x_{it}\beta + z_{it}^{obs}\gamma]$$

The measurement error  $u_{it}$  can be expressed as

$$u_{it} = \xi_{it} + \rho z_{it}^{obs}$$

where  $\xi_{it} = u_{it} - \rho z_{it}^{obs}$  is not correlated with  $z_{it}^{obs}$  and  $\rho$  is a constant. This yields

$$z_{it}^* = z_{it}^{obs}(1 - \rho) - \xi_{it}$$

Thus,

$$E[y|x_{it}, z_{it}^{obs}, z_{it}^*] = \exp[x_{it}\beta + z_{it}^{obs}(1 - \rho)\gamma - \xi_{it}\gamma]$$

and

$$E[y|x_{it}, z_{it}^{obs}] = E_{\xi}[\exp[x_{it}\beta + z_{it}^{obs}(1 - \rho)\gamma - \xi_{it}\gamma]]$$

An indication of the combined effect of measurement error and multicollinearity can be obtained from this last expression. For instance, if  $x_{it}$  and  $z_{it}^*$  are independent and both  $z_{it}^*$  and  $u_{it}$  are normally distributed, then  $z_{it}^{obs}$  and  $\xi_{it}$  are independent and

$$E[y|x_{it}, z_{it}^{obs}] = \exp[x_{it}\beta + z_{it}^{obs}\gamma(1 - \rho)]E_{\xi}\exp[-\xi_{it}\gamma]$$

Due to independence,  $E_{\xi}\exp[-\xi_{it}\gamma]$  is not a function of  $x_{it}$  or  $z_{it}^{obs}$ . Thus, estimation yields consis-

tent estimates of slope coefficients in  $\beta$  and the coefficient  $\gamma(1 - \rho)$ .

Another situation is when  $x_{it}$  and  $z_{it}^*$  are independent, but the distributions of  $u_{it}$  and  $z_{it}^*$  are such that  $z_{it}^{obs}$  and  $\xi_{it}$  are dependent. Then  $E_{\xi}\exp[-\xi_{it}\gamma]$  is a function of  $z_{it}^{obs}$ , in general, and  $\gamma(1 - \rho)$  is inconsistently estimated while the estimator of the slopes in  $\beta$  is consistent. A third case is when  $z_{it}^*$  and  $x_{it}$  are correlated. This will cause both  $\beta$  and  $\gamma(1 - \rho)$  to be inconsistently estimated, in general, because the conditional expectation  $E_{\xi}\exp[-\xi_{it}\gamma]$  is a function of both  $z_{it}^{obs}$  and  $x_{it}$ .

Another model is when the measurement errors can be written as

$$z_{it}^* = z_{it}^{obs} + u_{it}$$

which is a simple form of what is often called the Berkson model (Carroll et al. 1995). This model is applicable, for example, in a controlled experiment where the administered doses are fixed but the actual uptake can vary randomly. Another example is if a distance variable is measured as the distance between two points on a map, while the relevant distance is the road distance.

If  $u_{it}$  is independent of  $z_{it}^{obs}$ , then the mean of  $y$  conditional on  $x_{it}$  and  $z_{it}^{obs}$  is

$$E[y|(x_{it}, z_{it}^{obs})] = \exp[x_{it}\beta + z_{it}^{obs}\gamma]E_u[\exp[u_{it}\gamma]]$$

The Berkson measurement error model often makes it reasonable to assume that  $z_{it}^{obs}$  and  $u_{it}$  are independent, which allows unbiased estimates of  $\beta$  and  $\gamma$ . However, if  $u_{it}$  and  $z_{it}^{obs}$  are not independent, for example if  $z_{it}^{obs}$  is constant for all observations of an individual,  $z_{it}^{obs} = z_i^{obs}$  and  $u_{it}$  are dependent on an individual, then  $\gamma$  will not be possible to estimate without bias.

## TESTS OF RANDOM EFFECTS

In the case of repeated measurements, the assumption of independence may not be feasible, and observations within individuals tend to be correlated. One approach for modeling a correlation structure in repeated measurements data is to include a random individual specific component (random effect) in the linear predictor function. That is, let  $\alpha_i$  denote the

random effect component, then the mean function is written as

$$\mu_{it} = \exp(x'_{it}\beta + \alpha_i)$$

(see Lindsey 1995). Note that this model is equivalent to a model with a multiplicative random effect (see Cameron and Trivedi 1998) because

$$\mu_{it} = \exp(x'_{it}\beta + \alpha_i) = \exp(\alpha_i)\exp(x'_{it}\beta)$$

Inclusion of a random effects component into the model makes efficient estimation more complicated. The contribution to the likelihood from an individual is

$$L_i(\beta; y_i) = \int \prod_{t=1}^T f(y_{it} | \alpha_i; \beta) b(\alpha_i) d\alpha_i$$

where  $b(\alpha_i)$  denotes the distribution of the random effects. In general, this integral is not analytically solvable. One exception is obtained if  $b(\alpha_i)$  is the gamma distribution, which is conjugate to the Poisson distribution. The integral can then be solved and standard methods of estimation are available. For other choices of  $b(\alpha_i)$  several analytical as well as simulation-based approximations have been suggested (see Cameron and Trivedi 1998). However, if the distribution of the random effects is treated as a nuisance component in the model, standard ML methods yield consistent but inefficient estimates of the coefficients in  $\beta$ , except for the intercept term (Liang and Zeger 1986).

Several tests of random effects have been proposed. Breusch and Pagan (1980) derive a score test (the BP test) for the linear regression model with normally distributed disturbances and random individual effects. Honda (1985) proposes the signed square root of the BP test statistic as a new statistic for the test of random effects. Honda's test is robust against nonnormality and is more powerful than the original BP test. Jacqmin-Gadda and Commenges (1995) propose a score test of random effects in GLMs.

For the Poisson regression model, the test statistic proposed by Jacqmin-Gadda and Commenges (1995) is

$$H_{S_1} = \left( \sum_i \sum_{t \neq t'} \hat{\mu}_{it} \hat{\mu}_{it'} \right)^{-\frac{1}{2}} \sum_i \sum_{t \neq t'} \hat{u}_{it} \hat{u}_{it'}$$

where

$$\hat{u}_{it} = y_{it} - \hat{\mu}_{it}$$

To obtain a statistic that is robust to overdispersion when  $\phi$  is unknown, Jacqmin-Gadda and Commenges (1995) suggest the statistic

$$H_\phi = \hat{\phi}^{-1} H_{S_1}$$

where

$$\hat{\phi} = \sum_i \sum_t \hat{u}_{it}^2 \left[ \sum_i \sum_t \hat{u}_{it} \right]^{-1}$$

is a measure of overdispersion.

For the linear regression model, Häggström Lundevaller and Laitila (2002) propose the test statistic

$$W_n = \hat{V}^{-\frac{1}{2}} \sum_i \sum_{t \neq t'} \hat{u}_{it} \hat{u}_{it'}$$

where

$$\hat{V} = 2 \sum_i \sum_{t \neq t'} \hat{u}_{it}^2 \hat{u}_{it'}^2$$

The test statistic is designed to be robust against potential heteroskedasticity. The simulation results reveal that the test works well for testing of random effects in the Poisson regression models considered here. All three statistics,  $H_{S_1}$ ,  $H_\phi$ , and  $W_n$  are compared with the standard normal distribution, and the null hypothesis is rejected for large positive values.

These tests are derived to detect correlation that appears in random effects models. However, they are sensitive to correlation within individual observations regardless of the causes. Model misspecification that leads to correlation structures are likely to affect the tests. Here, the tests are applied to detect negative correlation.

## SIMULATION STUDY

To evaluate the effect of measurement errors on bias of parameter estimates and tests of random effects, a simulation study was done. The idea of the simulations is to take a random sample from a large travel survey database and record the explanatory variables for the sampled observations. These sampled explanatory variables are then used to simulate

**TABLE 1 Maximum Likelihood Estimates of the Model Assuming No Random Effects (Using the Complete Dataset)**

Parameters	Estimate	Standard error	Estimate
			Standard error
<b>Purchase parameters</b>			
$\beta_0$ (Intercept)	-1.5350	0.3795	-4.05
$\beta_1$ (Age)	0.0064	0.0007	9.41
$\beta_2$ (Gender)	0.1562	0.0265	5.89
$\beta_3$ (Income in millions SKR)	0.2488	0.1625	1.53
$\beta_4$ (Petrol price)	-0.0006	0.0005	-1.19
<b>Recreational parameters</b>			
$\beta_5$ (Intercept)	-0.7762	0.2481	-3.13
$\beta_6$ (Age)	-0.0048	0.0005	-10.76
$\beta_7$ (Gender)	-0.0629	0.0171	-3.67
$\beta_8$ (Income in millions SKR)	-0.3640	0.1070	-3.41
$\beta_9$ (Petrol price)	0.0005	0.0003	1.50

new response variables employing estimates from the whole dataset as “true” parameters. Random individual effects and measurement errors can be introduced in the simulated model. The model is then re-estimated using simulated data and the effect of measurement error can be evaluated because we know the “true” values of parameters, random effects, and the measurement error.

### Simulation Design

The data used in this simulation study are taken from the national travel survey (Statistics Sweden 1999); this survey is based on telephone interviews of samples of the Swedish population. The data were collected between April 1994 and December 1998 on a daily basis; 37,754 observations were recorded. The data collected consist of variables related to travel. The simulation study uses only observations with no partial nonresponse and with annual incomes of less than 900,000 SEK. The final dataset contains 30,775 observations.

Observations in the simulation study were obtained from the frequencies of purchase and recreational trips reported by the respondents. The individuals are indexed with  $i$  and trip purpose is indexed with  $t$ , where  $t = 1$  denotes a recreational trip and  $t = 2$  a purchase trip. Age (*Age*), gender (*Gen*), income (*Inc*), and a price index for petrol (*PP*) are used as explanatory variables. Both income

and the index of petrol price are deflated by the consumer price index.

To assess the magnitude of multicollinearity, the multiple  $R^2$  for each of these variables, using the other variables as explanatory variables is calculated. The results are 0.1838 (age), 0.0723 (gender), 0.2299 (income), and 0.0002 (petrol), which indicates a rather small multicollinearity problem. The model without measurement errors is as follows:

$$\exp(x'_{it}\beta + \alpha_i) = \exp(\beta_0 D_P + \beta_1 D_P \text{Age} + \beta_2 D_P \text{Gen} + \beta_3 D_P \text{Inc} + \beta_4 D_P \text{PP} + \beta_5 D_R + \beta_6 D_R \text{Age} + \beta_7 D_R \text{Gen} + \beta_8 D_R \text{Inc} + \beta_9 D_R \text{PP} + \alpha_i)$$

where  $D_R$  is a dummy variable for recreational trips,  $D_P$  is a dummy variable for purchase trips, and  $\alpha_i$  is the random effects component. The variables  $D_P \text{Age}$ ,  $D_R \text{Age}$ , and so on denote the original variables multiplied by the dummy variable. The estimates of the model, assuming no random effects, obtained from the complete dataset are reported in table 1.

In the simulations, the parameters given in table 1 are used as the true parameters. A sample of 1,000 individuals was taken with replacement from the whole dataset, and the explanatory variables for these individuals are recorded. In the case of no measurement error, an observation of the response variable,  $y_{it}$  is created by calculating

$$\tilde{\mu}_{it} = \exp(x'_{it}\tilde{\beta} + \alpha_i)$$

where

$\tilde{\beta}$  is the vector containing the estimates in table 1,

$x_{it}$  is the vector with the explanatory variables drawn from the dataset, and

$\alpha_i$  denotes random individual effects that are generated from the normal distribution with mean zero and standard deviation  $\sigma_{\alpha_i}$ .

The levels of the standard deviation considered are  $\sigma_{\alpha_i} = (0, 0.2, 0.4, 0.6)$ . The value  $\tilde{\mu}_{it}$  is then used as the mean in a Poisson distribution from which an observation  $y_{it}$  is generated by simulation.

In the case of classical measurement errors, the procedure for generating data in the simulations is similar to the one described. However, the value of the explanatory variable is contaminated with an additive random error after the response variable  $y_{it}$  has been generated.

Two of the explanatory variables are considered with measurement errors: the income and the petrol price index variables. The measurement errors for the petrol price index are generated as  $N(0, \sigma_p)$

where

$$\sigma_p = (0.1\mu_p, 0.2\mu_p, 0.3\mu_p)$$

where  $\mu_p = 760.3$  is the mean of the petrol price index over the observations. The measurements for income are generated as  $N(0, \sigma_I)$

where  $\sigma_I = (0.1\mu_I, 0.2\mu_I, 0.3\mu_I)$  and

$$\mu_I = 148534.6.$$

In the case of Berkson measurement errors, the value of the explanatory variable generated by sampling an observation from the dataset is stored and used in estimation. However, the simulated responses  $y_{it}$  are generated by adding a random error term to the explanatory variable.

Applying the tests of random effects described earlier to the original dataset and the estimates given in table 1 yield the test statistic values  $H_{S_1} = -8.84$  and  $W_n = -8.72$ . Both these statistics are to be compared with the standard normal distribution. Because the tests are one-sided where evidence against the null hypothesis is found in large positive values, the null hypothesis of no random effects is not rejected. However, the test statistics are negative and indicate a negative correlation between the two response variables. A negative correlation can be

motivated by, for example, time budget constraints (see Feather 1995).

For the study of the properties of tests of random effects under negatively correlated responses, one set of simulations are carried out where the random effect  $\alpha_i$  is added to the linear predictor for shopping trips,  $x'_{i1}\tilde{\beta} + \alpha_i$ , and the same value is subtracted from the linear predictor for recreation trips,  $x'_{i1}\tilde{\beta} - \alpha_i$ . Alternative models for generating negatively correlated responses could be used, but the chosen one is simple and is sufficient for the purposes of this study.

## Results

The results of the simulations are summarized in tables 2 and 3, which show the bias of the parameter estimates when measurement error exists. The sign of the true parameters is shown in a separate column. The results are for the case with no individual effects. The results observed with individual random effects, which are not shown here, are similar to those with no individual random effects. The measure of bias used is the mean of  $100 (\hat{\beta} - \beta) / |\beta|$  over the 2,000 replications. Parameters  $(\beta_0 - \beta_4)$  are related to purchase trips and  $(\beta_5 - \beta_9)$  to recreational trips and are taken from the estimation results using the whole table 1 dataset.

As expected, the simulation results shown in table 2 indicate only small bias under the Berkson measurement errors. In a few cases, especially for the estimates of the income and petrol price parameters, the values are not close to zero but there is no systematic pattern. This can also be seen in table 3 under the Berkson measurement errors where the bias estimates are generally close to zero with the exception of a few cases.

The results under classical measurement errors with measurement errors in petrol price in table 2 show a clear effect on the parameter estimates. The bias measures indicate that the estimates of the petrol price parameter are close to zero when measurement error exists. The intercept terms are also affected. These biased parameter estimates severely reduce the validity of the estimated model. The effect of classical measurement errors in income is less obvious (table 3). Here, a small tendency for the

**TABLE 2 The Bias Measure\* When Measurement Errors Exist in the Petrol Price Index**

Parameters	Classical					Berkson			
	$\sigma_P$ sign	0	$0.1\mu_P$	$0.2\mu_P$	$0.3\mu_P$	0	$0.1\mu_P$	$0.2\mu_P$	$0.3\mu_P$
<b>Purchase</b>									
$\beta_0$ Intercept	(-)	0	-27	-31	-29	1	-2	7	-1
$\beta_1$ Age	(+)	1	0	1	0	1	1	-1	1
$\beta_2$ Gender	(+)	-5	-1	2	-2	-3	2	5	1
$\beta_3$ Income	(+)	-7	0	-6	-18	-15	-3	14	-25
$\beta_4$ Petrol price	(-)	-1	89	102	99	-4	2	-29	5
<b>Recreation</b>									
$\beta_5$ Intercept	(-)	5	43	45	47	-2	0	0	0
$\beta_6$ Age	(-)	0	1	-3	-1	1	1	0	2
$\beta_7$ Gender	(-)	2	-5	3	-1	-3	-5	9	-5
$\beta_8$ Income	(-)	-1	-6	-4	-7	-7	-5	-2	-1
$\beta_9$ Petrol price	(+)	-13	-91	-95	-100	3	-1	0	0

\* This is the mean of  $100 (\hat{\beta} - \beta) / |\beta|$  over the replications.

**TABLE 3 The Bias Measure\* When Measurement Error Exists in the Income Variable**

Parameters	Classical					Berkson			
	$\sigma_I$ sign	0	$0.1\mu_I$	$0.2\mu_I$	$0.3\mu_I$	0	$0.1\mu_I$	$0.2\mu_I$	$0.3\mu_I$
<b>Purchase</b>									
$\beta_0$ Intercept	(-)	2	5	-2	2	-1	1	-2	2
$\beta_1$ Age	(+)	-1	1	1	2	-1	1	0	-1
$\beta_2$ Gender	(+)	3	-2	-1	-2	1	-2	2	2
$\beta_3$ Income	(+)	2	-13	-32	-37	3	-26	-15	-2
$\beta_4$ Petrol price	(-)	-8	-19	6	-6	0	-3	5	-8
<b>Recreation</b>									
$\beta_5$ Intercept	(-)								
$\beta_6$ Age	(-)	-4	1	4	5	-7	7	-3	8
$\beta_7$ Gender	(-)	1	0	-2	0	1	-1	-1	0
$\beta_8$ Income	(-)	6	-3	1	7	8	-2	1	-1
$\beta_9$ Petrol price	(+)	-2	2	13	21	0	-1	8	0

\* This is the mean of  $100 (\hat{\beta} - \beta) / |\beta|$  over the replications.

parameter estimates of the income variable to be closer to zero can be seen. The intercept terms are not much effected.

Table 4 shows the percentages of rejections observed at the 5% nominal significance level for the random effects test statistics in the case of explanatory variables with classical measurement errors. The distributions of measurement errors with the largest variances are compared with the

case of no measurement errors. The results for the other levels of measurement error variances are similar and are not shown.

The results indicate that measurement error does not seriously affect the properties of the test statistics considered. The tests have estimated sizes close to the nominal sizes, and the estimated powers are high and increase with the variance of the random effects component.

**TABLE 4 Percentage of Rejected Replications with Different Levels of Effects and Measurement Error (n = 2,000)**

error	$\sigma_\alpha$	0			$0.3\mu_p$			0		
		$H$	$H_\phi$	$W_n$	$H$	$H_\phi$	$W_n$	$H$	$H_\phi$	$W_n$
C	0	5.8	5.6	4.3	4.4	4.6	3.6	5.6	5.6	4.4
C	0.2	12.4	12.2	9.4	11.5	11.6	8.9	11.1	10.6	8.8
C	0.4	57	52.7	47.5	58.1	54	48.6	56.2	52.2	48
C	0.6	99.4	98.7	98.2	99.2	98.5	97.8	99	98.4	97.9
B	0	5.6	5.8	4.6	3.2	3.2	2.6	5.5	5.4	4.5
B	0.2	12	11.6	9	9.8	9.4	7.4	11.8	11.2	9.3
B	0.4	58.4	54.8	49.1	50.2	47	42.3	58.6	53.7	48.4
B	0.6	99.2	98.6	98.1	98.8	97.9	97	99.2	98.6	98.1

Key: C = classical measurement errors; B = Berkson measurement errors.

Note: The null hypothesis was rejected for large values of the test statistic at 5% level (one-tailed test).

For a study of the performance of the test statistics under negatively correlated responses, the tests were changed to employ double-sided alternative hypotheses. Results for the double-sided tests are shown in table 5, where rejection frequencies under negatively correlated response variables are considered. The table shows that the test statistics performed well according with rejection frequencies close to the nominal level when  $\sigma_\alpha = 0$  and an increasing power when  $\sigma_\alpha = 0$ . The results show slightly higher rejection frequencies for the statistic  $W_n$  when  $\sigma_\alpha$  is 0.2 or 0.4.

## CONCLUSIONS

In this paper, the effect of measurement errors in explanatory variables in a travel frequency model is studied. Of major interest is the degree of bias introduced in the estimates of the Poisson regression. The derivations show that the effects of classical measurement errors are potentially more severe than those obtained from the Berkson type measurement errors. This result is also confirmed by the simulation results where only small relative biases are observed when Berkson errors are introduced.

The results are different for classical measurement errors. In this case, the intercept terms and the parameter estimates for the variable effected by measurement error are influenced by the measurement errors. This means that the parameter estimates for these variables are in serious doubt if classical measurement errors are suspected. The

results for both Berkson and classical measurement errors confirm the findings in Zidek et al. (1996).

Problems due to the combination of multicollinearity and measurement errors are not observed in the results. The  $R^2$  value for the regression of petrol price on the other explanatory variables is low, we do not expect any difficulty. The  $R^2$  value for the income variable regressed on the other variables are a bit higher (0.2299), but still no effect can be observed.

Another problem that is addressed is the performance of tests of random effects in models of repeated measurement data under measurement errors in explanatory variables. The results suggest that the properties of the tests of random effects considered are not severely effected by measurement errors. The measurement errors are here assumed independent of the true explanatory variables.

The tests are also indicated to be potential candidates for tests of correlation of a more general form than the one obtained by the random effect specification. The properties of such tests under negative correlation among responses was studied. All statistics performed well in the simulations of negatively correlated response variables indicating that the tests can be used to test for negative correlation even though they have been suggested for positive correlation. Of special interest is the good performance of the test,  $W_m$ , proposed by Häggström Lundvall and Laitila (2002). This test was initially derived for the linear regression case, but these

**TABLE 5 Percentage of Rejected Under Negatively Correlated Individual Effects and Different Levels of Individual Effects and Measurement Errors**

error	$\sigma_\alpha$	0			0.3 $\mu_P$			0		
		$H$	$H_\phi$	$W_n$	$H$	$H_\phi$	$W_n$	$H$	$H_\phi$	$W_n$
C	0	4.7	4.8	5.4	4.8	5	5	5	4.8	5.6
C	0.2	6	6	7.2	5.8	5.2	7.2	5.8	5.4	7.8
C	0.4	37	30.7	40	36.4	31.6	39.7	36.4	31.8	39.6
C	0.6	96	90.8	95	95.6	91.6	94.7	96	91	94.6
B	0	5	5.2	5.2	5.4	5.5	6.6	5.4	5.2	5.6
B	0.2	7	6.8	8.6	8.4	7.4	10.4	6.2	6	8
B	0.4	37.8	31.8	39.8	42.2	35.6	43.8	38.4	32.4	40.6
B	0.6	95.7	90.3	94.4	96.8	92.8	96.2	96	92.2	95

Key: C = classical measurement errors; B = Berkson measurement errors.

Note: The null hypothesis was rejected for large and small values of the test statistic at the 5% level (two-tailed test).

results indicate that it can be used for Poisson regression also. However, further studies on the properties of the test in nonlinear models is needed.

## ACKNOWLEDGMENTS

The author gratefully acknowledges the financial support from the Swedish Communications and Transportation Research Board and the Swedish Agency for Innovation Systems.

## REFERENCES

Breusch, T.S. and A.R. Pagan. 1980. The Lagrange Multiplier Test and Its Applications to Model Specification in Econometrics. *Review of Economic Studies* 47:239–253.

Cameron, A.C. and P.K. Trivedi. 1998. *Regression Analysis of Count Data*. Cambridge, England: Cambridge University Press.

Carrol, R.J., D. Ruppert, and L.A. Stefanski. 1995. *Measurement Error in Nonlinear Models*. London, England: Chapman & Hall.

Feather, P. 1995. A Discrete-Count Model for Recreational Demand. *Journal of Environmental Economics and Management* 29:214–227.

Fuller, W. 1987. *Measurement Error Models*. New York, NY: Wiley.

Häggström Lundevaller, E. and T. Laitila. 2002. Test of Random Subject Effects in Heteroskedastic Linear Models. *Biometrical Journal* 44(7):825–834.

Hausman, J.A., G.K. Leonard, and D. McFadden. 1995. A Utility Consistent, Combined Discrete Choice and Count Data Model Assessing Recreational Use Losses Due to Natural Resource Damage. *Journal of Public Economics* 56:1–30.

Honda, Y. 1985. Testing the Error Components Model with Non-Normal Disturbances. *Review of Economic Studies* 52:681–690.

Jacqmin-Gadda, H. and D. Commenges. 1995. Tests of Homogeneity for Generalized Linear Models. *Journal of the American Statistical Association* 90:1237–1246.

Le Cessie, S. and H.C. van Houwelingen. 1995. Testing the Fit of a Regression Model Via Score Tests in Random Effects Models. *Biometrics* 51:600–614.

Liang, K.-Y. and S.L. Zeger. 1986. Longitudinal Data Analysis Using Generalized Linear Models. *Biometrika* 73:13–22.

Lindsey, J.K. 1995. *Modelling Frequency and Count Data*. New York, NY: Oxford University Press.

Maddala, G. 1983. *Limited Dependent and Qualitative Variables in Econometrics*. Cambridge, England: Cambridge University Press.

McCullagh, P. and J.A. Nelder 1989. *Generalized Linear Models*. London, England: Chapman & Hall.

Statistics Sweden, Swedish Institute for Transport and Communications Analysis. 1999. National Travel Survey.

Zidek, J.V., H. Wong, N.D. Le, and R. Burnett. 1996. Causality, Measurement Error and Multicollinearity in Epidemiology. *Environmetrics* 7:441–451.



# Measuring Variability in Urban Traffic Flow by Use of Principal Component Analysis

THEODORE TSEKERIS  
ANTONY STATHOPOULOS\*

---

Department of Transportation  
Planning and Engineering  
School of Civil Engineering  
National Technical University of Athens  
5 Iroon Polytechniou  
157 73 Athens, Greece

---

## ABSTRACT

This paper presents a new approach for the spatio-temporal analysis of variation in traffic flow. Traffic detectors located in several arterial links of an extended urban network yield the time series of aggregate data used in the approach, which is based on the Principal Component Analysis (PCA) of these time series spanning several weeks. The analysis demonstrates the small variability in traffic flow over the whole network. The statistical analysis of common sources of temporal variation in traffic flow provides considerable insight into the properties of long-term flow dynamics. The approach was found to be capable of identifying the location and the impact of extreme events in the network.

## INTRODUCTION

The increasing availability of traffic flow information from archived and readily available sources, such as inductive loop detectors, prompts the ongoing development of data fusion and processing techniques for the fast and efficient analysis of network congestion problems. A major issue in tackling such problems is the measurement of the spatial and temporal variations in traffic flow. These variations are exceedingly useful as input to a wide variety of

---

Email addresses:

\*Corresponding author—astath@transport.ntua.gr  
T. Tserkeris—fabtse@central.ntua.gr

---

KEYWORDS: Traffic flow variability, urban networks, principal component analysis, smoothing models.

applications. Such applications include advanced systems that provide traffic information to travelers, the identification of erroneous traffic forecasts and extreme events (outliers) such as incident detection, the validation of traffic simulation models, and network capacity planning. Other applications refer to the design and evaluation of traffic management strategies, including traffic control and pricing policies, and the assessment of their environmental effects.

Nonetheless, existing applications of methods used for measuring traffic variability are mostly focused on freeways (Rakha and Van Aerde 1995), and they typically refer to a short temporal scale of analysis, ranging from a few seconds to several hours (Treiber and Helbing 2002). Moreover, the usage of correction factors to measure daily or monthly traffic variations based on annual average daily traffic (AADT) estimates (Sharma et al. 1996; Davis 1997), as obtained from traffic counts of medium-time period (usually of 24-hour period), cannot provide an indepth explanation of the sources contributing to the variability in urban traffic.

The investigation of traffic variability in urban arterial networks over long periods of analysis, spanning several weeks or months, can provide promising insight to the potential of the aforementioned applications to alleviate increasing congestion problems. Stathopoulos and Karlaftis (2001) first examined the spatio-temporal variations of traffic flow in a real urban network, the road network of the Greater Athens Area (GAA), Greece, by presenting an exploratory analysis of the distribution characteristics of a set of traffic measurements collected over a period of several months. Also, Weijermars and van Berkum (2004) presented an analysis of variance (ANOVA) of traffic flow along an urban route across a series of weekdays, based on the assumption that flows follow a normal distribution.

This paper describes a novel, interpretive approach for the simultaneous modeling of network-wide traffic flow time series collected over a one-month period from traffic detectors located at major arterial links. The approach, which is based on the general theory of linear algebra, explicitly recognizes the fact that some of these time series are both tempo-

rally and spatially correlated in the network, without relying on any a priori assumption concerning the distribution of traffic flows. More specifically, the method of Principal Component Analysis (PCA), also known as Singular Value Decomposition (SVD) (Meyer 2000), is applied in order to disentangle the intricate sources of long-term traffic dynamics manifested in large-scale urban networks, such as the GAA network.

This is achieved by identifying common underlying sources of temporal variability in traffic flow, which are obtained by estimating the eigenflows, originally defined in (Lakhina et al. 2004) to describe variations in origin-destination (OD) flows of Internet networks. An eigenflow is a time series that captures a common pattern (or source) of temporal variability in traffic flow at the network level. Each traffic flow time series is expressed as a weighted sum of eigenflows and the corresponding weights reflect the extent to which each source of temporal variability is present in the given traffic flow. The method of PCA in the context of traffic flows is analytically described in the second section.

The third section presents the traffic detector data used for the purposes of analysis. The fourth section describes how PCA can be employed to measure the variability of individual traffic flows and of aggregate network traffic, and implications of this measurement for traffic data reconstruction and traffic flow prediction. The fifth section provides a method for decomposing eigenflows to identify different sources of variability in traffic flow. Applications of this method for traffic modeling and incident detection in extended urban networks are also reported. The final section concludes the findings of the study.

## **METHOD OF PRINCIPAL COMPONENT ANALYSIS**

The method of PCA provides the transformation (or mapping) of a dataset onto a new set of principal axes or components. These axes are ordered by the amount of variation (or energy) that they capture in the data. Namely, the first principal axis captures the maximum amount of variation that is possible to represent on a single axis. Each of the remaining

principal axes captures sequentially the maximum residual variation not captured by the preceding axes. In this way, the PCA offers a powerful tool for analyzing the total traffic variability in an urban-scale network composed of a large number of dimensions by approximating it within a lower dimensional structure that preserves its important properties.

Let  $m$  be the number of traffic detectors located on a subset of the total set of arterial links of an urban network,  $t$  be the number of successive days (e.g., the respective periods) in which the detector data are collected, and  $\tau$  be the number of time intervals wherein each day is partitioned. The present study refers to realistic large-scale networks composed of thousands of links servicing hundreds of thousands of travelers. Such networks typically involve hundreds of detectorized links with traffic detector data aggregated over small time intervals, such as 15 minutes. Then, a matrix  $X$  can be defined, referred to here as measurement matrix, with  $p$  rows and  $m$  columns, where  $p = \tau \times t$ . Therefore, each column  $i$  of matrix  $X$  denotes the  $i$ -th traffic flow time series, represented by the column vector  $X_i$ , and each row  $j$  denotes the particular point in the time series in which traffic flows have been collected at interval  $j$ .

The calculation of the  $i$ -th principal component,  $v_i$ , is carried out through the spectral decomposition of the matrix  $X^T X$ , which provides a measure of the covariance between traffic flows, as follows:

$$X^T X v_i = \lambda_i v_i, \quad i = 1, \dots, m \quad (1)$$

where  $\lambda_i$  is the non-negative real scalar, known as the eigenvalue, corresponding to principal component  $v_i$ . By convention, the eigenvalues are arranged in order of magnitude, from large to small, so that  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$ . By solving equation (1), the maximum variation of the measurement matrix  $X$  is captured by the first principal axis  $v_1$ . Proceeding recursively, once the first  $i-1$  principal components have been determined, the  $i$ -th principal component corresponds to the maximum variation of the residual, that is the difference between the original data and the data mapped onto the first  $i-1$  principal axes. The arrangement of the set of principal components  $\{v_i\}_{i=1}^m$  in such an order, as columns, results in the

principal matrix  $V$ , which has size  $m \times m$ . By definition, the columns of  $V$  matrix have unit norm, which means that the length of each column vector, as defined by the square root of the sum of squares of all entry values, is equal to unity.

Because the principal axes are arranged in order of contribution to the overall variation, the time-varying trend common to all flows along principal axis  $i$  can be represented through a column vector  $u_i$  with size  $p$ , referred to as the eigenflow of the  $i$ -th principal axis, as follows:

$$u_i = \frac{X \cdot v_i}{\sigma_i}, \quad i = 1, \dots, m \quad (2)$$

where  $\sigma_i = \sqrt{\lambda_i}$  is the singular value corresponding to the  $i$ -th principal axis. The magnitude of singular values demonstrates the overall variation attributable to each particular principal component and, hence, the potential to reconstruct total traffic data using a smaller number of dimensions. The arrangement of the set of eigenflows  $\{u_i\}_{i=1}^m$  as columns in order of decreasing strength of the common temporal trends results in the eigenflow matrix  $U$ , which has size  $p \times m$ . Based on equation (2), it can be shown that each traffic flow time series  $X_i$ , when normalized by the singular value  $\sigma_i$ , is a linear combination of the eigenflows, weighted by the associated principal component. More specifically, the relationship between matrices  $X$ ,  $U$ , and  $V$  can be represented as follows:

$$\frac{X_i}{\sigma_i} = U (V^T)_i, \quad i = 1, \dots, m \quad (3)$$

where  $(V^T)_i$  is the  $i$ -th row of matrix  $V$ . By assuming that only a small number of  $q < m$  singular values is non-negligible (see 2.2), or, in other words, only a small set of  $q$  eigenflows contributes to the bulk of temporal variability in traffic flow, then, the original data, that is, measurement matrix  $X$ , can be approximated as follows:

$$X' \approx \sum_{i=1}^q \sigma_i u_i v_i^T \quad (4)$$

The spatio-temporal reconstruction of matrix  $X$  by use of a lower number of dimensions can enhance the interpretability of the long-term dynamics of

each traffic flow, particularly for the case of large-scale urban networks, as shown in the fourth section.

### TRAFFIC DETECTOR DATA

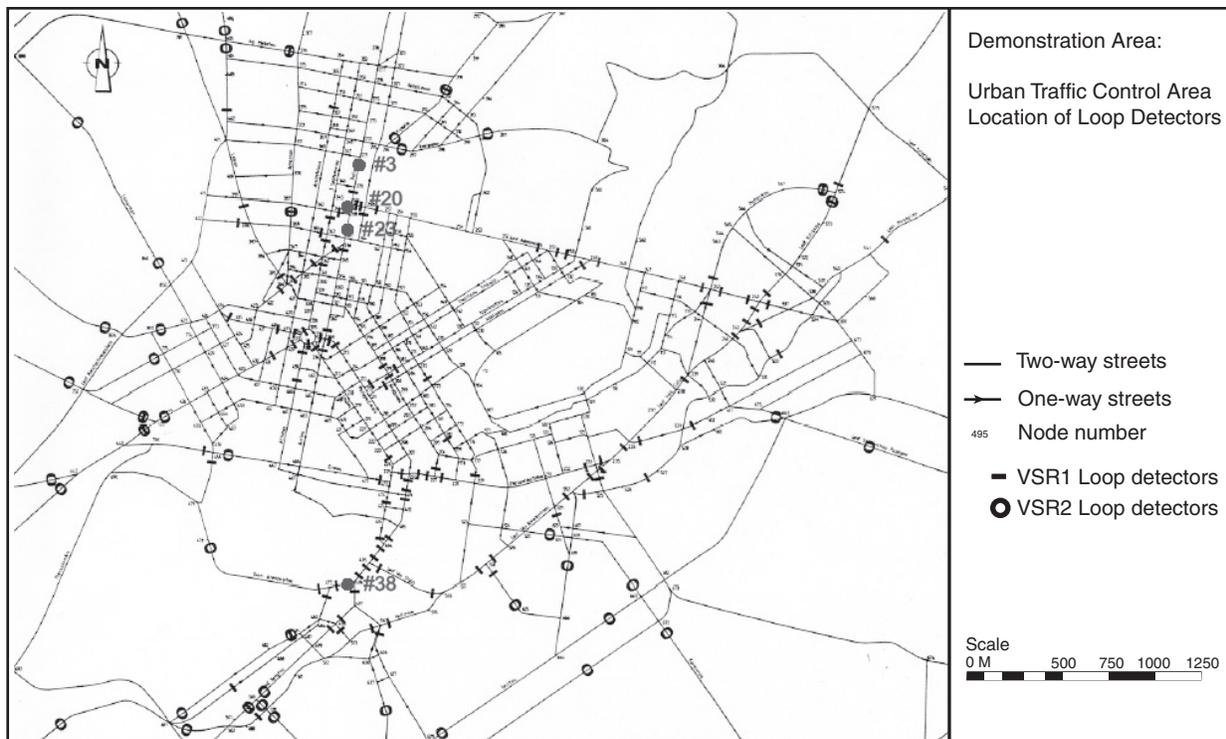
The traffic data used here for analysis purposes are automatically collected using loop detectors at 140 key locations around the urban road network of the Greater Athens Area (GAA), as illustrated in figure 1. These real-time data are stored at the end of every 90-sec signalization cycle and aggregated at time intervals of 15 minute duration. The traffic counting system provides an appropriate data quality control by performing screening and data repair functions so as to identify and exclude or smooth data from malfunctioning detectors. The dataset includes measurements corresponding to the first 28 of the 29 days in February 2000. Each day covers 16 time intervals of the period spanning between 6:00 am and 10:00 am. This yields a total number of 62,720 measurements, that is, a time series of 448 measurements for each of the 140 detector locations.

### USE OF PCA FOR MEASURING TRAFFIC FLOW VARIABILITY

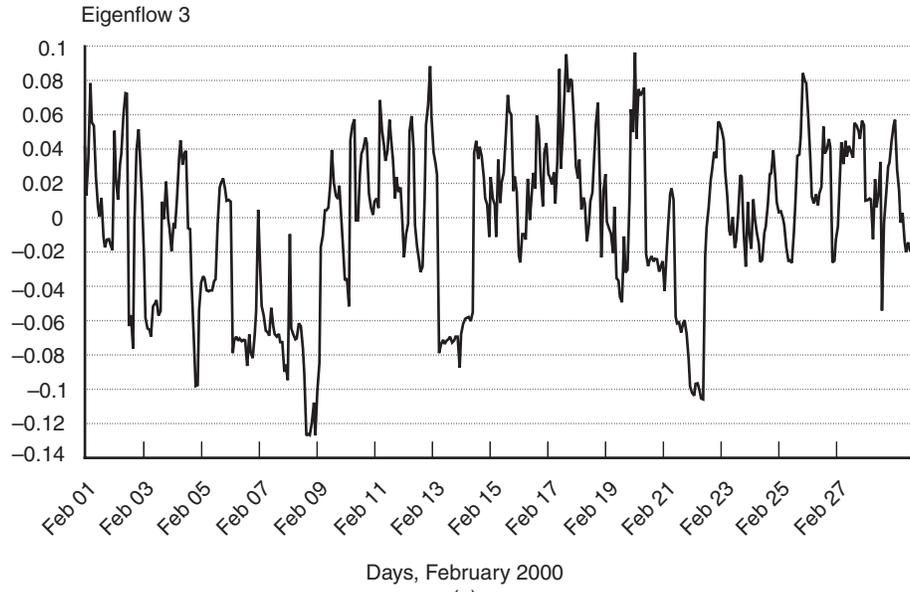
#### Spatio-Temporal Representation of Traffic Variations

Figure 2 presents a typical example of an eigenflow  $u_i$  ( $i=3$ ) and its corresponding principal axis  $v_i$ , as calculated with the PCA of the four-week traffic detector data of the Greater Athens Area. Figure 2(a) demonstrates the representation of a pattern of temporal variation common to all traffic flow time series through eigenflow 3. A later section provides a systematic way for distinguishing different types of this temporal variation. Figure 2(b) illustrates the extent to which this particular temporal pattern is present in each traffic flow, through the entries of the corresponding principal component. Eigenflow 3 is most strongly present at traffic flow measurement point number 38, as shown in figure 2(b), namely, in the time series measured at detector number 38, whose location in the network is marked on figure 1. The immediately next strongest

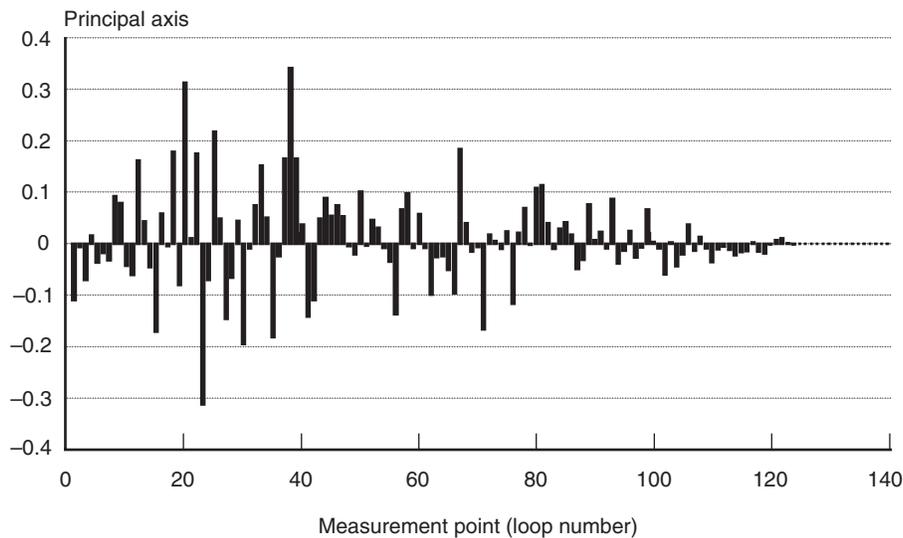
**FIGURE 1** Illustration of the Greater Athens Area (GAA) Network and Configuration of the Location of Loop Detectors



**FIGURE 2 Graphical Representation of (a) an Eigenflow and (b) its Corresponding Principal Axis**



(a)



(b)

temporal trends are those corresponding to traffic flow at number 20 and traffic flow at number 23 (see figure 1).

Based on the definition of eigenflows and principal components (axis) in the second section, the negative sign of many entries in some eigenflows, such as in eigenflow 3, denotes that the corresponding common temporal variation pattern is negatively correlated with some of the measured traffic flow time series. Respectively, the negative sign of many entries in the corresponding principal components indicates the negative value of the covariance

of the traffic flow measured at a specific traffic flow measurement point, such as the measurement point number 23 (see figure 2(b)), with the other traffic flows. Namely, an increase of the traffic flow rate at measurement point number 23 would result in the reduction of the traffic flow rate in the other measurement points. This kind of analysis helps identify different locations in the network as well as periods of the day, days, or weeks wherein a particular traffic flow has a large impact on the aggregate network traffic conditions.

## Measuring the Variability of Individual Flows

Based on the definition given in the first section, each eigenflow can be considered as a building block of the overall dynamics pertaining to each traffic flow. Thus, the variability of individual flows can be determined with regard to the number of significant eigenflows that constitute them. The number of significant eigenflows refers to the number of entries in the corresponding rows of the principal matrix  $V$  that are significantly different from zero. There exists a threshold that is equal to  $1/\sqrt{m}$  when a row of  $V$  has all entries equal, which implies a perfectly equal mixture of all eigenflows, taking into account that the columns of  $V$  have unit norm (Lakhina et al. 2004). Then, the number of significant eigenflows is obtained by counting how many entries in each row of matrix  $V$  exceed this threshold in absolute value. This approach allows determining the least required number of significant eigenflows, dependent on the sample size, which can provide a plausible reconstruction of each traffic flow (see below), based on equation (4).

Figure 3(a) illustrates the number of significant eigenflows that constitute traffic flows, as this is expressed by the Cumulative Density Function (CDF) of the number of entries per row of  $V$  that exceed the above threshold. The curve indicates that no traffic flow is composed of more than 45 significant eigenflows. In particular, it can be observed that 50% of traffic flows are composed of less than 30 significant eigenflows, and more than 30% of traffic flows are composed of less than 20 significant eigenflows. In addition, figure 3(b) presents the histogram of significant eigenflows that constitute traffic flows. This histogram shows that the class interval containing up to 5 significant eigenflows appears most frequently among traffic flows, with the class intervals containing 6 to 10 and 11 to 15 significant eigenflows to follow in order. These results clearly demonstrate that the temporal evolution of most traffic flows can be explained by only a small number of common underlying sources of variability.

Figure 4 shows the number of significant eigenflows with respect to the monthly average daily (for the respective period) traffic flow rate measured over the different detector locations. By and large,

the results demonstrate that there is a relationship between the size of a traffic flow and the eigenflows that comprise it. More specifically, the larger traffic flows tend to be composed primarily of a large number ( $>20$ ) of significant eigenflows (see cluster at the right-hand side of figure 4 separated by a dashed line), in comparison to the smaller traffic flows, which are basically composed of a small number ( $<20$ ) of significant eigenflows (see cluster at the left-hand side of figure 4). Consequently, the temporal variation of the larger flows has the most significant contribution to the long-term dynamics of the aggregate network traffic in relation to the variation of the other flows.

## Measuring the Variability of Aggregate Network Traffic

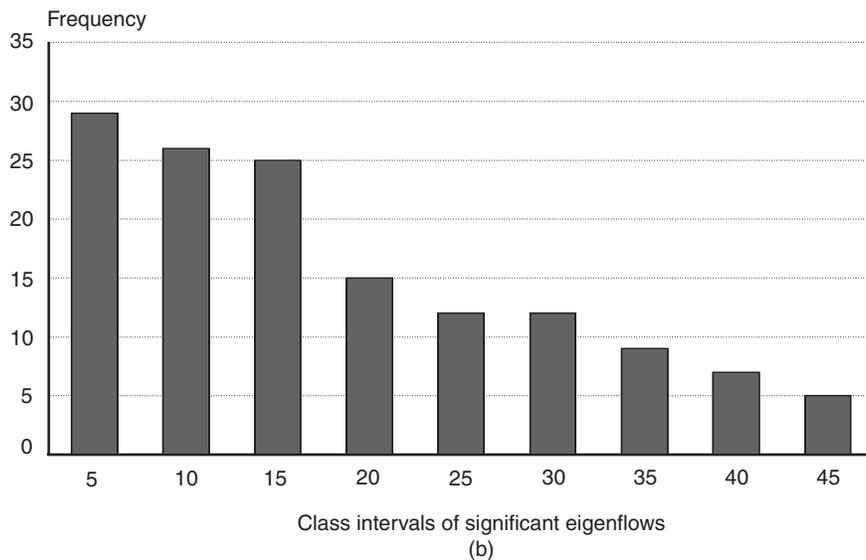
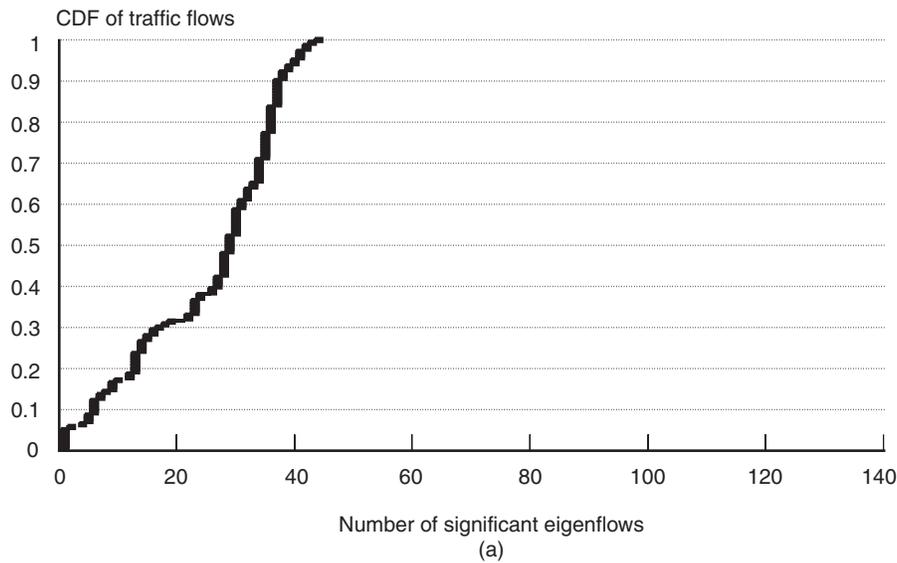
As explained in an earlier section, the singular values denote the overall variation attributable to each particular principal component. Hence, the order of magnitude of singular values can provide a plausible measure of the extent of variability in aggregate network traffic. Figure 5(a) shows the plot of singular values, in order of decreasing magnitude, corresponding to each traffic flow. This plot clearly demonstrates that the variability in traffic flow can be attributed to only a very small number of eigenflows, that is, common patterns of temporal variation. More specifically, the vast majority of traffic variability is contributed by the first few eigenflows, as signified by the sharp knee of the curve between the third and the eighth singular value. This result provides evidence of the small variability (spread) of the aggregate network traffic in the long run.

Given the effect of the size of traffic flow on the variability of individual traffic flows, as described in the previous subsection, the effect of the mean traffic flow rate on the small variability of the aggregate network traffic is also investigated here. For this purpose, a zero-mean normalization is applied, which denotes that all measurements of each time series  $X_i$  are subtracted from the corresponding sample mean so that their average is zero, as follows:

$$X'_i = X_i - \mu_i, \quad i = 1, \dots, m \quad (5)$$

where  $\mu_i = \mu(X_i)$  is the sample mean of time series  $X_i$ . Figure 5(b) shows the plot of singular

**FIGURE 3** (a) Number of Significant Eigenflows, in Terms of the Cumulative Density Function (CDF) of the Number of Entries in Each Row of the Principal Matrix That Exceeded the Threshold and (b) Histogram of Significant Eigenflows

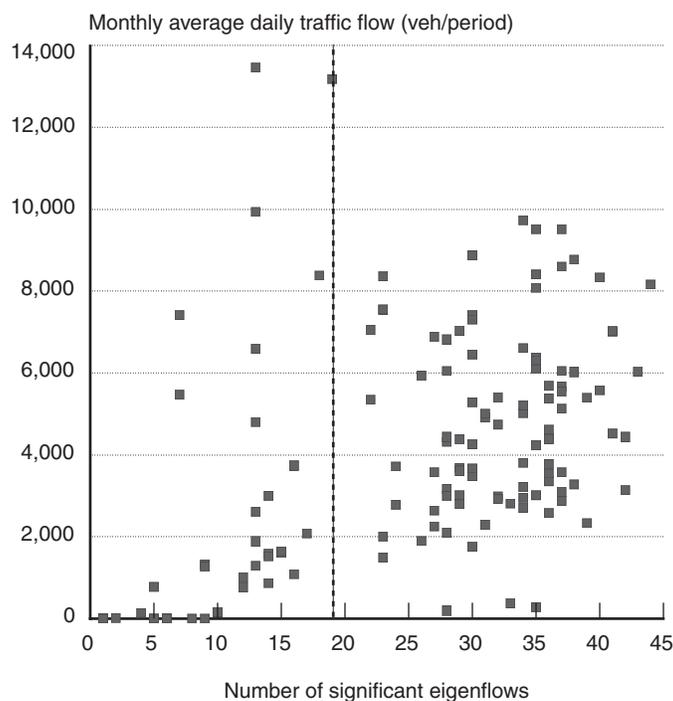


values corresponding to each traffic flow, based on the normalization of traffic flows, as indicated above. In contrast to the case of using the original traffic flows, in the case of using the normalized traffic flows the bulk of variability is signified by a less sharp knee of the curve between the 7th and the 20th singular value. Namely, the relative significance of the first few eigenflows has diminished. This effect can be explained by the fact that a large diversity in the magnitude of the mean traffic flow rate can render the variation of those traffic flows with increased size dominant in com-

parison to the variation of the other flows. In turn, this leads to a larger diversity between the first few singular values and the remaining singular values.

On the other hand, the fact that the profound majority of variability in traffic flow can still be attributed to only a very small number of eigenflows indicates the dominant role of the remaining effect, which is the effect of the correlation between temporal variation patterns in comparison to the effect of differences in flow size. Therefore, the process of normalization can ensure that the representation of

**FIGURE 4 The Number of Significant Eigenflows with Respect to the Monthly Average Daily Traffic Flow Rate**



these correlations by eigenflows is not skewed due to differences in the mean traffic flow rate.

**Implication of Variability Measurement for Traffic Data Reconstruction**

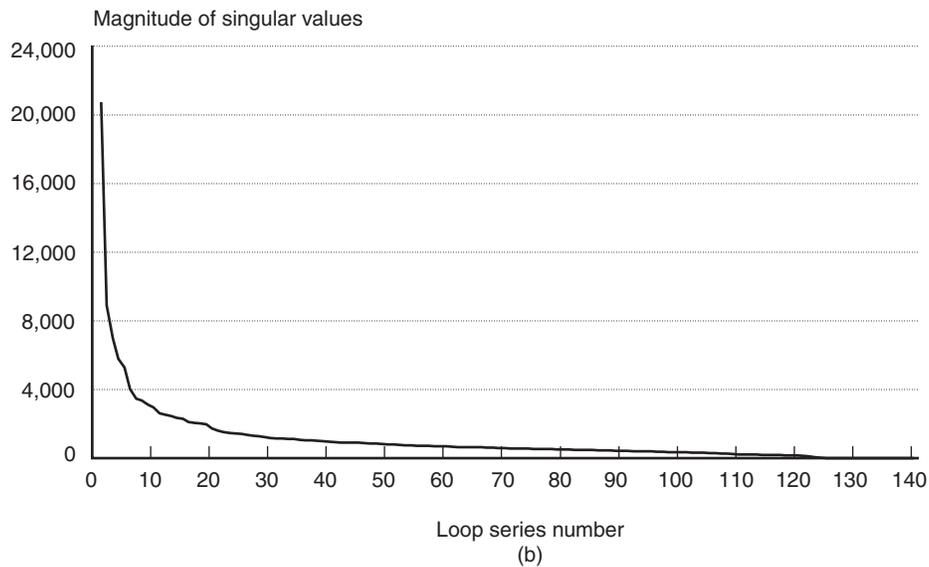
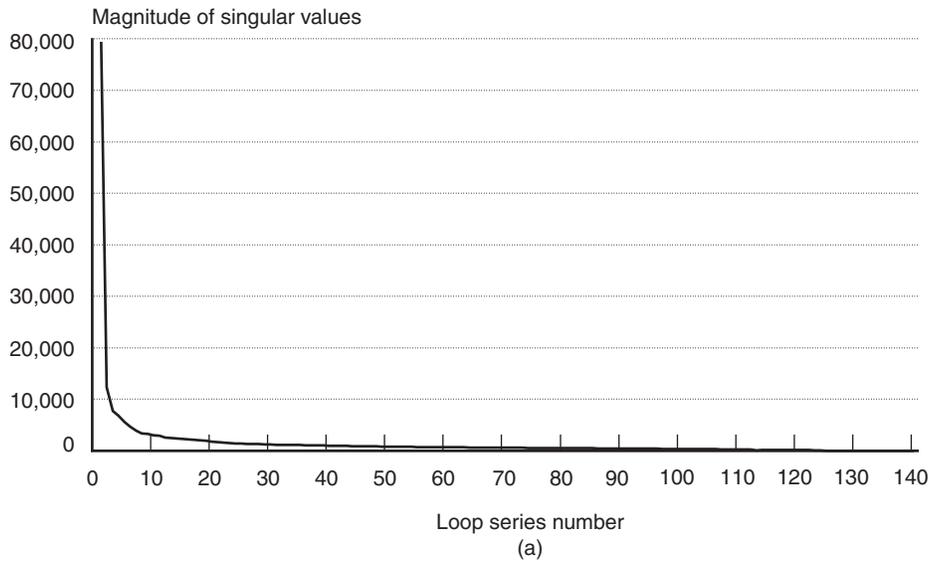
The fact that only a few singular values can depict the largest portion of the overall variation in aggregate network traffic demonstrates the potential to reconstruct traffic flows or approximate each column of the measurement matrix  $X$ , using a considerably smaller number of dimensions. The traffic flows reconstructed by using the whole set of significant eigenflows, on the basis of equation (4), are found to approximate the original (normalized) traffic flows without statistically significant differences at least at the 95% confidence level of the Student  $t$ -test statistic. This outcome indicates the correctness of the previously described method for selecting threshold values to determine the number of significant eigenflows composing each traffic flow. Moreover, traffic flow at measurement point number 3 (see figure 1) is randomly selected here to be approximated by using a number of  $q=5$  dimensions (see figure 6(a)). This traffic flow is composed of 30 significant eigenflows. The graphical analysis

shows that the temporal pattern of the reconstructed traffic flow remarkably resembles the temporal pattern of the original traffic flow.

Figure 6(b) shows the statistical analysis of the regression of the reconstructed traffic flow to the original traffic flow. The reconstructed flow generally underestimates the original flow, as it denotes the value of the slope of the linear trend line, which is lower than unity ( $<1.0$ ). This underestimation refers mainly to traffic flows of lower size ( $<400$  veh/15-min), as this is implied by the outliers corresponding to such flow sizes. This outcome indicates that the first 5 (most significant) eigenflows, which are employed in the reconstruction process, can better capture the temporal variation of larger traffic flows in comparison to the remaining eigenflows.

On the other hand, the  $R^2$  value, which represents the squared multiple correlation between the two datasets, indicates that the reconstructed flow data can capture approximately 80% of the systematic variation contained in the original flow data. The above results emphasize the ability of the proposed method to concentrate on a very small set of common sources of temporal variability in order to describe the complexity of traffic flow. In turn, this

**FIGURE 5 Plot of Singular Values for (a) Traffic Flows and (b) Normalized Traffic Flows**



facilitates the deeper understanding and a more plausible interpretation of the factors contributing to the long-term evolution of the main characteristics of urban network traffic.

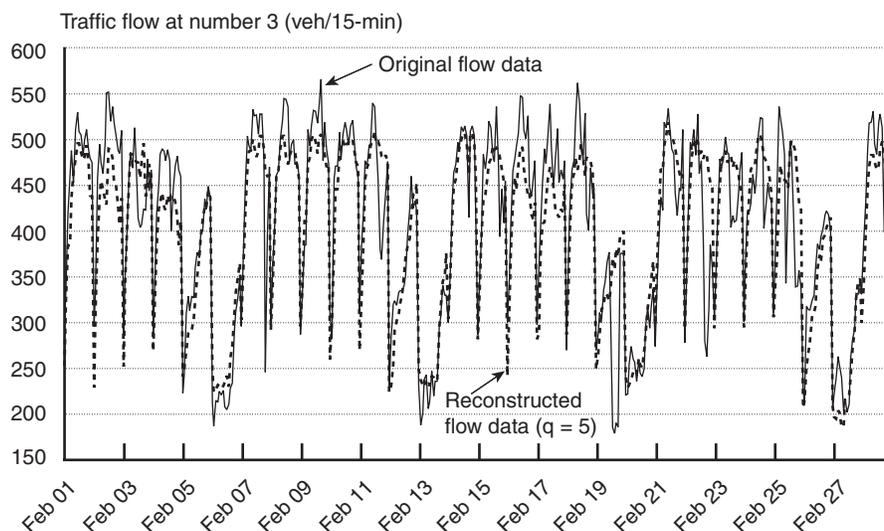
**Implication of Variability Measurement for Traffic Flow Prediction**

The outcome of the previous subsection (that only a very small set of eigenflows is sufficient for the plausible reconstruction of a traffic flow) emphasizes the need for investigating the potential of the PCA method to approximate future traffic flows. This task is addressed by analyzing data that were not part of the input to the PCA procedure. More spe-

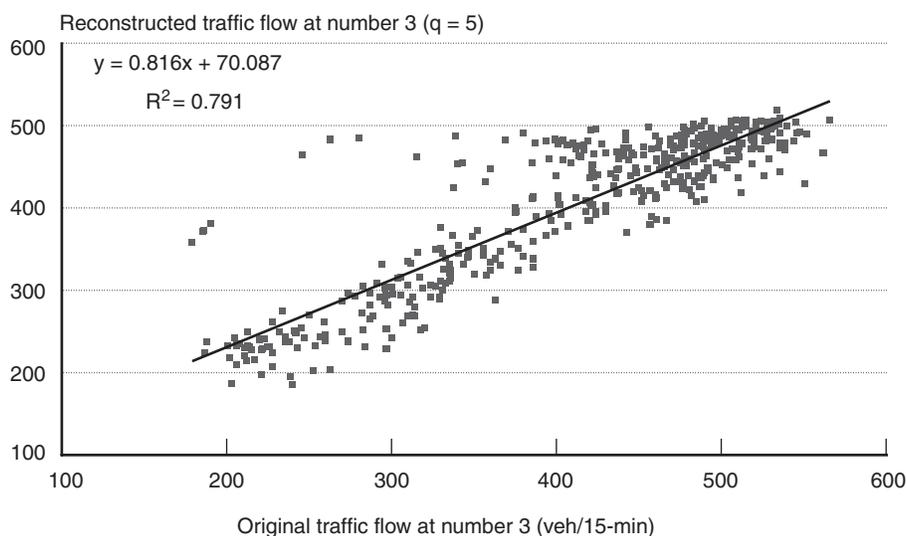
cifically, the PCA method is applied to the traffic data, denoted as  $X^{w1}$ , measured over a time period spanning the week Monday, Feb. 7, 2000, through Sunday, Feb. 13, 2000, to obtain the principal components  $\{\nu_i\}_{i=1}^m$ . Subsequently, these principal components are also used to approximate (predict) traffic flow data, denoted as  $X^{w2}$ , over the next week spanning Monday, Feb. 14, 2000, through Sunday, Feb. 20, 2000.

The error of approximating a typical traffic flow, that is, traffic flow at number 3, as used in the previous subsection, is investigated for both the first and second week, based on the principal components obtained from the PCA of the first-week data. In

**FIGURE 6 (a) Reconstruction of a Typical Traffic Flow Using 5 Principal Components and (b) Statistical Analysis of the Reconstructed Traffic Flow**



(a)



(b)

order to investigate the degree to which the flows of the second week preserve the hierarchical structure of temporal variability pattern pertaining to the flows of the first week, the approximation is carried out by using different dimensions (amounts of principal components), that is,  $q=5$ ,  $q=10$ ,  $q=20$  and  $q=40$ . The traffic flow at number 3 corresponding to the first week, denoted as  $X_3^{w1}$ , is composed of 33 significant eigenflows, while the traffic flow at number 3 corresponding to the second week, denoted as  $X_3^{w2}$ , is composed of 32 significant eigenflows. The

approximation error is measured through (a) the Mean Relative Error (MRE) of the approximation corresponding to each week, which is given as  $|X_3^{w1} - \tilde{X}_3^{w1}| / X_3^{w1}$  and  $|X_3^{w2} - \tilde{X}_3^{w2}| / X_3^{w2}$  respectively, where  $\tilde{X}_3^{w1}$  and  $\tilde{X}_3^{w2}$  are the reconstructed flows for the first and the second week, and (b) the value of the  $R^2$  coefficient.

Table 1 shows the approximation error, as expressed by the measures of MRE (%) and  $R^2$ , for the reconstructed flows of the first and the second week, based on the principal components obtained

**TABLE 1 Values of MRE Resulting from the Reconstruction of a Typical Traffic Flow (in percent)**  
(Values are across 2 successive weeks using principal components of the first week)

Number of dimensions $q$	5	10	20	40
First-week data	<sup>1</sup> 6.40 (0.935)	<sup>2</sup> 5.64 (0.952)	<sup>3</sup> 5.04 (0.965)	3.25 (0.988)
Second-week data	<sup>1</sup> 7.39 (0.904)	<sup>2</sup> 6.50 (0.926)	<sup>3</sup> 6.38 (0.934)	6.26 (0.941)

<sup>1, 2, 3</sup> Pairs with no statistically significant differences at the 95% confidence level of the  $t$ -test statistic

Note: Values in parentheses indicate  $R^2$ .

from the PCA of the first-week data. The results demonstrate that, when using the same number of dimensions ( $q=5$ ), the application of the PCA on a shorter term dataset, such as that of one week, results in a more accurate approximation of the original flows ( $R^2=0.935$ ) in comparison to the application on a longer term dataset, such as that of four weeks ( $R^2=0.791$ ). This outcome can be attributed to the fact that longer term data typically involve more sources (larger spread) of temporal variability in the network.

The approximation of  $\tilde{X}_3^{w2}$  based on the first-week principal components, as well as the approximation of  $\tilde{X}_3^{w1}$ , resulted in low MRE values (<10.0%) and high  $R^2$  values (>0.90). The differences between the MRE values that resulted from the two approximations were not found to be statistically significant at the 95% confidence level of the  $t$ -test statistic for all sets of principal components used, except for the case using 40 principal components (see footnote of table 1). Hence, the first-week principal components can be well used to approximate traffic flows of the second week, such as  $\tilde{X}_3^{w2}$ . Moreover, the loss of the predicting power of the first-week principal components can be attributed to the last few eigenflows, which are mostly related to smaller size and higher variability traffic flows. The results generally provide evidence of the increased temporal stability of the hierarchical pattern of traffic variations from one week to the next. Thus, they indicate the potential of using the first few eigenflows of the previous week to consistently reproduce, with a reasonable accuracy, most systematic features of the traffic flow of the next week.

## DECOMPOSITION OF EIGENFLOWS AND APPLICATIONS

### Method for Decomposing Eigenflows

Each eigenflow can be decomposed into nonstationary and stationary components, according to the nature of variability in traffic flow. For this purpose, each eigenflow is modeled here as an unobserved-components time series in state space form (Koopman et al. 1999), so that they enable the smoothing estimation of both nonstationary and stationary components. The nonstationary component refers to nonstationary variation (or changes) of the eigenflow mean, and it reflects periodicities, namely periodic trends, in traffic flow. These time-varying trends are due to diurnal cycles in travel demand, differences in traffic conditions among weekdays, as well as between weekdays and weekends. This component is calculated here with the state smoothing of each eigenflow, which captures changes in the level or trend of traffic variability in the long run.

The stationary components refer to structural breaks and outliers. These are typically expressed with isolated values, which are located outside a band of, for example,  $\pm 2$  standard errors (SE) from the trend line, that is, the smoothed eigenflow mean. The structural breaks reflect occasional bursts and dips in the level of traffic variability. These breaks correspond to stochastic and transient changes, such as traffic phase transitions, pertaining to the physical dynamics of (recurrent) congestion conditions. The outliers reflect noise, that is the remaining random variation in traffic data. They can be attributed to extreme or unusual traffic-related events, principally related to nonrecurrent congestion dynamics, such

as demonstrations, road works, sportive events, emergency situations, accidents, or other incidents.

The existence, in terms of their statistical significance, of each of these two types of stationary variation is identified here by applying the  $t$ -test statistic to the results obtained from the disturbance smoothing of each eigenflow. The solution of both the state smoothing model and the disturbance smoothing model is carried out by using an appropriate maximization routine written in the O $\alpha$  matrix programming language (Doornik 2002). Further information on the analysis of time series using state and disturbance smoothing models can be found in Durbin and Koopman (2001).

### **Practical Demonstration of the Method**

For demonstration purposes, the proposed method for the decomposition of eigenflows is implemented here for the same typical eigenflow used in the previous section, that is, eigenflow 3, at a finer level of temporal resolution, that is, a period of one week spanning from Monday, Feb. 14, 2000, to Sunday, Feb. 20, 2000. Based on the results of the previous section, the normalized traffic flows are used for the estimation of the temporal trend of the eigenflow. The usage of these data prevents the effect of possible bias caused by differences in flow size between various time intervals of the day as well as periods of successive days-of-the-week. Figure 7a presents the estimation of the temporal trend of the given eigenflow together with the corresponding band of  $\pm 2SE$  from the estimated trend line. In addition, figure 7b illustrates the distinction of structural breaks and outliers, which correspond to the given eigenflow, and their statistical significance, as determined by the range defined between the upper and the lower confidence level, through the  $t$ -test statistical analysis of these two types of stationary variation.

The process which was adapted here, referred to as temporal trend thresholding, through the suitable selection of a band with magnitude  $\pm 2SE$ , appears to provide a reasonable distinction between stationary variations and nonstationary changes from the eigenflow mean. This is clearly demonstrated by the fact that the selection of such a threshold or bandwidth can capture the existence of breaks and outliers because the values that are kept out of this band

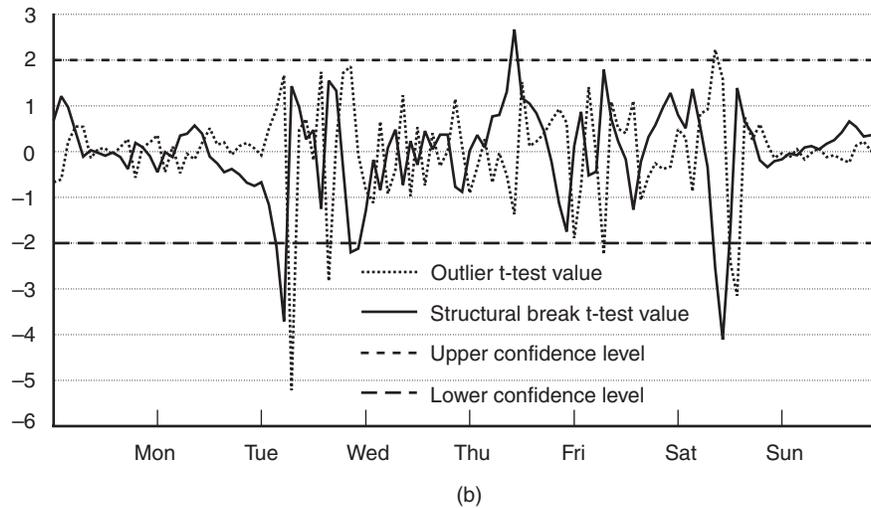
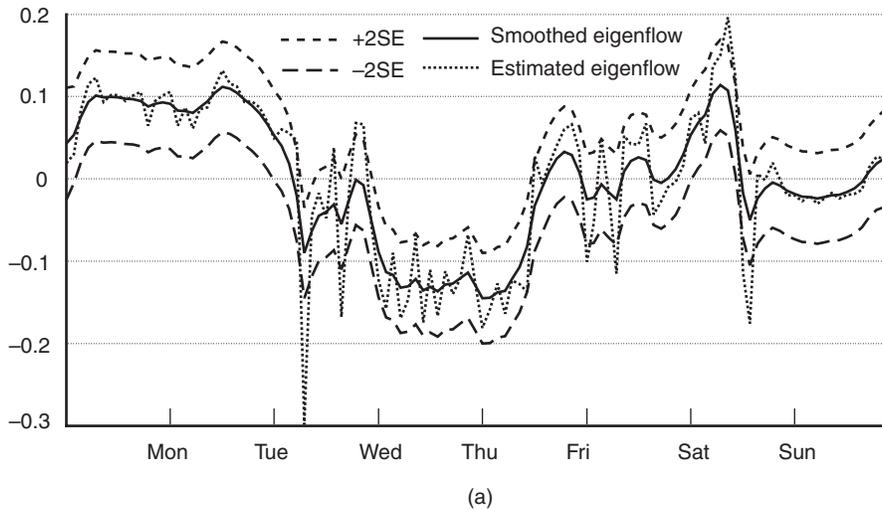
(see figure 7a) correspond to statistically significant stationary variations (figure 7b), based on the  $t$ -test statistic. In the present case, these short-lived, statistically significant outliers are mainly due to bus breakdowns and local accident situations. The above process can also provide information on the extent to which statistically significant breaks and outliers affect changes in the long-term trend of a common variation pattern.

This type of analysis is particularly helpful for understanding the characteristics of urban traffic, provided that most traffic flows can be sufficiently represented through only a small number of eigenflows, as shown in the previous section. On the one hand, the temporal trend thresholding is suitable for the long-term analysis of the expected or predictable variations of traffic, that is, those variations restricted within the band of  $\pm 2SE$ . The determination of the periodic trends of each eigenflow makes it possible to identify the extent to which the variability in traffic flow is predictable within time periods, such as those from week to week.

Furthermore, these results have implications for the verification or validation of traffic assignment and simulation models used to provide traffic predictions. The temporal trend thresholding enables the definition of bounds in which a long-term prediction can be considered as statistically reliable, or, otherwise, an extreme or erroneous forecast, which should be ignored or set equal to the upper or lower statistical bound of the corresponding point in time. In addition, this procedure can help identify whether traffic predictions in some links are more error-prone than others.

On the other hand, the process described here proposes a new scheme for identifying the location and the time of occurrence of statistically significant stationary variations in traffic flow, related to unusual or extreme events, as described in the previous subsection. The operation of the proposed scheme can be updated periodically (e.g., from week to week) in an automated manner by simultaneously reprocessing traffic flows measured over different locations in the network while it enables the statistical analysis of different types of stationary variation, such as structural breaks and outliers. For these reasons, this scheme can be considered as a more rigor-

**FIGURE 7 Decomposition of a Typical Eigenflow by Use of (a) Temporal Trend Thresholding and (b) *t*-test Statistical Analysis of Structural Breaks and Outliers**



ous and practically useful approach than the method of detecting outliers through simply comparing individual flows to some average traffic pattern obtained from past measurements over a given location and period of time (e.g., a week). Moreover, this process introduces a methodology for deriving a set of different models for local traffic prediction across multiple timescales by taking into account the fact that the traffic in different parts or links of the network may experience different rates and types of variation as time progresses.

## CONCLUSIONS

This paper describes the analysis and interpretation of the variability in urban traffic flow by processing

one-month traffic detector data corresponding to a realistic large-scale arterial network. The method of principal component analysis was found to provide a plausible and powerful tool for the purposes of the present study. Specifically, the PCA enables the identification of eigenflows, which denote common patterns of temporal variability, according to their contribution to the aggregate network traffic. Despite the underlying complexity in the phenomenology of urban traffic structure, the findings suggest that the spatio-temporal variation in traffic flow in such a network can be represented by only a small set of eigenflows. This small variability can be attributed to the increased correlation between temporal variation patterns and the presence of periodic

trends in these patterns, and it was found to carry useful implications for the updated prediction of traffic flow patterns, for example, from one week to the next.

Moreover, the statistical analysis of the calculated eigenflows allows for the presence of stationary variations, namely, breaks and outliers. The identification of such variations is particularly valuable for supporting real-time network operations, including detection of traffic anomalies and incidents. In addition, the proposed methodology offers a valuable tool to manage stored aggregate traffic flow data in large-scale urban networks for planning purposes. Such purposes can encompass the assessment of traffic responsive control strategies, the verification of traffic assignment and simulation models used to represent the variation in traffic patterns, the evaluation of the traffic network performance, and its impact on the environment.

## REFERENCES

- Davis, G.A. 1997. Accuracy of Estimates of Mean Daily Traffic: A Review. *Transportation Research Record* 1593:12–16.
- Doornik, J.A. 2002. *Object-Oriented Matrix Programming Using Ox*, 3rd ed. London, England: Timberlake Consultants Press. Available at: <http://www.doornik.com>.
- Durbin, J. and S.J. Koopman. 2001. *Time Series Analysis by State Space Methods*. Oxford, England: Oxford University Press.
- Koopman, S.J., N. Shephard, and J.A. Doornik. 1999. Statistical Algorithms for Models in State Space Using SsfPack 2.2. *Econometrics Journal* 2(1):113–166.
- Lakhina, A., K. Papagiannaki, M. Crovella, C. Diot, E.D. Kolaczyk, and N. Taft. 2004. Structural Analysis of Network Traffic Flows. *Proceedings of the International Conference on Measurements and Modeling of Computer Systems SIGMETRICS*. Edited by E.G. Coffman Jr., Z. Liu, and A. Merchant. New York, NY: ACM.
- Meyer, C.D. 2000. *Matrix Analysis and Applied Linear Algebra*. Philadelphia, PA: SIAM.
- Rakha, H. and M. Van Aerde. 1995. Statistical Analysis of Day-to-Day Variations in Real-Time Traffic Flow Data. *Transportation Research Record* 1510:26–34.
- Sharma, S.C., B.M. Gulati, and S.N. Rizak. 1996. Statewide Traffic Volume Studies and Precision of AADT Estimates. *Journal of Transportation Engineering* 122(6):430–439.
- Stathopoulos, A. and M.G. Karlaftis. 2001. Temporal and Spatial Variations of Real-Time Traffic Data in Urban Areas. *Transportation Research Record* 1768:135–140.
- Treiber, M. and D. Helbing. 2002. Reconstructing the Spatio-Temporal Traffic Dynamics from Stationary Detector Data. *Cooperative Transportation Dynamics* 1:3.1–3.24.
- Weijermars, W.A.M. and E.C. van Berkum. 2004. Daily Flow Profiles of Urban Traffic. *Urban Transport X*. Edited by C.A. Brebbia and L.C. Wadhwa. Southampton, UK: WIT Press.

# Frequency and Severity of Belgian Road Traffic Accidents Studied by State-Space Methods

ELKE HERMANS

GEERT WETS\*

FILIP VAN DEN BOSSCHE

---

Transportation Research Institute  
Hasselt University, Campus Diepenbeek  
Wetenschapspark 5 bus 6  
3590 Diepenbeek, Belgium

---

## ABSTRACT

In this paper we investigate the monthly frequency and severity of road traffic accidents in Belgium from 1974 to 1999. We describe the trend in the time series, quantify the impact of explanatory variables, and make predictions. We found that laws concerning seat belts, speed, and alcohol have proven successful. Furthermore, road safety increases with freezing temperatures while sun has the opposite effect, and precipitation and thunderstorms particularly influence accidents with light injuries. Economic conditions have a limited impact. State-space methodology is used throughout the analysis. We compared the results of this study with those of earlier research that applied a regression model with autoregressive moving average errors on the same data. Many similarities were found between these two approaches.

## INTRODUCTION

Every year, Belgium has about 70,000 road deaths and injuries (BIVV 2001). During the past decade, the steady increase in traffic volume has resulted in a steady growth in traffic problems. The negative impact of these problems on our society highlights the need for an effective road safety policy.

---

**KEYWORDS:** Road safety, time series, trend, seasonal, explanatory model, state-space methodology, prediction.

---

Email addresses:

\* Corresponding author—geert.wets@uhasselt.be

E. Hermans—elke.hermans@uhasselt.be

F. Van den Bossche—filip.vandenbossche@uhasselt.be

In order to take appropriate actions that will increase the level of road safety, we need to understand the underlying processes that result in traffic problems and their causes. This requires gathering extensive and reliable data over a long time period, together with modeling techniques suitable for describing, interpreting, and forecasting safety developments (EC 2004, 7). We studied the frequency and severity of traffic accidents in Belgium from 1974 through 1999.

Data in economics, engineering, and medicine are often collected in the form of time series—a sequence of observations taken at regular intervals of time (Peña et al. 2001, 1). This data collection method was also used here. From the broad category of time series model construction methods, we applied state-space methods in this study. This methodology will be explained in detail later in this paper. However, it is important to note here that one of the key characteristics of state-space time series models is that observations are regarded as comprising distinct components, such as trend, seasonal, and regression elements, each of which is modeled separately (Durbin and Koopman 2001, vii) and has a direct interpretation. Furthermore, the components are allowed to change in time, and the stationarity of the series is not required.

## **TIME SERIES APPLICATIONS IN ROAD SAFETY**

The increasing interest in road safety is evident in the literature. An important class of road safety models is based on time series analysis. The succession of data points in time is a fundamental aspect in this analysis. Models are used to describe the behavior of the data, to explain the behavior of the time series in terms of exogenous variables, and for forecasting (Aoki 1987, v). The most relevant ideas highlighting developments in road safety inside this movement are described in the COST329 report of the European Commission (2004).

In addition to giving a description of the trend in traffic data, many models test the influence of explanatory factors. A simple, well-known example of such a time series model is the classical linear regression, which assumes a linear relationship between a criterion or dependent variable ( $y_t$ ) and

one or more predictor or independent variables ( $x_t$ ). Explanatory models describe how the target variable depends on the explanatory variables and interventions. One special and prominent class of explanatory models in road safety analysis is known as the DRAG (Demand Routière, les Accidents et leur Gravité) family, extensively described in Gaudry and Lassarre (2000). DRAG models are structural explanatory models that include a relatively large number of explanatory variables whose partial effects on the exposure, the frequency, and the severity of accidents are estimated by means of econometric methods (EC 2004, 174).

The COST329 report (EC 2004, 47) mentions two main classes of univariate dynamic models: ARIMA models studied by Box and Jenkins; and unobserved components models, which are called structural models by Harvey. In a structural model, each component or equation is intended to represent a specific feature or relationship in the system under study (Harvey and Durbin 1986, 188). The models used here, state-space methods, belong to the latter group. To date, Box-Jenkins methods for time series analysis are applied more widely and are more popular than state-space methods, but this study will show the strengths of the state-space methodology.

Both classes are concerned with the decomposition of an observed time series into a certain number of components. ARMA models decompose the series into an autoregressive (AR) process, a moving average (MA) process, and a random process. Unobserved components models decompose a series in a trend, a seasonal, and an irregular part. An important characteristic is that the components can be stochastic. Moreover, explanatory variables can be added and intervention analysis carried out. The principal structural time series models are, therefore, nothing more than regression models in which the explanatory variables are functions of time and the parameters are time-varying (Harvey 1989, 10). The key to handling structural time series models is the state-space form, with the state of the system representing the various unobserved components. Once in state-space form, the Kalman filter (Kalman 1960) may be applied and this in turn leads to estimation, analysis, and forecasting.

Harvey (1989, 22–23) wrote comprehensively on structural time series models (primarily applied to economic time series), presenting an historical overview of the technique. A rapid growth of interest has ensued in recent years. Nowadays, the technique of unobserved components models is used in several studies: Flaig (2002) applied it to quarterly German Gross Domestic Product (GDP), Cuevas (2002) to real GDP and imports in Venezuela, and Orlandi and Pichelmann (2000) to unemployment series. Other than those economic applications, this technique (more specifically an intervention analysis) was also used in traffic-related research (Balkin and Ord 2001; Harvey and Durbin 1986). The state-space methodology forms a well-used approach in modeling road accidents in a number of countries, for example, the Netherlands (Bijleveld and Commandeur 2004), Sweden (Johansson 1996), and Denmark (Christens 2003). This paper presents the results of the first state-space analysis on Belgian data.

## DATA

The data used in this study are monthly observations from January 1974 through December 1999; 12 observations each year over a period of 26 years equals 312 observations. All data have been gathered from governmental ministries and official documents published by the Belgian National Institute for Statistics. In addition to four dependent traffic-related variables, we studied the effect of 16 independent variables. These 16 explanatory factors can be divided into 3 groups: juristic, climatologic, and economic variables. Table 1 gives an overview of all the variables used in this study.

The four dependent variables in our data are the number of accidents with persons killed or seriously injured (NACCKSI), the number of accidents with minor injuries (NACCLI), the number of persons killed or seriously injured (NPERKSI), and the number of persons with minor injuries (NPERLI). The evolution in time of these variables is displayed in figures 1a and 1b. In order to make a comparison between the results of the state-space method and the regression model with ARMA errors, the same variables, data, and time periods were used. In accor-

dance with the study of the regression model with ARMA errors, the logarithm of the dependent variables were modeled and written respectively as LNACCKSI, LNACCLI, LNPERKSI, and LNPERLI.

As figure 1a reveals, the variables concerning killed or seriously injured persons (NACCKSI and NPERKSI) show a decreasing trend over the period. This is less obvious in the case of lightly injured casualties (figure 1b). Another aspect is the recurring pattern in the data. Thirdly, some months have an extremely low value.

The first group of explanatory variables contains laws and regulations. Five dummy variables were included in the model to study the effect of policy measures introduced in Belgium at a certain date within the scope of our analysis. These variables are equal to zero before the introduction and have a value of one from the moment of introduction. Table 1 describes the laws. Weather conditions form the second group of explanatory factors. All meteorological variables were gathered by the Belgian Royal Meteorological Institute and published by the National Institute for Statistics. The quantity of precipitation (in mm) was measured as an average for the whole country. The other variables were measured in the climatologic center in Ukkel (in the center of Belgium). Thirdly, the influence of four indicators of the economic climate will be investigated.

According to several studies (e.g., Fridstrøm et al. 1995, 12; OECD 1997, 16), exposure is a key variable in traffic research. In this study, the frequency and severity of accidents will be explained by many variables, but the impact of exposure is not measured. We cannot describe this effect because adequate monthly data of the total number of kilometers covered on the whole Belgian road system are not available. Population-related exposure statistics could be a solution, but these data are only available on a yearly basis, and no distribution code is at hand. Although we are aware that this is a serious limitation, even without an exposure variable valid models can be constructed and a good fit obtained. (For more details, refer to Van den Bossche et al. 2005). Other factors possibly omitted are assumed to be taken into account to some extent by the unobserved components framework.

**TABLE 1** Dependent and Independent Variables

Category	Variable	Name
<b>Independent variables</b>		
Number of accidents	The (log-transformed) number of accidents with persons killed or seriously injured	<b>LNACCKSI</b>
	The (log-transformed) number of accidents with lightly injured persons	<b>LNACCLI</b>
Number of casualties	The (log-transformed) number of persons killed or seriously injured	<b>LNPERKSI</b>
	The (log-transformed) number of persons lightly injured	<b>LNPERLI</b>
<b>Dependent variables</b>		
Laws and regulations	Law of June 1975 → mandatory seat belt use in the front seats	<b>Law0675</b>
	Law of November 1988 → introduction of zones with a speed limit of 30 km/h	<b>Law1188</b>
	Law of January 1992 → a.o. speed limit of 50 km/h in urban areas and 90 km/h at road sections with at least 2 by 2 lanes without separation	<b>Law0192</b>
	Law of December 1994 → 0.5‰ blood alcohol concentration imposed and higher fines in case of 0.8‰ or higher	<b>Law1294</b>
	Law of April 1996 → a vehicle driver should give right of way to pedestrians who are (or have the intention to) crossing the street	<b>Law0496</b>
	Weather conditions	The quantity of precipitation (in mm)
The monthly percentage of days with precipitation		<b>Pdayprec</b>
The number of sunlight hours		<b>Hrssun</b>
The monthly percentage of days with sunlight		<b>Pdaysun</b>
The monthly percentage of days with frost/freezing temperatures		<b>Pdayfrost</b>
The monthly percentage of days with snow		<b>Pdaysnow</b>
The monthly percentage of days with thunderstorm		<b>Pdaythun</b>
Economic variables	The percentage inflation	<b>Inflat</b>
	The (log-transformed) number of unemployed persons	<b>Lnunemp</b>
	The (log-transformed) number of car registrations	<b>Lncar</b>
	The percentage of second-hand car registrations	<b>Poldncar</b>

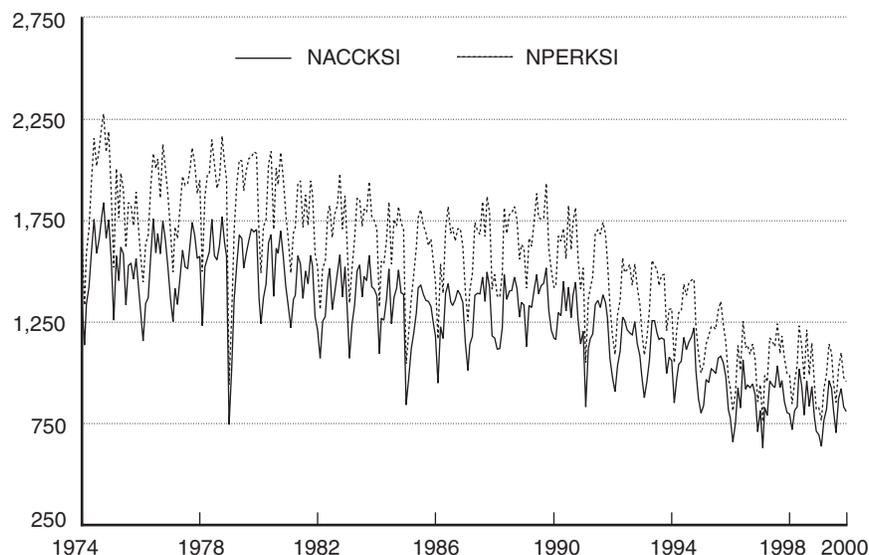
## METHODOLOGY

In this study, state-space models are constructed using *STAMP* software (Koopman et al. 2000). With state-space models, we were able to obtain an explicit description of the series in terms of trend and seasonal. It was also possible to quantify the impact of explanatory factors. For example, the effect of road safety measures over time can be checked by adding so-called intervention variables to the model. Apart from these purposes, state-space models can

easily be used for forecasting. (For a technical discussion of state-space models, see to the methodological appendix at the end of this paper.)

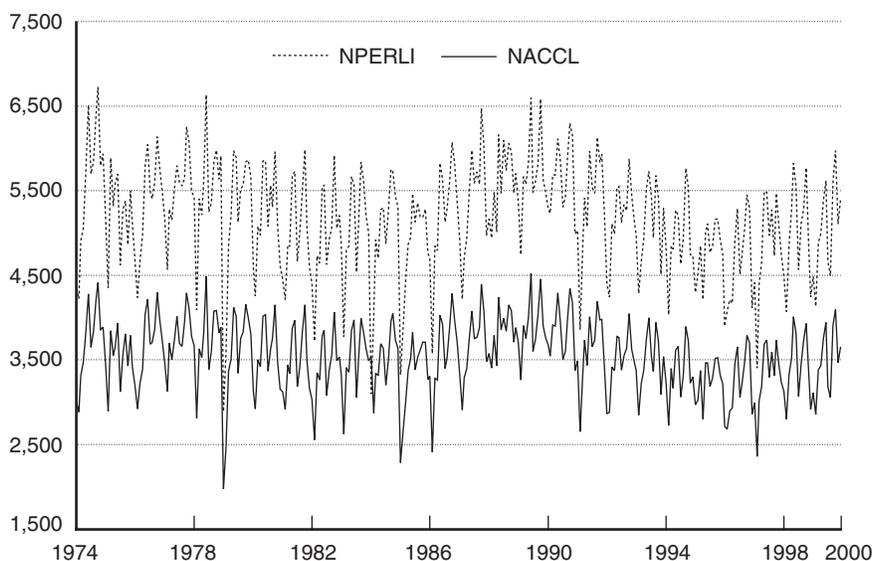
The objective here is to find the model that best describes the data. For each of the four dependent variables, we constructed several state-space models, each with their specific components. To be able to choose the best model, we used the Akaike Information Criterion (AIC), a measurement of fit that takes the number of parameters into account (Akaike 1973, 267–281; Koopman et al. 2000, 180).

**FIGURE 1 Actual Monthly Accident and Casualty Observations: 1974–1999**  
**A: Accidents Involving Persons Killed or Seriously Injured**



Key: NACCKSI = number of accidents, NPERKSI = number of persons

**B: Accidents Where Persons Sustained Minor Injuries**



Key: NACCL = number of accidents, NPERLI = number of persons

We conclude this section with the discussion of some of the advantages of state-space models compared with classical regression. An interesting characteristic of state-space methods is the possibility of modeling stochastically the variation in the estimation of the various components. Contrary to classical regression models, where components are fixed or unchangeable in time, a component can also vary in time. This is an advantage because variation in

time makes it easier to follow the fluctuations in the data. Secondly, when the time dependency between observations is taken into account (which is not the case in classical regression analysis), the observation errors will mostly be situated more closely to independently random values. This makes significance tests of explanatory variables more reliable. Furthermore, state-space methods can easily handle missing observations, multivariate data, and (sto-

chastic) explanatory variables. A last advantage is that the components can be modeled separately and interpreted directly.

## RESULTS

Not all numerical outcomes of the different models will be presented here. However, this section reports and discusses the most essential results of the analysis. It is divided into four parts. First, the outcomes of the descriptive analysis are presented, followed by an interpretation of the explanatory analysis. Next, the forecasting capacity is evaluated. Finally, we compare our results with those obtained by the regression model with ARMA errors and deduce the most important similarities and differences between these two methodologies.

### Description

Based on AIC, we chose the model that best describes the accident data. For each of the four variables the same model resulted in the best fit. This contains a stochastic trend (that adapts every time period) and a deterministic or fixed recurring seasonal pattern.

The interpretation of the seasonal coefficients shows that October and June are the most unsafe road traffic months of the year. During these months, respectively, approximately 13% and 11% more accidents happen than on average. The October percentage can be partly explained by the fact that it is a long month (31 days) without holidays; it is autumn and there is the transition from Central European Summer Time to Central European Time; and it is the start of the academic year. Possible explanations are not apparent for the large number of accidents during June.

### Explanation

To look at the explanatory objective, we tested the effect of 16 independent variables. In order to obtain more reliable results (which implies normally distributed residuals), we added correction variables to the model.

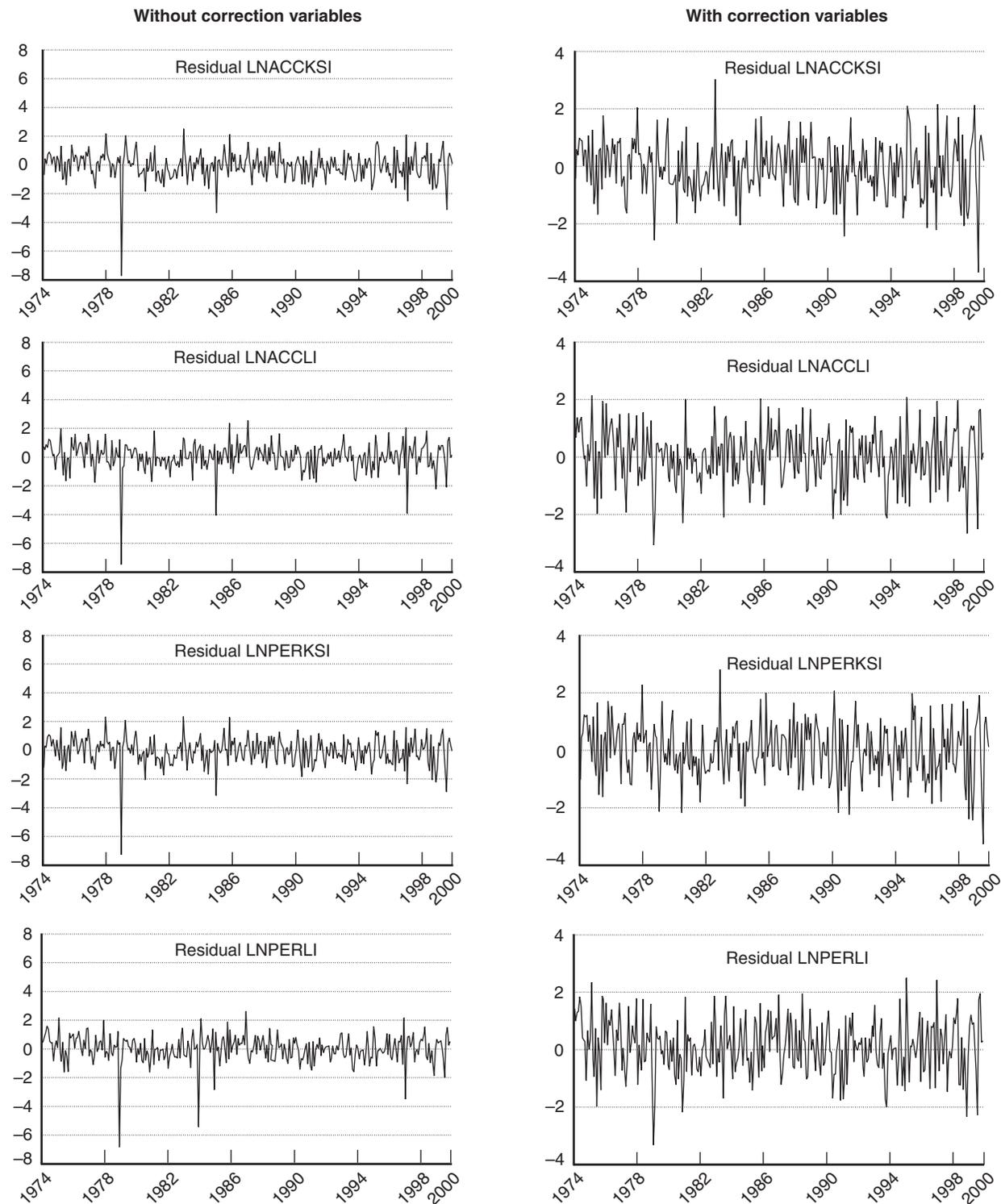
The inclusion of correction variables has algebraically been presented in the model formulation (see the methodological appendix). In general, two main

intervention effects can be distinguished (Sridharan et al. 2003), namely a pulse intervention and a step intervention. The first effect is used to capture single special events because they may cause outlying observations that the pulse regression variable accounts for. The variable takes value 1 if  $t$  is the month that needs correction for a special event and has value 0 otherwise. The second intervention—called a step intervention or level shift—is added to the model to capture events such as the introduction of new policy measures. Laws and regulations can be incorporated in a model as this second type of intervention. Before its introduction, the variable has value 0, but from the moment of introduction it has value 1. Our focus is on the first type, the temporal pulse intervention.

As could be seen on the graphs of the actual data (figures 1a and 1b) as well as on the graph of the residuals (figure 2), the number of accidents and casualties was unexpectedly low during some months. Either these months indeed had extremely low values or some registration error was left in the accident statistics. The following are extreme values for which correction is necessary. January 1979, January 1984 (only for LNPERLI, so a registration error probably occurred here), January 1985, and February 1997 are outliers. There are some indications for a very severe winter in 1979 and 1985 (BIVV 2001, 5). We explicitly correct for those four months by adding pulse intervention variables to the model, which are coded one during the month they represent and zero elsewhere. We are convinced that the most striking shocks must be excluded in order to fulfill the error terms conditions: no autocorrelation, homoscedasticity, and normality. In the end, we want to obtain a correct parameter interpretation. The inclusion of these correction variables lowers the difference between the predicted and the real series and thus improves the quality of the estimations. All tested correction variables are highly statistically significant. The exact  $t$ -values are given in table 2 under “correction variables.” Taking these outliers into account, the fit of the models improves.

The last step in the construction of the final model consists of the significance tests of the explanatory variables. An explanatory variable

**FIGURE 2 Residuals of the Model With Stochastic Trend, Deterministic Seasonal, and Explanatory Variables for the Four Dependent Variables**



must have a significant influence at least at the 90% confidence level to be included in the final model. Each model was re-estimated after dropping the nonsignificant variables such that the ultimate model for every dependent variable consists of a sto-

chastic level, a deterministic seasonal, and significant correction and explanatory variables. The addition of significant explanatory variables further improves the fit. Table 2 gives an overview of all significant combinations of variables. The parameter

**TABLE 2 Overview of the Significant Explanatory and Correction Variables for Each Dependent Variable**

	LNACCKSI		LNACCLI		LNPERKSI		LNPERLI	
	State space	Regression	State space	Regression	State space	Regression	State space	Regression
<b>Laws</b>								
LAW0675	-0.132 (-3.31)	-0.135 (-3.82)	-0.062 (-2.13)	-0.070 (-2.40)	-0.180 (-4.07)	-0.159 (-4.24)	-0.109 (-3.35)	-0.120 (-3.75)
LAW0192		-0.062 (-1.87)	-0.063 (-2.17)	-0.071 (-2.50)	-0.087 (-2.09)	-0.076 (-2.11)		-0.064 (-2.03)
LAW1294	-0.109 (-2.72)	-0.105 (-3.12)			-0.104 (-2.50)	-0.093 (-2.57)		
<b>Weather conditions</b>								
QUAPREC			0.001 (4.12)	0.001 (4.71)			0.001 (3.85)	0.001 (4.39)
PDAYPREC		0.001 (3.34)	0.001 (2.60)	0.001 (2.84)	0.001 (4.04)	0.001 (4.02)	0.001 (2.69)	0.001 (2.88)
HRSSUN	0.000 (3.92)	0.001 (5.35)	0.001 (4.99)	0.001 (5.38)	0.001 (4.48)	0.001 (4.80)	0.000 (3.82)	0.000 (4.17)
PDAYFROST	-0.002 (-5.93)	-0.001 (-5.77)	-0.001 (-5.73)	-0.001 (-5.63)	-0.001 (-4.82)	-0.001 (-4.81)	-0.001 (-4.44)	-0.001 (-4.36)
PDAYTHUN			0.001 (2.12)	0.001 (2.06)			0.001 (2.03)	0.001 (1.93)
<b>Economic variables</b>								
LNUNEMP					0.117 (1.69)			
LNCAR					-0.078 (-2.11)			
<b>Correction variables</b>								
Jan-79	-0.553 (-10.02)	-0.574 (-12.05)	-0.456(-10.02)	-0.441 (-10.40)	-0.558 (-9.72)	-0.550 (-10.75)	-0.464 (-9.51)	-0.458 (-9.97)
Jan-84							-0.450 (-9.28)	-0.449 (-9.87)
Jan-85	-0.277 (-5.04)	-0.282 (-5.85)	-0.2711 (-6.00)	-0.2852 (-6.80)	-0.297(-5.22)	-0.289 (-5.62)	-0.276 (-5.67)	-0.294 (-6.46)
Feb-97	-0.118 (-2.15)	-0.128 (-2.66)	-0.2033 (-4.50)	-0.2181 (-5.08)	-0.130 (-2.28)	-0.124 (-2.31)	-0.214 (-4.41)	-0.238 (-5.09)
<b>Forecasting measure</b>								
MSE(12)	0.005	0.0046	0.0034	0.0052	0.007	0.0063	0.0036	0.0052

Note: Calculated according to the state space method and the regression model with ARMA errors, together with the coefficients and t-values (in parentheses) and the MSE values of the predictions for 2000.

estimates and the  $t$ -statistics (between brackets) of the significant explanatory and correction variables according to the state-space method on the one hand and the regression model with ARMA errors on the other hand are presented.

At first sight, there are a lot of similarities between the results of the two methods. Note that the majority of explanatory variables is statistically significant at least at the 95% confidence interval. In the remainder of this section, we will interpret the significant explanatory variables according to the state-space method per category.

The results of *laws and regulations* are instructive and interesting. Three of the five variables originally included in the model proved to be significant for at least two dependent variables. Their introduction has been of major importance for road safety. This is reflected by the magnitude of the coefficients. The negative signs are as expected because laws are established to enhance road safety. The introduction of the law of June 1975 (LAW0675)—the mandatory seat belt use in the front seats—resulted in a considerable and highly significant increase in road safety. This law reduced all kinds of accidents and casualties. Several empirical studies (Hakim et al. 1991, 392; Harvey and Durbin 1986) have shown that seat belt legislation significantly reduces the number of fatalities and the severity of injuries. The introduction of a speed limit of 50 km/h in urban areas and 90 km/h at road sections with at least 2 x 2 lanes without separation (LAW0192) seemed significant for two dependent variables. The literature verifies the positive effect on road safety in case of a reduction in speed limit. Severity of injuries appears to be positively related to the allowed speed (Van den Bossche and Wets 2003, 15; Hakim et al 1991, 390). Yet another promising effect can be noted for the regulations and fines on the maximum blood alcohol concentration (LAW1294). They played an important role in the decrease in the number of serious accidents and the number of persons killed or seriously injured. The results confirm the hypothesis that drunk drivers often cause serious or fatal accidents. Amongst others, Gaudry (2000, 1-36) studied the effect of the consumption of alcohol on road safety and found that the relative accident probabili-

ty, as a function of blood alcohol concentration, is J-shaped.

In our models, it is assumed that the introduction of a law results in a sudden and permanent decrease in the dependent variable. This assumption of a step-based intervention is not always a natural one (Van den Bossche et al. 2004, 8). The significant impact of laws and regulations may be better described as “something changed at that time,” instead of attributing the whole effect to the law itself. Nevertheless, it makes sense to test whether these changes are indeed substantial.

As one would expect intuitively, the weather plays an important role in explaining the number of accidents and casualties (especially for the variables concerning lightly injured persons). In terms of direction, we can make a distinction between precipitation, sun, and thunderstorms on the one hand and freezing temperatures on the other hand. In addition to precipitation (QUAPREC and PDAYPREC) and thunderstorms (PDAYTHUN), the sun (HRSSUN) is a factor tied to an increase in accidents. It is plausible to assume reduced visibility in stormy weather and on sunny days, a greater likelihood of blinding by the sun. The only weather variable that has a positive effect on road safety is the monthly percentage of days with freezing temperatures (PDAYFROST). A possible explanation is that drivers adjust their driving habits—steer more slowly and prudently and concentrate more—because they perceive driving in freezing conditions as dangerous (which is not the case with rain and thunderstorms). Thus, it seems like road users compensate for the higher risk imposed by freezing temperatures. This result is in line with other studies (Fridstrøm et al. 1995, 9) wherein it is mentioned that exposure to traffic is lower in winter and the average driving capacity increases because less proficient drivers prefer to avoid driving on slippery roads.

The impact of freezing road conditions (PDAYFROST) and sun (HRSSUN) is noticeable for all dependent variables. The quantity of precipitation (QUAPREC) and the monthly percentage of days with thunderstorms (PDAYTHUN) are only relevant for the variables concerning lightly injured casualties. Eisenberg (2004, 641) noticed that in adverse weather conditions, persons possibly drive

more slowly and therefore, on average, accidents are less severe.

Concerning the quantity of precipitation (QUAPREC) (on the killed or seriously injured outcomes) it is possible that two effects canceled out each other. As also found in Gaudry and Lassarre (2000, 67–96), the onset of rain has a larger and more general impact than the amount of rain (habituation can lead to more risky driving behavior). A conclusive remark on the explanatory capacity of weather conditions is that the effect of weather data is strongly related to the geographical properties of the area of concern and the level of aggregation.

Concerning the economically related variables, two *economic indicators* happened to be significant, namely the number of unemployed (LNUNEMP) and the number of car registrations (LNCAR) for the variable LNPERKSI. They have an opposite sign and both imply that a better economy—with less unemployment and more car registrations—decreases the number of killed or seriously injured casualties. In the literature the findings about the direction of this effect are very diverse (Hakim et al. 1991, 384). In this study, the number of car registrations is used as one of the indicators for the economic climate. The assumption we make is that when the economy goes well more cars will be bought, and the average quality of the vehicles on the road increases. In the future, more variables (e.g., disposable income) should be included in the analysis to better assess the explanatory capacity of economic variables and their impact.

## Prediction

The third objective of this study is predicting accident data with state-space methods for the years 2000 and 2001. Future values of the explanatory variables are available. Only the values of QUAPREC and PDAYTHUN for 2001 have to be estimated. This is done with a simple univariate state-space model based on the data from 1974 through 2000.<sup>1</sup>) We use the final model—which contains a

---

<sup>1</sup> One could question the correctness of using estimated values in the prediction, but we can assume that the estimates of these two weather variables will be in line with the actual unknown values due to little variation from year to year and the strong seasonal pattern.

stochastic level, a deterministic seasonal, and significant explanatory and correction variables—to forecast the values of the out-of-sample dataset for 2000 and 2001 and compare them to the actual observations. To depict possible uncertainty, 95% prediction intervals are provided. The graphs (see figure 3) show us that the predictions are close to the actual observations. So we are able to capture a great part of the fluctuations in the series. Only a few points lie outside the prediction intervals.

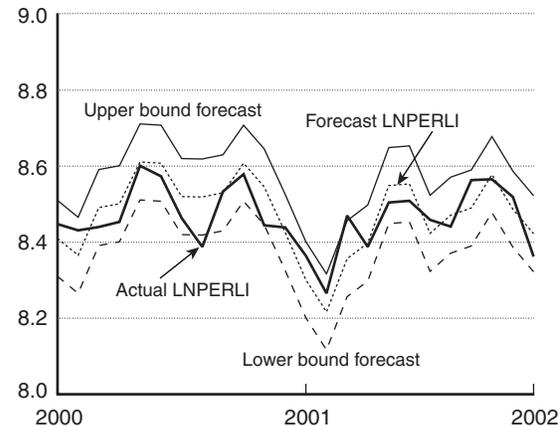
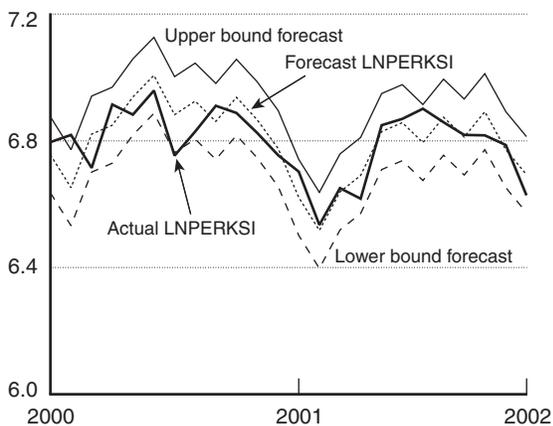
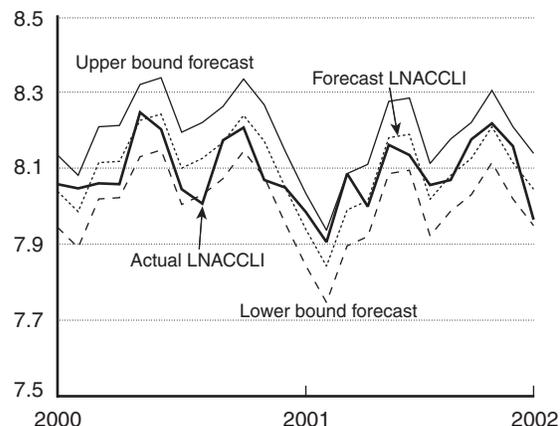
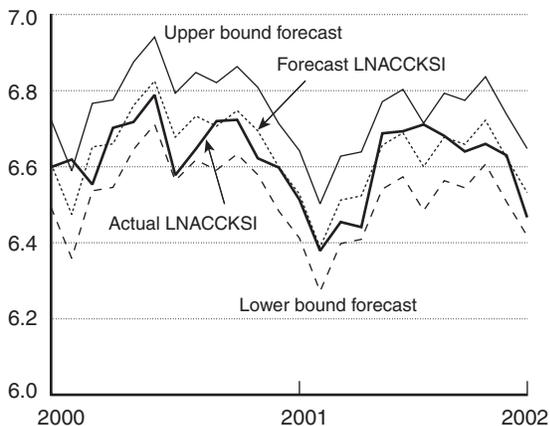
Apart from a visual presentation, we also quantified the forecasting precision. We interpreted the results of the Failure Chi-squared test and computed the mean squared error (MSE). Those tests confirmed our conclusion of accurate predictions.

## Comparison with ARMA regression model

In addition to the interesting characteristics of state-space models already mentioned in the methodology section, we discuss an important disadvantage of ARIMA models here. It is not possible to explicitly describe a time series in terms of the different components because ARIMA models require the time series to be stationary (Harvey and Durbin 1986, 188). In those models the trend and/or seasonal are treated as a problem and therefore removed from the series by a procedure called differencing (in order to transform the series into a stationary one) before any analysis can be performed. But few economic and social time series are stationary, and there is no overwhelming reason to suppose that they can necessarily be made stationary by differencing (Harvey and Shephard 1993, 266).

In 2003, a study on intervention time series analysis of crime rates (Sridharan et al.) showed that the results of a legislation on different kinds of crimes were very similar between the ARIMA model and the structural time series model. Both coefficients and *t*-values were very analogous. A comparison with the regression model with ARMA errors, however, showed different results. Earlier, Harvey and Todd (1983) compared the results of the prediction of a number of economic time series done by the basic structural model with those obtained using the Box-Jenkins models. They concluded that the forecasts given by both methods are comparable.

**FIGURE 3 Monthly Two-Year Ahead Predictions for the Four Dependent Variables**



In this study, we investigate the differences and similarities in explanatory and predictive analysis between the state-space method and the regression model with ARMA errors. Table 2 shows that the outcomes of these two approaches are comparable. The same correction variables seemed significant and the juristic and climatologic variables also matched quite well. Different from the results of the regression model with ARMA errors is the fact that two of the four economic variables are significant. A possible reason is that the evolution in economic factors is a very slow one. In case of a regression model with ARMA errors differences are taken, resulting in almost a constant. Differencing possibly cancels out the already little variation in time. Next, the estimated parameters of the two methods have the same (expected) sign and are of the same order of magnitude.

Both methods forecasted the data for the year 2000, so we are able to assess and compare the quality of the predictions. The measure used is MSE, and the values of the two methodologies for the four variables are reported in the last row of table 2. The lower MSE, the better the prediction. The values are of the same order of magnitude. The predictions for the two variables concerning killed or seriously injured persons from the regression model with ARMA errors are more accurate. The state-space method better predicts the values of the variables concerning lightly injured persons. In case of killed or seriously injured persons, the decreasing level is more important than the recurring seasonal pattern. In contrast, for light injuries with values fluctuating around the average, the seasonal effect is more important. Because the seasonal effect is explicitly modeled in the state-space model, this

model possibly predicts more accurately in case of lightly injured persons than the regression model with ARMA errors.

## CONCLUSIONS

In this study state-space models were elaborated to describe the developments in the frequency and severity of accidents and casualties in Belgium from 1974 through 1999. Furthermore, the impact of laws, weather, and economic conditions was measured. In the third place, an out-of-sample forecast of the dependent variables for 24 months was made. The results were compared with those obtained from a regression model with ARMA errors, based on the same data.

For each of the four dependent variables we built several models. The model that described all data best consisted of a level that is allowed to vary over time and a seasonal. Explanatory and correction variables were added to this descriptive model. The fact that accidents happen can to a certain extent be attributed to juristic, meteorological, and economic factors. Due to data and multicollinearity issues and for reasons of comparison, we tested the influence of 16 independent variables. Additionally, correction variables for January 1979, January 1984 (only for LNPERLI), January 1985, and February 1997 were significant.

From this study we can conclude that there is a lot of similarity between the results of the state-space method and the regression model with ARMA errors. Both methods labeled (more or less) the same explanatory variables as significant, and their influence was at all times in the same direction and of comparable magnitude. Several laws had a clear positive effect. Apart from those, the weather elements precipitation, sun, freezing temperatures, and thunderstorms were important. Nevertheless, note the difference between the two methods on the subject of the economic variables. The forecasting capacity of the methods was tested quantitatively and was shown to be approximately the same.

The models developed in this text show large potential for describing long-term trends in road safety. On the one hand, they can isolate the effect of phenomena that cannot be influenced, but certainly act on road safety (for example the weather).

Similarly, macroeconomic and sociodemographic evolutions could be added to the model. On the other hand, the efficiency of policy decisions (for example laws) or time-specific interventions can be tested. These are the direct tools for increasing the level of road safety. Moreover, forecasts can be made, uncertainty estimated, and ruptures in the time series detected. Furthermore, some advantages of state-space methods over regression and ARIMA models were reported.

We conclude with some topics for model improvement and further research. In this study the variable exposure was not included. In the future, monthly observations of the total mileage covered on the Belgian road system could be taken into account in order to measure this effect. Secondly, because the number of variables in our models is limited, the effect of more explanatory factors could be tested, for example income or public transportation. The elaboration of data quality and availability together with the development of extensive but statistically sound models should lead to high quality results.

## REFERENCES

- Akaike, H. 1973. Information Theory and an Extension of the Maximum Likelihood Principle. *Second International Symposium on Information Theory*. Edited by P.N. Petrov and F. Csaki. Budapest: Akadémiai Kiadó.
- Aoki, M. 1987. *State Space Modeling of Time Series*. New York, NY: Springer-Verlag.
- Balkin, S. and J.K. Ord. 2001. Assessing the Impact of Speed-Limit Increases on Fatal Interstate Crashes. *Journal of Transportation and Statistics* 4(1):1-26.
- Belgisch Instituut voor de Verkeersveiligheid (BIVV). 2001. Verkeersveiligheid Statistieken 2001. Available at <http://www.bivv.be/main/PublicatieMateriaal/Statistieken.shtml>.
- Bijleveld, F.D. and J.J.F. Commandeur. 2004. The Basic Evaluation Model, paper presented at the ICTSA meeting, INRETS, Arcueil, France, May 27-28, 2004.
- Christens, P.F. 2003. Statistical Modelling of Traffic Safety Development, IMM-PHD-2003-119. Available at <http://www.imm.dtu.dk>.
- Cuevas, M.A. 2002. Demand for Imports in Venezuela: A Structural Time Series Approach, World Bank Policy Research Working Paper No. 2825. Available at <http://ssrn.com/abstract=313423>.
- Durbin, J. and S.J. Koopman. 2001. *Time Series Analysis by State-Space Methods*. Oxford, England: Oxford University Press.

- Eisenberg, D. 2004. The Mixed Effects of Precipitation on Traffic Crashes. *Accident Analysis and Prevention* 36: 637–647.
- European Commission (EC). 2004. *COST Action 329: Models for Traffic and Safety Development and Interventions*. Luxembourg: European Communities.
- Flaig, G. 2002. Unobserved Components Models for Quarterly German GDP, CESifo Working Paper no. 681. Available at <http://www.cesifo.de>.
- Fridstrøm, L., J. Ifver, S. Ingebrigtsen, R. Kulmala, and L.K. Thomsen. 1995. Measuring the Contribution of Randomness, Exposure, Weather and Daylight to the Variation in Road Accident Counts. *Accident Analysis and Prevention* 27(1):1–20.
- Gaudry, M. and S. Lassarre. 2000. *Structural Road Accident Model: The International DRAG Family*. Oxford, England: Elsevier Science Ltd.
- Hakim, S., D. Shefer, A.S. Hakkert, and I. Hocherman. 1991. A Critical Review of Macro Models for Road Accidents. *Accident Analysis and Prevention* 23(5):379–400.
- Harvey, A.C. 1989. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge, England: Cambridge University Press.
- Harvey, A.C. and J. Durbin. 1986. The Effect of Seat Belt Legislation on British Road Casualties: A Case Study in Structural Time Series Modeling. *Royal Statistical Society A* 149(3):187–227.
- Harvey, A.C. and N. Shephard. 1993. Structural Time Series Models. *Handbook of Statistics* 11:261–302.
- Harvey, A.C. and P.H.J. Todd. 1983. Forecasting Economic Time Series with Structural and Box-Jenkins Models: A Case Study. *Journal of Business and Economic Statistics* 1(4):299–315.
- Johansson, P. 1996. Speed Limitation and Motorway Casualties: A Time Series Count Data Regression Approach. *Accident Analysis and Prevention* 28(1):73–87.
- Kalman, R.E. 1960. A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering D* 82:35–45.
- Koopman, S. J., A.C. Harvey, J.A. Doornik, and N. Shephard. 2000. *Stamp: Structural Time Series Analyser, Modeller and Predictor*. London, England: Timberlake Consultants Ltd.
- Organization for Economic Cooperation and Development (OECD). 1997. *Road Safety Principles and Models: Review of Descriptive, Predictive, Risk and Accident Consequence Models*. Paris, France.
- Orlandi, F. and K. Pichelmann. 2000. Disentangling Trend and Cycle in the EUR-11 Unemployment Series, ECFIN/27/2000-EN, No. 140. Available at <http://europa.eu.int>.
- Peña, D., G.C. Tiao, and R.S. Stay. 2001. *A Course in Time Series Analysis*. New York, NY: John Wiley & Sons.
- Sridharan, S., S. Vujic, and S.J. Koopman. 2003. Intervention Time Series Analysis of Crime Rates, TI 03-040/4. Available at <http://www.tinbergen.nl>.
- Van den Bossche, F. and G. Wets. 2003. Macro Models in Traffic Safety and the DRAG Family: Literature Review, RA-2003-08. Available at <http://www.steunpunt verkeersveiligheid.be/en>.
- Van den Bossche, F., G. Wets, and T. Brijs. 2004. A Regression Model with ARMA Errors to Investigate the Frequency and Severity of Road Traffic Accidents. Proceedings of the 83rd Annual Meetings of the Transportation Research Board, Washington, DC, January 11–15, 2004: 1-15.
- \_\_\_\_\_. 2005. The Role of Exposure in the Analysis of Road Accidents: A Belgian Case-Study. Proceedings of the 84th Annual Meetings of the Transportation Research Board, Washington, DC, January 9–13, pp.1–16.

## METHODOLOGICAL APPENDIX

In this appendix state-space models are discussed in more detail. The overall objective of the state-space analysis is to study the development of the state over time using observed values (Durbin and Koopman 2001, 11). More specifically, we want to obtain an adequate description of this development and to find explanations hereof. Furthermore, these models have the ability to predict developments of a series into the future.

The state is the unobserved value of the true development at time  $t$ . The gathering (or space) of possible values of the state is called the state-space of the process. The state consists of several components: on the one hand a *level*, *slope*, and *seasonal* that give a description of the time series and on the other hand *explanatory* and *intervention variables* that give an explanation about the actual development in the series.

A state-space model consists of an observation or measurement equation and one or more state equations (depending on the number of components). The first one contains the unobserved state at time  $t$  and an observation residual ( $\varepsilon_t$ ), which is white noise. In the state equation, time dependencies in the observed time series are dealt with by letting the state at time  $t+1$  be a direct function of the state at time  $t$ , and the state error is also white noise. Algebraically, the final state-space model used in this analysis can be written as:

$$\text{(Eq. 1) } y_t = \mu_t + \gamma_t + \sum_{j=1}^k \beta_{jt} x_{jt} + \sum_{i=1}^l \lambda_{it} w_{it} + \varepsilon_t$$

$$\varepsilon_t \sim \text{NID}(0, \sigma_\varepsilon^2)$$

$$\text{(Eq. 2) } \mu_{t+1} = \mu_t + \eta_t \quad \eta_t \sim \text{NID}(0, \sigma_\eta^2)$$

$$\text{(Eq. 3)} \quad \gamma_{t+1} = \gamma_t + \omega_t \quad \omega_t \sim NID(0, \sigma_\omega^2)$$

$$\text{(Eq. 4)} \quad \beta_{j,t+1} = \beta_{jt} + \tau_{jt} \quad \tau_{jt} \sim NID(0, \sigma_\tau^2)$$

$$\text{(Eq. 5)} \quad \lambda_{i,t+1} = \lambda_{it} + \xi_{it} \quad \xi_{it} \sim NID(0, \sigma_\xi^2)$$

for  $t = 1, \dots, n$ ;  $j=1, \dots, k$  and  $i=1, \dots, l$ .

The observation equation (Eq. 1) relates the values of the dependent variable  $y_t$  to the level  $\mu_t$ , the seasonal component  $\gamma_t$ , explanatory variables  $x_{jt}$  ( $j = 1, \dots, k$ ), intervention variables  $w_{it}$  ( $i = 1, \dots, l$ ), and an observation error  $\varepsilon_t$ . Each component has its state equation (Eq. 2 till 5 respectively). All (observation and state) errors are assumed to be mutually independent and normally distributed with mean zero and variances  $\sigma_\varepsilon^2$ ,  $\sigma_\eta^2$ ,  $\sigma_\omega^2$ ,  $\sigma_\tau^2$  and  $\sigma_\xi^2$  respectively.  $\beta_j$  is the unknown regression coefficient of the  $j$ th explanatory variable. One type of intervention is the temporal pulse intervention. Only during one time point a correction of an unusual high or low value occurs. In this paper, four correction variables of this type were used. Concerning these variables,  $w_{it} = 1$  if  $t$  is the month of correction, and 0 otherwise.  $\lambda_i$  is the coefficient of the  $i$ th correction variable.

The error variances are used in order to obtain the most parsimonious model that describes the data best. Each component can be chosen deterministically or stochastically. Deterministic implies one parameter estimate during the whole time period while stochastic implies that the estimate will be adapted every time point. However, this last option requires more parameters. Whether a state component should be treated deterministically or stochastically can be determined by evaluating the error variance of the component when analyzed stochastically. If the error variance of the stochastic component is very small (i.e., almost zero), this indicates that the corresponding state component should be handled deterministically. Because we consider only deterministic explanatory variables, the corresponding errors ( $\tau_{jt}$ ) are equal to zero.

In state-space methods the value of the unobserved state at the beginning of the time series ( $t = 1$ ) is unknown. Using diffuse initialisation (Durbin and Koopman 2001, 28) estimates for the unknown parameters are obtained. Also none of the observation and state error variances are known. The estimation of all these parameters can be obtained with an iterative process using the maximum likelihood principle.

## **REVIEWERS** for 2006

### **A**

Fariba Alamdari                      Cranfield University, Bedford, England

### **B**

Sandy Balkin                          Pfizer Pharmaceuticals Inc., New York, NY

### **D**

Raymond J. Dezzani                University of Idaho, Moscow, Idaho

### **G**

Laurie A. Garrow                    Georgia Institute of Technology, Atlanta, Georgia

### **O**

Keith Ord                              Georgetown University, Washington, DC

### **T**

Darren Timothy                      Federal Highway Administration, USDOT,  
Washington, DC  
Tomer Toledo                        Massachusetts Institute of Technology, Cambridge, MA  
Victor Torres                        Office of Productivity and Technology, Bureau of  
Labor Statistics, Washington , DC  
Martin Treiber                        Dresden University of Technology, Dresden, Germany

### **W**

Simon Washington                 Arizona State University, Tempe, AZ  
Jack Wells                            Office of the Secretary of Transportation, USDOT,  
Washington DC

### **Y**

Theodore Younglove                University of California, Riverside, CA



# INDEX FOR VOLUME 9\*

## A

Australia  
Adelaide, Australia, Vol.9(1):17, 19, 21  
Air Travel Price Index (ATPI), Vol.9(1):8, 12  
Air quality, Vol.9(1):29–31, 37–38  
Analysis of variance (ANOVA), Vol.9(1):50  
Annual average daily travel (AADT), Vol.9(1):37, 50  
ARTPLAN, Vol.9(1):30  
Average effective speed (AES), Vol.9(1):31, 36

## B

Bayesian analysis, Vol.9(1):18, 22  
Berkson measurement error, Vol.9(1):39, 41, 44  
Bureau of Labor Statistics (BLS), Vol.9(1):6–8, 10  
Bureau of Public Roads (BPR), Vol.9(1):30–31  
Bureau of Transportation Statistics (BTS), Vol.9(1):4–5, 8, 13

## C

Case-control, Vol.9(1):17–21, 23–25, 27  
Changes in productivity, Vol.9(1):2  
Chi-squared, Vol.9(1):72  
Clean Air Act, Vol.9(1):29  
Consumer Federation of America (CFA), 10  
Consumer Price Index (CPI), Vol.9(1):6–8, 10  
Consumer Price Index for Air Transportation (CPI-AT), Vol.9(1):6  
Correlation, Vol.9(1):5, 39–42, 44, 46, 55–56, 61  
Cumulative Density Function (CDF), Vol.9(1):54  
Curve data, Vol.9(1):36

## D

Deviance information criteria (DIC), Vol.9(1):24

## E

Eigenflow, Vol.9(1):50–62

## F

Federal Highway Administration (FHWA), Vol.9(1):34  
Forecasting, Vol.9(1):30, 64, 66, 68, 72, 74  
Free-flow speed (FFS), Vol.9(1):30–32, 36

## G

Greater Athens Area (GAA), Vol.9(1):50, 52

## H

Highway Economic Requirement System (HERS), Vol.9(1):31–32, 34–38  
Highway Performance Measurement System (HPMS), Vol.9(1):31–32, 35–38  
Hirschman-Herfindahl Index (HHI), Vol.9(1):10

## I

Increased productivity, Vol.9(1):2, 13  
Industry employees, Vol.9(1):1, 12  
Industry profits, Vol.9(1):1–2, 5, 12

## K

Kentucky  
Christian County, Kentucky, Vol.9(1):32

## L

Labor productivity, Vol.9(1):2–5, 7, 12  
Linear regression, Vol.9(1):40, 42, 46, 64  
Logit model, Vol.9(1):19

## M

Maximum allowable speed on a curve (VCURVE), Vol.9(1):36  
Maximum speed resulting from the speed limit (VSPLIM), Vol.9(1):36  
Mean Relative Error (MRE), Vol.9(1):58–59  
Minnesota run-off-road data, Vol.9(1):24  
MOBILE6, Vol.9(1):29–31, 37–38  
Models, Vol.9(1):61, 64–68, 71–72, 74–75  
ARIMA models, Vol.9(1):64  
Bayesian statistical models, Vol.9(1):27  
generalized linear models (GLM), Vol.9(1):40  
internal models, Vol.9(1):31  
frequency data, Vol.9(1):40  
linear models, Vol.9(1):40  
nonlinear models, Vol.9(1):40, 47  
quadratic models, Vol.9(1):24–26  
random effects model, Vol.9(1):42  
regression model, Vol.9(1):40, 42, 64, 67  
simulation models, Vol.9(1):60, 62  
smoothing models, Vol.9(1):60  
state-space models, Vol.9(1):66–67, 74  
time series models, Vol.9(1):64–65  
traffic simulation models, Vol.9(1):50  
travel demand models (TDM), Vol.9(1):29–30  
WinBUGS models, Vol.9(1):22  
Monte Carlo simulations, Vol.9(1):25  
Motor Vehicle Emission Simulator (MOVES), Vol.9(1):38  
Multifactor productivity (MFP), Vol.9(1):4–5

## N

National Highway Traffic Safety Administration (NHTSA), Vol.9(1):21  
North America Industry Classification System (NAICS), Vol.9(1):2

## O

Office of Airline Information (OAI), Vol.9(1):4–5, 14  
Origin-destination (OD), Vol.9(1):50

\* A complete index of all volumes of the journal is available online at [www.bts.gov/jts](http://www.bts.gov/jts).

**P**

Poisson regression, Vol.9(1):39–42, 46  
Producer Price Index (PPI), Vol.9(1):6–8  
Principal Component Analysis (PCA), Vol.9(1):49–52, 57,  
59, 61

**R**

Railway Labor Act, Vol.9(1):12  
Random coefficients, Vol.9(1):39  
Regression analysis, Vol.9(1):5, 40, 67  
Road Accident Research Unit (RARU), Vol.9(1):19–22,  
24–27

**S**

Sampling, Vol.9(1):19, 44  
    randomly, Vol.9(1):21–22  
Sensitivity analysis, Vol.9(1):37  
Simulation, Vol.9(1):39–40, 42–44, 46, 50, 62  
Singular Value Decomposition (SVD), Vol.9(1):50  
Smoothing models, Vol.9(1):49, 60  
Speed estimation, Vol.9(1):29–31, 37  
Speed limits, Vol.9(1):17–19, 21, 33, 36

**T**

Time series, Vol.9(1):49–54, 59–60, 63–65, 72, 74–76  
TranStats, Vol.9(1):4–5, 13  
Travel Agent Commission Overrides (TACOS), Vol.9(1):9

**U**

Urban networks, Vol.9(1):49–50, 52, 62  
U.S. Environmental Protection Agency (EPA), Vol.9(1):30,  
34, 38  
U-shaped relationship, Vol.9(1):17–18, 20, 23, 26–27

**V**

Vehicle-miles of travel (VMT), Vol.9(1):30

**W**

Weighted design speed, Vol.9(1):36–37









U.S. Department of Transportation  
Research and Innovative Technology Administration  
Bureau of Transportation Statistics

# JOURNAL OF TRANSPORTATION AND STATISTICS

Volume 9 Number 1, 2006  
ISSN 1094-8848

## CONTENTS

**ANTHONY D. APOSTOLIDES** Impacts of Productivity Changes in Air Transportation on Profits, Prices, and Labor Compensation: 1990–2001

**GARY A. DAVIS, SUJAY DAVULURI + JIANPING PEI**  
Speed as a Risk Factor in Serious Run-off-Road Crashes: Bayesian Case-Control Analysis with Case Speed Uncertainty

**HUAFENG GONG, MEI CHEN, JESSE MAYES + ROB BOSTROM**  
Speed Estimation for Air Quality Analysis

**ERLING HÄGGSTRÖM LUNDEVALLER**  
Measurement Errors in Poisson Regressions: A Simulation Study Based on Travel Frequency Data

**THEODORE TSEKERIS + ANTONY STATHOPOULOS**  
Measuring Variability in Urban Traffic Flow by Use of Principal Component Analysis

**ELKE HERMANS, GEERT WETS + FILIP VAN DEN BOSSCHE**  
Frequency and Severity of Road Traffic Accidents Studied by State Space Methods