M I N E T A   T R A N S P O R T A T I O N   I N S T I T U T E

Net Effects of Gasoline Price Changes on Transit Ridership in U.S. Urban Areas

MTI Report 12-19

MINETA TRANSPORTATION INSTITUTE
MTI

MINETA TRANSPORTATION INSTITUTE

The Norman Y. Mineta International Institute for Surface Transportation Policy Studies was established by Congress in the Intermodal Surface Transportation Efficiency Act of 1991 (ISTEA). The Institute's Board of Trustees revised the name to Mineta Transportation Institute (MTI) in 1996. Reauthorized in 1998, MTI was selected by the U.S. Department of Transportation through a competitive process in 2002 as a national "Center of Excellence." The Institute is funded by Congress through the United States Department of Transportation's Research and Innovative Technology Administration, the California Legislature through the Department of Transportation (Caltrans), and by private grants and donations.

The Institute receives oversight from an internationally respected Board of Trustees whose members represent all major surface transportation modes. MTI's focus on policy and management resulted from a Board assessment of the industry's unmet needs and led directly to the choice of the San José State University College of Business as the Institute's home. The Board provides policy direction, assists with needs assessment, and connects the Institute and its programs with the international transportation community.

MTI's transportation policy work is centered on three primary responsibilities:

Research

MTI works to provide policy-oriented research for all levels of government and the private sector to foster the development of optimum surface transportation systems. Research areas include: transportation security; planning and policy development; interrelationships among transportation, land use, and the environment; transportation finance; and collaborative labor-management relations. Certified Research Associates conduct the research. Certification requires an advanced degree, generally a Ph.D., a record of academic publications, and professional references. Research projects culminate in a peer-reviewed publication, available both in hardcopy and on TransWeb, the MTI website (http://transweb.sjsu.edu).

Education

The educational goal of the Institute is to provide graduate-level education to students seeking a career in the development and operation of surface transportation programs. MTI, through San José State University, offers an AACSB-accredited Master of Science in Transportation Management and a graduate Certificate in Transportation Management that serve to prepare the nation's transportation managers for the 21st century. The master's degree is the highest conferred by the California State University system. With the active assistance of the California Department of Transportation, MTI delivers its classes over a state-of-the-art videoconference network throughout the state of California and via webcasting beyond, allowing working transportation professionals to pursue an advanced degree regardless of their location. To meet the needs of employers seeking a diverse workforce, MTI's education program promotes enrollment to under-represented groups.

Information and Technology Transfer

MTI promotes the availability of completed research to professional organizations and journals and works to integrate the research findings into the graduate education program. In addition to publishing the studies, the Institute also sponsors symposia to disseminate research results to transportation professionals and encourages Research Associates to present their findings at conferences. The World in Motion, MTI's quarterly newsletter, covers innovation in the Institute's research and education programs. MTI's extensive collection of transportation-related publications is integrated into San José State University's world-class Martin Luther King, Jr. Library.

REPORT 12-19

NET EFFECTS OF GASOLINE PRICE CHANGES ON TRANSIT RIDERSHIP IN U.S. URBAN AREAS

Hiroyuki Iseki, Ph.D.
Rubaba Ali

December 2014

A publication of

Mineta Transportation Institute
Created by Congress in 1991

College of Business
San José State University
San José, CA 95192-0219

TECHNICAL REPORT DOCUMENTATION PAGE

<table>
<tr><td>1. Report No.
CA-MTI-14-1106</td><td>2. Government Accession No.</td><td>3. Recipient's Catalog No.</td></tr>
<tr><td colspan="2" rowspan="2">4. Title and Subtitle
Net Effects of Gasoline Price Changes on Transit Ridership in U.S. Urban Areas</td><td>5. Report Date
December 2014</td></tr>
<tr><td>6. Performing Organization Code</td></tr>
<tr><td colspan="2">7. Authors
Hiroyuki Iseki, Ph.D. and Rubaba Ali</td><td>8. Performing Organization Report
MTI Report 12-19</td></tr>
<tr><td colspan="2" rowspan="2">9. Performing Organization Name and Address
Mineta Transportation Institute
College of Business
San José State University
San José, CA 95192-0219</td><td>10. Work Unit No.</td></tr>
<tr><td>11. Contract or Grant No.
DTRT12-G-UTC21</td></tr>
<tr><td colspan="2" rowspan="2">12. Sponsoring Agency Name and Address
California Department of Transportation
Office of Research—MS42
P.O. Box 942873
Sacramento, CA 94273-0001

U.S. Department of Transportation
Research & Innovative Technology Admin.
1200 New Jersey Avenue, SE
Washington, DC 20590</td><td>13. Type of Report and Period Covered
Final Report</td></tr>
<tr><td>14. Sponsoring Agency Code</td></tr>
<tr><td colspan="3">15. Supplemental Notes</td></tr>
</table>

16. Abstract

Using panel data of transit ridership and gasoline prices for ten selected U.S. urbanized areas over the time period of 2002 to 2011, this study analyzes the effect of gasoline prices on ridership of the four main transit modes—bus, light rail, heavy rail, and commuter rail—as well as their aggregate ridership. Improving upon past studies on the subject, this study accounts for endogeneity between the supply of services and ridership, and controls for a comprehensive list of factors that may potentially influence transit ridership. This study also examines short- and long-term effects and non-constant effects at different gasoline prices.

The analysis found varying effects, depending on transit modes and other conditions. Strong evidence was found for positive short-term effects only for bus and the aggregate: a 0.61-0.62 percent ridership increase in response to a 10 percent increase in current gasoline prices (elasticity of 0.061 to 0.062). The long-term effects of gasoline prices, on the other hand, was significant for all modes and indicated a total ridership increase ranging from 0.84 percent for bus to 1.16 for light rail, with commuter rail, heavy rail, and the aggregate transit in response to a 10 percent increase in gasoline prices. The effects at the higher gasoline price level of over $3 per gallon were found to be more substantial, with a ridership increase of 1.67 percent for bus, 2.05 percent for commuter rail, and 1.80 percent for the aggregate for the same level of gasoline price changes. Light rail shows even a higher rate of increase of 9.34 percent for gasoline prices over $4. In addition, a positive threshold boost effect at the $3 mark of gasoline prices was found for commuter and heavy rails, resulting in a substantially higher rate of ridership increase.

The results of this study suggest that transit agencies should prepare for a potential increase in ridership during peak periods that can be generated by substantial gasoline price increases over $3 per gallon for bus and commuter rail modes, and over $4 per gallon for light rail, in order to accommodate higher transit travel needs of the public through pricing strategies, general financing, capacity management, and operations planning of transit services.

<table>
<tr><td>17. Key Words
Gasoline prices; Elasticity of transit ridership; Ten urbanized areas; Panel data regression analysis; Short-term, long-term, and non-constant effects</td><td colspan="3">18. Distribution Statement
No restrictions. This document is available to the public through The National Technical Information Service, Springfield, VA 22161</td></tr>
<tr><td>19. Security Classif. (of this report)
Unclassified</td><td>20. Security Classif. (of this page)
Unclassified</td><td>21. No. of Pages
125</td><td>22. Price
$15.00</td></tr>
</table>

Form DOT F 1700.7 (8-72)

Library of Congress Catalog Card Number:
2014956982

To order this publication, please contact:

Mineta Transportation Institute
College of Business
San José State University
San José, CA 95192-0219

Tel: (408) 924-7560
Fax: (408) 924-7565
Email: mineta-institute@sjsu.edu

transweb.sjsu.edu

120514

ACKNOWLEDGMENTS

The views presented and any errors or omissions are solely the responsibility of the authors and not the funding institution. The authors acknowledge the assistance in data collection, processing, analysis, and report preparation from the following graduate student assistants: Robert Jones, Naka Matsumoto, and Yuchen Cui of the Urban Planning and Design Ph.D. program at the University of Maryland, College Park; Ashley Sampson and Angela Martinez of the Urban Studies and Planning Master's program at the University of Maryland, College Park; and Julie Huang and Aliza Paz at San Jose State University. The authors also would like to thank Dr. Charles Rivasplata for his valuable comments and inputs in the preparation of this report. The authors wish to thank anonymous reviewers for their valuable comments and transit agencies that responded to the request for information.

The authors also thank MTI staff, including Executive Director Karen Philbrick, Ph.D.; Director of Communications and Technology Transfer Donna Maurillo; Research Support Manager Joseph Mercado; and Webmaster Frances Cherman, who also provided additional editorial and publication support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

Between 1999 and 2011 consumers in the U.S. experienced an unprecedented increase in and fluctuation of gasoline prices. In July 2008, gasoline prices exceeded $4 per gallon, marking the highest price in real value in U.S. history. In the same year, the nation's transit ridership reached 10.7 billion trips, the highest level since the Federal-Aid Highway Act of 1956.

The rising gasoline prices were considered to have resulted in substantial changes in travel behavior in terms of trip taking, choices of travel destinations, selection of vehicles for higher fuel efficiency, or travel mode. A change in travel mode from driving to transit results in a higher level of transit demand and ridership for transit agencies. With this background, gasoline price increases in the last decade have generated substantial interest in developing a better understanding of how people respond to fluctuations in gasoline prices—particularly with respect to switching modes from driving to public transit—so that transit agencies can better prepare for higher demand for their services during periods of increased gasoline prices.

The extensive literature review conducted for this study revealed that estimated values of elasticity obtained in the previous studies varied by geographic area, transit mode, travelers' demographic characteristics, trip characteristics, types of data, and analytical method. In particular, types of ridership data used in the previous studies included cross-sectional, time-series, pooled, and panel datasets of ridership for one or multiple agencies. The literature review revealed several important data and methodological issues that should be addressed in an analysis in addition to several different types of effects of gasoline prices on transit ridership.

Based on the literature review, the study has made improvements in developing four specifications of panel data regression analysis to analyze the net effect of gasoline prices on ridership in ten major urbanized areas (UAs) in the U.S. over a ten-year period. First, this study used monthly data on gasoline prices and ridership in order to gauge the effects in the short and long term. Monthly gasoline price data were collected from the Energy Information Administration, Department of Energy, for ten major UAs over the period of 2002 to 2011; monthly transit ridership data by agency were obtained from the National Transit Database and processed to obtain data for the ten UAs. Transit ridership data included the four main modes of transit—bus, commuter rail, light rail and heavy rail—and the aggregate of these four modes. The use of panel data allowed us to simultaneously take into account temporal and cross-sectional variation to obtain more robust, generalizable results.

To minimize the effects of omitted variables, a regression analysis was used to comprehensively control for a set of variables that potentially affect transit ridership, including factors both internal and external to transit services, such as, in the former case, transit fare and service frequency, and, in the latter, economic conditions and socioeconomic characteristics of potential travelers. In addition to the baseline specification model that simply examines potentially influential factors of transit ridership, the instrumental variable (IV) method was employed to address simultaneity between the supply of service and ridership (IV model), which may cause a bias in estimated coefficients of other independent variables, including gasoline prices. Comparison of the results from these first two models confirmed that there is no substantial difference in the estimated coefficients for gasoline prices. Thus, the two models that examine short- and long-term effects and non-constant elasticity were specified based on the baseline specification.

The main findings of this study are:

- The short-run elasticity of bus ridership to gasoline price (i.e., the cross-price elasticity) is 0.06, indicating a 0.6 percent increase in ridership in response to a 10 percent increase in the current gasoline prices. The short-run elasticity was about the same level for the aggregate transit ridership (0.5-0.6), but was not significantly different from zero for the three rail modes.

- The long-run cross-price elasticity, on the other hand, was significant for all modes and ranged from 0.084 for bus to 0.116 for light rail, with commuter rail, heavy rail, and the aggregate transit in between. In other words, a total change in ridership ranges from 0.84 percent to 1.16 percent in response to a 10 percent increase in gasoline prices. Higher values of elasticity were found for gasoline prices higher than $4 for light rail and higher than $3 for the other modes. A percent increase in ridership in response to a 10 percent increase in gasoline prices exceeds 1 percent for bus (1.67 percent), commuter rail (2.05 percent), and the aggregate transit (1.80 percent). Similarly, light rail shows a very high rate of 9.34 percent for the same level of increase over $4 of gasoline prices.

- Threshold boost effects of gasoline prices were found at the $3 mark for commuter rail and heavy rail, resulting in a substantially higher rate of ridership increase: 5.27 percent for commuter rail and 4.85-6.15 percent for heavy rail in response to a 10 percent increase in gasoline prices that crosses the $3 mark.

- The fare elasticity of transit ridership (i.e., the own-price elasticity) was generally found to be greater than the gasoline price elasticity and is consistent with findings from previous studies.

While the effects of gasoline prices on transit ridership obtained in this study are generally modest, compared to some of the findings in the other studies on the subject, the implication of more substantial effects found for gasoline prices over $3 is important. As it is likely that gasoline prices will remain above $3 per gallon and possibility increase in the future due to a market price increase and/or an increase in fuel taxes and potential carbon taxes, the effects of gasoline prices will be on the higher end of this study’s findings or even higher. Furthermore, while a ridership increase may be good news for transit agencies during the off-peak periods, even a small percentage of ridership increase can require a substantial increase in service supply and facility capacity during the peak periods when the service level is at or near the maximum supply capacity for transit agencies.

This study provides a more comprehensive understanding of the net effects of gasoline prices on transit ridership, which gives insight and guidance for how transit agencies will plan and prepare for accommodating higher transit travel needs of the public through pricing strategies, general financing, capacity management, and operations planning for different transit modes during times of substantial gasoline price increases.

I. INTRODUCTION

Between 1999 and 2011 consumers in the U.S. experienced an unprecedented gasoline price increase. Although gasoline prices sharply increased during the oil crisis due to Organization of the Petroleum Exporting Countries' (OPEC) oil embargo in 1973, the Iran-Iraq war in 1981, and Iraq's invasion of Kuwait in 1990, gasoline prices gradually declined thereafter and became stable (Figure 1). Gasoline prices spiked again in 2005 due to the confluence of a number of factors, including new, major, oil-consuming nations, aging U.S. refining infrastructure, and increased demand, and were further accentuated by the Hurricane Katrina disaster (Figure 2, Bomberg and Kockelman, 2007). In July 2008 gasoline prices exceeded $4 per gallon in nominal value and marked the highest price in real value in U.S. history (Figure 1 and Figure 2).

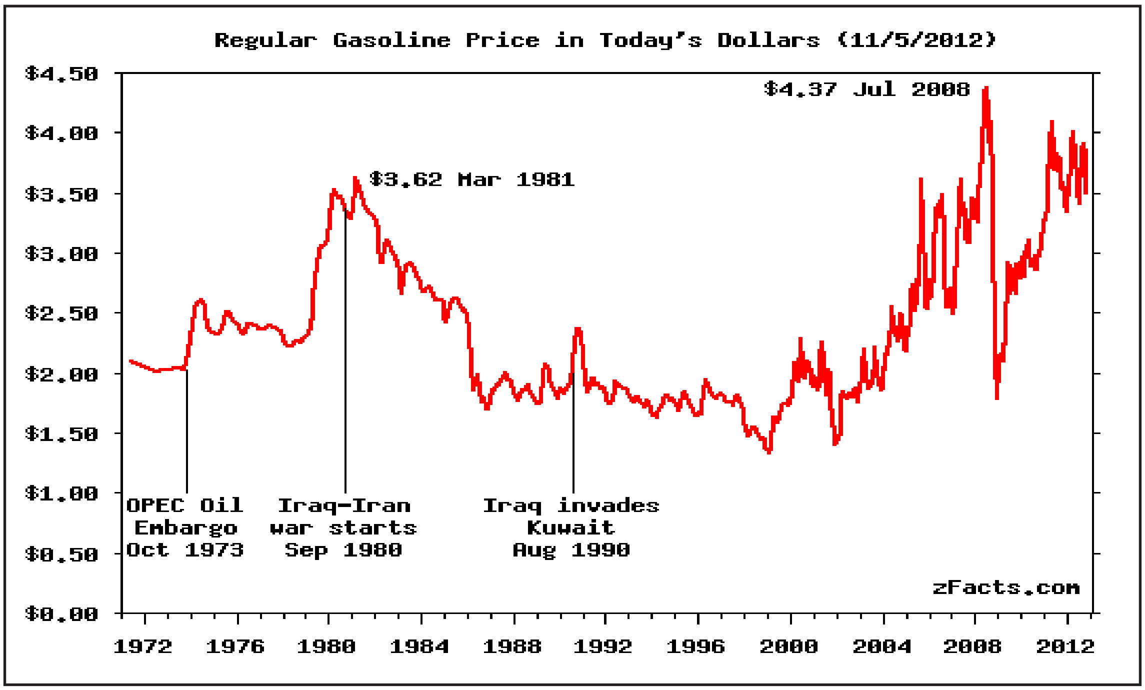


Figure 1. Historic Price of Regular Gasoline in the U.S. (in 2011 Dollars)

Source: Facts, http://zfacts.com/p/35.html

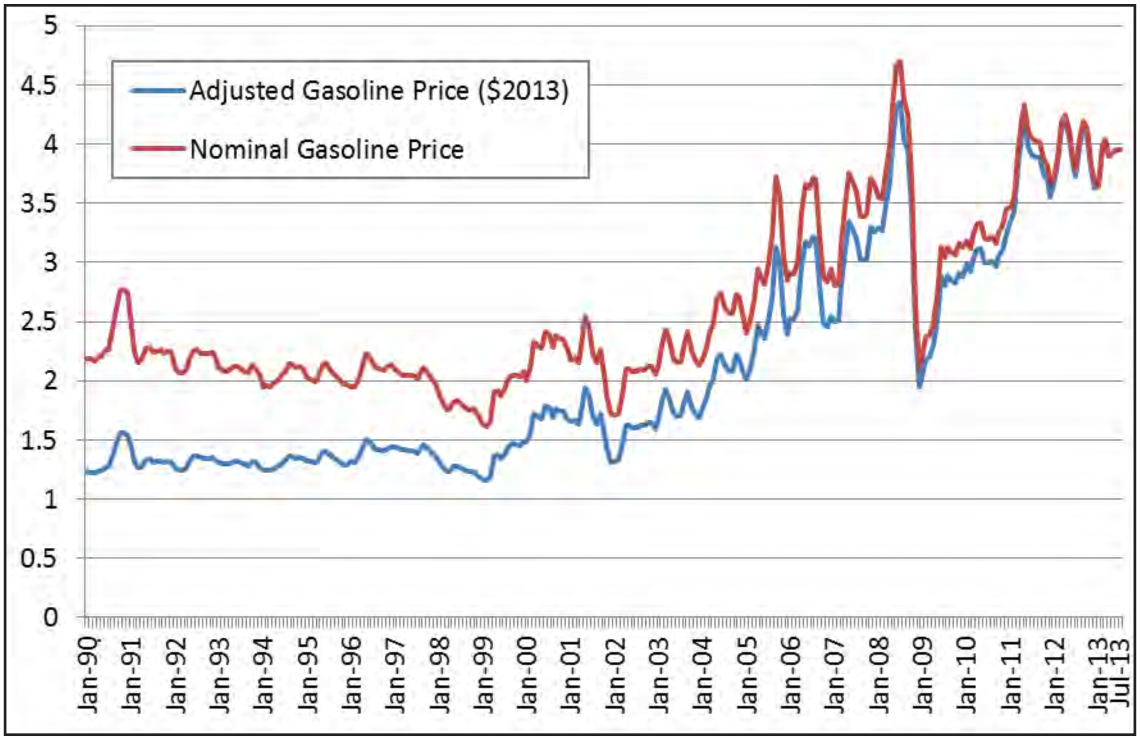


Figure 2. Nominal and Inflation-adjusted Gasoline Prices in the U.S. between 1990 and 2013 (adjusted in 2013 Dollars)

Source: Author's graph based on data from the Energy Information Administration, Department of Energy.

When gasoline prices substantially increased in 2008, drivers were reported to have adjusted their travel behavior by driving less and using transit (APTA 2012; Cooper 2009). High gasoline prices were also reported to have prompted drivers to shift to more fuel-efficient cars (Korkki 2009; Busse, Knittel, and Zettelmeyer 2009). The American Public Transportation Association (APTA) (2012) reported that total driving declined by 56 billion vehicle miles traveled (VMT) (1.9 percent) or by 91 billion person miles of travel (1.8 percent) between 2007 and 2008. APTA also reported that transit ridership rose by 5.2 percent during the second quarter of 2008 compared to the prior year, after an increase of 3.4 percent in the first quarter of 2008. Transit ridership in 2008 peaked with 10.7 billion trips, the highest level since the Federal-Aid Highway Act of 1956 (Cooper 2009). APTA attributed the decline in driving and the increase in transit ridership to the gasoline price increase, although it did not take into account other factors, such as economic conditions. In April 2011, gasoline prices in many urban areas surpassed the $4-per-gallon mark again and raised serious concerns among motorists.

Gasoline price increases in the last decade have generated interest in gaining insight through rigorous research into how people respond in their travel behavior to the fluctuation of gasoline prices. A substantial increase in travel cost due to rising gasoline prices can affect motorists' travel behavior—whether or not to take a trip, which place to travel, which mode of travel to take, and which route to take—in an effort to reduce expenditures on fuel. For transit agencies, the way people respond to gasoline prices means potential changes in transit service demand, as well as an increase in operating costs—particularly for bus services. The magnitude of change in transit ridership in response to a change in gasoline price is measured by *elasticity*, which is defined as the ratio of a percentage change in one variable to a percentage change in another variable. A comprehensive understanding of elasticity of ridership to gasoline prices for different modes of transit is important to guide transit agencies' preparation in terms of pricing strategies, capacity management, and supply of different modes of transit services during times of such gasoline price changes.

Given the importance of the subject, this study uses panel data of transit ridership and gasoline prices from ten major urbanized areas (UA) in the United States for the maximum of a ten-year period and controls for a comprehensive set of factors to estimate the short- and long-term effects of the price of gasoline on transit ridership for bus, light rail, heavy rail, commuter rail, and these four modes combined. An analysis using panel data allows us to simultaneously take into account temporal and cross-sectional variation to obtain more robust, generalizable results (Greene 2012).

The remainder of the report is organized as follows. Section 2 reviews the recent studies that analyzed the effect of gasoline prices on transit ridership with a focus on types of data and analytical methods used. Improving upon the past studies reviewed, Section 3 presents the panel data regression methods applied in this study. Section 4 describes data and data sources. Section 5 reports results from a series of panel data regression analyses. Section 6 provides a discussion of analysis findings and concludes with implications for transit planning, as well as potential improvements for future research.

II. LITERATURE REVIEW

The gasoline price increase in the U.S. fostered research that yielded a broad literature on the gasoline price elasticity of travel demand. The literature review in this section pays particular attention to limitations of recent studies and highlights some of the improvements that need to be made in the analytical method used to estimate the net effects of gasoline prices on transit ridership, which is measured by the elasticity of transit ridership to the price of gasoline.[1] In this case, the value of elasticity is the ratio of the percent change in transit ridership to the percent change in gasoline price. For example, an elasticity value of 0.10 indicates transit ridership increases by 1 percent in response to a 10 percent increase in the price of gasoline.

Some of the more recent studies specifically examined the effect of changes in the price of gasoline on public transit use. (Blanchard 2009; Currie and Phung 2007; Maley and Weinberger 2009). There is another group of studies that included an analysis of the effect of gasoline price along with other factors on transit use (Bomberg and Kockelman 2007; Chen, Varley, and Chen 2010; Kain and Liu 1999; Lane 2010; Mattson 2008; Novak and Savage 2013; Stover and Bae,2011; Taylor et al. 2009; Yanmaz-Tuzel and Ozbay 2010). The results obtained in recent literature on the subject of gasoline price elasticity of transit ridership significantly vary by location of study, mode of transit, type of data used, type of effect estimated (i.e., short-term or long-term effect), and estimation method used, as discussed in this review.

Table 1 shows that recent studies used a variety of locations and time periods to examine the causal effect of the price of gasoline on transit ridership. Some of these recent studies focus specifically on a city in the United States (Bomberg and Kockelman, 2007; Chen, Varley, and Chen 2010; Maley and Weinberger 2009; Yanmaz-Tuzel and Ozbay 2010) but their findings may not be generalized due to a lack of external validity. External validity becomes an issue when the results from data for one city may not be comparable to results from data for another city because the cities' characteristics differ from one another. Other studies compare the gasoline price elasticity of transit ridership in a few cities (Currie and Phung 2008; Kain and Liu 1999) while some others analyze transit ridership in a group of cities or urban areas (Blanchard 2009; Lane 2010; Haire and Machemehl 2007; Storchmann 2001; Taylor et al. 2009). Among the recent studies that study transit ridership only in cities or urban areas, Mattson (2008) is an exception because of its geographic focus on urban and rural areas in the U.S. Upper Midwest and Mountain States.

Table 1. List of Location and Years Analyzed in Recent Studies of Gasoline Price Elasticity of Transit Ridership

Study	Location	Year
Bomberg and Kockelman (2007)	Austin, Texas	February and April 2006
Kain and Liu (1999)	San Diego, CA; Houston, TX	1980; 1990
Taylor, Miller, Iseki and Fink (2009)	265 US urbanized areas	2000
Currie and Phung (2007)	US	1998-2005

Table 1, Continued

Study	Location	Year
Haire and Machemehl (2007)	5 US cities: Atlanta, Dallas, Los Angeles, San Francisco, and Washington DC	1999-2006
Maley and Weinberger (2009)	Philadelphia	January 2001-June 2008
Lane (2010)	9 US metropolitan areas: Boston; Chicago; Cleveland; Denver; Houston; LA; Miami; San Francisco; Seattle	January 2002/ June 2003-April 2008
Mattson (2008)	Urban and rural areas in upper midwest and mountain states: Duluth, MN; St. Cloud, MN; Rochester, MN; Sioux Falss, SD, Fargo, ND, Billings, MT, Grand Forks, ND; Missoula, MT; Great Falls, MT; Rapid City, SD; Cheyenne, WY; Logan, UT.	Time-series analysis: monthly data from January 1999-December 2006. Panel data analysis: annual data from 1997-2006.
Yanmaz-Tuzel and Ozbay (2010)	Northern New Jersey, with one line running between Atlantic City and Philadelphia.	1980 to 2008
Chen, Varley, and Chen (2010)	New Jersey and New York City	January 1996-February 2009
Storchmann (2001)	Germany, public transportation in urban areas of Germany	1980-1995
Curie and Phung (2008)	Melbourne, Brisbane, and Adelaide in Australia	Melbourne (January 2002-December 2005), Brisbane (July 2004-November 2006), Adelaide (January 2002-November 2006).
Blanchard (2009)	218 US cities	2002-2008
Stover and Bae (2011)	11 counties in Washington State	January 2004-November 2008
Nowak and Savage (2013)	Chicago metropolitan area	January 1999 and December 2010

EFFECTS OF GEOGRAPHIC LOCATION AND SCALE ON GASOLINE PRICE ELASTICITY

The diversity of geographic location and scale of studies by Bomberg and Kockelman (2007), Chen, Varley, and Chen (2010), Maley and Weinberger (2009), Yanmaz-Tuzel and Ozbay (2010), Currie and Phung (2008), Kain and Liu (1999), Blanchard (2009), Lane (2010), Haire and Machemehl (2007), Storchmann (2001), and Taylor et al. (2009) raises the question of how the gas price elasticity of transit ridership varies by city size and location (i.e., urban or rural or mix of urban and rural).

There are certain conditions that influence the magnitude of elasticity. First, the traveler must have the option to either drive or take public transit for his/her trip. A majority of current transit users in the U.S. are transit-dependent and do not have access to private automobiles. Zero-vehicle households represent the largest share of the transit market, accounting for 48.5 percent of trips while persons living in households with inadequate vehicles access account for an additional 17.1 percent (Chu 2012). For all income groups, transit use increases in urban areas (Pucher and Renne 2001). In the urban center of metro areas like New York, Washington DC, Chicago and San Francisco, individuals may choose to ride public transportation for convenience. For these individuals who do not own personal vehicles, travel behavior is not affected by gas prices. On the other hand, those who reside in suburban areas of large metropolitan areas, small cities, and rural areas are more likely to own a private car and potentially be a transit rider by choice (Pucher and Renne 2005). Second, to cause a switch in travel modes, the effect of a gasoline price increase on the generalized costs of making a trip (e.g., commuting trip, social trip) has to be substantial. Following this, it is likely that those who usually drive relatively long distances may be more sensitive to a gasoline price hike and may switch to public transit to avoid additional financial burden (Maley and Weinberger 2009; Currie and Phung 2007). At the same time, taking transit should not impose substantial non-monetary burdens, such as longer travel time and inconvenience—at least no more of a burden than an increase in monetary costs due to a fuel cost increase. These conditions lead to a variance in response based on residential location, mode of transit, and demographic characteristics.[2]

VARIATION IN GASOLINE PRICE ELASTICITY BY TRANSIT MODE

Table 2 shows that studies vary widely in the modes examined; some focus exclusively on rail (Chen, Varley, and Chen 2010) or bus (Kain and Liu 1999; Mattson 2008); some analyze the total ridership of multiple modes and multiple transit systems combined for a large geographic area, such as an urbanized area or metropolitan area (Taylor et al. 2009; Yanmaz-Tuzel and Ozbay 2010), while others consider each mode separately, as well as all modes combined (Currie and Phung 2007; Haire and Machemehl 2007; Lane 2010). Studies that focus only on bus or rail ridership may not be generalizable to other modes in these areas, as studies have asserted that the gasoline price elasticity of ridership varies by mode (Maley and Weinberger 2009). On the other hand, studies that use aggregate data for the entire transit system obtain the average effect of a change in the price of gasoline on all modes without distinguishing variance among different modes. Given that the characteristics of trips and travelers vary by mode—for example, travel distance of rail trips is usually longer than that of local bus trips, and bus riders are more likely to be transit-dependent for financial reasons than are rail riders[3]—it is important to conduct an analysis of the price of gasoline elasticity to transit ridership by mode.

Table 2. Type of Data and Mode Analyzed in Recent Studies of Gasoline Price Elasticity of Transit Ridership

Study	Type of Data	Mode	Level of Aggregation
Bomberg and Kockelman (2007)	Cross-section	Bicycle, driving, and transit	Individual household
Kain and Liu (1999)	Cross-section	Bus	Metro service area in Houston, and MTS service area in San Diego.
Taylor, Miller, Iseki and Fink (2009)	Cross-section	Total level of transit service provided by all transit agencies in an urbanized area	Urbanized area level
Currie and Phung (2007)	Time series	Bus, light rail, heavy rail, total for all modes combined	By mode for all of US
Haire and Machemehl (2007)	Time series	Bus, light rail, heavy rail, commuter rail	By mode for each of 5 US cities

Table 2, Continued

Study	Type of Data	Mode	Level of Aggregation
Maley and Weinberger (2009)	Time series	Rail provided by Regional Rail Division of SEPTA. Bus service, nine light rail or street car route service and two subway route service provided by City Transit Division of SEPTA.	By mode
Lane (2010)	Time series	Bus, rail, bus and rail combined	For nine cities combined, monthly data analysis by mode: bus, rail, and bus and rail combined. For each of 9 cities separately he analyzed monthly data by mode: bus, rail and bus and rail combined.
Mattson (2008)	Both time-series data, and panel data were used.	Bus	For time-series analysis he divided monthly data from upper midwest and mountain states into 4 groups of metropolitan areas based on population size. Four groups are above 2 million population, 500 thousand to 2 million, 100 thousand to 500 thousand, and below 100 thousand. For panel analysis, he used annual ridership data for each transit system.
Yanmaz-Tuzel and Ozbay (2010)	Time series	Overall New Jersey transit ridership	Monthly data for all modes combined in New Jersey
Chen, Varley, and Chen (2010)	Time series	New Jersey commuter rail	Monthly data on New Jersey commuter rail ridership
Storchman (2001)	Time series	All urban public transportation: bus, tram and underground	He ran the regressions at the mode of transport level for a given purpose of travel for example, work, leisure etc. using annual data
Curie and Phung (2008)	Time series	Rail, Australian (bus rapid transit) BRT, and bus	Using monthly data they ran regressions for each city separately after aggregating transit usage for all modes. They also ran city wide regression disaggregating at the rail, bus and bus rapid transit level.
Stover and Bae (2011)	Time-series, panel data	Aggregate transit ridership	Regress aggregate ridership for each county separately
Nowak and Savage (2013)	Time-series	City heavy rail, city bus and suburban bus and suburban rail	Regress ridership for each model separately
Blanchard (2009)	Panel data	Commuter rail, heavy rail, light rail and bus	Regress separately for each mode: motorbus, light rail, heavy rail, commuter rail using monthly data for 218 cities

A wide variety of studies analyze transit ridership either by mode or by regional system, and using different geographic scales (i.e., cities, regions or the entire country). This variation generates inconsistencies. For example, Currie and Phung (2007) show that national light rail ridership in the United States has the highest elasticity, with values ranging from

0.27 to 0.38; heavy rail follows, with elasticities from 0.17 to 0.19; and bus ridership has elasticities from 0.04 to 0.08. Currie and Phung (2007) speculate that a higher share of “choice” riders—those who own or could easily own an automobile—choosing light rail could explain the high values of gasoline price elasticity for light rail. Haire and Machemehl (2007) estimated the gasoline price elasticity of ridership in five U.S. cities for four different modes—bus, light rail, heavy rail, and commuter rail—and obtained results different from Currie and Phung’s study, with the lowest elasticities for light rail (0.07), followed by heavy rail (0.26), commuter rail (0.27), and bus (0.24).

Lane (2010) analyzed the effect of the price of gasoline on transit ridership for both bus and rail modes, as well as the total ridership for the two modes combined, in nine U.S. metropolitan areas. Lane (2010) found that in some cities gasoline price had a positive effect on bus ridership but no statistically significant effect on rail ridership; while in a few other cities it had a positive effect on rail ridership but no effect on bus ridership. These results further indicate that transit ridership elasticity varies by mode and by geographic location, and that the use of different data—in terms of both geographic location and scale—will yield different estimates of elasticity for each mode.[4]

VARIATION IN GASOLINE PRICE ELASTICITY BY TRIP CHARACTERISTIC

Previous studies show that the gasoline price elasticity of transit ridership varies significantly by trip purpose, which is another way of grouping transit trips. Storchmann (2001) found that the cross-price elasticity for public transit in Germany varies by trip purpose: 0.202 for work-related trips, 0.12 for school trips, 0.05 for leisure trips, 0.03 for shopping trips, and 0.02 for holiday trips. Travel distance also affects gasoline price elasticity of transit ridership as found by Currie and Phung (2008). For example, they found that the gasoline price elasticity of transit ridership is higher for longer distance travel in Melbourne. This finding is explained by the higher cost savings accrued by a mode shift from automobile to public transit for long-distance trips. These results suggest that it is important to take into account trip characteristics, such as transit mode, trip purpose, and travel distance, when analyzing the gasoline price elasticity of transit ridership. Information on trip purpose, however, typically has been omitted in regression analysis due to limited data availability. To address this limitation, surveys of transit riders need to be conducted to collect data on whether riders use different modes for different trip purposes and, if so, which mode they ride for which purpose.[5]

NON-CONSTANT ELASTICITY OF GASOLINE PRICE

Two recent studies raised a question about more complex effects of gasoline price on transit ridership. Chen, Varley, and Chen (2010) examined symmetry of price elasticity of transit ridership—whether the magnitude of elasticity is the same depending on an increase or decrease in gasoline price—and found that the ridership elasticity to a rise in the gasoline price is higher than the elasticity to a fall in the gasoline price. Maley and Weinberger (2009) suggest different levels of transit elasticity by the level of gasoline price; travelers may be much more sensitive to a gasoline price change between $2 and $3 per gallon, although they may not be sensitive to a change in a lower price range. When gasoline prices are in the low range of $2 to $3, travelers may be less conscious of their spending on gasoline since the overall expenditure is low; hence, they may be less sensitive to a change in price. In other words, the gasoline price elasticity of ridership is not constant and possibly has a threshold effect at a price of about $3 per gallon of gasoline. This idea motivated Maley and Weinberger to add a squared term for gasoline price along with a linear term in their regression analysis; however, they did so in an ad hoc manner without providing any theoretical basis for adding the squared term. The implication of these complex effects of gasoline price on transit elasticity for an analytical methodology is that the most commonly used simple log-log regression model, which assumes a constant elasticity regardless of the value of gasoline price or ridership, is not adequate and should be modified to capture this complexity in elasticity.

LAGGED EFFECTS OF GASOLINE PRICE CHANGES

Distinction between short- and long-term elasticity is important. Travelers may adjust to gasoline price hikes by making personal budget adjustments, decreasing non-work, discretionary travel, or linking discretionary trips together in the short term. In addition, it is less likely that they change travel mode in response to a gasoline price hike that does not continue for a particular period of time (Horowitz 1982; Keyes 1982; Yanmaz-Tuzel and Ozbay 2010). To address this difference between short- and long-term elasticity, a time-series data analysis for a city or transit system over longer time horizons is better suited, as it can capture temporal variation in gasoline prices (Maley and Weinberger 2009; Yanmaz-Tuzel and Ozbay 2010; Curie and Phung 2008).

As seen in Table 3, some studies use time-series data to analyze only short-term, instantaneous effects of a gasoline price change (i.e., the effect of a change in the price of gasoline measured over different time periods, either monthly or yearly) on transit ridership (Currie and Phung 2008; Maley and Weinberger 2009; Storchmann 2001), while other studies use time-series data to examine both short- and long-term effects (Chen, Varley, and Chen 2010; Mattson 2008; Yanmaz-Tuzel and Ozbay 2010). A few studies find that long-run effects of gasoline price change are statistically significant. Mattson (2008) used monthly data and a polynomial distribution lag model with 15 lags of gasoline price to analyze the long-term effects of changes in gasoline price on ridership and found that coefficients for gasoline price up to the seventh lag (i.e., the seventh month) were statistically significant. Studies by Keyes (1982), Litman (2004), and Yanmaz-Tuzel and Ozbay (2010) concluded that the long-term elasticity is larger than short-term elasticity, as a more lasting increase in gasoline price could provide a stronger incentive to switch travel modes and result in a higher demand for public transit trips.

Recent studies have used different types of data—cross-sectional, time-series, and panel (as shown in Table 2)—resulting in some variation in estimated values of elasticity. Studies that analyze cross-sectional data for households from a particular area (e.g., Bomberg and Kockelman 2007) are certainly important to transit service planning in that area, but the findings of these studies are not generalizable to transit systems in other areas due to lack of external validity. Cross-sectional studies that estimate the average gasoline price elasticity of many areas at a point in time (Kain and Liu 1999; Taylor et al. 2000) are inadequate to examine short- and long-term impacts of gasoline price on transit ridership, while they allow for the control of many other variables that could affect transit ridership.

As the effect of gasoline price change on ridership is inherently temporal, time-series data analysis has advantages over cross-sectional analysis in examining the effects. As previously mentioned, it is likely that results from a time-series analysis in a particular city or on a particular transit system may not be applicable to other cities and transit systems. Panel data analysis is advantageous, as it allows researchers to simultaneously take into account temporal and cross-sectional variation to obtain more robust, generalizable results over time (Greene 2012). Mattson (2008), however, conducts panel data analysis using yearly data for each transit agency, which does not allow examination of the long-term effect of a gasoline price change on public transit ridership within 12 months. Since it is possible to detect the full effect of a change in the price of gasoline in less than a year (Yanmaz-Tuzel and Ozbay 2010), use of yearly data in the panel data analysis by Mattson can limit the usefulness of this study.

Blanchard (2009) conducted a panel data analysis to analyze short-term and long-term effects of the change in gasoline prices on ridership of commuter rail, heavy rail, light rail, and bus, using data from 218 U.S. cities to show that the elasticity for light rail is the highest. This study found that long-term elasticities were higher than short-term elasticities for almost all modes. This finding is consistent with assertions made by the earlier studies (Chen, Varley, Chen 2010).

Table 3. Dependent and Independent Variables and Analytical Methods Used in Studies on the Gasoline Price Elasticity of Transit Ridership Using Time-series Data

Study	Dependent Variable	Independent Variables	Empirical Specification and Strategy
Currie and Phung (2007)	Log of national (US) transit ridership	Log of gas price, log of gas price interacted with dummies for 9/11 incident, the Iraq war and Hurricane Katrina, month dummies	Simple OLS, regressing log of dependent variable on log of independent variables
Haire and Machemehl (2007)	Change in ridership over two consecutive months	Price of gasoline	Simple OLS, regressing level of dependent on level of independent variables
Maley and Weinberger (2009)	Monthly ridership	Gas price, monthly dummies to control for seasonality	Simple OLS, regressing level of dependent on independent variables
Lane (2010)	Monthly unlinked passenger trips for bus, rail and rail and bus combined	Current gas price, one month lagged gas price, standard deviation of monthly gas price for each month, time trend, seasons such as fall, spring, summer, supply of transit variables such as vehicle revenue miles operated, vehicles operated in maximum service	Simple OLS, regressing level of dependent variable on level of independent variables
Mattson (2008)	Log of monthly ridership	For time-series data analysis: 15 lags of gas price, yearly dummy. For panel data analysis: Size of labor force, unemployment level, transit service and fare, time trend interacted with dummy indicating transit system, and dummy variables indicating whether there have been events to create demand shocks for any specific transit system.	For time-series analysis: polynomial distribution lag model to analyze long term effect of gas prices on ridership. He used a log-log model where the lagged gas prices are also logged. Panel data analysis: Simple OLS, regressing log of ridership on log of independent variables, which did not include lagged gas prices.
Yanmaz-Tuzel and Ozbay (2010)	Monthly transit ridership (in thousands)	Total monthly employment in New Jersey and New York City (in thousands), average monthly gasoline prices, lagged monthly gasoline prices, average NJ transit fare, vehicle revenue hours in thousands, month dummies.	Simple OLS, regressing log of dependent variable on log of independent variables
Chen, Varley, and Chen (2010)	Number of New Jersey commuter rail trips to and from New York City	They control for lagged ridership, positive and negative changes in gasoline price and transit fare, labor force and service level measured as vehicle revenue miles and its fourth lag, seasonal dummies (captured using monthly dummies).	They regress change in transit ridership between period t and t-1 on change in ridership between period t-1 and t-2 and change in gasoline price interacted with a dummy equal to 1 if the price change is non-negative and equal to 0 otherwise; similarly, they control for negative changes in prices by interacting the price with a dummy equal to 1 if the price change is negative, and 0 otherwise.

Table 3, Continued

Study	Dependent Variable	Independent Variables	Empirical Specification and Strategy
Storchman (2001)	Number of trips for work, school, shopping, business, leisure, and holiday, by mode. Average distance travelled each trip purpose by mode.	In the equation where they estimate choice of mode of transport for each travel purpose, they control for demographic variables, income, and a dummy indicating German unification in 1991. In the estimation of distance travelled using public transport for each purpose they control for gas price, stock of public transport, income and transit fare, available public infrastructure (such as railroads or road network) and German unification. In the estimation of demand for passenger kilometers, they control for purpose of trip, distance travelled, seats per vehicle, average peak seat load factor during peak period and average speed during peak periods.	He estimates a system of equations. Estimate how demography and German unification in 1991 affected number of trips taken for each of these travel purposes: work, school, shopping, business, leisure, holiday. Then estimate average distance of trip for each of the purposes listed above are affected by stock of cars, transportation prices, available railroads, or road network, and German unification. Then he estimates the public transit vehicle demand as a function of peak passenger kilometers, seats per vehicle, average peak seat load factor during peak period and average speed during peak periods. Then he estimates the cross-price elasticity for public transportation demand.
Curie and Phung (2008)	Per capita validations (which is equivalent to per capita transit usage)	Gasoline price, interest rate, and monthly dummy variables to indicate seasonality	Simple OLS, regressing log of per capita transit usage on log of gasoline prices, absolute level of interest rate, and monthly time dummies
Stover and Bae (2011)	Unlinked revenue trips	Gas price, transit fare, supply of transit, unemployment rate, size of labor force, season dummies	Simple OLS, regressing log of ridership on log of independent variables
Nowak and Savage (2013)	Unlinked trips for CTA bus, count of passengers entering stations for CTA rail, number of ticket sales for Metra, number of boardings for Pace	Gas price, gas price interacted with dummy that is equal to one if gas price is more than $3, gas price interacted with dummy that is equal to one if gas price is more than $4, average daily transit bus miles, transit fare, unemployment rate, proportion of weekdays in month, dummy variable for leap year	Simple OLS, regressing log of ridership on log of independent variables
Blanchard (2009)	Ridership measured as unlinked passenger trips by mode: commuter rail, heavy rail, light rail, motorbus	Supply of transit, gasoline price, and lagged gasoline prices, monthly dummies, year dummies	Simple OLS, regressing log of ridership on log of current and past gas prices

DISCUSSION OF OMITTED VARIABLES

Several recent studies of time-series analysis (Chen, Varley, and Chen 2010; Currie and Phung 2007; Lane 2010; Maley and Weinberger 2009; Mattson 2008; Yanmaz-Tuzel and Ozbay 2010; Currie and Phung 2008), as shown in Table 3 may suffer from omitted variables bias. Omitted variable bias arises when studies do not comprehensively account for the effect of changes in external and internal factors on transit ridership in a regression analysis. With this bias, it is not possible to isolate how much the fluctuation in gasoline prices alone contributes to changes in ridership (i.e., measure the net effect of gasoline price changes on transit ridership).

Some studies listed in Table 3 simply analyze the change in transit ridership as a result of a change in the price of gasoline without controlling for factors either external or internal to transit agencies (Haire and Machemehl 2007; Maley and Weinberger 2009), while other studies control for external factors but not internal factors (Bomberg and Kockelman 2007). External factors refer to factors outside the control of transit agencies, such as the regional economy, demographic changes, changes in highway infrastructure, and availability of parking, while internal factors are those over which transit agencies have a certain degree of control, such as fare levels, service coverage, operating hours, frequency (or headway), and service. As Kain and Liu (1999) analyzed the factors that affect transit ridership by selectively controlling for some internal and external factors, their analysis could face the omitted variables bias because the factors controlled are not comprehensive and because they do not control for gasoline price.

Table 4 lists studies on gasoline price elasticity of transit ridership that have used cross-sectional data. Comparing the past studies on transit ridership in Table 3 and Table 4, Taylor et al. (2009) included the most comprehensive list of influential factors in its cross-sectional regression analysis, investigating how each of the internal and external factors affects total urbanized area ridership and per capita ridership. The authors attempted to isolate the degree of change in ridership attributable to the fluctuation in gasoline prices alone, (i.e., the net effect of gasoline price changes on transit ridership) by controlling for a comprehensive list of variables that could also influence transit ridership. These variables included internal factors, such as fares, frequency of service, hours of service, on-time performance, service coverage, and quality of service, and external factors, such as measures of regional economic activity, population, population density, labor market, availability of parking in the CBD, and socioeconomic demographics of the population (age, income, vehicle ownership, etc.). Understanding the relative importance of these various factors and the interaction between them is very important since transit agencies could possibly control internal factors in order to achieve their goals and objectives while they cannot affect external factors (Taylor et al. 2009).

Although panel data analysis by Blanchard (2009) provides a methodological step in the right direction, the estimated values of the gasoline price elasticity obtained in this study are also likely to suffer from omitted variable bias. Blanchard did not attempt to include a set of dummy variables indicating city (i.e., city fixed effects). Such variables could have controlled for some of unobserved characteristics that do not vary over time but vary among geographic locations.

In addition, given that transit ridership could be influenced by changes in the level of potential rider activities, such as schooling and touring, some studies employ dummy variables to represent quarters of the year (Lane 2010) or months (Blanchard 2009; Chen, Varley, and Chen 2010; Currie and Phung 2008; Maley and Weinberger 2009; Yanmaz-Tuzel and Ozbay 2010) to account for seasonal or monthly variation in transit ridership, respectively. Since the use of monthly dummy variables allows variation over a shorter time period, it is considered a more general approach than the use of quarterly dummies. Unlike quarterly dummies, monthly dummies allow controls for factors that change on a monthly basis.

Table 4. Dependent and Independent Variables and Analytical Methods Used in Studies on Gasoline Price Elasticity of Transit Ridership Using Cross-sectional Data

Study	Dependent Variable	Independent Variables	Empirical Specification and Strategy
Bomberg and Kockelman (2007)	Shopping around for gas, overall driving, chaining activities, carpooling, transit use, and bicycle trips	Respondent's transportation needs, demographic attributes such as age, gender, income, student or not, household size, number of vehicle per driver. Neighborhood/local characteristics such as local population, whether or not the area is residential, or commercial area, retail employment, service employment, total employment in the area, distance to CBD, bus stop density, zone density. Gas expenditure, fuel economy of all household vehicles, no. of non-work related trips, whether or not works from home, whether household has children going to school.	They used ordered probit models to examine likelihood of respondents increasing trip chaining or reducing their driving, and taking public transit in response to the 2005 gas price spike
Kain and Liu (1999)	Log of ridership	Standard Metropolitan Statistical Area employment, central city population, bus and rail miles supplied by the transit system in the area, real fares.	Using simple OLS model, they regressed log of ridership on log of independent variables
Taylor, Miller, Iseki and Fink (2009)	Total urbanized area ridership, Per capita ridership	Geographic land area, total population, population density, regional dummy, median household income, ratio of unemployed to labor force, ratio of enrolled college students an total population, ratio of population in poverty to total population, ratio of immigrant population to total population, percent of votes cast for democratic party in 2000 presidential election, freeway lane miles, average gas price per gallon of gas, ratio of sum of non-transit and non-SOV commutes to all commutes, ratio of household with no vehicle to total household, total lane miles, daily vehicle miles travelled per capita. They also control for transit system characteristics, such as transit fares, headways/ service frequency.	They used two-staged least squares estimation strategy and instrumented supply of transit, measured as total urbanized area transit service vehicle revenue hours with total population, percent voting Democrat in 2000 presidential election

ENDOGENEITY BETWEEN TRANSIT SERVICE SUPPLY AND RIDERSHIP

To analyze the effect on ridership of internal factors, such as transit service supply, it is necessary to recognize that the level of transit service supply is highly correlated with ridership. The level of transit consumption can significantly affect the supply of transit service, as transit agencies adjust the supply of transit service within financial constraints in response to ridership levels (Taylor et al. 2009). The provided level of transit service directly influences the consumption of transit trips. Where a portion of the high demand for trips is not accommodated by the existing level of service, greater supply of service in the form of higher service frequency or extended operating hours leads to higher ridership. At the same, transit ridership can influence the level of supply as transit agencies increase or decrease the supply in response to fluctuations in ridership as well as cost to provide service and available funding. In other words, while transit supply affects transit demand and ridership, transit demand and ridership simultaneously affect transit supply. This leads to the endogeneity bias. While conceptually straightforward, most studies do not account for this potential endogeneity bias that arises from the simultaneity between transit supply and demand/consumption. Although cross-sectional data analysis by Kain and Liu (1999) and time-series data analyses by Chen, Varley, and Chen (2010), Lane (2010), Mattson (2008), and Yanmaz-Tuzel and Ozbay (2010), included a variable for transit service supply to examine metro ridership, these analyses did not take into account the fact that transit supply is endogenous to transit ridership.

The dependent variable in a regression analysis—ridership—is not only influenced by the explanatory variables of service supply, it also may influence them. In this case, a more appropriate econometric framework is to simultaneously estimate ridership and transit service supply. This leads to approaches such as the two-stage least squares (2SLS) or simultaneous equation structure to cope with the potential biases in estimated coefficients in regression. The study by Taylor et al. (2009) is one that accounted for endogeneity between ridership and transit service supply by using total population and percentage of the population voting Democrat in the 2000 Presidential Election as instrumental variables (IVs) that predict an urban area's level of transit supply—measured as total vehicle revenue hours (or in-service vehicle hours). They find that transit supply and external factors, such as the metropolitan economy, regional geography, population characteristics, and highway system characteristics, affect transit ridership.

In theory, as long as the instrumental variables are valid, this method enables the authors to obtain an unbiased estimate of the effect of supply on ridership. However, the IVs used by Taylor et al. (2009) may violate the assumption of exclusion restriction—one of the two key assumptions in the IVs estimation method—which requires at least one of the instruments, total population or percentage voting Democrat, to affect ridership only indirectly through their effects on transit supply but not transit ridership directly (Wooldridge 2002). In addition to the high likelihood that population directly affects transit ridership, it may be the case that those who vote for Democrats likely use public transit more (Florida 2013). This could be part of the reason that Democratic leaders may provide more funding for public transit, compared to Republican leaders.

SUMMARY OF LITERATURE REVIEW

To review, as demographic characteristics of transit riders and their trip characteristics can differ substantially by mode of transit, it is important to conduct an analysis of gasoline price elasticity of transit ridership by mode when data are available. Results from studies that focus on a particular transit system or a geographic area may not be transferable to other systems or locations due to the absence of external validity.

This literature review has found two studies—those by Blanchard (2009) and Mattson (2008)—that used panel data analysis. These studies allow researchers to simultaneously take into account temporal and regional variation to obtain more robust, generalizable results; however, they have their limitations. Mattson (2008) is not able to distinguish between short-term and long-term effects due to the use of annual data, and Blanchard (2009) omits some key variables from his analysis and does not account for the endogeneity of transit service supply with ridership. Most of the studies using time-series data do not control comprehensively for external and internal variables as Taylor et al. (2009) does. Although Taylor et al. (2009) controlled for a comprehensive set of factors, the instruments used in their study may not be valid.

In short, this literature review reveals several shortcomings in past studies that analyze the causal effect of a change in gasoline price on transit ridership.

III. ANALYTICAL METHODOLOGY

This study improves upon research methods used in past studies by addressing important econometric issues in estimating elasticity of transit ridership to gasoline prices.

This study uses the panel data analysis method, as it allows us to simultaneously take into account temporal and cross-sectional variation to obtain more robust, generalizable results than studies that use either cross-sectional or time-series data (Greene 2012). As the effect of gasoline price change on ridership is inherently temporal, time-series data analysis has advantages over cross-sectional analysis in examining temporal and lagged effects of changes in gasoline prices (i.e., a ridership increase with a time delay). However, results from a time-series analysis on a particular city or a particular transit system also lack generalizability due to issues of external validity and transferability. This is because transit services, the built environment, and sociodemographic characteristics of residents and workers could substantially vary from one city to another, and these conditions are difficult to capture with typical variables. Such conditions include travel distance, duration of each trip, level of accessibility to various socioeconomic opportunities, and the level of service coordination among multiple transit agencies. Results from this study are more generalizable compared to studies that focus specifically on one or a few U.S. cities. In addition, unlike studies that focused only on either one mode or the aggregate ridership, this study examines four modes—bus, commuter rail, light rail and heavy rail—as well as the aggregate of all modes.

The following sub-sections describe econometric specifications within the framework of panel data analysis that address potential biases due to omitted variables and lack of consideration for the endogeneity of transit service supply. Additional econometric specifications were also considered to examine lagged effects and the nonlinear nature of effects that gasoline prices may have on transit ridership.

BASELINE SPECIFICATION

Equation (1) expresses a baseline specification in a series of analyses using panel data. It controls for a wide range of variables, including regional geography, metropolitan economy, population characteristics, highway system, and transit system characteristics.

$$R_{it}^{M} = \alpha_0 + y_{it}^{M}{}'\alpha + \delta P_{it} + X_{it}^{'}\beta + \mu_i + \eta_t + \kappa_y + \varepsilon_{it}^{M} \qquad \text{Equation (1)}$$

where

R_{it}^{M} denotes transit ridership in urbanized area (UA) i at time t for mode M (bus, light rail, heavy rail, commuter rail, and the aggregate).

y_{it}^{M} denotes a vector of internal influential factors that include the levels of transit service supply measured as vehicle revenue hours (VRH), service frequency, and average fare for mode M, in UA i at time t.

An apostrophe, (‘) in an equation represents the transpose of a vector.

P_{it} denotes the price of gasoline in UA i at time t.

X'_{it} is a vector of external influential factors that include total population, number of recent immigrants, federal highway miles, mean household income, unemployment rate, and percent of households with no vehicle in UA i at time t.

μ_i denotes UA level fixed effects.

η_t denotes monthly fixed effects to represent seasonal variation in transit ridership.

κ_y denotes yearly fixed effects.

ε_{it}^{M} denotes the stochastic error term corresponding to the regression for mode *M.*

α_0 denotes the intercept parameter.

α and *β* are vectors of slope parameters associated with vectors of internal and external influential factors respectively.

δ denotes the slope parameter associated with the price of gasoline.

The baseline specification regresses the log of ridership (R_{it}^{M}) on a set of multiple independent variables that include gasoline prices. While we examined many other variables, such as level of transit service supply, which is measured by vehicle revenue hours (VRH), and average fare, which is obtained by dividing total fare revenue by ridership, the actual set of variables varies by each model due to considerations for multicollinearity and a parsimonious specification. Many of these independent variables are in the natural logarithmic form of an original variable except for the unemployment rate and the proportion of households with no vehicle, which are percentages.

The baseline specification also includes area-level fixed effects (μ_i) to control for unobserved time-invariant factors that affect transit ridership in each UA, monthly dummy variables (η_t) to account for the seasonal variation in transit ridership, and yearly dummy variables (κ_y) to account for macroeconomic effects that change from year to year and affect ridership of transit systems in all UAs in the same manner. This baseline specification assumes no endogeneity between transit service supply and ridership. Results from this specification will be compared to those obtained from the following specification that applies instrumental variables to address the potential endogeneity problem.

MODEL FOR INSTRUMENTAL VARIABLES METHOD

The instrumental variables estimation method is used to obtain an unbiased estimate of the effect of transit service supply on ridership with the presence of endogeneity between transit service supply and ridership. Equations (2) and (3) present the model specification that accounts for simultaneity in transit service supply and ridership.[6] Specifically, Equation (2) shows a specification of the first stage of the two-stage regression analysis in which a variable of transit service supply measured by VRH is regressed on instrumental variables and other variables that are included in the second stage. Equation (3) represents the second stage in which transit ridership is estimated using the predicted value of VRH from Equation (2).

$$y_{it}^{M} = \alpha_1 + Z_{it}^{'}\gamma + \delta_1 P_{it} + X_{it}^{'}\theta + \breve{y}_{it}^{M'}\pi + \mu_i + \eta_t + \kappa_y + u_{it}^{M} \qquad \text{Equation (2)}$$

$$R_{it}^{M} = \alpha_0 + \hat{y}_{it}^{M}\alpha + \delta P_{it} + X_{it}^{'}\beta + \breve{y}_{it}^{M'}\omega + \mu_i + \eta_t + \kappa_y + \varepsilon_{it}^{M} \qquad \text{Equation (3)}$$

where

R_{it}^{M}, P_{it}, $X_{it}^{'}$, μ_i, η_t, κ_y, α_0, β and δ are the same as defined in Equation (1).

y_{it}^{M} denotes the supply of transit service measured as VRH for mode *M* in UA i at time t.

$\hat{y}_{it}^{M}$ denotes the predicted value of transit supply from the first stage.

$\breve{y}_{it}^{M}$ denotes the vector of variables of transit fare and service frequency for mode *M* in UA i at time t.

α_1 denotes the intercept parameter in the first stage estimation (Equation 2).

α denotes the slope parameter associated with the predicted value of transit supply.

Z_{it}' denotes the vector of excluded instruments that affect only transit supply but does not affect ridership in UA i at time t.

π denotes the vector of slope parameters associated with the vector of variables of transit fare and service frequency for mode *M* in UA i at time t in the first-stage estimation (Equation (2)).

ω denotes the vector of slope parameters associated the vector of variables of transit fare and service frequency for mode M in UA i at time t in the second stage estimation (Equation (3)).

θ denotes the vector of slope parameters associated with vectors of internal factors in the first stage estimation represented in Equation (2).

γ denotes the vector of slope parameters associated with the vector of excluded instruments.

δ_1 denotes the slope parameter associated with price of gasoline in the first stage estimation shown in Equation (2).

u_{it}^{M} and ε_{it}^{M} denote the stochastic error terms corresponding to the first and second stages respectively.

The instrumental variables in Equation (2) are: (1) total number of employees, (2) total fleet measured as the total seating and standing capacity of transit vehicles, and (3) total funds available for transit agencies in each UA in a particular year, combining local, state and federal funds. For these three instrumental variables, two conditions—instrument relevance and instrument exogeneity—need to be met. The first condition, instrument relevance, requires that the covariance between the instruments and supply of transit service cannot be zero. In other words, the three instruments should significantly affect supply of transit services. The three instruments selected in this study meet the condition of instrument relevance because the three instruments are jointly statistically significant as demonstrated by the large F-statistics reported later in Table 9.

The second condition, instrument exogeneity, requires that the instruments cannot directly affect transit ridership, but do so only by affecting the supply of transit services. The three instruments selected in this study satisfy the condition of instrument exogeneity, as all of them directly affect the supply level of transit service. In other words, these variables determine how much service transit agencies can produce. In addition, it is unlikely that typical riders would even notice or pay attention to conditions of these three variables prior to their use of transit, and it is also unlikely that their decision whether to take transit is influenced by these factors. Therefore, these three excluded instruments do not directly affect ridership.

Also in the analysis using the instrumental variables method, a log-log model is used, taking the natural logarithmic form of most of the independent variables in Equations (2) and (3) except the unemployment rate and the proportion of households with no vehicle, so that estimated coefficients can be interpreted as elasticity.

MODEL FOR TESTING SHORT- AND LONG-TERM EFFECTS

Changes in gasoline prices may influence transit ridership, not immediately but with some time lag or possibly in the long-term. A finite distributed lag model is used to analyze whether there are lagged effects of changes in gasoline prices, as shown in Equation (4). This equation (4) is a modified version of Equation (1) and it includes monthly lagged variables of gasoline prices.

$$R_{it}^{M} = \alpha_0 + y_{it}^{M}{}'\alpha + \delta_0 P_{it} + \delta_1 P_{i,t-1} + \ldots \delta_N P_{i,t-N} + X_{it}^{'}\beta + \mu_i + \eta_t + \kappa_y + \varepsilon_{it}^{M} \quad \text{Equation (4)}$$

where

$R_{it}^{M}, y_{it}^{M}, X_{it}^{'}, \mu_i, \eta, \kappa_y, \alpha_o, \alpha, \beta$ and ε_{it}^{M} are the same as defined in Equation (1).

$P_{it}, P_{i,t-1}$ and $P_{i,t-N}$ denote the gasoline price in UA i at time t, t-1, and t-N, respectively.

δ_0, δ_1 and δ_N denote the coefficients associated with the gasoline price in UA i at time t, t-1, and t-N, respectively.

In the models that analyze short- and long-term effects, specifications with different combinations of monthly lagged gasoline prices were tested to find a set of the variables with statistically significant coefficients. As the lagged gasoline prices are included in a logarithmic form in the model, estimated coefficients are interpreted as elasticity.

MODEL FOR TESTING NON-CONSTANT ELASTICITY

Boost effect and unequal elasticity in different price range

When gas prices exceeded $4 in 2008, popular media reported that people switched from driving to taking public transit. Similarly, Maley and Weinberger (2009) suggested that the gasoline price elasticity of transit ridership might be non-constant, including some boost effects at particular values of gasoline price. Following these, this study investigates whether there are any boost effects at gas prices of $2, $3, and $4 that result in an increase in transit ridership in addition to an increase represented by elasticity.

In addition, this study further investigates if the gas price elasticity of transit ridership varies for different ranges of gasoline prices. Equation (5) below shows a specification of regression analysis to examine: 1) boost effects at $2, $3, and $4 marks, and 2) different values of the gas price elasticity of transit ridership for the gasoline price ranges of less than $2, $2-$3, $3-$4, and greater than or equal to $4.

$$R_{it}^{M} = \alpha_0 + y_{it}^{M}\alpha + \partial_1 D_{it}^{2} + \partial_2 D_{it}^{3} + \partial_3 D_{it}^{4} + \delta_1 P_{it} + \delta_2 P_{it}^{2} + \delta_3 P_{it}^{3} + \delta_4 P_{it}^{4} + X_{it}^{'}\beta + \mu_i + \eta_t + \kappa_y + \varepsilon_{it}^{M}$$

Equation (5)

where

R_{it}^{M}, y_{it}^{M}, $X_{it}^{'}$, μ_i, η_t, κ_y, α_0, α and β are the same as defined in Equation (1).

D^{2}_{it} denotes a dummy variable that is equal to one when gasoline price is equal to or higher than $2 (≥ $2).

D^{3}_{it} denotes a dummy variable that is equal to one when gasoline price is equal to or higher than $3 (≥ $3)

D^{4}_{it} denotes a dummy variable that is equal to one when gasoline price is equal to or higher than $4 (≥ $4).

∂_1, ∂_2, and ∂_3 denote estimated coefficients for gasoline price and the three dummy variables described above.

P^{2}_{it} denotes a variable that has a value of zero when gasoline price is lower than $2 and has a value of the natural logarithm of gasoline price minus the natural logarithm of 2 when gasoline price ≥ $2.

P^{3}_{it} denotes a variable that has a value of zero when gasoline price is lower than $3 and has a value of the natural logarithm of gasoline price minus the natural logarithm of 3 when gasoline price ≥ $3.

P^4_{it} denotes a variable that has a value of zero when gasoline price is lower than $4 and has a value of the natural logarithm of gasoline price minus the natural logarithm of 4 when gasoline price ≥ $4.

δ_1, δ_2, δ_3, and δ_4 denote estimated coefficients for gasoline price and three dummy variables described above.

By subtracting the natural logarithm of 2, 3, and 4 from the natural logarithm of gasoline price, estimated coefficients for D^2_{it}, D^3_{it}, and D^4_{it} can be converted, using an exponential function (i.e. e^x). The resulting value minus 1 can be interpreted as a boost effect. For example, assuming the estimated coefficient of a dummy variable is 0.007528, a boost effect can be calculated as [$e^{0.007528}$ – 1 = 1.007556], which means it increases ridership by 0.7556 percent.

METHODOLOGICAL APPROACHES SPECIFIC TO THIS STUDY

In summary, this study applies panel data analysis methods to obtain results that are more generalizable than those obtained from time-series analysis, taking into account the inherent temporal property of the effects of gasoline prices on transit ridership. This study also examines four different modes of transit in addition to the aggregate of these modes.

The panel data set analyzed in this study contains a more comprehensive set of internal and external factors that influence transit ridership in order to address the omitted variable bias in estimated coefficients, mirroring the cross-sectional data analysis by Taylor et al. (2009). When such a bias is present, it is not possible to isolate the net effect a change in gasoline price has on transit ridership. This study also examines a specification of instrumental variables regression method to address potential bias in the estimated coefficients due to the simultaneity issue between transit service supply and ridership. In addition, urbanized area dummy variables are included to control for some of the unobserved characteristics that do not vary over time but vary between geographic locations. Following several past studies, monthly dummy variables are included to account for the temporal variation in transit ridership related to factors that exhibit seasonal patterns, including schooling and touring. Finally, this study also examines the possibility of lagged and nonlinear effects of gasoline prices on ridership using panel data analysis. Nonlinear effects include the boost effect of $2, $3, and $4 price points and different values of elasticity in different ranges of gasoline price.

IV. DATA AND DATA SOURCES

The analysis in this study uses data on monthly average gasoline prices based on weekly prices of three different types of gasoline—regular, midgrade, and premium—collected from the U.S. Energy Information Administration from January 2002 to December 2011 for ten urbanized areas (UAs) in the United States—namely Boston, Chicago, Cleveland, Denver, Houston, Los Angeles, Miami, New York, San Francisco and Seattle. Table 5 shows these ten UAs and the proportion of the total number of observations from each UA.

Table 5. Distribution of Observations across Urbanized Areas by Mode

	Proportion of Observations from Each UA in a Dataset for Each Mode (%)				
Urbanized Area	**Bus**	**CR**	**LR**	**HR**	**All Transit**
Boston, MA	9.2	13.1	10.8	12.0	9.1
Chicago, IL	10.6	15.2	-	15.6	10.6
Cleveland, OH	9.2	-	10.8	13.5	9.1
Denver, CO	10.6	-	14.2	-	10.6
Houston, TX	10.6	-	11.4	-	10.6
Los Angeles, CA	10.6	15.2	12.7	15.6	10.6
Miami, FL	9.2	13.1	0.8	12.0	9.1
New York, NY	10.6	15.2	12.7	15.6	10.6
San Francisco, CA	10.6	15.2	14.2	15.6	10.6
Seattle, WA	9.2	13.1	12.3	-	9.1
Total Number of Observations	1,126	789	840	761	1,132

The analysis uses monthly data on unlinked passenger trips for bus, commuter rail, light rail, and heavy rail, collected from the National Transit Database (NTD) for the transit ridership variable. Unlinked passenger trips for the four different modes were also aggregated to generate the total transit ridership for each UA. Other variables used in the analysis include monthly data for vehicle revenue hours (VRH) and vehicle revenue miles (VRM) as the supply of transit services, average service frequency (VRM divided by route miles), total number of employees (number of full-time employees + 0.5 * number of part-time employees), total fleet capacity (seating and standing capacity) by mode, and total available funds in thousands of dollars (sum of state, local and federal funds available to transit agencies). All of these variables were collected from the NTD and processed to obtain data for each of the ten UAs on an annual basis from 2002 to 2011. Demographic and socioeconomic variables, which were obtained from the American Community Survey (ACS) 1-year estimates between 2005 and 2011, include total population, number of recent immigrants, mean household income, unemployment rate, percent of households with no vehicle, number of workers that carpool, number of people in different age groups, number of people working in different industries, and college and graduate school enrollment in each UA. The analysis also uses annual data on federal highway miles from 2002 through 2010 collected from the Highway Statistics Series prepared by the U.S. Department of Transportation Federal Highway Administration.

Table 6 shows the summary statistics of the variables used in this study's analysis.

Table 6. Descriptive Statistics of Explanatory Variables Considered for Regression Analysis

Variables Used in Regression	Mean	Standard Deviation
Unlinked passenger trips: bus	25,494,332	29,98,5045
Unlinked passenger trips: commuter rail	4,651,073	6,944,772
Unlinked passenger trips: light rail	2,209,016	2,132,060
Unlinked passenger trips: heavy rail	34,233,693	67,177,982
Unlinked passenger trips: all modes combined	55,095,686	93,021,065
Average of regular midgrade and premium gasoline price	2.61	0.74
Bus fare	0.81	0.26
Commuter fare	3.54	1.18
Light rail fare	0.79	0.28
Heavy rail fare	1.08	0.64
Transit fare	0.94	0.28
Vehicle revenue hours: bus	598,822	579,093
Vehicle revenue hours: commuter rail	94,055	153,020
Vehicle revenue hours: light rail	24,324	22,571
Vehicle revenue hours: heavy rail	384,020	771,348
Vehicle revenue hours: all modes combined	956,911	1,335,703
Frequency of service: bus	975	290
Frequency of service: commuter rail	2,458	2,025
Frequency of service: light rail	6,173	3,184
Frequency of service: heavy rail	26,842	24,711
Frequency of service: all modes combined	1,536	1,001
Total standing+seating capacity for bus	201,780	200,463
Total standing+seating capacity for commuter rail	138,715	182,406
Total standing+seating capacity for light rail	21,033	21,588
Total standing+seating capacity for heavy rail	200,532	353,223
Total standing+seating capacity for all modes combined	455,390	659,999
Total number of employees	14,990	22,284
Total funds available to transit agencies (in thousands of dollars)	2,441	1,701
Total population	7,327,483	5,068,340
Foreign-born population - naturalized U.S. citizen	883,610	840,711
Federal highway miles (urban and rural)	284,425	142,052
Mean household income	79,854	11,387
Unemployment rate	8.11	2.31
Households with no vehicle	4.15	3.19

Figures 3-6 show the relationship between unlinked passenger trips—representing ridership of bus, commuter rail, light rail and heavy rail in Boston—and four different gasoline prices: regular, midgrade, premium, and the average of the three. Appendix B contains graphs for all of the UAs. It is important to note that not all UAs offer all modes of transit service. While buses are available in all ten UAs, commuter rail is not available in

Cleveland, Denver or Houston; light rail is not available in Chicago; and heavy rail is not available in Denver, Houston or Seattle.

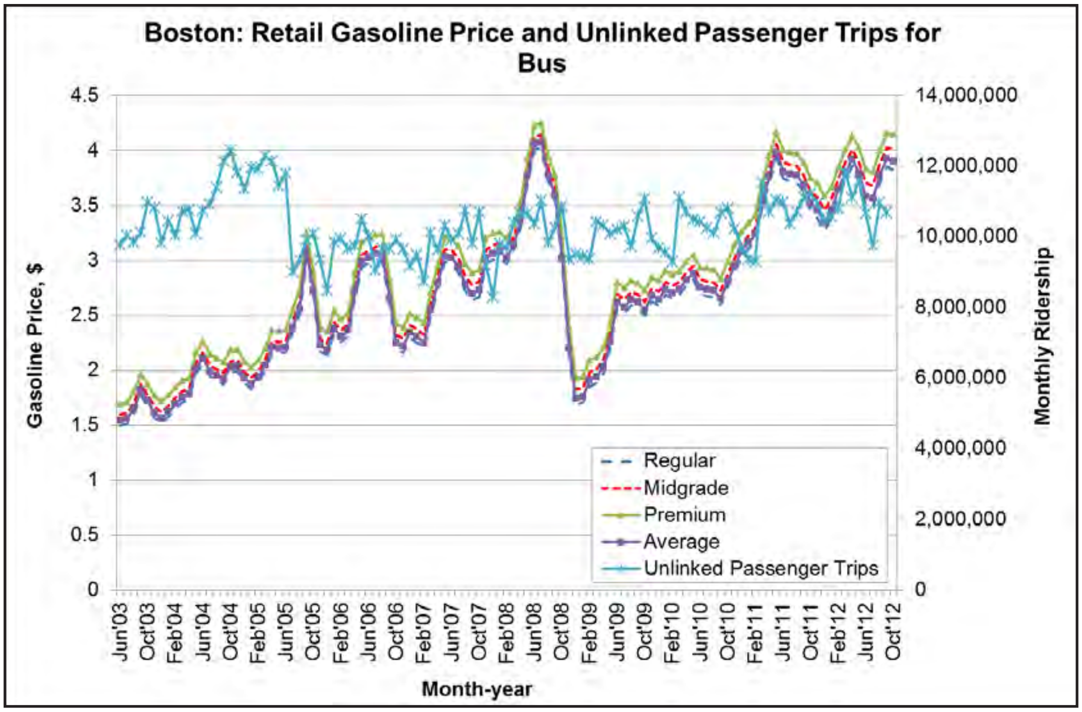


Figure 3. Boston: Retail Gasoline Price and Unlinked Passenger Trips for Bus

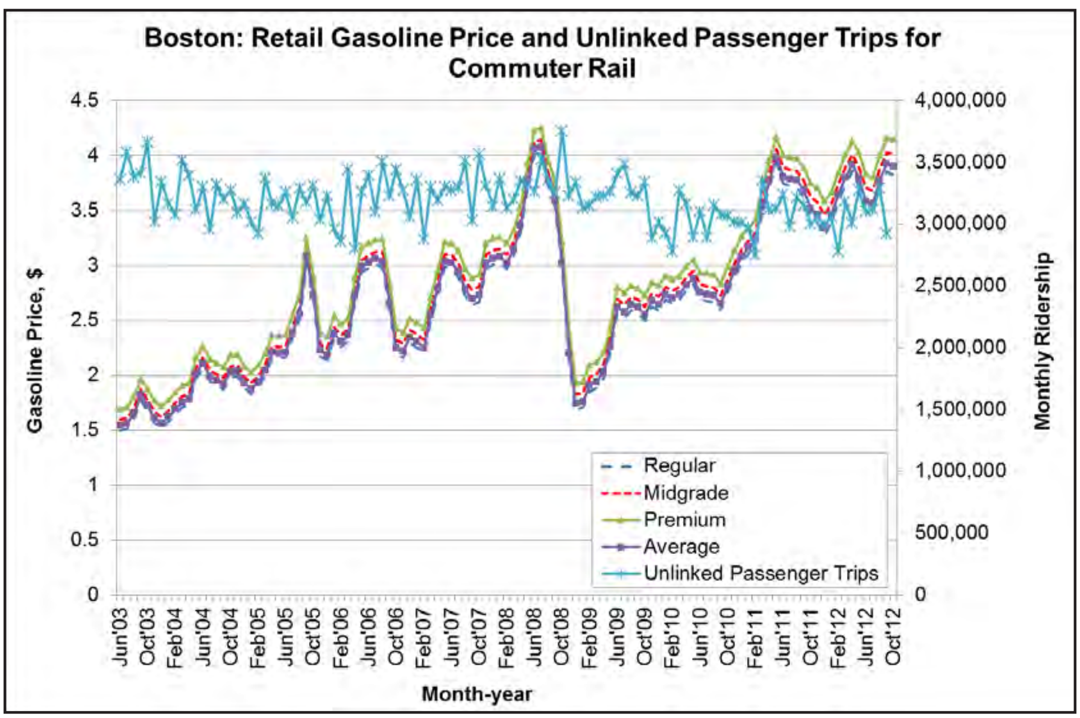


Figure 4. Boston: Retail Gasoline Price and Unlinked Passenger Trips for Commuter Rail

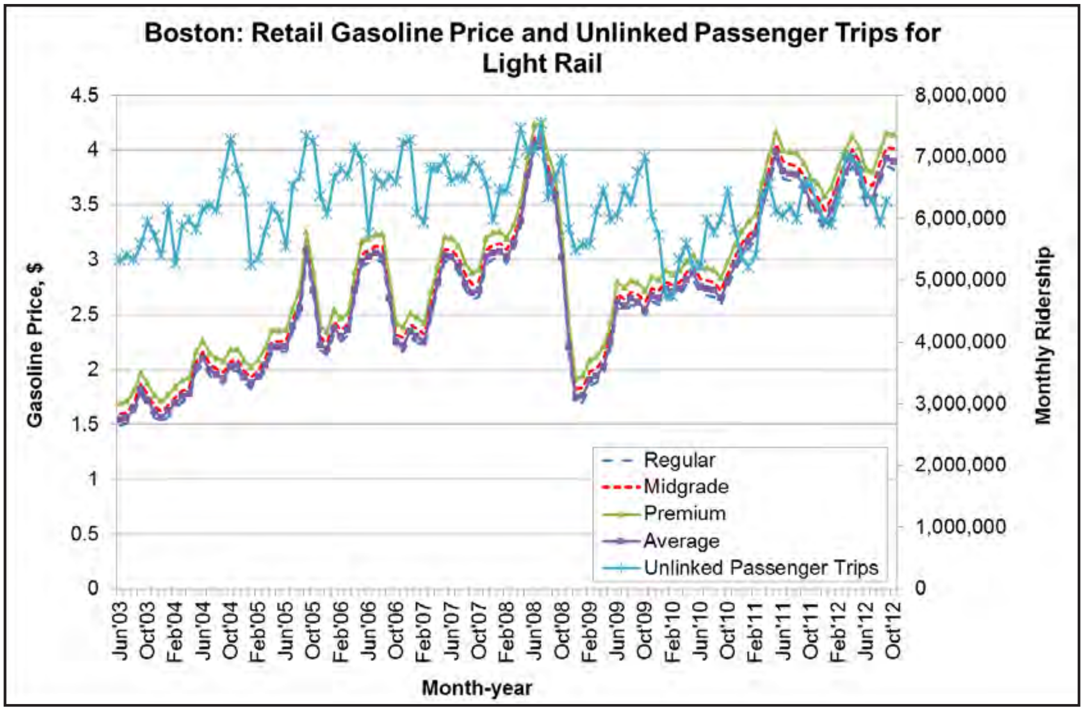


Figure 5. Boston: Retail Gasoline Price and Unlinked Passenger Trips for Light Rail

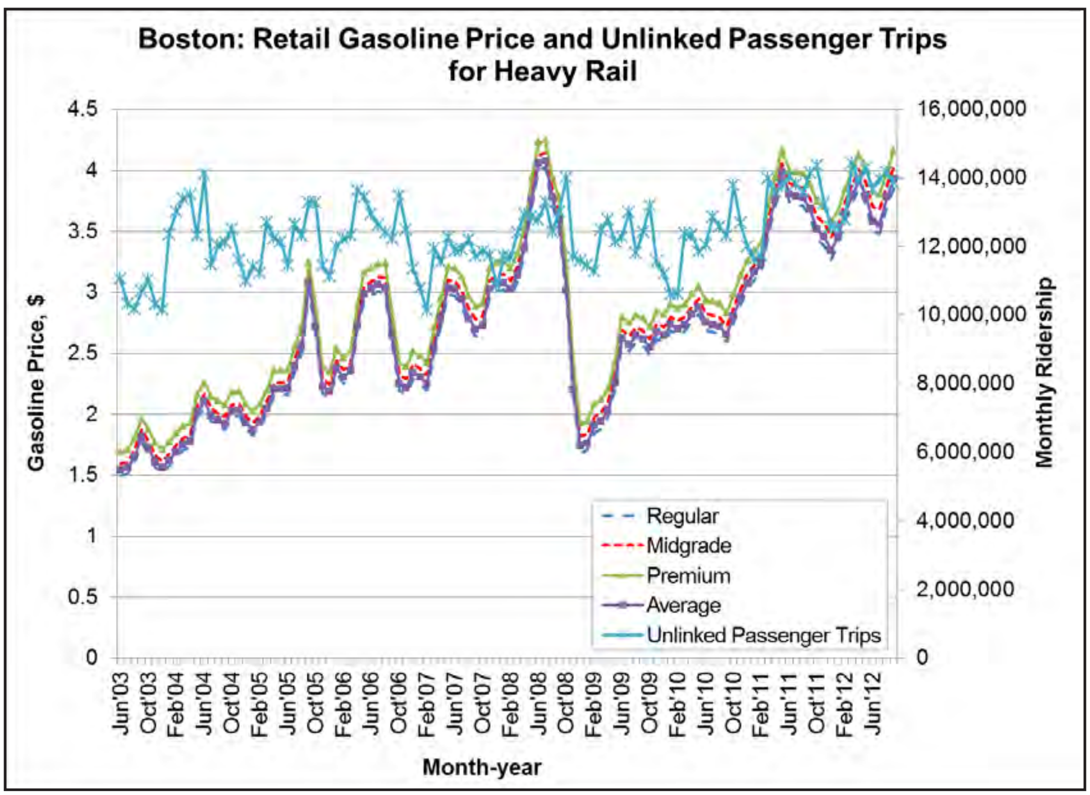


Figure 6. Boston: Retail Gasoline Price and Unlinked Passenger Trips for Heavy Rail

Table 7 (a) through (e) show correlations between independent variables that were considered in regression specifications using data only from 2007.[7] These correlations provide insight into the pairs of variables that could cause collinearity in the regression analysis. In Table 7 (a) through (e), a correlation greater than 0.7 is shown in bold type, while a correlation greater than 0.4 and smaller than 0.7 is shown in italic type. A threshold value of 0.7 is typically used as an indicator for when collinearity possibly begins to cause severe distortion of model estimation and subsequent prediction, while a value of 0.4 is a more restrictive, less commonly used indicator value (Dormann et al 2007).

For bus mode, VRH, frequency of service, total population, and naturalized citizen are highly correlated. In addition, transit fare has high correlation with percent of households with no vehicle. Percent of households with no vehicle and mean household income have a moderate level of correlation with several variables (Table 7). For commuter rail, total population is highly correlated with transit fare and VRH. Frequency of service has high correlation with VRH, as does total population and naturalized citizens. For light rail, the following pairs have high correlation: transit fare and federal highway miles, total population and naturalized citizens, naturalized citizens and mean household income, and unemployment rate and mean household income. Heavy rail shows a more complex pattern in pairs with high correlation. VRH, frequency of service, total population, and naturalized citizen are highly correlated. Frequency of service has high correlation with mean household income and percent of households with no vehicle. Unemployment rate has high correlation with naturalized citizen and mean household income. Lastly, VRH and percent of households with no vehicle also show high correlation. For all modes combined, VRH, frequency of service, total population, and naturalized citizen are highly correlated. Percent of households with no vehicle shows high correlation with VRH and frequency of service, and transit fare shows the same with frequency of service.

Taking into account this multicollinearity, the significance level of estimated coefficients, and the R-squared value, the following process was used to select a set of independent variables to create the most parsimonious model. Starting with a full specification model with all variables, the variable with the estimated coefficient that has the least statistical significance was removed to get a new set of variables for running another model. In some cases, due to multicollinearity among more than two variables, a few specifications were examined before choosing which variable to remove. By selecting a set of independent variables in this way, a more parsimonious model with a high R-squared value and no substantial collinearity was obtained for each mode and each specification.

Table 7. Correlations using the 2007 Data

(a) Bus

Bus	**[1]**	**[2]**	**[3]**	**[4]**	**[5]**	**[6]**	**[7]**	**[8]**	**[9]**	**[10]**
Log of average of regular midgrade and premium gas price [1]	1.000									
Log of fare [2]	0.037	1.000								
Log of vehicle revenue hours [3]	0.231	*0.490*	1.000							
Log of frequency of service [4]	0.264	0.365	**0.896**	1.000						
Log of total population [5]	0.127	0.241	**0.906**	**0.794**	1.000					
Log of naturalized citizen [6]	0.206	0.175	**0.886**	**0.825**	**0.943**	1.000				
Log of federal highway miles [7]	0.074	-0.404	0.171	0.398	0.207	0.254	1.000			
Log of mean household income [8]	0.251	0.265	*0.576*	*0.442*	*0.513*	*0.593*	-0.058	1.000		
Unemployment rate [9]	-0.098	-0.045	-0.176	-0.102	-0.069	-0.246	0.133	-0.587	1.000	
Percent of households with no vehicle [10]	-0.025	**0.750**	*0.665*	*0.431*	*0.588*	*0.517*	-0.191	*0.406*	-0.129	1.000

(b) Commuter Rail

Commuter Rail	**[1]**	**[2]**	**[3]**	**[4]**	**[5]**	**[6]**	**[7]**	**[8]**	**[9]**	**[10]**
Log of average of regular midgrade and premium gas price [1]	1.000									
Log of fare [2]	0.034	1.000								
Log of vehicle revenue hours [3]	-0.119	*0.622*	1.000							
Log of frequency of service [4]	0.004	*0.504*	**0.908**	1.000						
Log of total population [5]	-0.025	**0.764**	**0.795**	*0.673*	1.000					
Log of naturalized citizen [6]	0.076	*0.596*	*0.547*	*0.562*	**0.898**	1.000				
Log of federal highway miles [7]	0.339	0.177	-0.064	0.095	0.362	*0.591*	1.000			
Log of mean household income [8]	0.180	*0.440*	0.362	*0.626*	0.137	0.160	0.112	1.000		
Unemployment rate [9]	-0.077	0.099	*0.631*	*0.462*	*0.540*	0.323	0.165	-0.222	1.000	
Percent of households with no vehicle [10]	-0.165	*0.531*	*0.646*	*0.652*	*0.593*	*0.526*	-0.110	0.317	-0.017	1.000

Table 7, Continued

(c) Light Rail

Light Rail	**[1]**	**[2]**	**[3]**	**[4]**	**[5]**	**[6]**	**[7]**	**[8]**	**[9]**	**[10]**
Log of average of regular midgrade and premium gas price [1]	1.000									
Log of fare [2]	0.176	1.000								
Log of vehicle revenue hours [3]	0.203	0.251	1.000							
Log of frequency of service [4]	0.002	0.092	*0.541*	1.000						
Log of total population [5]	0.130	-0.195	0.395	-0.042	1.000					
Log of naturalized citizen [6]	0.224	-0.128	*0.449*	-0.026	**0.969**	1.000				
Log of federal highway miles [7]	0.076	-0.717	-0.190	-0.415	0.195	0.245	1.000			
Log of mean household income [8]	0.306	0.321	*0.631*	0.287	*0.555*	**0.711**	-0.048	1.000		
Unemployment rate [9]	-0.154	-0.236	-0.251	-0.576	-0.185	-0.313	0.113	-0.709	1.000	

(d) Heavy Rail

Heavy Rail	**[1]**	**[2]**	**[3]**	**[4]**	**[5]**	**[6]**	**[7]**	**[8]**	**[9]**	**[10]**
Log of average of regular midgrade and premium gas price [1]	1.000									
Log of fare [2]	0.249	1.000								
Log of vehicle revenue hours [3]	0.015	0.368	1.000							
Log of frequency of service [4]	0.053	0.263	**0.950**	1.000						
Log of total population [5]	0.112	-0.098	**0.702**	**0.837**	1.000					
Log of naturalized citizen [6]	0.205	0.089	*0.612*	**0.807**	**0.918**	1.000				
Log of federal highway miles [7]	0.327	0.190	-0.119	-0.081	0.156	0.275	1.000			
Log of mean household income [8]	0.245	*0.666*	*0.676*	**0.723**	*0.574*	*0.661*	0.021	1.000		
Unemployment rate [9]	-0.186	-0.325	-0.355	-0.575	-0.501	-0.744	0.073	-0.737	1.000	
Percent of households with no vehicle [10]	-0.111	0.012	**0.770**	**0.773**	*0.599*	*0.502*	-0.164	0.384	-0.286	1.000

Table 7, Continued

(e) All Modes (Aggregate)

All Modes	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
Log of average of regular midgrade and premium gas price [1]	1.000									
Log of fare [2]	0.190	1.000								
Log of vehicle revenue hours [3]	0.190	*0.599*	1.000							
Log of frequency of service [4]	0.167	**0.724**	**0.951**	1.000						
Log of total population [5]	0.127	0.308	**0.904**	**0.822**	1.000					
Log of naturalized citizen [6]	0.206	0.297	**0.857**	**0.784**	**0.943**	1.000				
Log of federal highway miles [7]	0.074	*-0.455*	0.025	0.010	0.207	0.254	1.000			
Log of mean household income [8]	0.251	*0.658*	*0.626*	*0.640*	*0.513*	*0.593*	-0.058	1.000		
Unemployment rate [9]	-0.098	-0.078	-0.135	-0.047	-0.069	-0.246	0.133	*-0.587*	1.000	
Percent of households with no vehicle [10]	-0.025	*0.508*	**0.766**	**0.786**	*0.588*	*0.517*	-0.191	*0.406*	-0.129	1.000

V. ANALYSIS RESULTS

This section presents results from a series of panel data regression analyses. It begins with the baseline specification models. Then, using the instrumental variables models, it examines whether the simultaneity issue between transit service supply and ridership has any significant effect on estimated coefficients for gasoline prices. As the instrumental variables models showed no indication of significant biases in the estimated coefficients of gasoline prices, models that examine short- and long-term effects and non-constant elasticity were specified based on the baseline specifications.

In the following discussion, most of estimated coefficients from regression are interpreted directly as elasticity. When the natural logarithm of values of both an explanatory variable and a dependent variable are taken, an estimated coefficient can be interpreted as elasticity. For example, the estimated coefficient for gasoline prices shows an elasticity of 0.061, which indicates a 0.61 percent increase in ridership in response to an increase in gasoline prices by 10 percent (or 6.1 percent increase in ridership in response to an increase in gasoline prices by 100 percent). There are a few exceptions to this, including independent variables expressed in terms of percentages, such as unemployment rate and the percentage of households with no vehicle, and dummy variables that have either a value of 0 or 1. Interpretation of estimated coefficients for these variables is provided in a different way.

New York City is known to account for approximately 40 percent of the nation's fixed-route transit trips.[8] It also comprises a significantly different transit environment than other regions of the United States. Since a significant difference was observed in the analysis for heavy rail with or without the New York urbanized area (UA), results from both cases are reported and discussed.[9]

BASELINE SPECIFICATION MODEL RESULTS

Table 8 shows the estimated coefficients, standard errors (in parentheses), and significance levels (indicated by asterisks *, **, or ***) obtained from the six different models of the baseline specification: three models for bus, commuter rail, and light rail; two models for heavy rail, with or without the New York UA; and one model for all the four modes combined (the aggregate). Estimated coefficients that are statistically significant at a significance level of 0.10 or higher are shown in bold. Table 8 also shows the number of observations; R-squared; number of urbanized areas; and inclusion of monthly, yearly, and urbanized area dummy variables for each model. The values of R-squared range from 0.431 for the model of heavy rail with the New York UA to 0.844 for the model of light rail.

Table 8. Results from the Baseline Specification Model

Variables	(1) Bus	(2) CR	(3) LR	(4-1) HR w/o NY	(4-2) HR w/ NY	(5) Transit
Log of monthly gasoline price	**0.0611*** **	0.0547	0.0330	-0.0217	-0.0276	**0.0494** **
	(0.0219)	(0.0387)	(0.0507)	(0.0436)	(0.0400)	**(0.0205)**
Log of fare	**-0.230*** **	**-0.372*** **	**-0.124*** **	**-0.284*** **	**-0.210*** **	**-0.340*** **
	(0.0207)	**(0.0616)**	**(0.0348)**	**(0.0324)**	**(0.0280)**	**(0.0240)**
Log of vehicle revenue hours	**0.263*** **	**0.283*** **	**0.786*** **	**0.305*** **	**0.00449**	**0.166*** **
	(0.0258)	**(0.0197)**	**(0.0221)**	**(0.0366)**	**(0.00815)**	**(0.0157)**
Log of frequency of service	**0.115*** **		**0.0998*** **			
	(0.0167)		**(0.0273)**			
Log of total population	**0.996*** **	**4.168*** **		**-1.422*** **	**-0.671***	**1.412*** **
	(0.152)	**(0.427)**		**(0.404)**	**(0.388)**	**(0.141)**
Log of federal highway miles	**0.0702*** **	**-0.133*** **	-0.0471**	**-0.0397***	**0.000188**	**0.0572*** **
	(0.0100)	**(0.0192)**	**(0.0210)**	**(0.0212)**	**(0.0204)**	**(0.00922)**
Log of mean household income		**2.896*** **				
		(0.390)				
Unemployment rate (%)	**0.0341*** **		**0.0442*** **	**0.0454*** **	**0.0304*** **	**0.0385*** **
	(0.00425)		**(0.0104)**	**(0.00978)**	**(0.00882)**	**(0.00396)**
Households with no vehicle (%)	**-0.0304*** **		**-0.0541** **			
	(0.0103)		**(0.0262)**			
Constant	**-4.247***	**-85.34*** **	**5.955*** **	**34.16*** **	**25.97*** **	**-8.310*** **
	(2.339)	**(6.105)**	**(0.314)**	**(6.225)**	**(6.110)**	**(2.189)**
Seasonal effects (month dummies)	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Urbanized area fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,126	789	840	669	789	1,132
R-squared	0.518	0.630	0.844	0.485	0.431	0.53
Number of urbanized areas	10	7	9	6	7	10

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Key to Variables: Bus = Unlinked passenger trips: bus; CR = Unlinked passenger trips: commuter rail; LR = Unlinked passenger trips: light rail; HR = Unlinked passenger trips: heavy rail; Transit = Unlinked passenger trips: All modes combined. The statistically significant coefficients are in bold.

The estimated elasticity of bus ridership to gasoline prices is 0.061 and is statistically significant at the 0.01 significance level. This indicates that a 10 (or 100) percent increase in the price of gasoline leads to 0.61 (or 6.1) percent increase in bus ridership. Checking the elasticity by mode, the aggregate ridership elasticity is 0.049 and is statistically significant at the 0.05 significance level. While the estimated elasticity is positive, as expected, these values are slightly lower than those found in previous studies that ranged from 0.10 to 0.30 (Litman 2004; Haire and Machemehl 2007; Currie and Phung 2008; Chen, Varley, and Chen, 2010). However, the elasticity for three other modes is not statistically significant.

These results indicate that an increase in gasoline prices has a positive impact on bus ridership and the aggregate ridership; people are more likely to take public transit—especially buses—in order to cope with the increased expense of traveling by private automobile due to higher gasoline prices. The results also indicate that bus riders are more sensitive to a change in gasoline prices. In a usual context, this could be explained by a relatively lower income level of bus riders, which leads to higher price sensitivity. However, the analysis in this study that uses aggregate data for UAs does not detect this effect of individual income levels on price sensitivity well. It is unlikely that commuter rail, light rail and heavy rail show a clear change in ridership due to the change in gasoline prices. However, an analysis of elasticity in different ranges of gasoline prices shows slightly different results, as discussed later in this chapter.

The three internal factors—fare, vehicle revenue miles (VRH), and frequency of service—generally have expected effects on transit ridership. The estimated elasticity of ridership to fare is negative for all modes, as well as the aggregate, as expected, ranging from -0.372 for commuter rail to -0.124 for light rail. A negative sign indicates that an increase in fare leads to a decrease in ridership (or a decrease in fare leading to an increase in ridership). While elasticity for bus and heavy rail lines is in the middle range, the elasticity for the aggregate shows a relatively large absolute value of elasticity, -0.340. The estimated elasticity values, except for light rail, are consistent with those reported in past studies ranging from -0.20 to -0.90 for the short-term effects (Litman 2004; Chen, Varley, and Chen 2010).

The higher elasticity of commuter rail may be because of a combination of a distance-based fare and relatively longer travel distances for this mode, which could result in a fare increase for each traveler, compared to a fare increase in other modes. For heavy rail, elasticity is higher for the model without the New York UA than with the New York UA, implying that the effect of heavy rail fare on ridership is relatively lower in the New York UA, taking into account the level of transit ridership. A large number of people in New York ride transit because they may not have access to private automobiles and/or because transit service is less costly and more convenient than driving, regardless of the level of gasoline prices. In addition, the estimated elasticity values from this baseline specification model show that ridership for all modes is more sensitive to transit fares than gasoline price. This might be because people have much more resistance against and more sensitivity to a change in transit fare, which is almost always an increase and more lumpy than a change in gasoline prices.

The positive signs of estimated elasticity for VRH indicate that an increase in the supply of transit service leads to an increase in ridership for all modes except heavy rail when including the New York UA. The value of elasticity for VRH ranges from 0.263 for bus to 0.786 for light rail, showing relatively greater effects of VRH on ridership than gasoline prices as expected. These estimated elasticity values are in line with those found in the literature, ranging from 0.3 to 1.14 (Litman 2004; Taylor et al. 2009, Chen, Varley, and Chen 2010). It is surprising to find that estimated elasticity was not statistically significant for heavy rail when including the New York UA. It may be because the New York transit system is accommodating people's trips well and an increase in VRH may not necessarily increase ridership. This increase may, however, reduce the level of crowding, increase service hours, and lead to improvements in the overall service quality from transit users' perspective.

The higher the service frequency, the higher transit ridership for bus and light rail is. This positive relationship between service frequency and ridership is expected. However, estimated elasticity values for the other modes were found statistically insignificant. A detailed examination of the data reveals little variance in the service frequency variable especially for commuter rail, which may explain the statistical insignificance.

The rest of the variables that remained in the baseline specification models are external factors that transit agencies do not have any control over. As other studies indicated, some of the estimated elasticity values of these variables are statistically significant and have a very substantial effect on ridership for some modes. Total population in an urbanized area has a positive effect on ridership for bus, commuter rail, and the aggregate, as expected. The larger the total population in the UA, the higher the transit ridership. Estimated elasticity is 0.996 for bus, 4.168 for commuter rail, and 1.412 for the aggregate. The effect of a population increase on commuter rail ridership is very significant in terms of the magnitude of effect as well as statistical significance. However, elasticity of total population for light rail was statistically insignificant, and those for heavy rail have negative signs (-1.422 and -0.671). Higher population is also related to a lower level of heavy rail ridership. These mixed results of the effect of population on transit ridership by mode are difficult to interpret. This may be caused by the population variable in the model that includes all of the population in the entire urbanized area, not only in the areas that are well served by service of each mode.

Federal highway miles in the UA has a positive elasticity value of 0.070 for bus ridership, but has a negative effect on ridership for commuter rail, light rail, and heavy rail (-0.133, -0.047, and -0.039 respectively). Overall, the effect of highway miles increase on bus ridership seems to dominate, as federal highway miles have a positive effect on aggregate transit ridership as indicated by the positive elasticity of 0.057.

The positive effect on bus ridership may be explained by bus service provided through highways that can be an effective way to increase ridership. In addition, the higher highway miles might mean less overall congestion, faster bus service, and shorter bus travel time—controlling for the amount and frequency of bus service, resulting in higher ridership. However, more travel speed on the road network works against rail services, resulting in less competitiveness for rail transit modes against driving modes. In addition, as the higher presence of highway miles implies a higher likelihood of traveling by automobile, it is fairly understandable to have a negative effect on rail ridership. The higher elasticity value for commuter rail may be explained by long-distance commuting trips, for which commuter rail could be more easily substituted by driving modes, especially when commuter rail lines run parallel to highways. Examining the values of elasticity for the two heavy rail models, the estimated elasticity for federal highway miles was found statistically insignificant in the model with the New York UA. This indicates that in the New York UA, federal highway miles doesn't affect heavy rail ridership.

Partly because of multicollinearity, mean household income only got into the model for commuter rail and has a positive elasticity of 2.896, indicating that the higher the mean household income in the UA, the higher the commuter rail ridership. This makes sense as commuter rail typically provides service to commuters in higher income groups from relatively affluent suburbs who travel to downtown and charges fares more in proportion with travel distance than other transit modes.

Unemployment rate was used in the most models as a statistically significant variable with a consistently positive effect on ridership, ranging from 0.030 for heavy rail with New York UA to 0.045 for heavy rail without New York UA. The only model that did not use unemployment rate was for commuter rail. In general, when the unemployment rate is high, it increases transit ridership. This positive correlation between unemployment rate and transit ridership is stronger than the negative effect that this variable has on the number of activities and trips as a regional economic indicator.

Note that the estimated coefficients for this unemployment rate and the following percent of households with no car are interpreted in a way that is different from the other variables, since these variables are measured as percentages. For example, the coefficient of the employment rate for bus, 0.034, indicates that an increase in the *percentage-wise* unemployment rate by 10 percent (e.g., 8 percent to 8.8 percent) results in an increase in bus ridership by 0.34 percent.

Lastly, the estimated coefficient for percent of households with no car was included and found statistically significant for only two models—for bus (-0.030) and light rail (-0.054). The reason that this variable did not appear in the models of the other modes is due to multicollinearity. In spite of testing various specifications, the negative sign for the estimated coefficients was found and is counterintuitive; the higher the percentage of households with no car is, the lower transit ridership is for bus and light rail, although one would expect people of no-car households to take public transportation more than their counterparts and increase transit ridership.

While the baseline specification models provide the results described and discussed above, estimated coefficients for explanatory variables—especially VRH—could be biased by the endogeneity problem, potentially arising from the simultaneity issue between transit service supply and ridership. The next section describes and discusses results obtained using instrumental variables model that can address this potential problem.

INSTRUMENTAL VARIABLES MODEL RESULTS

One of the requirements for valid instruments is that explanatory variables are statistically significant in predicting a dependent variable in the first stage of the IV model. The dependent variable from the first stage is used as an instrumental variable in the second stage. Table 9 shows the estimated coefficients of the variables—total number of employees, total fleet, and total fund available to transit agencies all in logarithmic form—in the first stage of the IV model that regresses log of vehicle revenue hours (VRH) on all of these three variables and all the variables included in the baseline specification for each mode.

Table 9. Results from the First Stage of the Instrumental Variables Model

	(1)	(2)	(3)	(4-1)	(4-2)	(5)
Variables	**Bus**	**CR**	**LR**	**HR w/o NY**	**HR w/ NY**	**Transit**
Log of total number of employees	**-0.0776*** **	-0.0354	0.0639	**-0.541*** **	-0.379	-0.00508
(full-time+part-time/2)	**(0.0191)**	(0.194)	(0.0451)	**(0.107)**	(0.427)	(0.0317)
Log of total fleet	**0.552*** **	**-0.0565*** **	**0.443*** **	**3.449*** **	**6.587*** **	**0.499*** **
(seating+standing capacity)	**(0.0296)**	**(0.0154)**	**(0.0159)**	**(0.403)**	**(1.555)**	**(0.0530)**
Log of total fund available to transit agencies	**0.0547*** **	-0.0286	**0.0939*** **	**0.146*** **	**0.700*** **	**0.0352** **
	(0.00988)	(0.0538)	**(0.0243)**	**(0.0304)**	**(0.121)**	**(0.0162)**
Seasonal effects (month dummies)	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Urbanized area fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,126	777	840	669	789	1,132
R-squared	0.529	0.581	0.829	0.434	0.347	0.300
Number of Urbanized Areas	10	7	9	6	7	10
F-stat	123.56	4.93	278.67	47.82	20.27	31.34
P-value	0.000	0.0022	0.000	0.000	0.000	0.000

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

In order to test the hypothesis that the coefficients of these three variables are jointly zero, whether or not F-statistics for the hypothesis are greater than the critical value was examined (F-test). The values of the F-statistics and associated P-values in Table 9 show that the set of three excluded instruments significantly affect VRH for all modes. Estimated coefficients show that the total funds availability variable significantly increases VRH of bus, light rail, heavy rail, and the aggregate transit. The number of employees affects VRH negatively with statistical significance only for bus and heavy rail without the New York UA. Seating and standing capacity affect VRH of bus, light rail, heavy rail, and aggregate transit positively, while it has a negative effect for commuter rail.

Table 10 shows the results from the second stage of the IV model, in which the estimated coefficient of VRH is expected to be unbiased. Most of the estimated coefficients in Table 10 are similar to those obtained from the baseline specification models.

Table 10. Results from the Second Stage of the Instrumental Variables Model

Variables	(1) Bus	(2) CR	(3) LR	(4-1) HR w/o NY	(4-2) HR w/ NY	(5) Transit
Log of monthly gasoline price	**0.0617*** **	0.0640	0.0253	-0.0393	-0.0303	**0.0573** **
	(0.0223)	(0.0428)	(0.0516)	(0.0530)	(0.0454)	**(0.0262)**
Log of fare	**-0.220*** **	**-0.444*** **	**-0.141*** **	**-0.261*** **	**-0.193*** **	**-0.319*** **
	(0.0212)	**(0.0940)**	**(0.0355)**	**(0.0418)**	**(0.0319)**	**(0.0308)**
Log of vehicle revenue hours	**0.407*** **	0.0625	**0.901*** **	-0.0185	0.0247	**0.577*** **
	(0.0519)	(0.152)	**(0.0314)**	(0.100)	(0.0201)	**(0.0712)**
Log of frequency of service	**0.0772*** **		0.0130			
	(0.0207)		(0.0324)			
Log of total population	**0.811*** **	**6.730*** **		-0.475	-0.542	**0.845*** **
	(0.164)	**(1.735)**		(0.550)	(0.422)	**(0.202)**
Log of federal highway miles	**0.0670*** **	**-0.116*** **	**-0.0675*** **	0.00797	-0.00115	0.0202
	(0.0102)	**(0.0240)**	**(0.0217)**	(0.0271)	(0.0219)	(0.0133)
Log of mean household income		1.441				
		(1.082)				
Unemployment rate (%)	**0.0331*** **		**0.0343*** **	**0.0253** **	**0.0305*** **	**0.0280*** **
	(0.00433)		**(0.0108)**	**(0.0121)**	**(0.00941)**	**(0.00533)**
Households with no vehicle (%)	**-0.0396*** **		**-0.0931*** **			
	(0.0109)		**(0.0277)**			
Constant	-2.874	**-107.5*** **	**6.107*** **	**22.0*** **	**23.6*** **	-4.301
	(2.411)	**(15.37)**	**(0.321)**	**(8.280)**	**(6.76)**	(2.867)
Seasonal effects (month dummies)	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Urbanized area fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,126	777	840	549	669	1,132
R-squared	0.504	0.662	0.839	0.423	0.487	0.237
Number of urbanized areas	10	7	9	5	6	10

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

The values of R-squared range from 0.237 for the model of the aggregate transit to 0.839 for the model of light rail. The aggregate transit model shows a much lower R-squared value, part of which is explained by the statistical insignificance of federal highway miles and the constant that are statistically significant in the baseline model. The models for the other modes have a comparable R-squared value between the baseline and IV models.

In this IV model, the estimated elasticity for gasoline prices is almost identical to the value previously estimated in the baseline specification models—0.061 for bus and 0.057 for the aggregate—while the one for commuter rail, light rail and heavy rail is found statistically insignificant. The values of fare elasticity in Table 10 are also very similar, both in terms of statistical significance and magnitude, to those found in the baseline specification models. However, the estimated elasticity values of VRH in the IV models (Table 10) are quite different from the baseline specification models in Table 8. The elasticity of VRH of commuter rail and heavy rail both with and without the New York UA is not statistically significant in the parsimonious IV model, while it is significant for all modes except for heavy rail including the New York UA in the baseline model.

As expected, the elasticity values of VRH for the three modes—bus, light rail, and the aggregate transit—are statistically significant in both the baseline and IV models but are quite different; the elasticity is generally higher in the IV model (0.407, 0.901, and 0.577) (Table 10) than in the baseline model (0.263, 0.786, and 0.166) (Table 8). These results indicate that the baseline model underestimates the effect of service supply of these modes by not accounting for the endogeneity of VRH. The estimated value of service supply elasticity was statistically insignificant for commuter rail and heavy rail in the IV models.

The estimated coefficients for service frequency show mixed effects on ridership depending on modes. Frequency of service has a positive and statistically significant impact on bus ridership; the estimated elasticity is slightly smaller in the IV model (0.077) than in the parsimonious baseline model (0.115). In contrast, statistically insignificant elasticity of service frequency was found for light rail in the IV model, while it was significant previously.

The effect of population on bus, commuter rail, and aggregate transit ridership remains statistically significant, but not for heavy rail. The estimated elasticity of population is lower for bus and the aggregate, but higher for commuter rail in the IV model than in the baseline model, implying the possibility that the elasticity of population in the baseline model could be biased.[10]

The elasticity for federal highway miles in the UA estimated in the IV model is lower for bus but higher for commuter rail and light rail than in the baseline model. However, the estimated elasticity was found statistically insignificant for both heavy rail cases with or without the New York UA. Although the mean income of households was included for commuter rail, following the result in the baseline model, its estimated coefficient was found statistically insignificant in the IV model.

For all modes but commuter rail, unemployment rate gives positive, statistically significant estimates of elasticity, which are consistent with results obtained in the baseline model. The worse the regional economy, the more people take public transit. This effect seems more substantial than the effect to reduce the overall trips, including transit trips, resulting from a bad economy. The percent of households with no vehicle shows a statistically significant negative coefficient only for bus (-0.039) and light rail (-0.093) in the IV model.

The results from the baseline specification models and from the IV models are very similar; if a variable has a positive effect on ridership in parsimonious baseline, it also has a positive effect in the parsimonious IV model (or vice versa). For this reason, the baseline specifications are used in the analysis of short- and long-term (lagged) effects of changes in gasoline prices and non-constant elasticity of gasoline price in the following sections.

SHORT- AND LONG-TERM (LAGGED) EFFECTS OF GASOLINE PRICES

In the analysis of potential short- and long-term (lagged) effects of gasoline prices on transit ridership, the results from the most parsimonious models are presented in Table 11, while many different combinations of lagged gasoline price variables, as well as control variables, were tested.

The values of R-squared range from 0.47 for the model of heavy rail with the New York UA, to 0.855 for the model of light rail. Compared to the results of the baseline specification models in Table 8, these R-squared values are comparable or even higher for the models of commuter rail, light rail, and the aggregate transit, and indicate the larger portion of variance is explained by the set of independent variables included in the models with the lagged gasoline prices, although the number of observations decreases due to the use of lagged variables. For example, the R-squared is 0.855 in the light rail model with the 13-month lagged gasoline price, compared to 0.844 in the baseline specification model, while the number of observations is 746, compared to 840.

Table 11. Results from the Model Estimating Short- and Long-term (Lagged) Effects

Variables	(1) Bus	(2) CR	(3) LR	(4-1) HR w/o NY	(4-2) HR w/ NY	(5) Transit
Log of current monthly gasoline price (GP)						0.0569***
						(0.0206)
Log of the 1-month lag GP (t-1)	0.0839***					
	(0.0267)					
Log of the 2-month lag GP (t-2)		0.106*				
		(0.0547)				
Log of the 13-month lag GP (t-13)			0.116*	0.0943*	0.0899*	0.0486*
			(0.0594)	(0.0543)	(0.0497)	(0.0250)
Log of fare	-0.230***	-0.344***	-0.0874**	-0.291***	-0.238***	-0.328***
	(0.0207)	(0.0621)	(0.0374)	(0.0345)	(0.0303)	(0.0243)
Log of vehicle revenue hours	0.261***	0.272***	0.747***	0.358***	0.173***	0.418***
	(0.0259)	(0.0198)	(0.0219)	(0.0387)	(0.0304)	(0.0275)
Log of frequency of service	0.118***		0.144***			
	(0.0172)		(0.0278)			
Log of total population	1.009***	4.084***			-0.728*	0.809***
	(0.152)	(0.424)			(0.430)	(0.151)
Log of federal highway miles	0.0699***	-0.131***		-0.0393*		0.0371***
	(0.0100)	(0.0189)		(0.0216)		(0.00904)
Log of mean household income		2.770***				
		(0.386)				
Unemployment rate (%)	0.0339***		0.0453***	0.0416***	0.0351***	0.0290***
	(0.00426)		(0.0105)	(0.0100)	(0.00879)	(0.00387)
Households with no vehicle (%)	-0.0298***					
	(0.0103)					
Constant	-4.437*	-82.57***	4.987***	11.21***	24.84***	-2.066
	(2.344)	(6.042)	(0.213)	(0.468)	(6.739)	(2.279)

Table 11, Continued

Variables	(1) Bus	(2) CR	(3) LR	(4-1) HR w/o NY	(4-2) HR w/ NY	(5) Transit
Seasonal effects (month dummies)	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Urbanized area fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,122	775	746	591	698	1,002
R-squared	0.517	0.705	0.855	0.470	0.417	0.582
Number of urbanized areas	10	7	8	6	7	10

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Different variables of gasoline prices were found statistically significant among the different modes (Table 11). The models for bus, commuter rail, and aggregate transit show the exactly same set of the control variables as the baseline specification model and estimated coefficients for the control variables were very similar (Table 8). However, the bus model shows that the 1-month lag variable of gasoline price replaced the current price variable. As lagged variables have high correlation to each other—for example, correlation of 0.96 between the 1-month lag variable of gasoline price and the current price variable in the entire sample set—two variables are not included in one model (Table 12). The higher elasticity of the 1-month lag variable indicates a time lag in changing travel modes selected by bus riders.

Table 12. Correlations between Current and Lagged Gasoline Price Variables

	Current	Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6	Lag 7	Lag 8	Lag 9	Lag 10	Lag 11	Lag 12	Lag 13
Current	1.000	-	-	-	-	-	-	-	-	-	-	-	-	-
Lag 1	0.959	1.000	-	-	-	-	-	-	-	-	-	-	-	-
Lag 2	0.880	0.959	1.000	-	-	-	-	-	-	-	-	-	-	-
Lag 3	0.804	0.882	0.959	1.000	-	-	-	-	-	-	-	-	-	-
Lag 4	0.742	0.807	0.883	0.960	1.000	-	-	-	-	-	-	-	-	-
Lag 5	0.689	0.745	0.809	0.885	0.961	1.000	-	-	-	-	-	-	-	-
Lag 6	0.654	0.692	0.747	0.811	0.887	0.962	1.000	-	-	-	-	-	-	-
Lag 7	0.639	0.656	0.694	0.749	0.815	0.889	0.962	1.000	-	-	-	-	-	-
Lag 8	0.642	0.642	0.658	0.697	0.754	0.818	0.891	0.963	1.000	-	-	-	-	-
Lag 9	0.658	0.645	0.645	0.663	0.702	0.758	0.820	0.892	0.963	1.000	-	-	-	-
Lag 10	0.679	0.662	0.649	0.650	0.668	0.707	0.761	0.823	0.894	0.964	1.000	-	-	-
Lag 11	0.693	0.685	0.668	0.656	0.657	0.675	0.712	0.766	0.826	0.896	0.964	1.000	-	-
Lag 12	0.690	0.698	0.691	0.675	0.664	0.665	0.681	0.718	0.770	0.829	0.896	0.965	1.000	-
Lag 13	0.680	0.695	0.704	0.697	0.682	0.671	0.671	0.687	0.723	0.774	0.831	0.899	0.966	1.000

Note: The entire sample data set is used.

The light rail model shows that the 2-month lag variable of gasoline prices is statistically significant at the level of 0.1. While it is weaker evidence, the estimated elasticity value of 0.106 indicates that a change in travel mode among current and potential commuter riders takes about two months after gasoline prices change. The longer time lag in a travel mode change of commuter rail riders, compared to bus riders, may be explained by certain characteristics of trips taken in these two modes. Commuter rail trips are usually taken for a long-distance commute from suburbs to a central business district. Although the financial impact of gasoline price increase can be more substantial for such long-distance trips, people may need more time to find an alternative mode of travel or to make arrangements for a private car or carpooling.

The aggregated transit model shows the current gasoline price variable is statistically significant at the level of 0.01, which is consistent with the baseline specification model, and that the 13-month lag variable is statistically significant at a level of 0.1. The estimated coefficient of the current gasoline price is 0.057 and is close to the value of 0.049 found in the baseline specification model. While the correlation between the 13-month lag variable of gasoline price and the current price variable is 0.680 (Table 12), these two variables were included in this model. These two elasticity values indicate that the aggregate transit ridership increases shortly after an increase in gasoline prices and then 13 months later. It is also possible that the 13-month lag effect could be contributed by residual that is not picked up by the month dummy variables (Chen, Varley, and Chen, 2010). Combining the two estimated coefficients of gasoline price variables, the total long-term elasticity is 0.105 for the aggregate transit model.

While the set of control variables is different for the models with lag variables and the baseline specification models for light rail and heavy rail, estimated coefficients for these control variables are not far off from the ones in the baseline specification models. These light rail and heavy rail models also found the 13-month lag variable has an estimated elasticity of 0.116, 0.094, and 0.089, for light rail, heavy rail without the New York UA, and heavy rail including the New York UA, respectively. As these elasticity estimates are statistically significant at the level of 0.1, they are not as strong as other estimates significant at the levels of 0.05 or 0.01, however, they suggest, for example, that a 10 percent increase in gasoline ridership will increase ridership by 1.16 percent, 0.94 percent, and 0.89 percent for light rail, heavy rail without the New York UA, and heavy rail with the New York UA, respectively.

As expected, elasticity is positive regardless of the length of time for which the effect of gasoline prices on transit ridership is measured. The estimated elasticity for aggregate transit, 0.057 in the short-term and 0.105 in the long-term, are relatively smaller than the estimates found in previous studies that range from 0.10 to 0.30 (Litman 2004; Haire and Machemehl 2007; Currie and Phung 2008; Chen, Varley, and Chen, 2010).

ANALYSIS OF NON-CONSTANT ELASTICITIES

Table 13 shows results from the models that examine non-constant elasticity, that is, threshold boost effects of particular gasoline price values ($2, $3, and $4) and different elasticity values for different ranges of gasoline prices. As discussed in Section 3, various specifications were examined using dummy variables to indicate gasoline prices higher than $2, $3, and $4, and their respective interaction terms with the gasoline price variable. However, only the variables shown in Table 13 were found statistically significant and remained in the final specifications; The figures for “Dummy for gas price > $3 (D$3)” represent effects of gasoline prices that exceed $3, and those for “[Log of GP – Log of $3]*D$3” and “[Log of GP – Log of $4]*D$4” indicate interactive terms that examine a different elasticity value for gasoline prices equal to $3 or higher and to $4 or higher, respectively. The values of R-squared range from 0.440 for the model of heavy rail with the New York UA to 0.844 for the model of light rail. These R-squared values are very similar to the values obtained in the baseline model, while the commuter rail model now has a higher value of 0.718, compared to 0.630 in the baseline specification model.

Table 13. Results from the Non-constant Elasticity Model

Variables	(1) Bus	(2) CR	(3) LR	(4-1) HR w/o NY	(4-2) HR w/ NY	(5) Transit
Log of monthly gasoline price	0.0282	**-0.0950***	0.0166	**-0.170*****	**-0.146*****	0.00970
(Log of GP)	(0.0252)	**(0.0510)**	(0.0515)	**(0.0561)**	**(0.0519)**	(0.0235)
Dummy for gas price>$3 (D$3)		**0.0415*****		**0.0539*****	**0.0439*****	
		(0.0147)		**(0.0171)**	**(0.0158)**	
[Log of GP-Log of $3]*D$3	**0.146*****	**0.308*****		**0.239****	**0.189***	**0.177*****
	(0.0552)	**(0.0974)**		**(0.109)**	**(0.101)**	**(0.0512)**
[Log of GP-Log of $4]*D$4			**0.920***			
			(0.532)			
Log of fare	**-0.228*****	**-0.353*****	**-0.124*****	**-0.283*****	**-0.217*****	**-0.340*****
	(0.0206)	**(0.0610)**	**(0.0347)**	**(0.0321)**	**(0.0277)**	**(0.0239)**
Log of vehicle revenue hours	**0.263*****	**0.294*****	**0.788*****	**0.299*****		**0.165*****
	(0.0258)	**(0.0196)**	**-0.0221**	**(0.0353)**		**(0.0157)**
Log of frequency of service	**0.112*****		**0.100*****			
	(0.0167)		**(0.0273)**			
Log of total population	**1.021*****	**4.013*****		**-1.610*****	**-0.756****	**1.446*****
	(0.151)	**(0.424)**		**(0.402)**	**(0.383)**	**(0.140)**
Log of federal highway miles	**0.0718*****	**-0.119*****	**-0.0464****			**0.0588*****
	(0.0100)	**(0.0192)**	**(0.0210)**			**(0.00919)**
Log of mean household income		**2.900*****				
		(0.385)				
Unemployment rate (%)	**0.0337*****		**0.0462*****	**0.0514*****	**0.0300*****	**0.0378*****
	(0.00425)		**(0.0104)**	**(0.00948)**	**(0.00842)**	**(0.00394)**
Households with no vehicle (%)	**-0.0302*****		**-0.0537****			
	(0.0103)		**(0.0262)**			
Constant	**-4.617****	**-82.78*****	**5.943*****	**36.69*****	**27.41*****	**-8.835*****
	(2.337)	**(6.052)**	**(0.314)**	**(6.202)**	**(6.064)**	**(2.184)**

Table 13, Continued

Variables	(1) Bus	(2) CR	(3) LR	(4-1) HR w/o NY	(4-2) HR w/ NY	(5) Transit
Seasonal effects (month dummies)	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Urbanized area fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,126	789	840	669	789	1,132
R-squared	0.521	0.718	0.844	0.495	0.440	0.535
Number of urbanized area	10	7	9	6	7	10

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 13 indicates, estimated coefficients for monthly gasoline prices are statistically significant for commuter rail and heavy rail with or without including the New York UA, but not for bus, light rail, and the aggregate models. The same two rail modes show statistical significance for estimated coefficients for the dummy variable for gasoline prices exceeding $3 with a positive sign. This variable was excluded from the models for bus, light rail, and the aggregate, due to statistical insignificance that indicates no boost effect for these three modes.[11] Estimated coefficients for gasoline prices exceeding $3 ([Log of GP – Log of $3]*D$3) are statistically significant with a positive sign for all cases except light rail. Finally, an estimated coefficient for gasoline prices exceeding $4 ([Log of GP – Log of $4]*D$4) is statistically significant with a positive sign for light rail. The substantial increase in the R-squared value obtained for the non-constant elasticity model for light rail represents an increase in the variance of ridership explained by the four variables related to gasoline prices, compared to the earlier result in the baseline specification model that shows statistical insignificance for the gasoline price variable. The sign and magnitude of other control variables are similar to what was found in the baseline specification models for all the modes, except federal highway miles, which was excluded in the model of heavy rail without the New York UA, as it became statistically insignificant.

Because of the inclusion of dummy variables and interactive terms of gasoline prices, the interpretation of estimated coefficients is more complex than it is for the baseline specification models. Table 14 summarizes non-constant effects of gasoline prices on transit ridership for three different levels of gasoline price increase—5, 10, and 20 percent. Estimated coefficients for the gasoline price variable (Log of GP) is included to measure total effects regardless of statistical significance, as it is conventionally done. In addition, estimated coefficients for threshold effects of gasoline prices of $3 (D$3), as well as different elasticity values in different price ranges of gasoline, are incorporated in the calculation of a percent increase in ridership, as in the case when a gasoline price increases from either $2.90 or $3.90. For example, "crossing the $3 mark from $2.90" under a 10 percent gasoline price increase means an increase from $2.90 to $3.19. The magnitude of effect varies depending on the gasoline price before and after a change.[12]

Table 14. Non-constant Effects of Gasoline Prices

Variables	(1) Bus	(2) CR	(3) LR	(4-1) HR w/o NY	(4-2) HR w/ NY	(5) Transit
	% Increase in Ridership					
5% increase in Gasoline Prices						
Within the GP range below $3	0.14%	-0.46%	0.08%	-0.83%	-0.71%	0.05%
Crossing the $3 mark (from $2.90)	0.36%	4.23%		5.04%	4.04%	0.31%
Within the GP range $3-$4	0.85%	1.04%		0.34%	0.21%	0.92%
Crossing the $4 mark (from $3.90)			2.27%			
Within the GP range over $4			4.68%			
10% increase in Gasoline Prices						
Within the GP range below $3	0.27%	-0.90%	0.16%	-1.61%	-1.38%	0.09%
Crossing the $3 mark (from $2.90)	1.17%	5.27%		6.15%	4.85%	1.19%

Table 14, Continued

Variables	(1) Bus	(2) CR	(3) LR	(4-1) HR w/o NY	(4-2) HR w/ NY	(5) Transit
	% Increase in Ridership					
Within the GP range $3-$4	1.67%	2.05%		0.66%	0.41%	1.80%
Crossing the $4 mark (from $3.90)			6.82%			
Within the GP range over $4			9.34%			
20% increase in Gasoline Prices						
Within the GP range below $3	0.52%	-1.72%	0.30%	-3.05%	-2.63%	0.18%
Crossing the $3 mark (from $2.90)	2.72%	8.25%		6.79%	5.24%	2.84%
Within the GP range $3-$4	3.23%	3.96%		1.27%	0.79%	3.46%
Crossing the $4 mark (from $3.90)			15.89%			
Within the GP range over $4			18.62%			

As shown in Table 14, the total effect varies mainly by: (a) different ranges of gasoline prices and (b) whether or not either the $3 or $4 mark is crossed by a gasoline price increase. In the following paragraphs, the effects of gasoline prices increase are discussed, taking a case of an increase by 10 percent in the middle of Table 14.

For bus and the aggregate transit, the effects of gasoline prices below $3 are positive but very small, with an elasticity value of 0.028 and 0.009. These values are translated into a ridership increase by 0.27 percent and 0.09 percent, respectively, in response to a gasoline price increase by 10 percent. The magnitude of the effect of a percent increase in gasoline price becomes larger when gasoline prices are higher to start with. When a gasoline price increases, crossing the $3 mark, the percent increase in ridership goes up to 1.17 and 1.19. Moreover, the effect increases further to 1.67 and 1.80 when gasoline prices exceed $3. These elasticity values for gasoline prices above $3—0.167 and 0.180—are substantially higher than the values found in the baseline model: 0.06 and 0.05. These results also indicate higher price sensitivity for transit users for gasoline prices above $3, and even the possibility that substantial changes in travelers' mode choice occur mostly when the price of gasoline is above $3 per gallon.

Commuter and heavy rails show a change in elasticity below and above $3, as is the case for bus and the aggregate, but elasticity is found to be negative, indicating a reduction in ridership (-0.90, -1.61, and -1.38 percent for commuter rail and heavy rail, with or without the NY UA, respectively, in response to a 10 percent increase in gasoline prices). By contrast, elasticity values become positive for gasoline prices over $3, indicating a ridership increase by 2.05, 0.66, and 0.41 percent, respectively. In addition, both commuter and heavy rails show a substantial positive boost effect of crossing the $3 mark (0.0415, 0.0539, and 0.439 in Table 13), which is included in the calculation of non-constant effects in Table 14—5.27, 6.15, and 4.85 percent,

Light rail presents a different case, as the elasticity remains the same in the gasoline price range under $4 but substantially increases above $4, indicating a 9.34 percent increase in ridership in response to a 10 percent increase in gasoline prices over $4. When gasoline prices increase from $3.90 to $4.29 (10 percent), the effect is a 6.82 percent increase in ridership.

Thus, the effects of gasoline prices on ridership are found to be higher for all modes when gasoline prices are higher than $4 for light rail or $3 for the other modes. The magnitude of these effects is more substantial in the higher range of gasoline prices for all modes, compared to what was found in the baseline case. In particular, light rail shows a very high elasticity value for gasoline prices over $4, while other modes show more modest change in their elasticity for gasoline prices over $3. This may be due to different demographics of riders and their trip characteristics by modes and will require a more detailed disaggregated analysis of individual travelers' mode choices. Logically, thinking of the relationship between gasoline prices and transit ridership, it is difficult to explain any of the negative effects found in the results.

VI. SUMMARY OF RESULTS, DISCUSSION, AND CONCLUSION

This study examined the net effects of gasoline prices on transit ridership by mode—bus, light rail, heavy rail, and commuter rail—and total ridership of these four modes combined—for the ten selected U.S. urbanized areas. The summary of elasticity of transit ridership to gasoline prices as well as other effects was estimated in the panel data regression analysis using four different specification models as shown in Table 15. Elasticity of gasoline prices on transit ridership is interpreted in the following manner: When an estimated elasticity value is 0.061, an increase in gasoline prices by 10 (or 100) percent is associated with a 0.61 (or 6.1) percent increase in ridership in response.

Table 15. Summary of Estimated Elasticity by Mode by Model

Model No.	Model Name	Variables	Motor Bus	Commuter Rail	Light Rail	Heavy Rail w/o NY	Heavy Rail w/ NY	Aggregate Transit
I	Baseline	Monthly gasoline price	0.061	n/s	n/s	n/s	n/s	0.049
II	Instrumental variables (IV) model	Monthly gasoline price	0.062	n/s	n/s	n/s	n/s	0.057
III	Short- and Long-term effects model	Monthly gasoline price	n/s	n/s	n/s	n/s	n/s	0.057
		1-month lag gasoline price	0.084	n/s	n/s	n/s	n/s	n/s
		2-month lag gasoline price	n/s	0.106	n/s	n/s	n/s	n/s
		13-month lag gasoline price	n/s	n/s	0.116	0.094	0.090	0.049
		Cumulative effect of gas price	0.084	0.106	0.116	0.094	0.090	0.106
IV	Non-constant elasticity model	Elasticity below $3	n/s	-0.095	n/s	-0.170	-0.146	n/s
		Elasticity between $3-$4	0.174	0.213		0.069	0.043	0.187
		Elasticity over $4			0.937			

Note: n/s indicates that an estimated coefficient is statistically insignificant at the 0.10 significance level.

Although there was a concern that the problem with endogeneity between transit service supply and ridership may cause biases in estimated coefficients of gasoline prices, the baseline and instrumental variables (IV) models (Models I and II in Table 15) produced very similar results for the estimated elasticity values. In contrast, different specifications that examine short- and long-term effects (Model III) and non-constant elasticity (Model IV) showed interesting results that were not captured by the first two models, which assumed one elasticity value across the entire range of gasoline prices. Table 15 also shows a large variance of gasoline price elasticity and other effects on transit ridership among the four different modes and the aggregate of these four modes.

To help interpret the results obtained in this study, Table 16 shows the effects on transit ridership in response to a 10 percent increase of gasoline price.

Table 16. Summary of the Effects on Transit Ridership in Response to a Gasoline Price Increase by 10 Percent by Mode by Model

Model No.	Model Name	Variables	Motor Bus	Commuter Rail	Light Rail	Heavy Rail w/o NY	Heavy Rail w/ NY	Aggregate Transit
I	Baseline	Monthly gasoline price	0.61%	-	-	-	-	0.49%
II	Instrumental variables (IV) model	Monthly gasoline price	0.62%	-	-	-	-	0.57%
III	Short- and Long-term effects model	Monthly gasoline price	-	-	-	-	-	0.57%
		1-month lag gasoline price	0.84%	-	-	-	-	-
		2-month lag gasoline price	-	1.06%	-	-	-	-
		13-month lag gasoline price	-	-	1.16%	0.94%	0.90%	0.49%
		Cumulative effect of gas price	0.84%	1.06%	1.16%	0.94%	0.90%	1.06%
IV	Non-constant elasticity model	Within the GP range below $3	-	-0.90%	-	-1.61%	-1.38%	-
		Crossing the $3 mark (from $2.90)	1.17%	5.27%	-	6.15%	4.85%	1.19%
		Within the GP range $3-$4	1.67%	2.05%	-	0.66%	0.41%	1.80%
		Crossing the $4 mark (from $3.90)	1.67%	2.05%	6.82%	0.66%	0.41%	1.80%
		Within the GP range over $4	1.67%	2.05%	9.34%	0.66%	0.41%	1.80%

The elasticity of bus ridership to gasoline prices was found consistently positive across all models. The baseline specification and IV models showed very close elasticity values—0.061 and 0.062, respectively—indicating a 0.61 to 0.62 percent increase in ridership in response to a 10 percent increase in the current gasoline prices. The short- and long-term effects model implies that a one-month lagged effect of gasoline prices can be more substantial, with a 0.84 percent increase at the same level of gasoline price change a month earlier. While the estimated elasticity is statistically insignificant for gasoline prices lower than $3, it increases to 0.174 for gasoline prices over $3. A change in gasoline prices at a higher price—above $3—has a more substantial impact on the mode choice of current and potential bus users. The study did not find any threshold boost effects of gasoline prices at the $2, $3, and $4 marks for the bus mode.

When elasticity is assumed to be constant across all gasoline prices in the baseline and models, the analysis did not find statistical evidence of non-zero elasticity for all three rail modes—commuter rail, light rail, and heavy rail. The results in the first two models for commuter rail are consistent with the conclusions in studies by Litman (2004) and Hanly and Dargay (1999), with the explanation that commute trips are generally less sensitive to a change in travel cost, including gasoline price and transit fare. However, the other two models found estimated elasticity with statistical significance. An elasticity value of 0.106 for the two-month lag indicates a longer delay effect of gasoline prices on current and potential commuter rail riders. The non-constant elasticity model for commuter rail showed negative elasticity for gasoline prices below $3 (-0.095) and positive elasticity for prices over $3 (0.213). A negative elasticity value was also found for heavy rail for gasoline prices below $3. It is not clear what leads to negative elasticity, which indicates that an increase in gasoline prices leads to a decline in transit ridership or a decline in gasoline prices leads to an increase in transit ridership. At the same time, commuter rail also shows more substantial effects in the long term and for higher gasoline prices, as is the case with the bus mode. The higher elasticity may be because potential commuter rail riders may stay on a current driving mode in the short term after a gasoline price hike and in the lower range of gasoline price but are affected significantly by the effects of gasoline prices in the long term and when over $3. Although it is much less in magnitude compared to commuter rail, light and heavy rails also show higher elasticity with a positive sign for higher gasoline prices. Boost effects were found at the $3 mark for commuter and heavy rails but not for other modes. The elasticity values of the 13-month lag for light and heavy rail modes mean a longer delay of the effect of gasoline prices on their ridership. These much-delayed effects on light and heavy rail are difficult to explain, while Chen, Varley, and Chen (2010) state that these may be due to picking up the residual seasonal effects.

Bus is the dominant mode of transit in eight of the studied urbanized areas out of ten. Buses carry a different proportion of riders in these areas, from 53 percent in Chicago to 82 percent in Los Angeles (in January 2012), while heavy rail has the highest ridership among all modes in Boston and New York. Perhaps because of this dominance of bus modes in these eight UAs, elasticity for the aggregate transit ridership is similar to the value for the bus mode.

This study showed that long-term elasticity is higher than short-term elasticity, while the length of delay varies by mode. This is an expected result since there are usually more travel options that can potentially change in the long term that were previously fixed in the short term. The finding in this study that short- and long-term elasticity values are close for bus and commuter rail modes means that few additional constraints for switching travel modes can be removed, even in the long term. In contrast, the elasticity in the long term for other rail modes may have to do with the median travel distance of riders—on average they are not as short as bus trips and not as long as commuter rail trips—which also leads to medium-level financial impacts of a gasoline price increase. These impacts, combined with a financial advantage for light and heavy rail, many of which use a flat fare system, may be felt by travelers only in the long term.

In this study, the estimated short-term elasticity values using only current gasoline prices were found statistically insignificant for light and heavy rail modes. An explanation for this lack of statistical significance requires a careful analysis of the demographics of local riders, what types of trips they are taking, and how they are served by each of the two transit modes. Cross elasticity of transit ridership to gasoline prices implies that a traveler has the option to switch between a driving mode and a transit mode when the financial impact of driving increases due to higher gasoline prices. Such a traveler is not characterized simply by a low income and/or a lack of vehicle access. Since a bus system generally has greater service coverage and travel time that is substantially longer than driving, many people that have access to bus service may prefer driving a car when they can afford it. It is also relatively easy to switch back to a bus. A commuter service, while it has limited service coverage, has a very distinct group of customers—commuters that tend to travel relatively long distances to major job centers, including downtowns. With some of these commuter services, users may switch to commuter rail when fuel costs increase due to the financial burden of long-distance driving. In contrast, light rail and heavy rail systems offer more limited service coverage than a bus system, but they serve more diverse types of travelers and trips than commuter rail. Thus, an analysis of light rail and heavy rail likely requires a much finer geographic unit of analysis within an urbanized area, carefully taking into account the alignment of rail lines.

The finding that elasticity values are larger for higher gasoline prices for all modes is consistent with the finding in the study by Chen et al. (2010). This is probably because there are more people at the margin of choosing either driving or transit as gasoline prices increase, reflecting the fact that the current dominant travel mode is driving, even in major U.S. urbanized areas. This finding implies that a further increase to gasoline prices from currently high prices, compared to increases to the lower prices of the past, will have more substantial negative impacts on travel budgets and that the elasticity of ridership to gasoline prices will likely become even higher if gasoline prices continue to increase in future.

One question that arises here is whether the reason for a gasoline price increase—e.g., a market price increase vs. a fuel tax increase—affects the value of elasticity of transit ridership. There is probably a large difference in sentiment among the public toward these two different causes because public opinion can influence policy decisions for a fuel tax increase. It is difficult, however, to assume that an economic effect of gasoline prices varies by such differences, as it is likely that people are concerned about an increase in gasoline prices because of its impact on their travel behavior, not because of the reasons for the increase. Results from this study are indicative of what is to be expected for transit ridership resulting from a fuel tax increase. The same discussion applies also to a tax on carbon dioxide, which would increase taxes on gasoline, as the public may become increasingly concerned with climate change in the future.

One additional finding in this study is that the magnitude of elasticity to transit fare was generally found larger than that of elasticity to gasoline prices. This indicates higher price sensitivity to transit fare among transit riders. This is also consistent with the findings from past studies (Litman 2004; Chen, Varley, and Chen 2010). The explanation may be that the elasticity of transit ridership to transit fare is more direct than elasticity to gasoline prices, which is a cross elasticity.

The values of elasticity of transit ridership to gasoline prices estimated in this study are generally lower than those of past studies, perhaps because half of the ten urbanized areas studied (Boston, Chicago, Cleveland, New York, and San Francisco) have a higher proportion of zero-car households than the national average (US Census Bureau 2010-2012). (By contrast, Denver, Houston, Los Angeles, Miami, and Seattle have a lower-than-average proportion.) These zero-vehicle households represent a relatively large number of (a) transit-dependent individuals who are less price-sensitive than discretionary riders or (b) individuals who choose not to own a private car and therefore are unlikely to change travel mode.

This study developed elasticities of transit ridership to gasoline prices that are more generalizable, by analyzing data for the ten selected urbanized areas. The obtained results indicate that transit agencies should consider preparing capacity management plans and increasing the supply of transit services to prepare for a likely ridership increase—particularly for bus, commuter rail, and light rail modes, for which a 10 percent increase in gasoline prices can cause a 1.7, 2.1, and 9.3 percent increase in ridership, respectively, when gasoline prices go over $3 or $4 per gallon. In addition, the boost effect at the $3 mark is so substantial for commuter rail and heavy rail, a gasoline prices increase crossing $3 requires special attention. The response time should be somewhat short for the bus and commuter rail modes, compared to light and heavy rail modes, according to the findings in this study. For light and heavy rail modes, it is likely that the impacts of gasoline prices on ridership occur more slowly, giving transit agencies more time to respond.

Taking into account the global political economy, in which politics and foreign affairs influence the market for oil, along with the geographically uneven distribution of fossil fuels, it is not likely that gasoline prices in the U.S. will fall below $3 in the future. Although the percent of ridership increases may appear small, they can have substantial negative impacts on transit operation and management if a ridership increase is concentrated during the peak periods, when service levels are at or near the maximum supply capacity for transit agencies. A ridership increase in the peak period requires a substantial increase in service supply and facility capacity, which in turn requires transit agencies to increase service inputs, such as capital and labor. For example, the Washington Metropolitan Area Transit Authority (WMATA) has been experiencing overcrowding during peak periods and is, in turn, increasing train frequency and expanding the capacity of transfer stations for heavy rail service. On the other hand, it should be noted that a ridership increase may also improve cost effectiveness during off-peak periods by increasing the ratio of riders to vehicle capacity.

While transit managers may be concerned that a reduction in gasoline prices may lower transit ridership in future, the study by Chen et al. (2010) showed a relatively lower sensitivity of transit ridership to a decrease in gasoline price; a decrease in gasoline prices does not discourage transit use as much as an increase to gasoline prices encourages it.

This study chose a panel data analysis method over a disaggregate analysis method for numerous reasons: (1) it is very difficult to obtain a large, comprehensive data set for a disaggregate model *over time—before and after a fuel price change*, and (2) elasticity measures a change that is inherently temporal. At the same time, as discussed above, the urbanized areas used in this study may have been too large of a geographic unit of analysis, taking into account the relatively smaller service coverage area of rail service, compared to an entire urbanized area. Therefore, the use of such a geographic unit for socioeconomic characteristics variables may potentially improve the quality of control variables. Furthermore, compared to an aggregate analysis, a disaggregate (discrete choice) analysis has advantages in that it can possibly incorporate details of socioeconomic characteristics and life conditions of individual travelers or households and trip characteristics, as well as other variables of aggregated levels that could be similar across travelers and trips. While it will require substantial resources and effort, it is important to collect data that contain a sufficient number of transit trips over time to examine the elasticity of transit ridership, which is inherently a question about a change in people's travel mode choice over time.

APPENDIX A: ANNOTATED BIBLIOGRAPHY

This annotated bibliography is a compilation of selected readings that provide the background for the report. Each article provides a brief summary of the issue examined in the study, methodology used, and findings, along with comments from the authors of this report.

The length of each review depends on the length and breadth of the article and its relevance to our report. A quick reference to the articles appears is presented in Tables 17, 18 and 19, for the articles. Studies of cross-sectional analysis are reviewed first, studies of time-series analysis second, and then studies of panel data analysis. The reviews appear in the same order as in the tables.

Table 17. List of Location, Year, and Data Sources Used in Studies on the Gasoline Price Elasticity of Transit Ridership

Study	Location	Year	Data source
Bomberg and Kockelman (2007)	Austin, Texas	February and April 2006	Data on fuel economy of respondent's vehicle from US Department of Energy and the US Environmental Protection Agency. Using GIS software, they match geocodes for repondents' home location to neighborhoods and obtain data on neighborhood characteristics such as density of bus stops, euclidean distance to the CBD, total zonal density measured as the ratio of total number of jobs and households per unit of area. Data on accessibility Indices (AIs) from the Gupta et al (2004) paper.
Kain and Liu (1999)	San Diego, CA; Houston, TX	1980, 1990	Demographic data from US Census of Population, 1980, 1990, data on bus and rail miles and fares from transit agencies Metro and Metropolitan Transit Development Board.
Taylor, Miller, Iseki and Fink (2009)	265 US urbanized areas	2000	National Transit Database (NTD) for ridership and transit related data. 2000 US Census for data on demographic and other factors external to transit agencies that affect ridership.
Currie and Phung (2007)	US	1998-2005	American Public Transport Association
Haire, Machemehl (2007)	5 US cities: Atlanta, Dallas, Los Angeles, San Francisco, and Washington DC	1999-2006	American Public Transport Association
Maley and Weinberger (2009)	Philadelphia	January 2001-June 2008	Gas price data from Phillygasprices.com, and ridership data from SEPTA reports
Lane (2010)	9 US metropolitan areas: Boston; Chicago; Cleveland; Denver; Houston; LA; Miami; San Francisco; Seattle	January 2002/ June 2003-April 2008	Ridership data from(NTD) 2008. Gas price data from Energy Information Administration of US Department of Energy, 2008
Mattson (2008)	Urban and rural areas in upper Mid-west and mountain states: Duluth, MN; St. Cloud, MN; Rochester, MN; Sioux Falss, SD, Fargo, ND, Billings, MT, Grand Forks, ND; Missoula, MT; Great Falls, MT; Rapid City, SD; Cheyenne, WY; Logan, UT.	For time-series analysis they use monthly data for January 1999-December 2006. For panel data analysis they use annual data from 1997-2006.	Aggregate national ridership data from American Public Transportation Association (APTA), gas price data from EIA. Obtained data on demand for and supply of transit from National Transit Database for panel analysis.
Yanmaz-Tuzel and Ozbay (2010)	Northern New Jersey, with one line running between Atlantic City and Philadelphia.	1980 to 2008	Data on NJ transit ridership from NJ Transit Market Analysis and Pricing Department, data on annual unemployment rate from Bureau of labor statistics

Table 17, Continued

Study	Location	Year	Data source
Chen, Varley, and Chen (2010)	New Jersey and New York City	January 1996-February 2009	New Jersey Transit provided data on monthly commuter rail trips. Data on Vehicle revenue miles from National Transit Database. Population data from US Census. Employment and size of labor force data from Bureau of Labor Statistics. Monthly gas price data from EIA was adjusted for inflation using CPI collected from BLS.
Storchman (2001)	Germany, public transportation in urban areas of Germany	1980-1995	Yearly statistical handbook from German Ministry of transportation
Curie and Phung (2008)	Melbourne, Brisbane, and Adelaide in Australia	Melbourne (Jan. 02-Dec. 05), Brisbane (July 04-Nov. 06), Adelaide (Jan. 02-Nov. 06).	Gas price data from Australian Automobile Association, Interest data from Reserve bank of Australia. Transit usage data from Brisbane, Melbourne and Adelaide transit authority.
Stover and Bae (2011)	11 counties in Washington state	January 2004-November 2008	Monthly ridership data from NTD, and gas price, unemployment and labor force data obtained from Bureau of Labor Statistics
Nowak and Savage (2013)	Chicago metropolitan area	January 1999 and December 2010	Ridership data from regional transportation asset management systems, gas price data from American Automobile Association-Chicago Motor Club, and unemployment data from Bureau of Labor Statistics
Blanchard (2009)	218 US cities	2002-2008	Monthly data on transit ridership and transit supply measured as vehicle revenue miles obtained from NTD. Gasoline price data from US EIA adjusted for inflation using GDP implicit price deflators. Data on vehicle miles traveled gathered from Federal Highway Administration.

Table 18. List of Mode, Type of Data, and Aggregation Level Used in Studies on Gasoline Price Elasticity of Transit Ridership

Study	Type of data	Mode	Level of Aggregation
Bomberg and Kockelman (2007)	Cross-section	Bicycle, driving, and transit	Individual household
Kain and Liu (1999)	Cross-section	Bus	Metro service area in Houston, and MTS service area in San Diego
Taylor, Miller, Iseki and Fink (2009)	Cross-section	Total level of transit service provided by all transit agencies in an urbanized area	Urbanized area level
Currie and Phung (2007)	Time series	Bus, light rail, heavy rail, total for all modes combined	By mode for all of US
Haire and Machemehl (2007)	Time series	Bus, light rail, heavy rail, commuter rail	By mode for each of 5 US cities

Table 18, Continued

Study	Type of data	Mode	Level of Aggregation
Maley and Weinberger (2009)	Time series	Rail provided by Regional Rail Division of SEPTA. Bus service, nine light rail or street car route service and two subway route service provided by City Transit Division of SEPTA.	By mode
Lane (2010)	Time series	Bus, rail, bus and rail combined	For nine cities combined, monthly data analysis by mode: bus, rail, and bus and rail combined. For each of 9 cities separately he analyzed monthly data by mode: bus, rail and bus and rail combined.
Mattson (2008)	Both time-series data, and panel data were used	Bus	For time-series analysis he divided monthly data from upper Mid-west and mountain states into 4 groups of metropolitan areas based on population size. Four groups are above 2 million population, 500 thousand to 2 million, 100 thousand to 500 thousand, and below 100 thousand. For panel analysis, he used annual ridership data for each transit system.
Yanmaz-Tuzel and Ozbay (2010)	Time series	Overall New Jersey transit ridership	Monthly data for all modes combined in New Jersey
Chen, Varley, and Chen (2010)	Time series	New Jersey commuter rail	Monthly data on New Jersey commuter rail ridership
Storchman (2001)	Time series	All urban public transportation: bus, tram and underground	He ran the regressions at the mode of transport level for a given purpose of travel for example, work, leisure etc. using annual data
Curie and Phung (2008)	Time series	Rail, Australian bus rapid transit) BRT, and bus	Using monthly data they ran regressions for each city separately after aggregating transit usage for all modes. They also ran city wide regression disaggregating at the rail, bus and bus rapid transit level.
Stover and Bae (2011)	Time series, panel data	Aggregate transit ridership	Regress aggregate ridership for each county separately
Nowak and Savage (2013)	Time-series	City heavy rail, city bus and suburban bus and suburban rail	Regress ridership for each model separately
Blanchard (2009)	Panel data	Commuter rail, heavy rail, light rail and bus	Regress separately for each mode: motorbus, light rail, heavy rail, commuter rail using monthly data for 218 cities

Table 19. List of Dependent and Independent Variables and Empirical Estimation Methods Used in Studies on Gasoline Price Elasticity of Transit Ridership

Study	Dependent Variable	Independent Variables	Empirical specification and Strategy
Bomberg and Kockelman (2007)	Shopping around for gas, overall driving, chaining activities, carpooling, transit use, and bicycle trips.	Respondent's transportation needs, demographic attributes such as age, gender, income, student or not, household size, number of vehicle per driver. Neighborhood/local characteristics such as local population, whether or not the area is residential, or commercial area, retail employment, service employment, total employment in the area, distance to CBD, bus stop density, zone density. Gas expenditure, fuel economy of all hh vehicles, no. of non-work-related trips, whether or not works from home, whether household has children going to school.	They used ordered probit models to examine likelihood of respondents increasing trip chaining or reducing their driving, in response to the 2005 gas price spike. While they use binary logit models for driving slower and driving at steadier speeds.
Kain and Liu (1999)	Log of ridership	SMSA employment, central city population, bus and rail miles supplied by the transit system in the area, real fares	Using simple OLS model, they regressed log of ridership on log of independent variables
Taylor, Miller, Iseki and Fink (2009)	Total urbanized area ridership, Per capita ridership	Geographic land area, total population, population density, regional dummy, median household income, ratio of unemployed to labor force, ratio of enrolled college students an total population, ratio of population in poverty to total population, ratio of immigrant population to total population, percent of votes cast for democratic party in 2000 presidential election, freeway lane miles, average gas price per gallon of gas, ratio of sum of non-transit and non-SOV commutes to all commutes, ratio of household with no vehicle to total household, total lane miles, Daily vehicle miles travelled per capita. They also control for transit system characteristics, such as transit fares, headways/service frequency.	They used two staged least squares estimation strategy and instrumented supply of transit, measured as total urbanized area transit service vehicle revenue hours with total population, percent voting Democrat in 2000 presidential election.
Currie and Phung (2007)	Log of national (US) transit ridership	Log of gas price, log of gas price interacted with dummies for 9/11 incident, the Iraq war and Hurricane Katrina, month dummies	Simple OLS based log-log model
Haire and Machemehl (2007)	Change in ridership over two consecutive months	Price of gasoline	Simple OLS, using level of dependent and independent variable
Maley and Weinberger (2009)	Monthly ridership	Gas price, monthly dummies to control for seasonality	They used simple ordinary least squares estimation model to regress gas price and monthly dummies on ridership

Table 19, Continued

Study	Dependent Variable	Independent Variables	Empirical specification and Strategy
Lane (2010)	Monthly unlinked passenger trips for all rail and bus modes, Monthly Unlinked passenger trips for bus, Monthly Unlinked passenger trips for rail	Current gas price, one month lagged gas price, standard deviation of monthly gas price for each month, time trend, seasons such as fall, spring, summer, supply of transit variables such as vehicle revenue miles operated, vehicles operated in maximum service	Level of dependent variables on level of independent variables using simple OLS
Mattson (2008)	Log of monthly ridership	For time-series analysis, he controlled for 15 lags of gas price data along with yearly dummy to control for time trend. For the panel regression he controlled for size of labor force and unemployment level, quantity of service provided, fare for that transit, time trend interacted with dummy indicating transit system, and dummy variables indicating whether there have been events to create demand shocks for any specific transit system.	For time-series analysis: polynomial distribution lag model to analyze long term effect of gas prices on ridership. He used a log-log model where the lagged gas prices are also logged. For the panel data analysis, used simple OLS model to regress log of ridership on log of independent variables, which did not include lagged gas prices.
Yanmaz-Tuzel and Ozbay (2010)	Monthly transit ridership (in thousands)	Total monthly employment in New Jersey and New York City (in thousands), average monthly gasoline prices, lagged monthly gasoline prices, average NJ transit fare, vehicle revenue hours in thousands, month dummies	They use a log-log model to estimate the gas price elasticities from the coefficient of gasoline price, and lagged gasoline price. Using the log-log model they estimate correlations between employment level and ridership.
Chen, Varley, and Chen (2010)	Number of New Jersey commuter rail trips to and from New York City	They control for lagged ridership, positive and negative changes in gasoline price and transit fare, labor force and service level measured as vehicle revenue miles and its fourth lag, seasonal dummies (captured using monthly dummies)	They regress change in transit ridership between period t and t-1 on change in ridership between period t-1 and t-2 and change in gasoline price interacted with a dummy equal to 1 if the price change is non-negative and equal to 0 otherwise, similarly they control for negative changes in prices by interacting the price with a dummy equal to 1 if the price change is negative and 0 otherwise. They adopt the same strategy to analyze the differential effects of positive or negative change in transit fare. ARFIMA (auto-regressive fractionally integrated moving average) to answer what factors affect transit ridership, and what are the short-term and long-term effects of various factors on transit ridership. They use AR(1) model to examine if there is asymmetry in transit ridership in response to rises and falls in gasoline price and transit fare.

Table 19, Continued

Study	Dependent Variable	Independent Variables	Empirical specification and Strategy
Storchman (2001)	Number of trips for work, school, shopping, business, leisure, and holiday, by mode. Average distance travelled each trip purpose by mode.	In the equation where they estimate choice of mode of transport for each travel purpose, they control for demographic variables, income, and a dummy indicating German unification in 1991. In the estimation of distance travelled using public transport for each purpose they control for gas price, stock of public transport, income and transit fare, available public infrastructure (such as railroads or road network) and German unification. In the estimation of demand for passenger kilometers, they control for purpose of trip, distance travelled, seats per vehicle, average peak seat load factor during peak period and average speed during peak periods.	He estimates a system of equations. Estimate how demography and german unification in 1991 affected number of trips taken for each of these travel purposes: work, school, shopping, business, leisure, holiday. Then estimate average distance of trip for each of the purposes listed above are affected by stock of cars, transportation prices, available railroads, or road network, and German unification. Then he estimates the public transit vehicle demand as a function of peak passenger kilometers, seats per vehicle, average peak seat load factor during peak period and average speed during peak periods. Then he estimates the cross-price elasticity for public transportation demand. Then using the gas price elasticities from these regressions, he simulates/calculates the effect of gasoline on revenue collection from taxes on gasoline and also revenue change of public transport sector.
Curie and Phung (2008)	Per capita validations (which is equivalent to per capita transit usage)	Gasoline price, interest rate, and monthly dummy variables to indicate seasonality	Using a simple OLS model, they regress a log of per capita transit usage on log of gasoline prices, absolute level of interest rate, and monthly time dummies
Stover and Bae (2011)	Unlinked revenue trips	Gas price, transit fare, supply of transit, unemployment rate, size of labor force, season dummies	Simple OLS, regressing log of ridership on log of independent variables
Nowak and Savage (2013)	Unlinked trips for CTA bus, count of passengers entering stations for CTA rail, number of ticket sales for Metra, number of boardings for Pace	Gas price, gas price interacted with dummy that is equal to one if gas price is more than $3, gas price interacted with dummy that is equal to one if gas price is more than $4, average daily transit bus miles, transit fare, unemployment rate, proportion of weekdays in month, dummy variable for leap year	Simple OLS, regressing log of ridership on log of independent variables
Blanchard (2009)	Ridership measured as unlinked passenger trips by mode: commuter rail, heavy rail, light rail, motorbus.	Supply of transit, gasoline price, and lagged gasoline prices, monthly dummies, year dummies	Simple OLS based log-log model to calculate the immediate elasticities and also used log of past gas prices to analyze long term effects of gasoline prices on ridership
Yanmaz-Tuzel and Ozbay (2010)	Monthly transit ridership (in thousands)	Total monthly employment in New Jersey and New York City (in thousands), average monthly gasoline prices, lagged monthly gasoline prices, average NJ transit fare, vehicle revenue hours in thousands, month dummies	They use a log-log model to estimate the gas price elasticities from the coefficient of gasoline price, and lagged gasoline price. Using the log-log model they estimate correlations between employment level and ridership.

Bomberg, Matthew, and Kara M Kockelman, “Traveler Response to the 2005 Gas Price Spike.” Proceedings of the 86th Annual Meeting of the Transportation Research Board, Washington DC (2007).

Bomberg and Kockelman (2007) analyzed the effect of a severe spike in gas prices that transpired in 2005. Using a survey of over 500 residents in Austin, Texas and ordered probit and binary logit models, this study analyzed the effect of a gas price hike on the change in people’s propensity to shop around for gas, drive automobile, chain activities, carpool, use public transit, and ride bicycles. In analyzing the effect, this study accounted for socioeconomic and demographic attributes of the respondent, as well as neighborhood/local characteristics, such as distance to the central business district (CBD), bus stop density, geographic zone density, etc. This study found that, all else being equal, individuals are most likely to increase trip chaining in response to the price spike if they live in or near the CBD, and also that higher income households are less likely to reduce driving in response to a gas price spike. This study analyzed data from a specific region, Austin, Texas, with specific metropolitan characteristics. Therefore, the results are based on data from a region that may not be comparable to other metropolitan or urban regions and, so, cannot be generalized to other areas with different socioeconomic, geographic, and neighborhood characteristics.

Kain, John F., and Zvi Liu, “Secrets of success: assessing the large increases in transit ridership achieved by Houston and San Diego transit providers.” *Transportation Research A*, 33 (1999): 601–624.

This is one of the earlier studies that analyzed factors affecting transit ridership in the U.S. Kain and Liu analyzed the effect of various factors on bus ridership in the Metropolitan Transit System (MTS) in San Diego, CA, and Metro in Houston, TX between 1980 and 1990. Their analysis compared how changes in ridership were affected by factors outside the control of the transit providers (i.e., service area characteristics) : Standard Metropolitan Statistical Area (SMSA) employment and central city population. The study also examined factors within the control of the transit agencies: bus and rail miles supplied by each system and real fares. They found that large service increases and fare reductions, as well as population growth and metropolitan employment, tended to increase ridership. This study suffers from omitted variable bias, as it does not control for gasoline prices. Changes in the price of gasoline can potentially affect ridership, and transit agencies may respond by changing their supply of bus services. The omitted variable—the price of gasoline—is correlated with bus ridership and bus supply and, therefore, biases the estimated coefficient for bus supply. Similarly, a gasoline price change can often alter the cost of supplying bus services, as many buses use gasoline for fuel, a factor that can prompt transit agencies to change their bus fares. Therefore, gasoline price is correlated with bus fare and ridership, which causes the estimated coefficient of bus fare to be biased.

In addition to omitted variable bias, the authors do not address the problem of simultaneity between supply and demand: Supply of buses affects demand for bus ridership because increased (or decreased) bus supply can improve (or worsen) the convenience of traveling by bus by altering frequency and duration of service, influencing riders to adjust demand accordingly. Bus ridership demand can also impact supply, therefore bus supply is endogenous. The authors treat the supply of buses as exogenou, and do not account for this endogeneity, causing the estimated coefficient of bus supply to be biased.

Taylor, Brian D., Douglas Miller, Hiroyuki Iseki, and Camille Fink, “Nature and/or Nurture? Analyzing the determinants of transit Ridership across US Urbanized Areas.” *Transportation Research* Part A, 43, (2009): 60-77.

Taylor et al. (2009) offers an improvement over Kain and Liu (1999), as it analyzes factors affecting transit demand using transit ridership data from the National Transit Database (NTD). This allows measurement of the impact of a wider array of factors affecting transit ridership in a larger sample of geographic regions. Taylor et al. (2009) conducted the most comprehensive cross-sectional study to date of 265 U.S. urbanized areas in the year 2000 to investigate how factors controlled by transit agencies (internal factors) and factors outside the control of transit agencies (external factors) affect total urbanized area ridership and per capita ridership. They were able to isolate how much change in ridership is attributable to the fluctuation in gasoline prices alone (net effect of gasoline price changes on transit ridership) by controlling for a variety of other variables that also influence transit ridership. These included such internal factors as fares, frequency of service, hours of service, on-time performance, service coverage and quality of service. They also controlled for external factors that included measures of regional economic activity, population, population density, labor market size, availability of parking in the CBD, and socioeconomic demographics of the population (age, income, vehicle ownership, etc.).

The study recognized that transit supply (availability of service in terms of area coverage and hours of operation) affects transit demand, but transit demand (due to an increase in gasoline prices, for example) simultaneously influences transit supply, making supply endogenous due to this simultaneity. To account for the endogeneity of supply, Taylor et al. (2009) instrumented for transit supply, measured here as total urbanized area transit service vehicle revenue hours, with total population and the percent voting Democrat in 2000 presidential election. This was an important improvement over past studies, which failed to account for endogeneity of transit supply. In addition, Taylor et al. (2009) distinguish between the effects of internal factors on transit ridership from the effects of external factors, allowing them to obtain an unbiased estimate of the effect of supply on ridership. They found that transit supply and external factors, such as metropolitan economy, regional geography, population characteristics, and automobile highway system characteristics affect total urbanized area ridership and per capita transit ridership.

Note that despite the strength of this paper, the authors cannot distinguish between the effects of the price of gasoline on each individual mode, as they aggregated all modes of transit ridership over urbanized regions. This type of cross-sectional analysis cannot distinguish between short-term and long-term impacts of gasoline prices on transit ridership, nor can it indicate whether gasoline price elasticity of ridership varies when there is an increase, as opposed to a decrease, in gasoline prices. Studies that span longer time horizons, such as those using time-series data to capture variability in prices, are more suited for such analysis.

Currie, Graham, and Justin Phung, "Transit ridership, auto gas prices, and world events: New drivers of change?" *Transportation Research Record: Journal of the Transportation Research Board* 1992 (-1), (2007): 3-10.

Curie and Phung (2007) analyzed the elasticity of U.S. transit ridership with respect to gas prices. The authors conducted a time-series data analysis using the monthly U.S. national transit ridership data from the American Public Transportation Association (APTA)—in total and by mode—between 1998 and 2005, with special attention to three "world events": September 11, 2001, the most recent Iraq War, and Hurricane Katrina. This study regressed the log of ridership relative to the log of gas price, and the log of gas price interacted with dummy variables, for these three events. The results show a variation in gas price elasticity of transit ridership by mode: light rail ridership has the highest elasticity, with values ranging from 0.27 to 0.38, heavy rail follows with an elasticity of 0.17 to 0.19, and bus ridership shows a very low elasticity of 0.04 to 0.08. The authors speculate that higher shares of "choice" riders (i.e., those who own, or could easily own, an automobile) choosing to use light rail could explain high gas price elasticity values for light rail. This study does not control for a full range of explanatory variables, such as demographic characteristics that affect ridership, an issue that needs to be addressed to improve the statistical performance of the modeling.

Haire, Ashley, and Randy Machemehl, "Impact of rising fuel prices on US transit ridership." *Transportation Research Record: Journal of the Transportation Research Board* 1992, no. 1 (2007): 11-19.

Haire and Machemehl (2007) investigated the effect of changes in the price of gasoline on changes in ridership for four different modes of transit—bus, light rail, heavy rail, and commuter rail—in five U.S. cities (Atlanta, Dallas, Los Angeles, San Francisco, and Washington DC), using the APTA monthly ridership data between 1999 and 2006. Based on the ratio of the percent change in ridership to percent change in price of gasoline, they obtained cross elasticity estimates by mode that are quite different from the results from Currie and Phung's study. Haire and Machemehl show the lowest elasticity for light rail (0.07), followed by heavy rail (0.26), commuter rail (0.27), and bus (0.24). This analysis did not account for any external or internal factors that affect ridership, such as demographic and regional characteristics and supply of transit services. These results indicate that the use of different analysis methods and data, in terms of both geographic location and scale, yield different estimates of elasticity. In addition, cross-price elasticity of gasoline with respect to ridership varies depending on whether data are analyzed for the entire transit system or by mode.

Maley, Donald W., and Rachel Weinberger, "Rising Gas Prices and Transit Ridership Case Study of Philadelphia." *Transportation Research Record, Journal of the Transportation Research Board* 2139, no. 1 (2009): 183-188.

Using weekly data between January 2001 and June 2008 for the SEPTA rail and bus transit system in Philadelphia, Maley and Weinberger (2009) analyzed the correlation between gas prices and transit ridership, accounting for seasonal differences. The authors performed a set of regressions by transit mode—bus and rail—to analyze the effect of gas price. They controlled for the month, using dummy variables to account for seasonal variation in ridership. They found a positive effect of gas price on ridership for both modes. Maley and Weinberger obtained an elasticity estimate of 0.12 for bus and of 0.22 for rail. In another set of regressions, they added a squared term for gasoline price along with gas price and monthly dummies, which produced a better R-squared, but they did not explain the theoretical basis for addition of the squared term. The study by Maley and Weinberger is too simplistic to fully examine the net effect, as it only isolates the correlation between gas price and ridership by controlling for seasonal effects but does not control for other external and internal factors, such as demographic, economic, and regional characteristics; transit fare; or transit service supply. Not controlling for other external and internal factors can lead to omitted variable bias in the estimated coefficients in the regression model. Another limitation of this paper is the use of data from only one city—its findings may not be generalized due to lack of external validity.

Despite its flaws, the paper adds value to the literature on transit ridership, as it mentions (in the conclusion) an interesting feature regarding the ridership elasticity to gas price that can be explored in future research. The authors posit that there may be a threshold effect at the $3 level of gasoline price and suggest examining the possibility of different elasticity values at different price levels. The implication is that the most commonly used simple log-log regression model, which assumes a constant elasticity regardless of the value of ridership or gasoline price, should be modified to capture this difference in elasticity. This suggestion is similar to the hypotheses posed in studies by Chen, Varley, and Chen (2010) and Maley and Weinberger (2009), which suspect that the elasticity for gas price increases may vary from the elasticity for gas price decreases.

Lane, Bradley W., “The Relationship Between Recent Gasoline Price Fluctuations and Transit Ridership in Major U.S. Cities.” *Journal of Transport Geography*, Vol. 18, No. 2, (2010): 214–225.

Lane (2010) analyzed the effect of the price and price variability of gasoline on transit ridership. This study conducted time-series data analysis using monthly data on transit ridership from NTD over the period June 2002 through April 2008 for nine U.S. metropolitan areas: Boston, Chicago, Cleveland, Denver, Houston, Los Angeles, Miami, San Francisco, and Seattle. Transit ridership for both bus and rail modes, as well as the total ridership for both modes combined, was analyzed for each city, resulting in twenty seven regression equations in total. Independent variables included the current gasoline price; the one-month-lagged gasoline price; the standard deviation of monthly gasoline price; year dummy variables; seasonal effects in fall, spring, and summer;[13] and transit supply variables measured in terms of vehicle revenue miles operated and vehicles operated in maximum service. Lane found that for most cities, the gasoline price variation (standard deviation of gasoline price) had no significant effect on either bus or rail ridership.

This study also found that in some cities gasoline price had a positive effect on bus ridership but no statistically significant effect on rail ridership, while in a few other cities it had a positive effect on rail ridership but no effect on bus ridership. Note that this large diversity of results for different cities may be due to the fact that the socioeconomic and demographic characteristics of transit users and residents of these cities are different; hence, the purpose of public transit trips may be so different that the elasticity of gasoline price to systemwide ridership varies significantly due to a variation of elasticity by trip purpose. Lane did not control for income, employment, and other important socioeconomic factors that can change over time and influence decisions about mode choice. Moreover, as with most studies in the literature, this study does not consider endogeneity of transit service supply to ridership. Also, note that given the NTD data used, the author could have conducted panel data analysis, as NTD has data for some cities over multiple periods of time. By using a fixed effects model with panel data, it would have been possible to control for time-invariant, unobserved factors that affect ridership and thus fix the omitted variable bias to some extent. The use of panel data analysis methods may address potential problems associated with statistical insignificance for many estimated coefficients in the city-level regressions in the study.

Mattson, Jeremy, “Effects of Rising Gas Prices on Bus Ridership for Small Urban and Rural Transit Systems.” *Small Urban and Rural Transit Center*, Upper Great Plains Transportation Institute, North Dakota State University, (2008).

Mattson (2008) analyzed the effect of gasoline prices on bus ridership. Using ridership data from APTA, Mattson conducted multiple analyses of bus ridership using time-series and panel data for twelve urban and rural areas in the Upper Midwest and Mountain States: Duluth, MN; St. Cloud, MN; Rochester, MN; Sioux Falls, SD; Fargo, ND; Billings, MT; Grand Forks, ND; Missoula, MT; Great Falls, MT; Rapid City, SD; Cheyenne, WY; and Logan, UT. For time-series analysis, the author used monthly data for the period January 1999 through December 2006 for four groups of metropolitan areas: those with a population of more than two million, 500,000 to 2 million, 100,000 to 500,000, and less than 100,000. For each group of metropolitan areas he used a polynomial distribution lag model to regress the log of monthly ridership on the log of 15 lags of gas price along with month and year dummy variables to analyze the long-term effects of gas price on ridership. This study uses month dummy variables to control for seasonality and dummy variables indicating year to account for trends in ridership demand in all areas. It found that coefficients for gas price up to the seventh lag were statistically significant. Because time-series data analysis of this study did not control for all socioeconomic and demographic variables, or any variables indicating supply of transit, it is likely that the results suffer from omitted variables bias.

As the effect of gas price change on ridership is inherently temporal, there are advantages to using longitudinal data to examine it; however, at the same time, results from a time-series analysis on a few particular transit systems may not be generalizable. Panel data analysis proves to be advantageous, allowing the researcher to simultaneously take into account temporal and regional variation to obtain more robust, generalizable results.

Using data from 1997-2006, Mattson conducted panel data analysis for which he aggregated the data to make yearly data for each transit agency. The author used both linear and double log functional form models of regression. The double log regression model included variables for the price of gasoline, labor force participation, unemployment, the quantity of transit service provided, and transit fares. Additional controls included year dummy variables interacted with dummy variables for transit systems, as well as dummy variables for events that created demand shocks for any specific transit system. The analysis also included fixed effects to take into account unobserved, time-invariant factors that can affect demand for transit. For example, the variable population density, a measure of urban sprawl (which is not controlled for and is likely to affect supply and demand for bus services), changes very slowly over time and can be considered time-invariant in this analysis because it spans nine years. However, the fixed effects model is still unable to account for the endogeneity bias that arises because supply affects demand and is simultaneously affected by it. Another shortcoming of this study is that it uses annual data for panel data analysis, and many previous analyses indicated that the full effect of a change in gas price is observed within one year (Yanmaz-Tuzel and Ozbay, 2010; Mattson, 2008).

Thus, the use of annual ridership data does not allow researchers to analyze how long-lasting the effect of gasoline price on public transit ridership is within 12 months. Mattson analyzed only bus ridership; hence, the results may not be generalizable to other modes in these areas, as people use each mode for different purposes. For this reason, the gas price elasticity may vary by mode.

Yanmaz-Tuzel, Ozlem and Kaan Ozbay, "Impacts of Gasoline Prices on New Jersey Transit Ridership." *Transportation Research Record: Journal of Transportation Research Board*, 2144, no. 1 (2010): 52-61.

Yanmaz-Tuzel and Ozbay (2010) analyze monthly transit ridership data for all modes in Northern New Jersey over the period 1980 to 2008, controlling for some of the same external and internal factors that were accounted for by Taylor et al. (2009). Yanmaz-Tuzel and Ozbay used a log-log model to examine how transit ridership is affected by total monthly employment in New Jersey and New York City, average monthly gasoline prices, lagged monthly gasoline prices, average New Jersey transit fare, and vehicle hours, while controlling for month dummy variables. The second lag and third lag of gasoline price have statistically significant cross-price elasticities to ridership in the range of 0.15 to 0.23 and 0.03 to 0.20 respectively. This indicates that travelers consider the trend in gasoline price before making a decision to switch their mode of travel, and there is a time lag between the change in the gasoline price and its effect on transit demand. Careful analysis in the future should control for long lags of gasoline price when analyzing gasoline price elasticity. However, this study is not without its flaws; external validity is an issue in this study, as it analyzed ridership using data from only one region. In addition, this study did not analyze the effects of economic factors on ridership over this study period. Controlling for a few supply variables is an improvement over earlier studies; however, the authors do not account for the fact that supply of transit is endogenous. The authors did not use the instrumental variables method to instrument for the supply of transit.

Chen, Cynthia, Don Varley, and Jason Chen, "What Affects Transit Ridership? A Dynamic Analysis involving Multiple Factors, Lags, And Asymmetric Behaviour." *Urban Studies*, Vol. 48, No. 9, July (2010): 1893-1908.

Chen, Varley, and Chen (2010) also analyzed monthly data for New Jersey commuter rail ridership[14] over the period 1996 to 2009 to examine whether gasoline price elasticity of ridership differs for a gasoline price increase vs. a gasoline price decrease. In addition, this study examines whether transit fare elasticity to ridership differs depending on whether a change in transit fare is an increase or decrease. The authors of this study created an interaction term between the change in gasoline price for the same two consecutive months and a dummy variable that indicates either an increase or decrease in gasoline price, and then regressed the change in ridership between two consecutive months on this interaction term. They adopted the same strategy to analyze whether the transit fare elasticity to ridership differs between an increase and a decrease in transit fare. They controlled for size of labor force, transit service supply, and seasonal effects (captured using month dummy variables). This study found that the gasoline price elasticity for a rise in gasoline price differs from the gasoline price elasticity for a fall in gasoline price. Similarly, the authors found this to be the case with transit fare. This study investigated an interesting feature related to gasoline price elasticity and transit fare elasticity of ridership; however, the analysis is flawed because it controls for the effect of transit supply on metro ridership but not for the fact that transit supply is endogenous. External validity also remains an issue.

Storchmann, Karl, "The impact of fuel taxes on public transport—an empirical assessment for Germany." *Transport Policy* 8 (1), (2001): 19–28.

Storchmann (2001) analyzed the effect of gasoline taxes in Germany that caused an increase in gasoline price using time-series data from 1980 to 1995. This study analyzed travel by different modes for different purposes. In addition, it analyzed how the fuel tax affected passenger kilometers (measured as the product of number of trips and distance traveled) by different modes of transport and for different purposes. The author used a system of equations and analyzed how demography and German unification in 1991 affected the number of trips taken and the average travel distance for each of the following six travel purposes: work, school, shopping, business, leisure, and holiday. In the estimation of the public transit demand as a function of various internal factors, such as peak passenger kilometers, seats per vehicle, average peak seat load factor during peak period and average speed during peak period, Storchmann (2001) found that cross-price elasticity of transit demand varies by purpose of trip, with 0.202 for work-related trips, 0.121 for school trips, 0.045 for leisure trips, 0.031 for shopping trips, and 0.016 for holiday trips. Thus, a large amount of variation in U.S. transit ridership elasticity with respect to the price of gasoline found in earlier studies may be explained by the residents' purposes for using public transit, but this kind of information has been typically omitted in regression analysis due to limited data availability. The analysis in this study showed that an increase in gasoline price resulting from a fuel tax increase had a positive impact on peak-hour transit use but not on transit use for leisure or during off-peak hours. The higher marginal cost of providing transit service during peak hours, which resulted from increased peak-hour demand, led to the seemingly counterintuitive prediction that higher fuel taxes could ultimately increase the transit system's budget deficit.

Currie, Graham, and Justin Phung, "Understanding Links Between Transit Ridership and Auto Gas Prices—US and Australian Evidence." Paper presented at the Transportation Research Board 87th Annual Meeting, Washington, DC, January 13-17, 2008.

Currie and Phung (2008) analyzed transit ridership for rail and bus alone and for all modes combined in Melbourne, Brisbane, and Adelaide, Australia. Using monthly time-series data with per capita transit usage as the dependent variable, this study analyzed the effect of gas price and home mortgage interest rates on transit ridership, controlling for seasonal effects but not for transit service supply. The authors found the gas price elasticity of ridership of all modes combined to be 0.22 in Melbourne, 0.22 in Adelaide, and 0.14 Brisbane. For Melbourne, the authors analyzed information on whether the tickets were used for long-distance travel or short-distance travel and found that the elasticity is higher for longer distance travel. This suggests that gas price elasticity of ridership varies by trip length. This study is subject to omitted variables bias, as it does not control for factors pertaining to transit supply, socioeconomics, or demographics. As this study analyzed transit ridership for each city independently, the results from data for one city may not be comparable to results from another city because the cities have different characteristics. Consequently, the results suffer from external validity.

Stover, Victor W., and C. H. Christine Bae, "Impact of Gasoline Prices on Transit Ridership in Washington State." *Transportation Research Record*, Vol. 2217, (2011): 11-18.

The authors analyzed gasoline price elasticity of aggregate transit ridership for transit agencies in 11 counties in Washington State from January 2004 to November 2008 using a time-series estimation method. This time-series estimation is useful in providing evidence of the effect of gasoline price changes on transit ridership for transit agencies in these specific counties. To estimate the more general effect—the average effect of gasoline price change on ridership—the authors used a panel estimation technique for these counties over the same time frame for aggregate transit ridership.

Using a log-log model, the paper estimates the effect of gasoline prices on aggregate transit and controls for internal factors—fare and supply measured as vehicle revenue hours—and external factors—population and unemployment rate. It also controls for seasonal variation using season dummies instead of month dummies, which accounts for seasonal variation in ridership demand. The authors estimated a static model—i.e., the contemporaneous effect of gasoline price and transit fare change on transit ridership—which is a shortcoming of the paper. Given the nature of the data used (time-series and panel), it would have been possible to estimate long-term effects. A shortcoming is that the paper does not present the estimates obtained from panel data analysis. In the time-series analysis, the authors find that population has a positive effect on ridership, which is as expected. In addition, they find a positive effect of unemployment rate on ridership, which they explain may be due to the difficulty of maintaining a vehicle without a source of income. In turn, this makes riding public transit more favorable for the unemployed.

Although this paper controls for both external and internal factors, it only controls for a subset of external and internal factors that affect ridership and omits some key variables that are likely to affect ridership, for example, frequency of transit service, household income, or the quantity or quality of local highways. The analysis in this paper may be useful for guiding transit agencies in Washington State but may be not generalized, as the data is from only one state.

Nowak, William P., and Ian Savage, "The cross elasticity between gasoline prices and transit use: Evidence from Chicago." *Transit Policy*, 29, (2013): 38-45.

This paper analyzed the cross-price elasticity between the price of gasoline and transit ridership in the Chicago metropolitan area using monthly data from January 1999 to December 2010 for city and suburban buses, city heavy rail, and suburban commuter rail. A log-log model was used to examine whether the cross-price elasticity of transit ridership is different among the gasoline price ranges of under $3, between $3 and $3.99, and over $4. To take seasonal variation into account the authors 12th differenced the data (i.e., subtracted from the log of each variable for each month the value of the log of that variable in that same month from the previous year). They also controlled for transit fare, unemployment rate, proportion of weekdays in the month, and a dummy variable indicating a leap-year February. They found evidence suggesting that elasticities are different for different ranges of gasoline price. They found that the unemployment rate has a negative effect on ridership, while transit supply has a positive effect. With their model, transit fare has a mixed effect on ridership for the different modes. For example, it has a positive significant effect for city bus, but for the other modes the effect is negative.

Blanchard, Christopher. *The Impact of Rising Gasoline Prices on US Public Transit Ridership*, Masters Thesis, Duke University; (2009).

Blanchard (2009) analyzed the impact of increasing fuel prices on public transit ridership in the United States. This study used panel data on ridership (measured as unlinked passenger trips) on four transit modes—commuter rail, heavy rail, light rail, and bus—from 218 U.S. cities between 2002 and 2008. The simple OLS regression models included variables for the log of transit supply, log of current price of gasoline and lagged prices of gasoline, monthly dummy variables, and year dummy variables. Current gasoline price and lagged gasoline prices were used to analyze, respectively, instantaneous and long-run effects of the change in gasoline price. Elasticity was found to be the highest for light rail, with a peak of 0.507 in Dallas, while ranging from -0.103 to 0.507 for other cities. Other modes had similar results, with ranges from -0.012 to 0.213 for commuter rail, to 0.137 to 0.377 for heavy rail, and 0.047 to 0.121 for bus. The counterintuitive negative elasticity estimates for commuter and light rail may indicate a bias in the estimate due to the omission of important factors pertaining to regional geography, the metropolitan economy, population characteristics, transit characteristics, and highway system characteristics that can change over time and affect ridership. The study does not include dummy variables for each city (i.e., fixed effects for cities), which could have controlled for some of the time-invariant, otherwise unobserved characteristics.

Blanchard's study (2009) also shows variation in elasticity by city population size. For example, the estimated elasticity for bus services was 0.09 for cities with a population greater than 2 million and 0.08 for cities with a population between 0.5 and 2 million. The variation in cross-price elasticity by mode and by city size may also be explained by the share of captive riders, availability of alternative modes, and high parking costs in the CBD. In particular, Blanchard, as well as Maley and Weinberger (2009), explains that higher elasticity for commuter trains is related to its service of relatively long-distance trips, a mode shift that would lead to higher gasoline cost savings.

APPENDIX B: RETAIL GASOLINE PRICE AND UNLINKED PASSENGER TRIPS FOR BUS

In this appendix, four sets of graphs present the relationship between gasoline prices and ridership of bus, commuter rail, light rail, heavy rail, and all of these models combined, for ten UAs. Specifically, the graphs show the relationship of unlinked passenger trips (i.e., ridership) with three different gasoline prices—regular, midgrade, and premium—as well as the average of these three prices (labeled “all”) on ridership.

It is important to note that all UAs do not have all modes of transit service. The first set of graphs shows the relationship between gasoline prices and bus ridership for ten UAs. The second set shows the relationship between gasoline prices and commuter rail in seven UAs. The third and fourth sets show the same for light rail in nine UAs and for heavy rail in seven UAs, respectively.

These graphs show that there is no clear pattern in the relationship between transit ridership and gasoline prices alone. Therefore, regression analysis that accounts for impact of other factors on ridership is necessary.

Figure 7. Retail Gasoline Price and Unlinked Passenger Trips for Bus: Ten Other Urbanized Areas

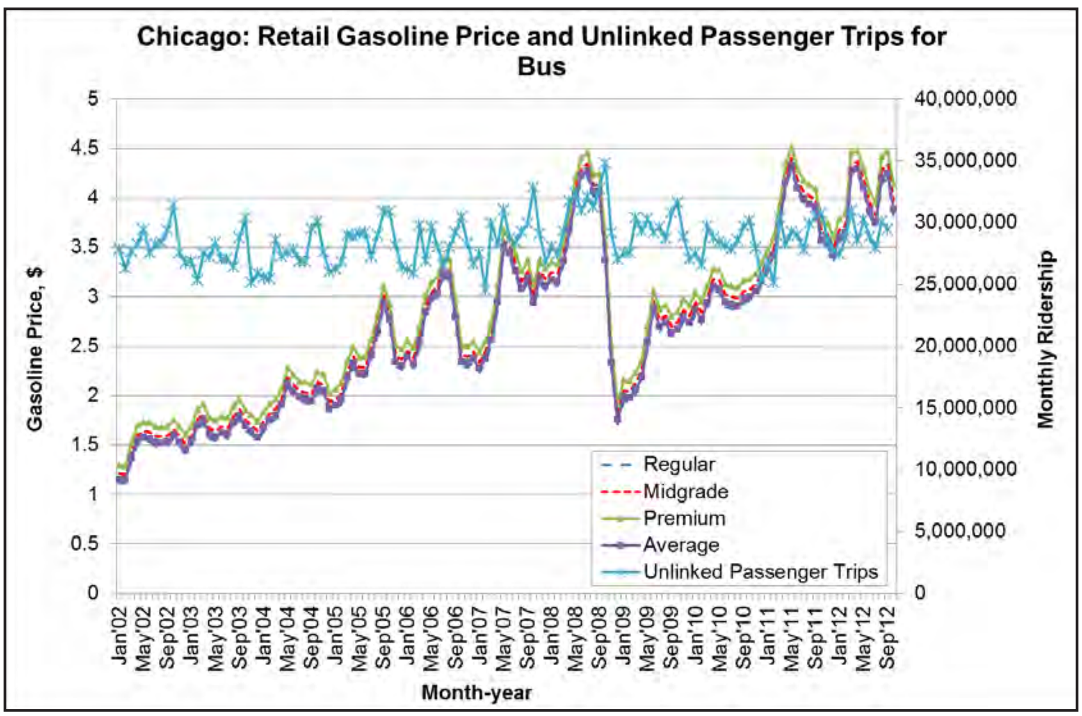
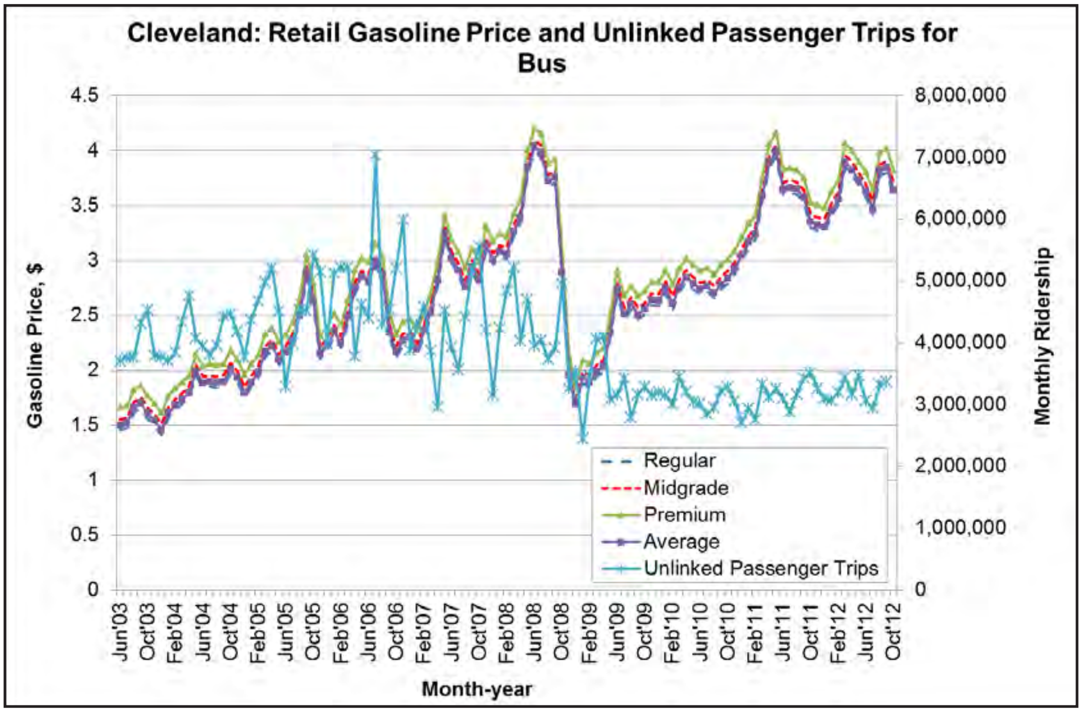


Figure 7 (Continued)



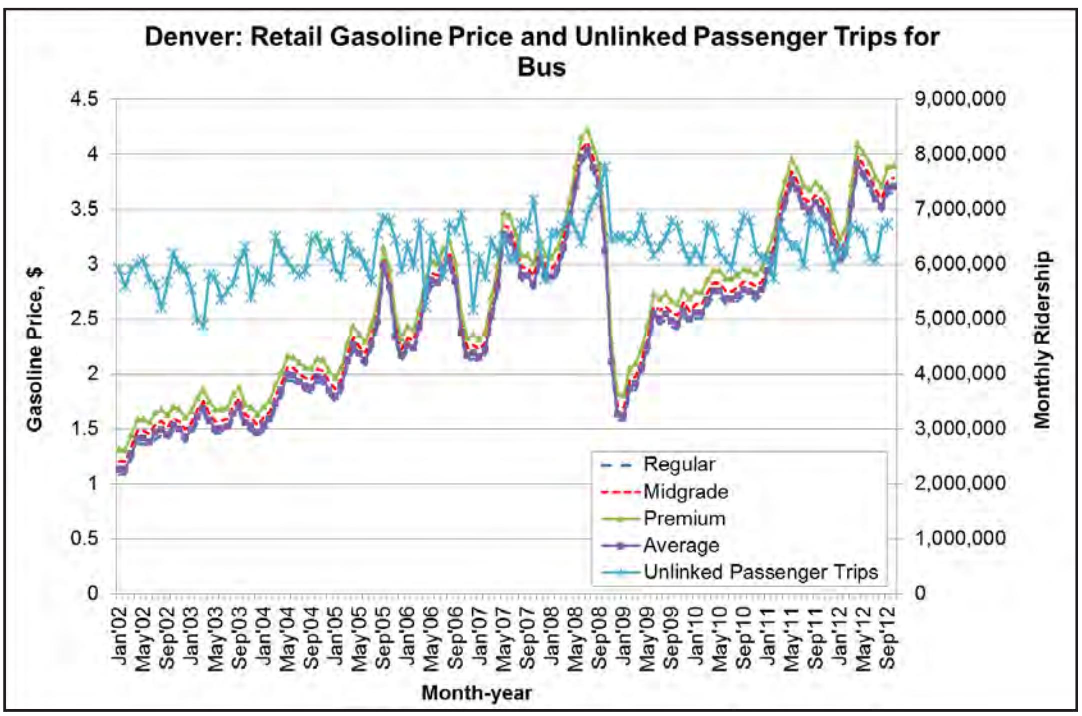
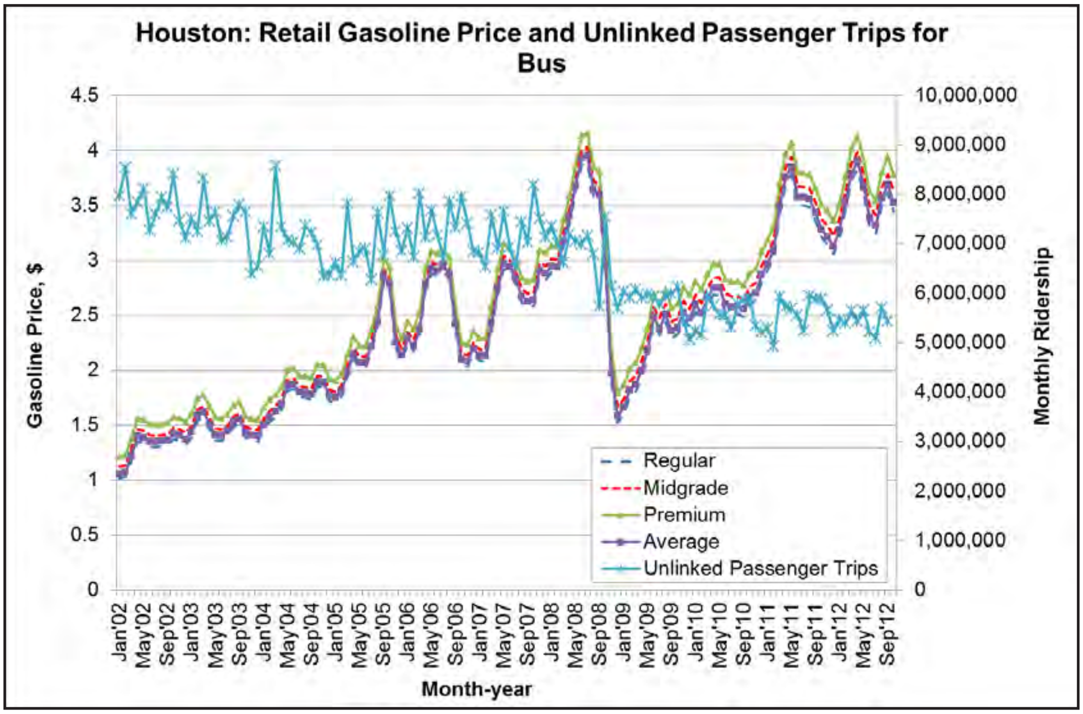


Figure 7 (Continued)



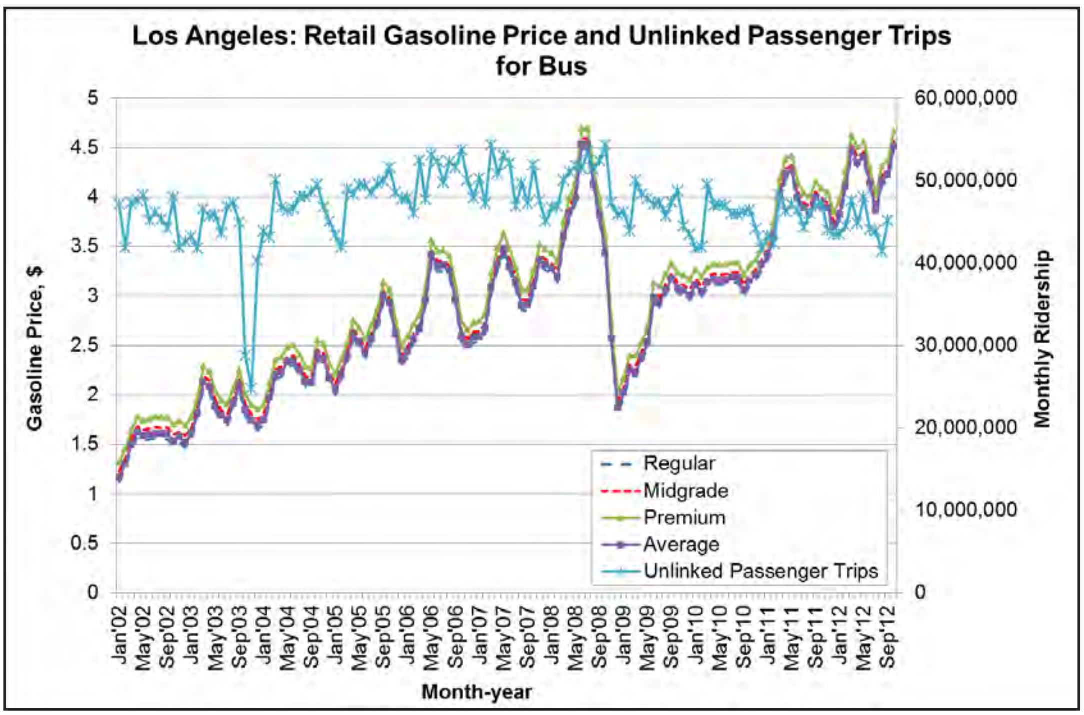
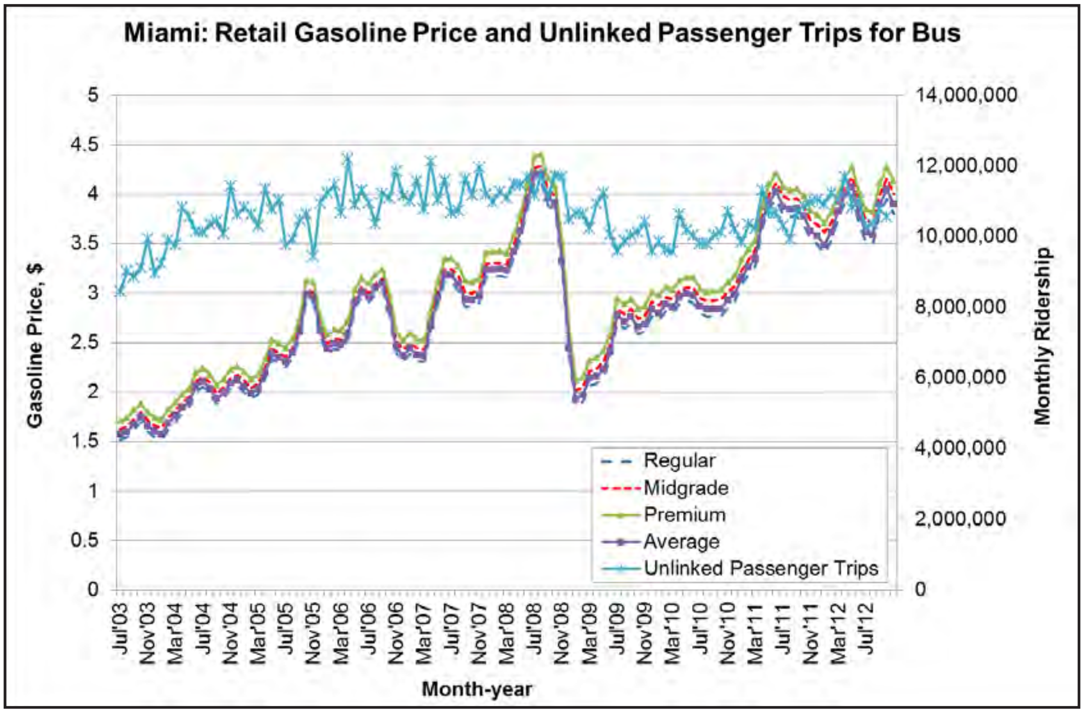


Figure 7 (Continued)



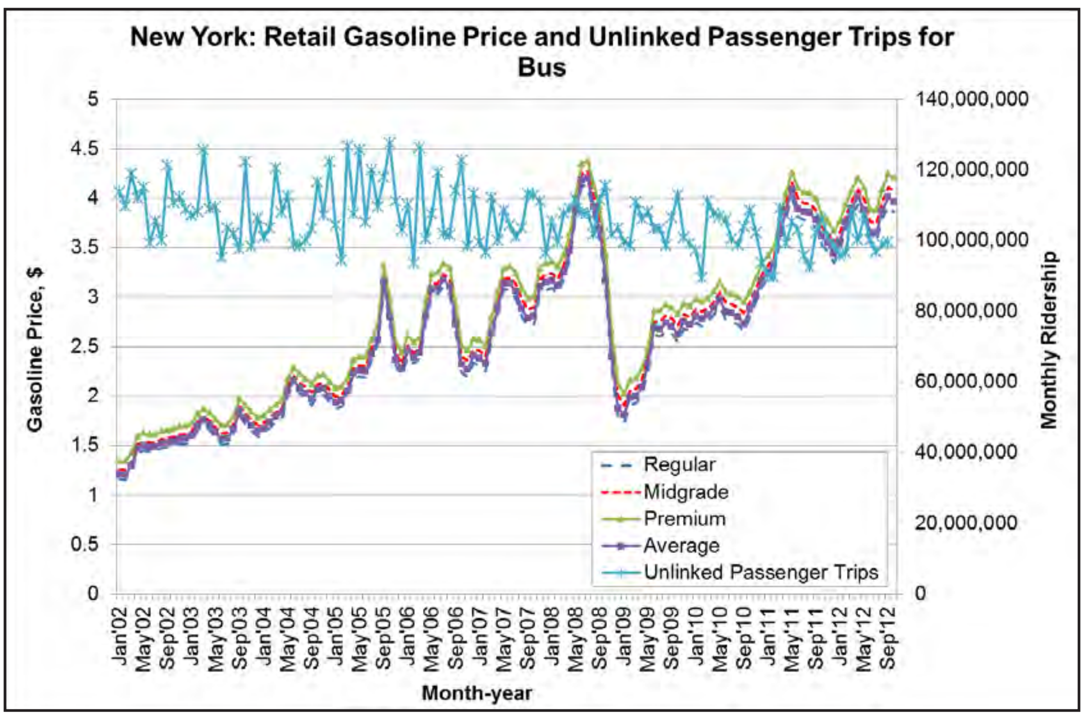
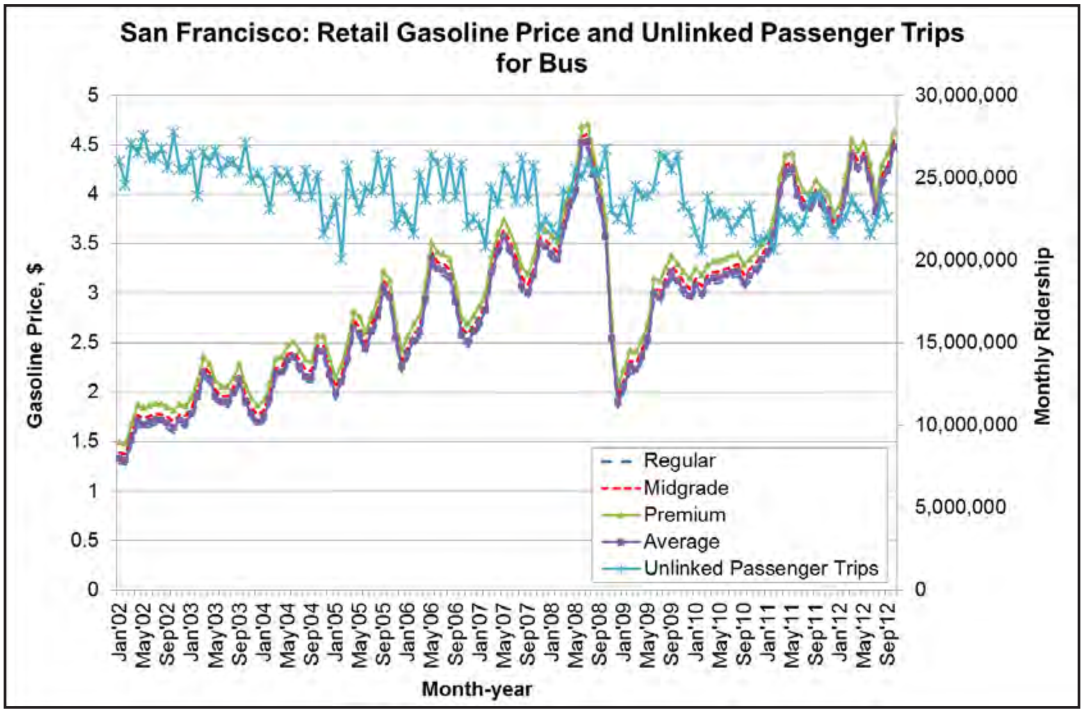


Figure 7 (Continued)



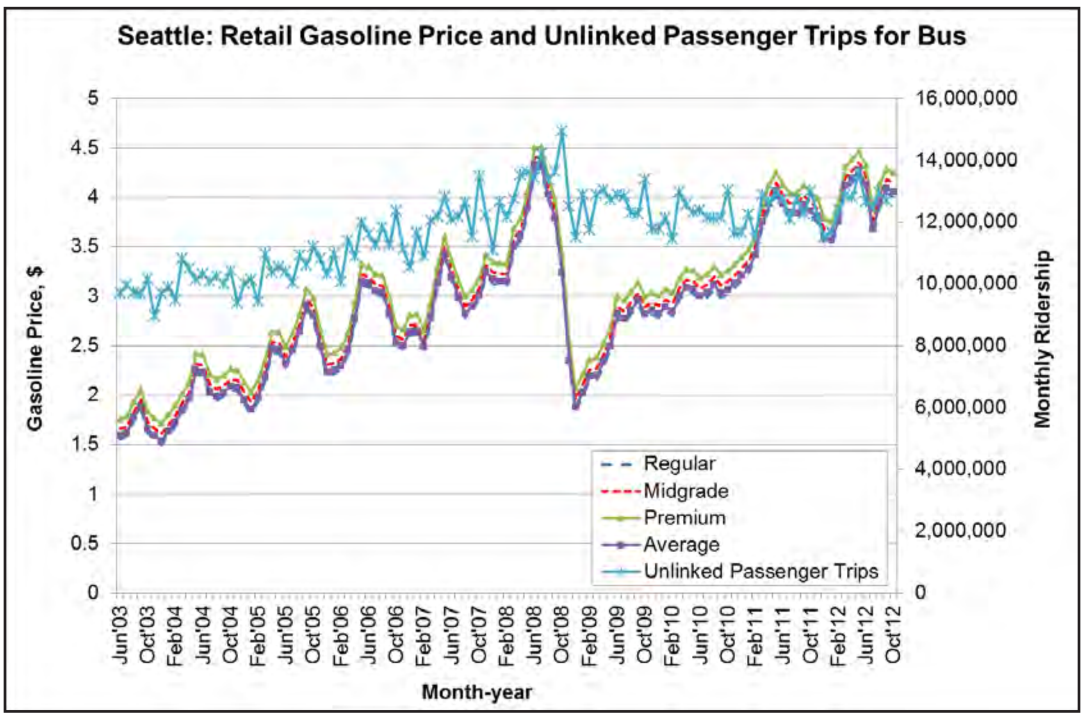
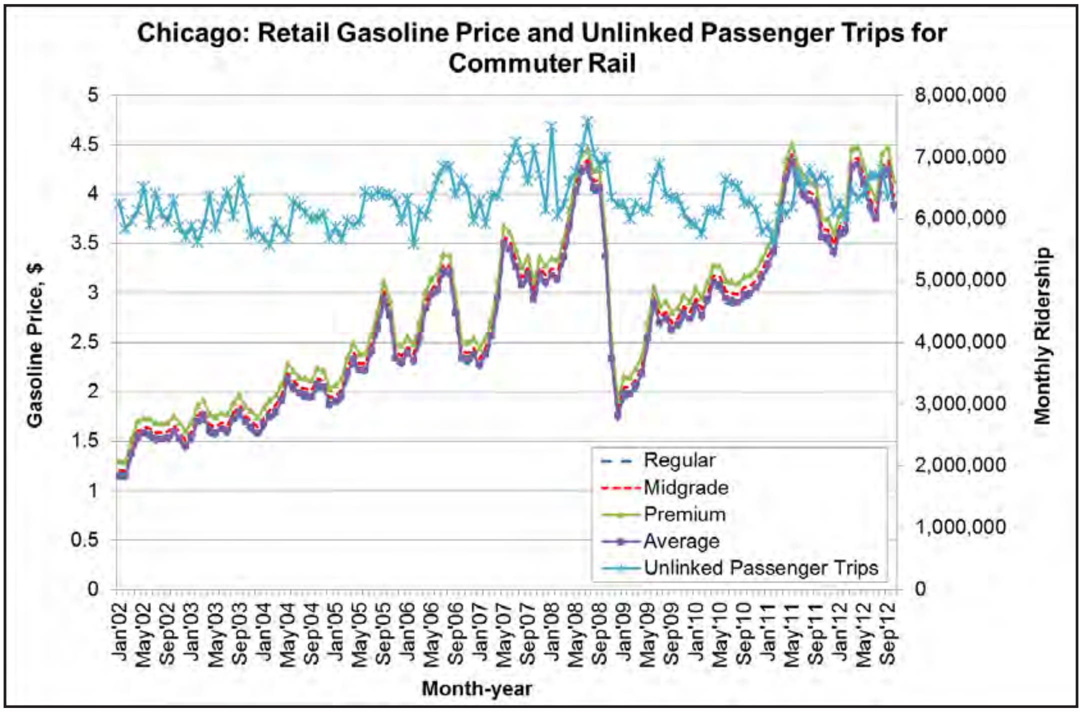


Figure 8. Retail Gasoline Price and Unlinked Passenger Trips for Commuter Rail: Six Other Urbanized Areas



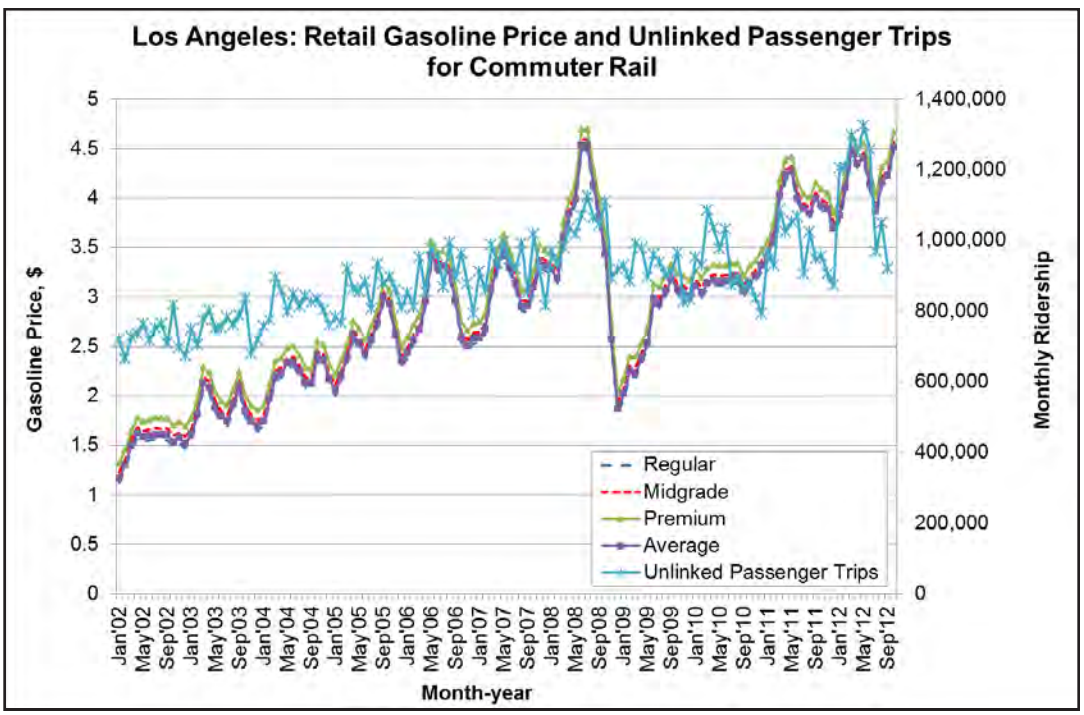
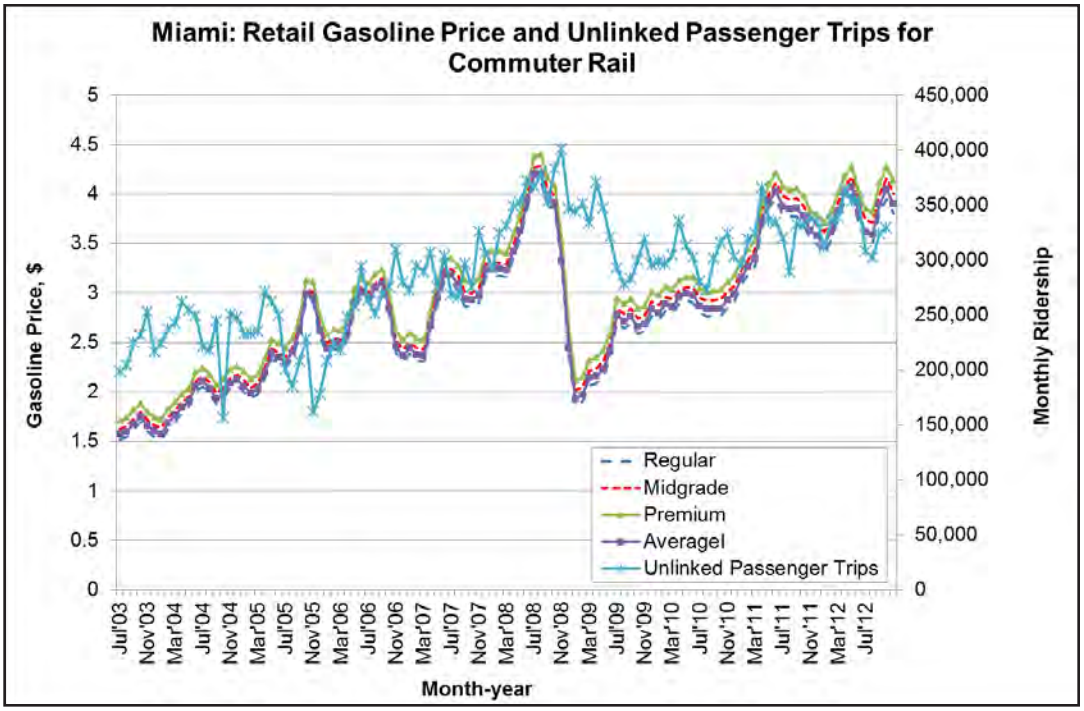


Figure 8 (Continued)



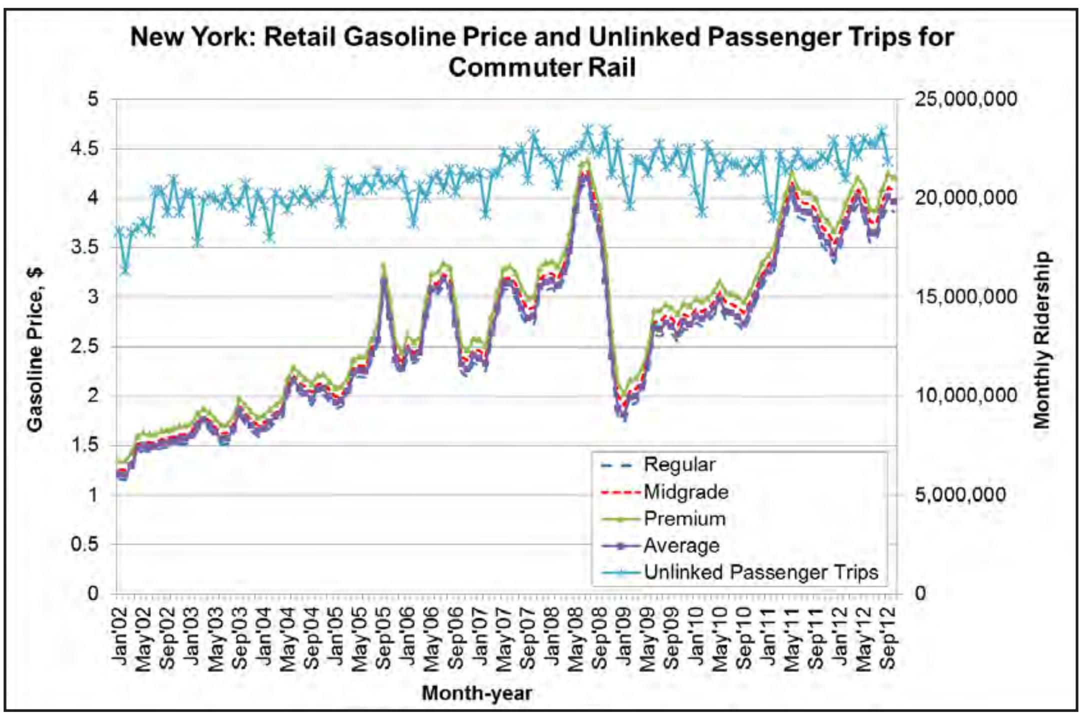
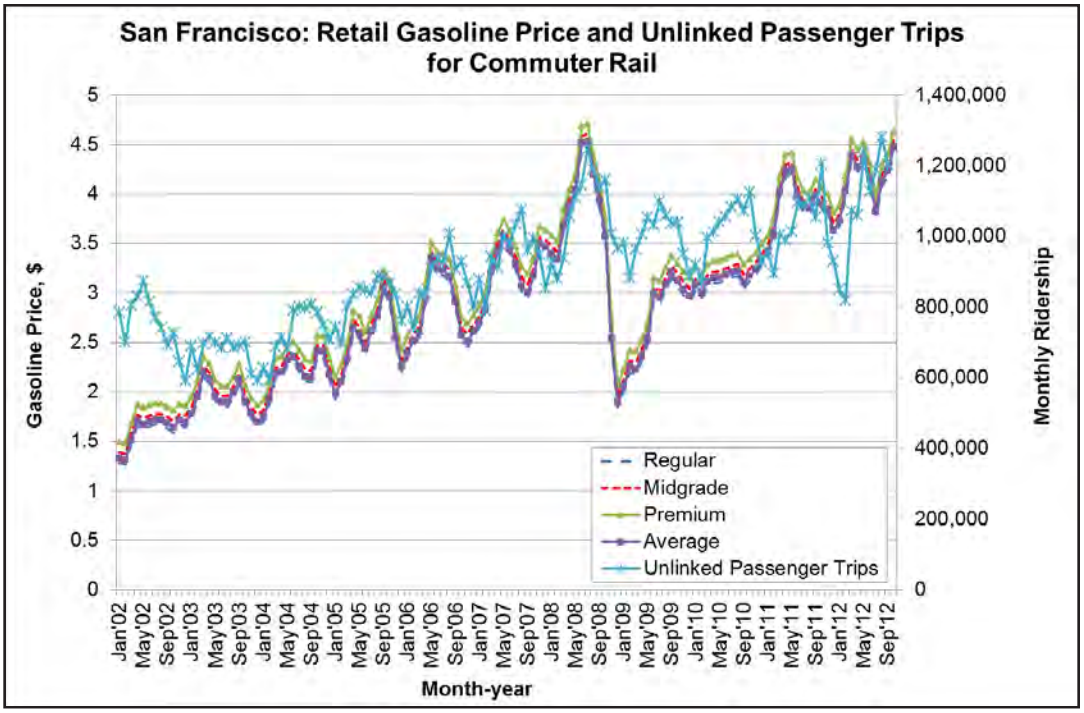


Figure 8 (Continued)



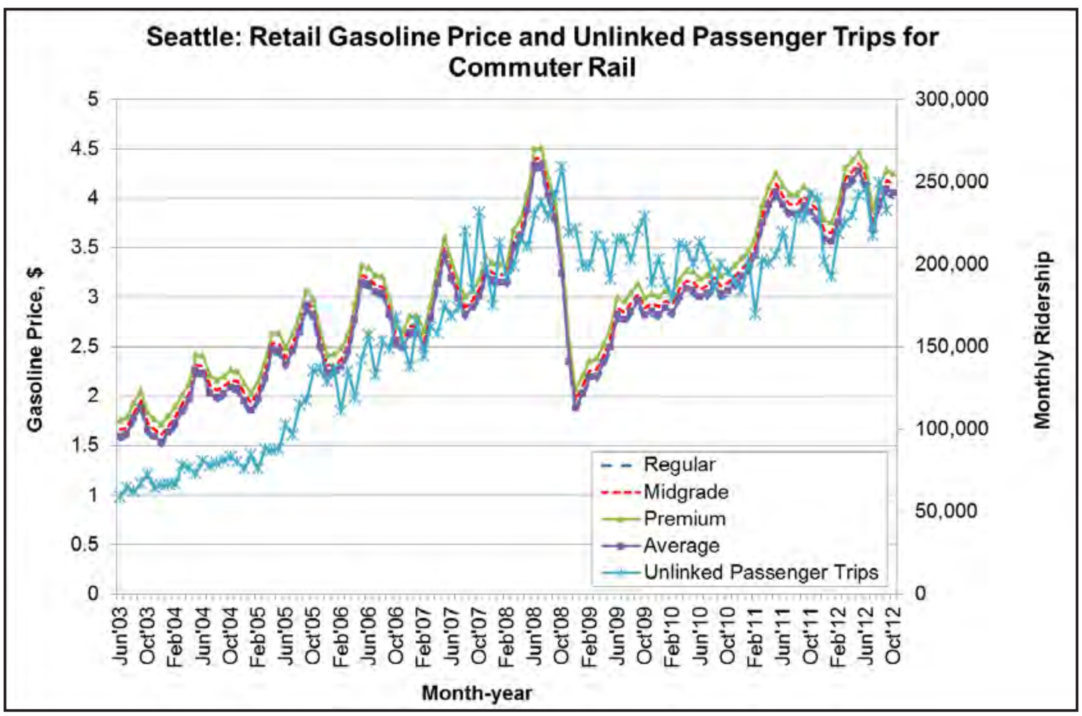
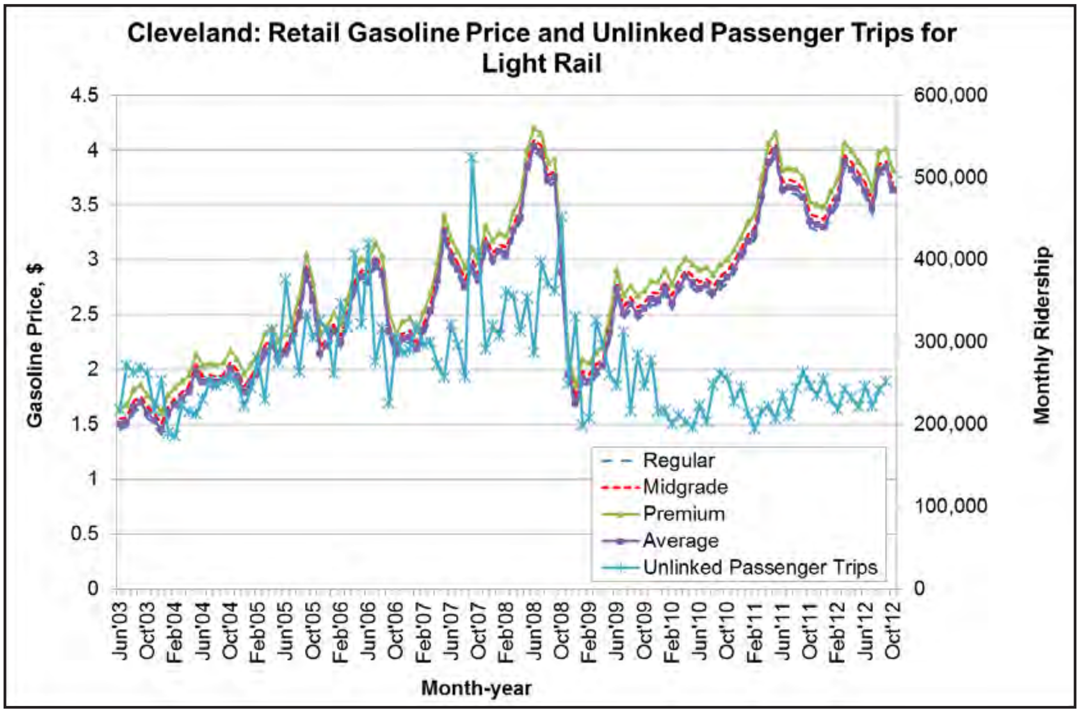


Figure 9. Retail Gasoline Price and Unlinked Passenger Trips for Light Rail: Eight Other Urbanized Areas



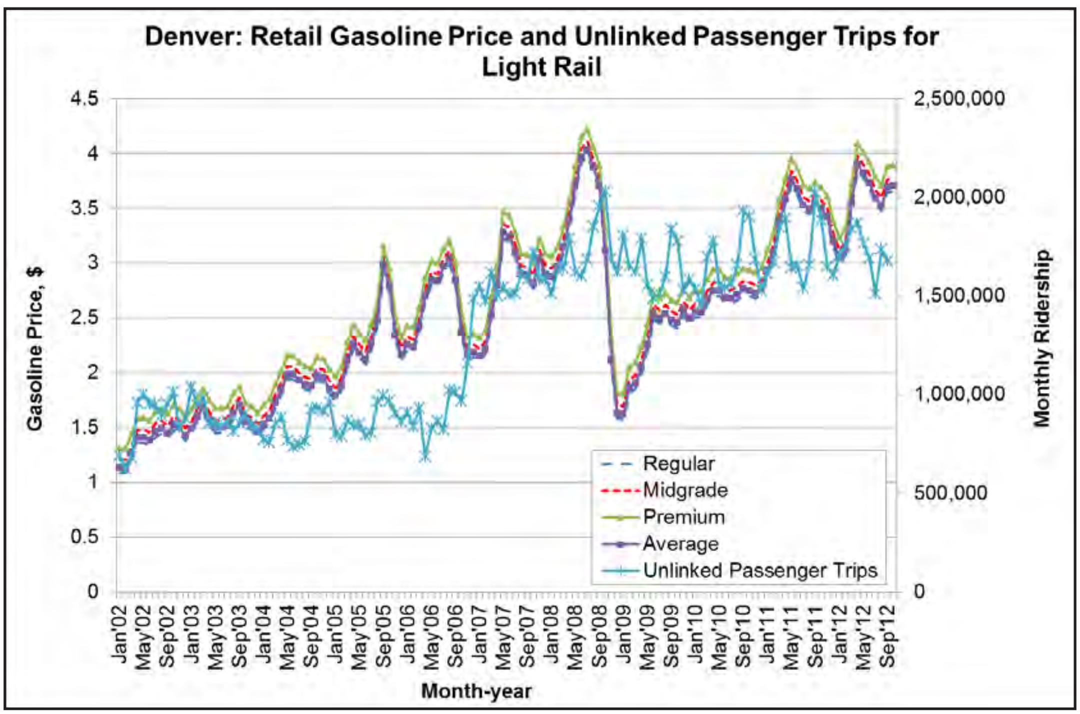
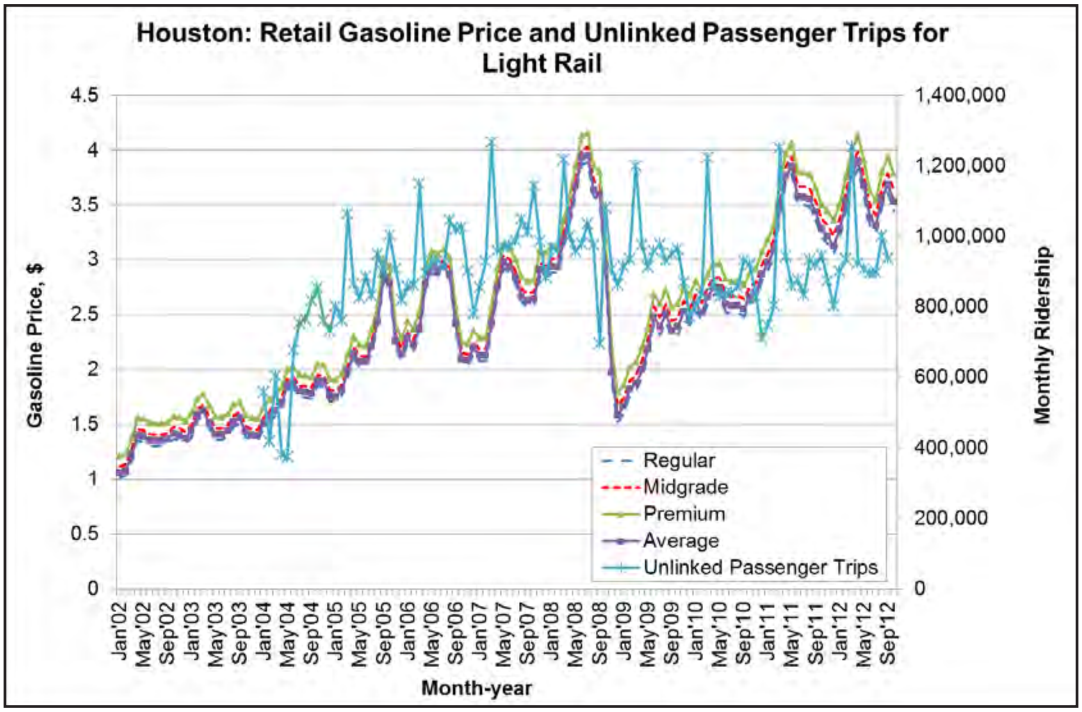


Figure 9 (Continued)



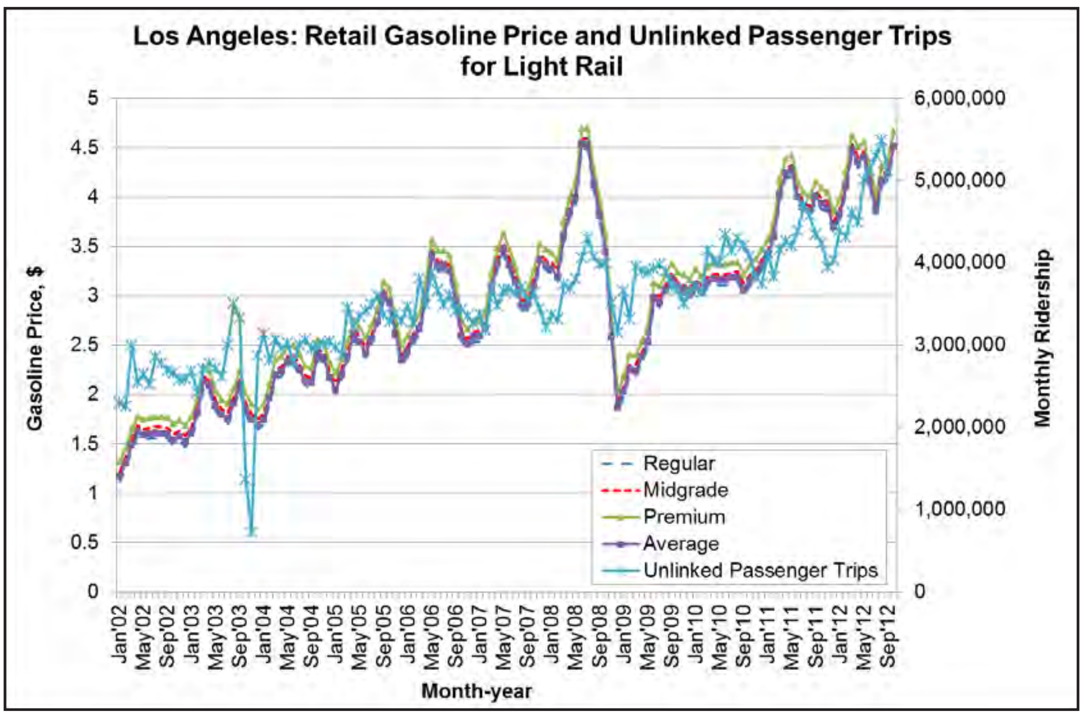
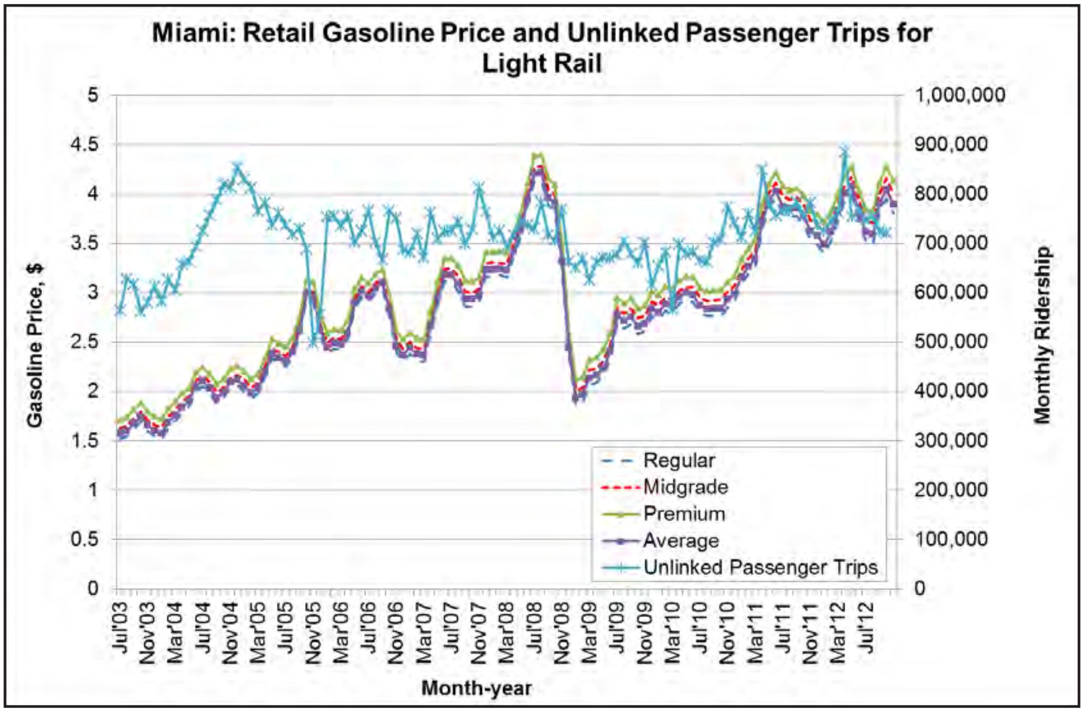


Figure 9 (Continued)



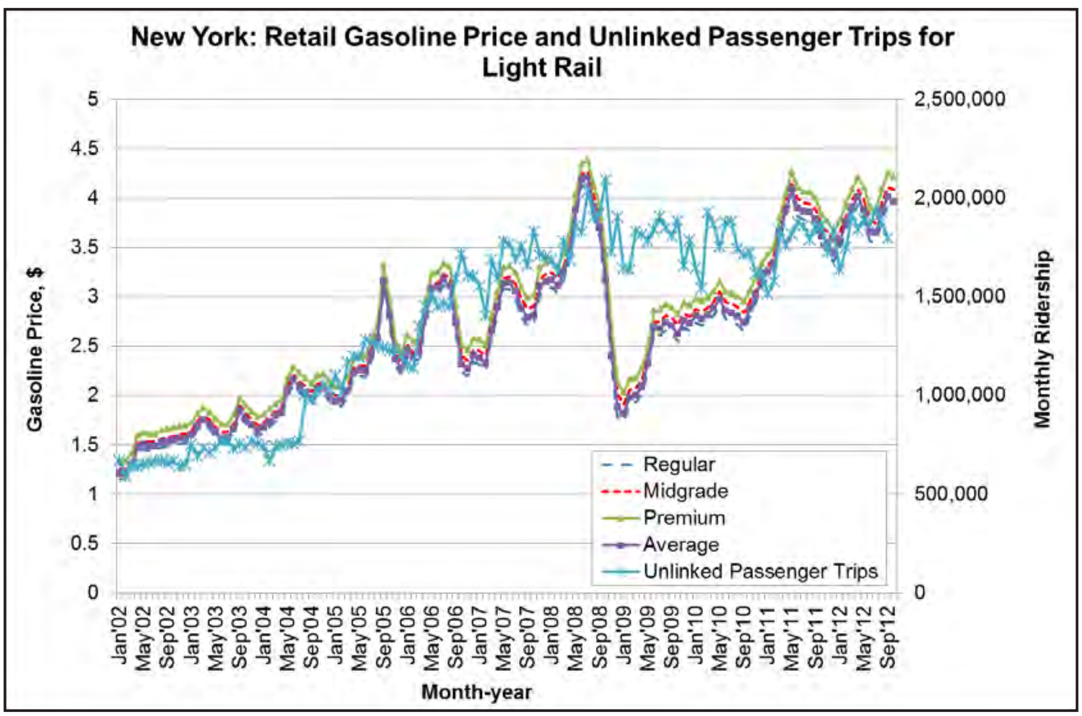
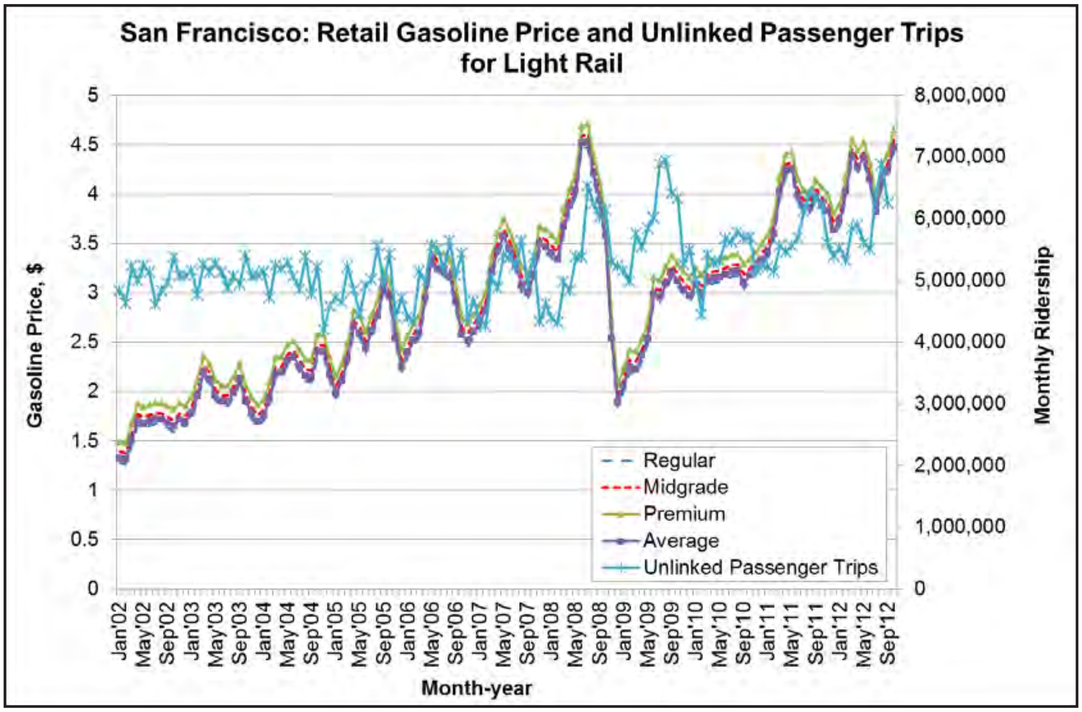


Figure 9 (Continued)



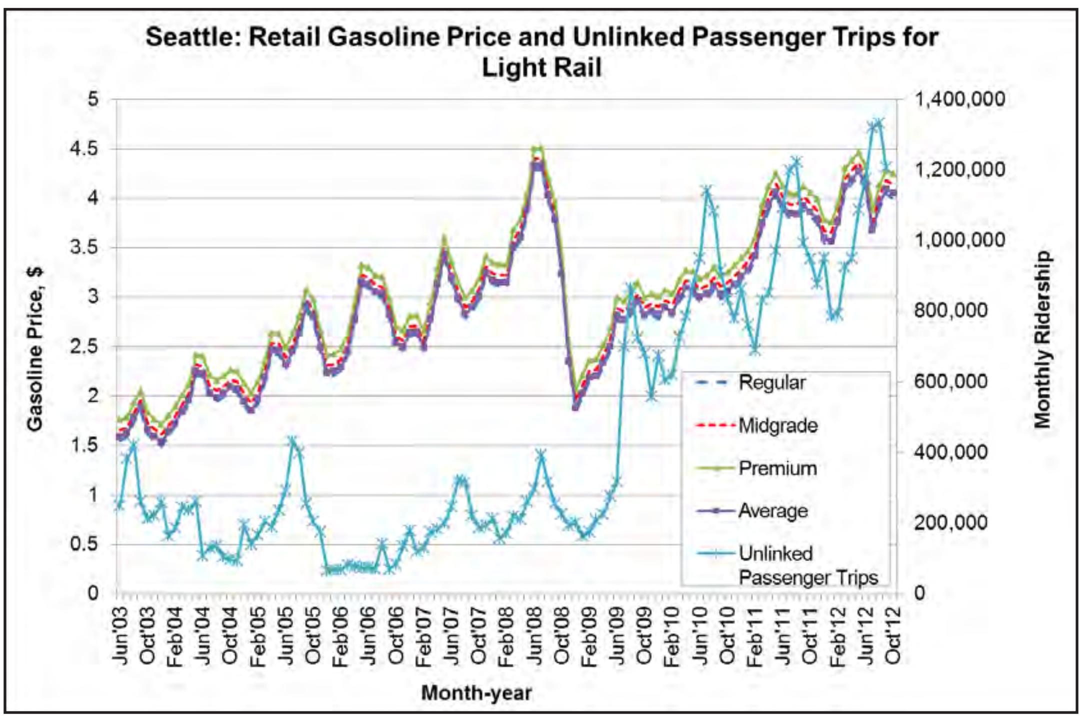
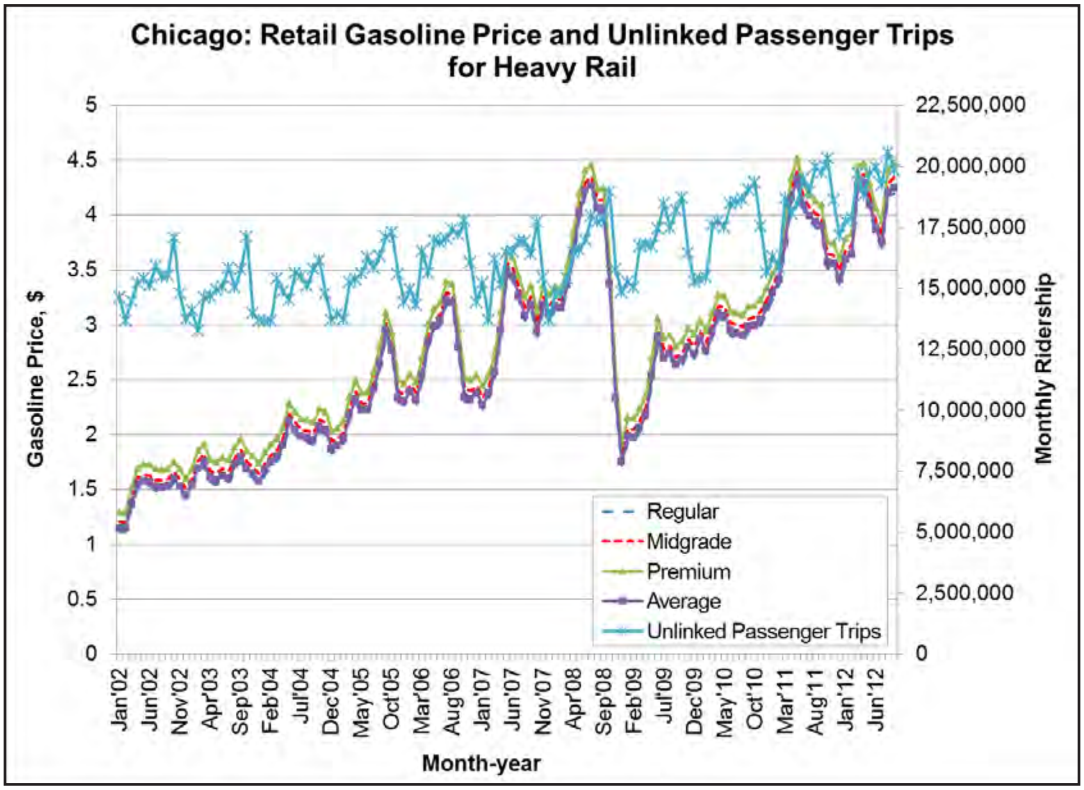


Figure 10. Retail Gasoline Price and Unlinked Passenger Trips for Heavy Rail: Six Other Urbanized Areas



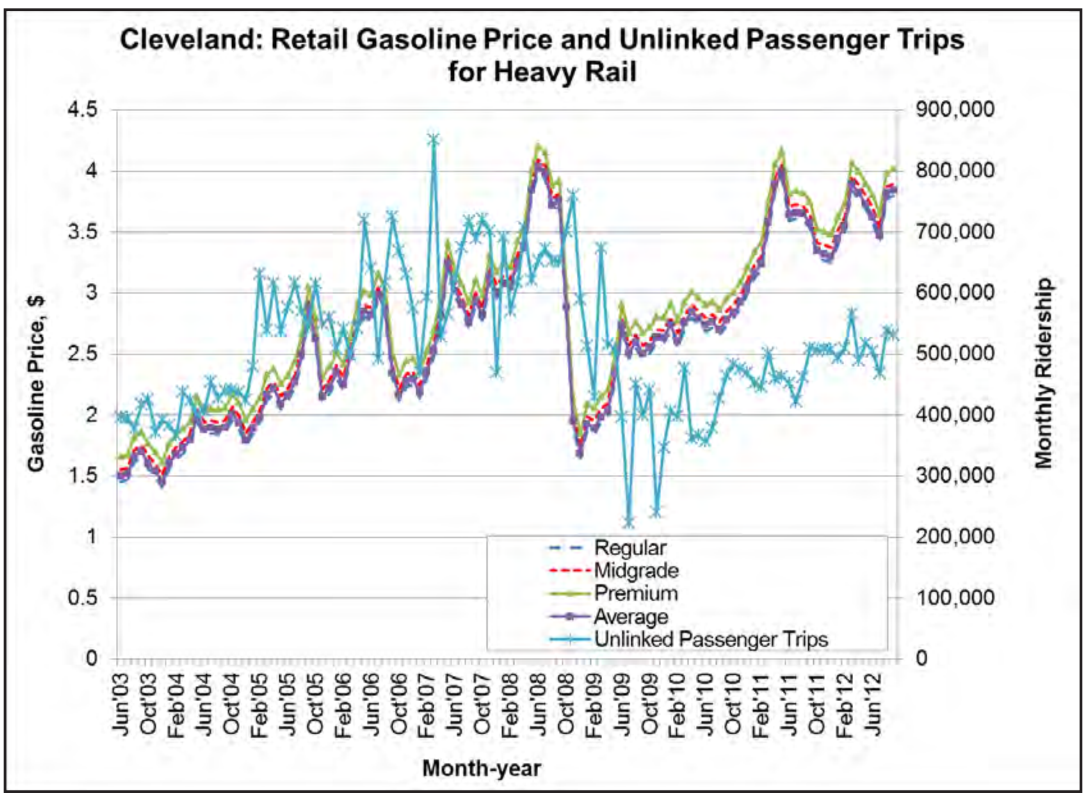
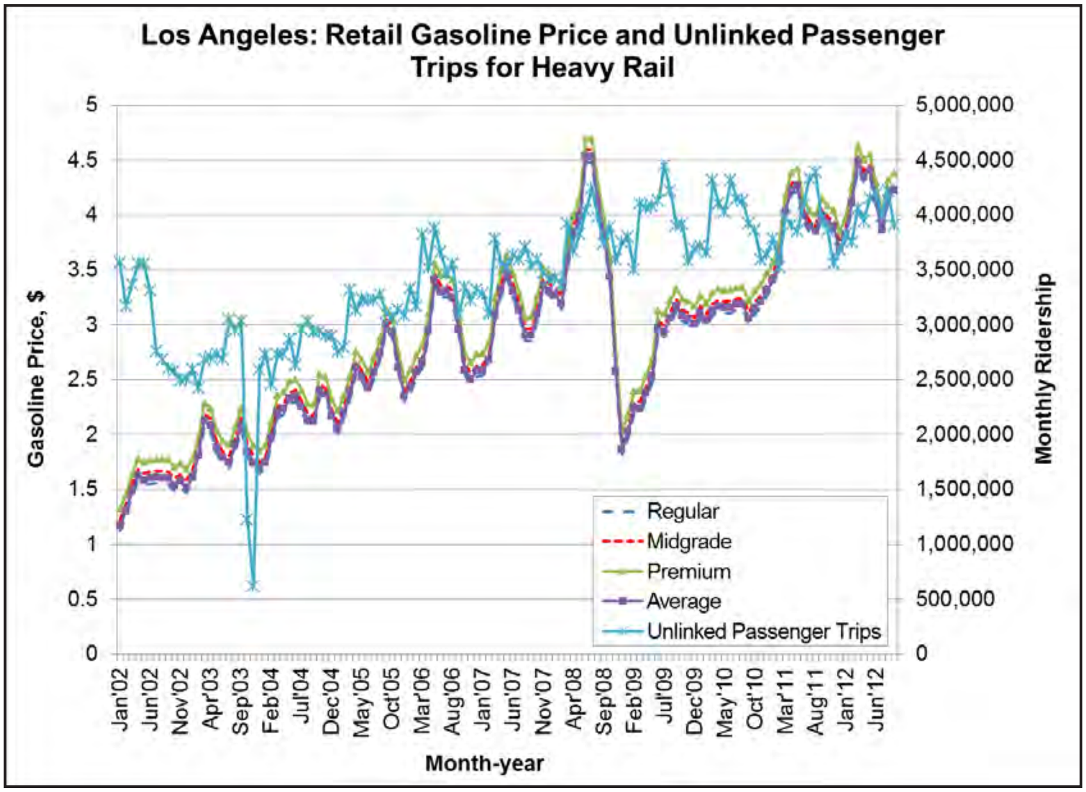


Figure 10 (Continued)



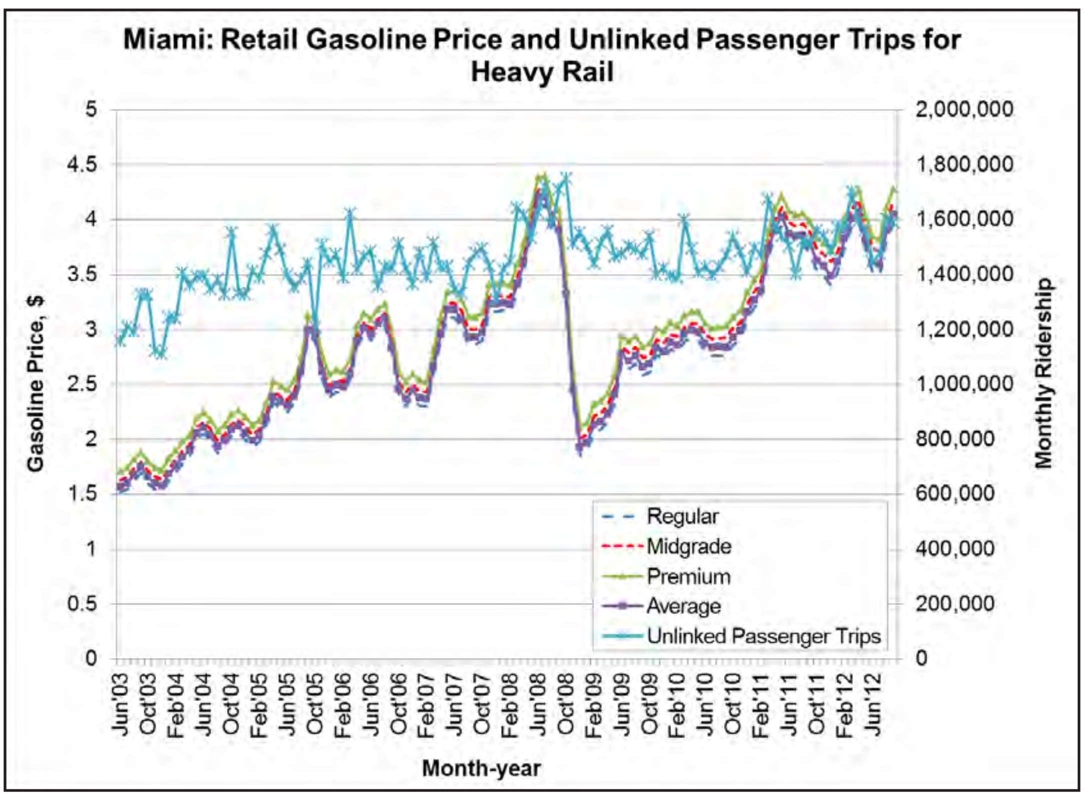
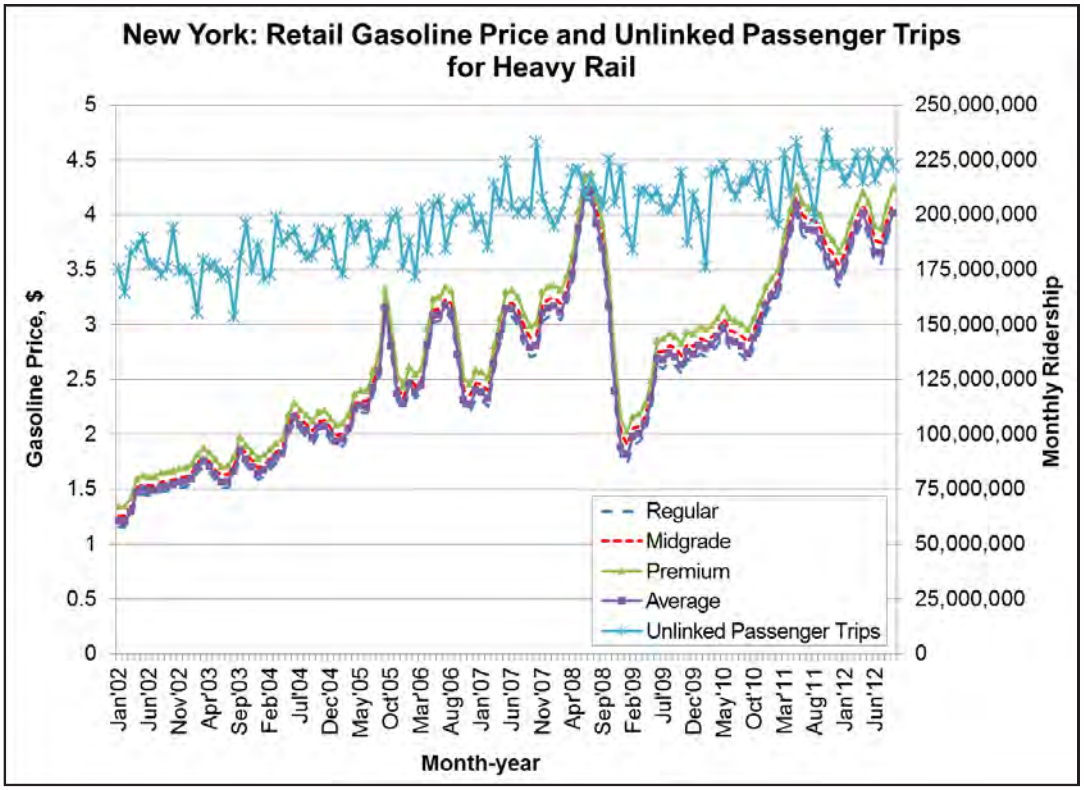
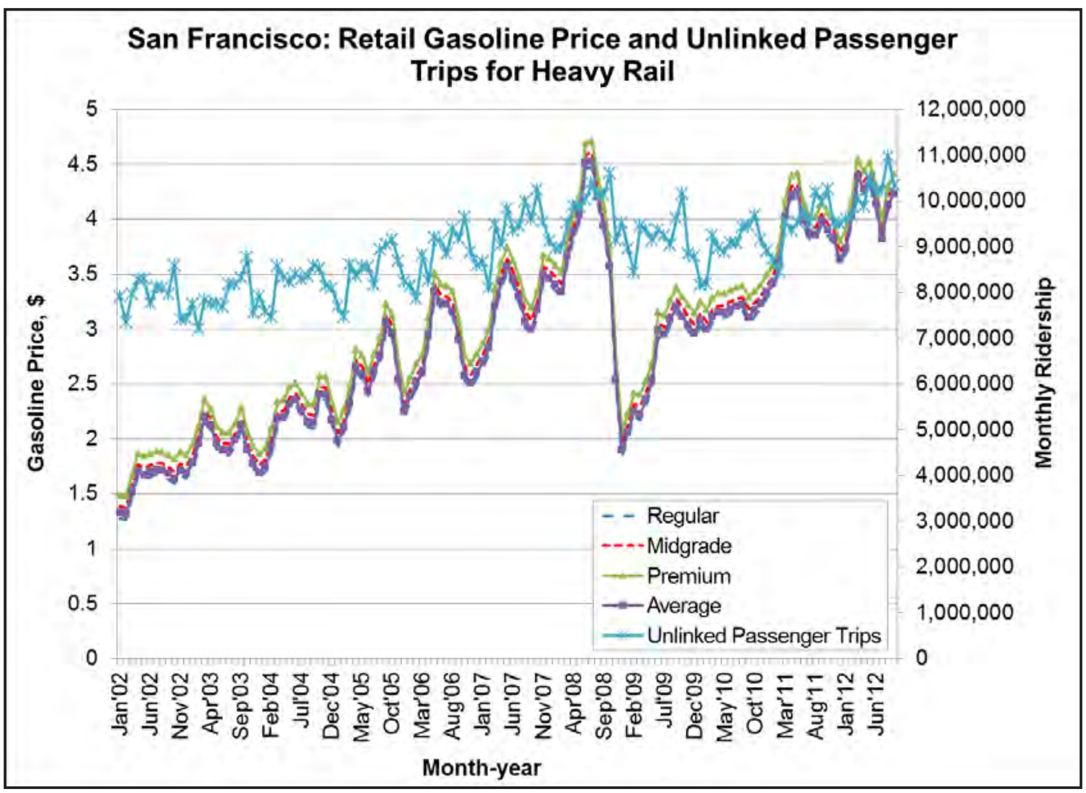


Figure 10 (Continued)





APPENDIX C: CORRELATION USING DATA FROM ALL YEARS

While Chapter IV provided correlation of the independent variables that were considered for regression specifications using the data only from 2007, this appendix provides correlations using the pooled data set of all years from 2002 to 2010. In the following tables, correlation greater than 0.7 is shown in bold type, and correlation greater than 0.4 and smaller than 0.7 is shown in italic type. A threshold value of 0.7 is typically used as an indicator for the point at which collinearity possibly begins to cause severe distortion of model estimation and subsequent prediction, while a value of 0.4 is a more restrictive less commonly used indicator value (Dormann et al. 2007). Overall, there is little difference in correlation values between the pooled data sets and the 2007 data set. Among pairs of variables that have a correlation value larger than 0.4 in both data sets, a difference in correlation values larger than 0.1 occurs in only three pairs for bus (VRH and frequency of service, frequency of service and total population, fare and percent household with no vehicles), only one pair for commuter rail (fare and naturalized citizens), and only one pair for all modes combined (fare and frequency of service).

Table 20. Correlations using All-Year Data

(a) Bus

Commuter Rail	**[1]**	**[2]**	**[3]**	**[4]**	**[5]**	**[6]**	**[7]**	**[8]**	**[9]**	**[10]**
Log of average of regular midgrade and premium gas price [1]	1.000									
Log of fare [2]	0.221	1.000								
Log of vehicle revenue hours [3]	-0.050	*0.575*	1.000							
Log of frequency of service [4]	0.045	*0.444*	**0.890**	1.000						
Log of total population [5]	-0.064	**0.789**	**0.775**	*0.600*	1.000					
Log of naturalized citizen [6]	0.002	*0.698*	*0.524*	*0.489*	**0.898**	1.000				
Log of federal highway miles [7]	-0.103	0.181	-0.073	0.005	0.307	*0.496*	1.000			
Log of mean household income [8]	0.323	*0.411*	0.357	*0.591*	0.117	0.125	0.029	1.000		
Unemployment rate [9]	0.266	0.191	0.093	0.151	0.098	0.145	-0.077	-0.044	1.000	
Percent of households with no vehicle [10]	-0.052	*0.510*	*0.646*	*0.640*	*0.603*	*0.546*	-0.084	0.313	-0.062	1.000

(b) Commuter Rail

Commuter Rail	**[1]**	**[2]**	**[3]**	**[4]**	**[5]**	**[6]**	**[7]**	**[8]**	**[9]**	**[10]**
Log of average of regular midgrade and premium gas price [1]	1.000									
Log of fare [2]	0.221	1.000								
Log of vehicle revenue hours [3]	-0.050	*0.575*	1.000							
Log of frequency of service [4]	0.045	*0.444*	**0.890**	1.000						
Log of total population [5]	-0.064	**0.789**	**0.775**	*0.600*	1.000					
Log of naturalized citizen [6]	0.002	*0.698*	*0.524*	*0.489*	**0.898**	1.000				
Log of federal highway miles [7]	-0.103	0.181	-0.073	0.005	0.307	*0.496*	1.000			
Log of mean household income [8]	0.323	*0.411*	0.357	*0.591*	0.117	0.125	0.029	1.000		
Unemployment rate [9]	0.266	0.191	0.093	0.151	0.098	0.145	-0.077	-0.044	1.000	
Percent of households with no vehicle [10]	-0.052	*0.510*	*0.646*	*0.640*	*0.603*	*0.546*	-0.084	0.313	-0.062	1.000

Table 20, Continued

(c) Light Rail

Light Rail	**[1]**	**[2]**	**[3]**	**[4]**	**[5]**	**[6]**	**[7]**	**[8]**	**[9]**	**[10]**
Log of average of regular midgrade and premium gas price [1]	1.000									
Log of fare [2]	0.273	1.000								
Log of vehicle revenue hours [3]	0.160	0.256	1.000							
Log of frequency of service [4]	-0.009	-0.037	*0.577*	1.000						
Log of total population [5]	0.060	-0.037	0.357	-0.033	1.000					
Log of naturalized citizen [6]	0.089	-0.071	*0.436*	-0.013	**0.965**	1.000				
Log of federal highway miles [7]	-0.075	-0.369	-0.068	-0.210	0.197	0.252	1.000			
Log of mean household income [8]	0.349	0.323	**0.700**	0.227	*0.496*	*0.629*	-0.026	1.000		
Unemployment rate [9]	0.237	0.194	0.037	-0.180	0.022	0.010	-0.098	-0.099	1.000	
Percent of households with no vehicle [10]	0.006	0.262	0.173	-0.293	*0.676*	*0.595*	-0.115	0.378	0.018	1.000

(d) Heavy Rail

Heavy Rail	**[1]**	**[2]**	**[3]**	**[4]**	**[5]**	**[6]**	**[7]**	**[8]**	**[9]**	**[10]**
Log of average of regular midgrade and premium gas price [1]	1.000									
Log of fare [2]	0.239	1.000								
Log of vehicle revenue hours [3]	0.021	0.379	1.000							
Log of frequency of service [4]	0.000	0.287	**0.903**	1.000						
Log of total population [5]	-0.016	-0.094	*0.643*	**0.810**	1.000					
Log of naturalized citizen [6]	0.043	0.097	*0.570*	**0.795**	**0.920**	1.000				
Log of federal highway miles [7]	-0.138	0.082	-0.067	-0.047	0.134	0.224	1.000			
Log of mean household income [8]	0.257	*0.629*	*0.638*	**0.718**	*0.511*	*0.603*	-0.006	1.000		
Unemployment rate [9]	0.268	0.121	-0.177	-0.269	-0.170	-0.169	-0.176	-0.208	1.000	
Percent of households with no vehicle [10]	-0.040	0.097	*0.681*	*0.694*	*0.587*	*0.493*	-0.141	0.367	-0.148	1.000

Table 20, Continued

(e) All Modes (Aggregate)

All Modes	**[1]**	**[2]**	**[3]**	**[4]**	**[5]**	**[6]**	**[7]**	**[8]**	**[9]**	**[10]**
Log of average of regular midgrade and premium gas price [1]	1.000									
Log of fare [2]	0.384	1.000								
Log of vehicle revenue hours [3]	0.074	*0.540*	1.000							
Log of frequency of service [4]	0.072	*0.622*	**0.903**	1.000						
Log of total population [5]	0.054	0.317	**0.902**	**0.786**	1.000					
Log of naturalized citizen [6]	0.104	0.307	**0.857**	**0.765**	**0.944**	1.000				
Log of federal highway miles [7]	-0.168	-0.358	0.071	0.075	0.185	0.227	1.000			
Log of mean household income [8]	0.334	*0.603*	*0.588*	*0.576*	*0.448*	*0.511*	-0.091	1.000		
Unemployment rate [9]	0.232	0.316	-0.028	0.006	0.040	0.046	-0.106	-0.128	1.000	
Percent of households with no vehicle [10]	0.022	*0.532*	**0.742**	**0.728**	*0.596*	*0.525*	-0.145	0.389	-0.048	1.000

APPENDIX D: MORE PARSIMONIOUS INSTRUMENTAL VARIABLES (IV) SPECIFICATIONS

An analysis of the more parsimonious specifications was conducted for all modes except bus by dropping variables that turned out to be statistically insignificant. In these models, the following variables were dropped from the specifications shown in Section 5: frequency of service for light rail, total population for both heavy rail cases, federal highway miles for both heavy rail cases and aggregate transit, mean household income for commuter rail, and unemployment rate for heavy rail without the New York UA. The following tables show results from the first and second stages of the more parsimonious Instrumental Variables specifications. Estimated coefficients for gasoline prices were very similar between these results and the results shown in Section 5.

Table 21. Results from the First Stage of the Instrumental Variables Model

	(1)	(2)	(3)	(4-1)	(4-2)	(5)
Variables	Bus	CR	LR	HR w/o NY	HR w/ NY	Transit
Log of total number of employees	-0.0776***	-0.0701	0.158***	-0.445***	-0.233	-0.0150
(full time+part-time/2)	(0.0191)	(0.205)	(0.0557)	(0.110)	(0.426)	(0.0321)
Log of total fleet	0.552***	-0.0708***	0.514***	3.050***	6.440***	0.494***
(seating+standing capacity)	(0.0296)	(0.0162)	(0.0192)	(0.417)	(1.516)	(0.0537)
Log of total fund available to	0.0547***	-0.177***	0.179***	0.174***	0.701***	0.0340**
transit agencies	(0.00988)	(0.0541)	(0.0296)	(0.0313)	(0.123)	(0.0165)
Seasonal effects (month dummies)	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Urbanized area fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,126	777	892	669	789	1,132
R-squared	0.529	0.533	0.715	0.444	0.346	0.280
Number of Urbanized Areas	10	7	9	6	7	10
F-stat	123.56	7.63	278.25	41.7	21.61	29.56
P-value	0.000	0	0.000	0.000	0.000	0.000

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 22. Results from the Second Stage of the Most Parsimonious Instrumental Variables Model

	(1)	(2)	(3)	(4-1)	(4-2)	(5)
Variables	**Bus**	**CR**	**LR**	**HR w/o NY**	**HR w/ NY**	**V. Transit**
Log of monthly gasoline price	**0.0617*** **	0.0689	0.0225	-0.0485	-0.0259	**0.0577** **
	(0.0223)	(0.0462)	(0.0498)	(0.0597)	(0.0457)	**(0.0255)**
Log of fare	**-0.220*** **	**-0.494*** **	**-0.156*** **	**-0.236*** **	**-0.180*** **	**-0.323*** **
	(0.0212)	**(0.0919)**	**(0.0351)**	**(0.0432)**	**(0.0324)**	**(0.0301)**
Log of vehicle revenue hours	**0.407*** **	-0.0581	**0.933*** **	**-0.284*** **	0.0349	**0.559*** **
	(0.0519)	(0.125)	**(0.0247)**	**(0.0862)**	(0.0235)	**(0.0705)**
Log of frequency of service	**0.0772*** **					
	(0.0207)					
Log of total population	**0.811*** **	**8.287*** **				**0.867*** **
	(0.164)	**(1.178)**				**(0.198)**
Log of federal highway miles	**0.0670*** **	**-0.0977*** **	**-0.0865*** **			
	(0.0102)	**(0.0215)**	**(0.0213)**			
Log of mean household income						
Unemployment rate (%)	**0.0331*** **		**0.0222** **		**0.0315*** **	**0.0287*** **
	(0.00433)		**(0.0102)**		**(0.00929)**	**(0.00522)**
Households with no vehicle (%)	**-0.0396*** **		**-0.111*** **			
	(0.0109)		**(0.0258)**			
Constant	-2.874	**-114.9*** **	**6.33*** **	**17.65*** **	**14.93*** **	-4.18
	(2.411)	**(17.7)**	**(0.292)**	**(0.919)**	**(0.254)**	(2.803)
Seasonal effects (month dummies)	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Urbanized area fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,126	777	892	549	669	1,132
R-squared	0.504	0.594	0.829	0.264	0.408	0.259
Number of urbanized areas	10	7	9	5	6	10

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

APPENDIX E: IMPACT OF GASOLINE PRICES AND OTHER FACTORS ON TRANSIT RIDERSHIP IN FIVE TRANSIT SYSTEMS: A TIME-SERIES ANALYSIS

PURPOSE OF STUDY

As a complimentary study of the panel data analysis, this part of the study uses time-series analysis to examine the elasticity of ridership with respect to gasoline price and the effects of particular “events” that took place specifically in areas of five selected transit systems, while controlling for seasonal changes on transit ridership as well as other main influential factors.

SELECTION OF TRANSIT SYSTEMS, DATA, AND DATA SOURCES

Some factors that could influence transit ridership, such as policy and operation changes and events, are not available in the NTD, so the research team directly contacted 68 agencies in the 10 urbanized areas covered by the panel data analysis. Of these, only five agencies provided the comprehensive information the team was seeking; thus, this time-series study focuses on those five transit systems. They include Broward County Transit Division, FL; Fort Wayne Citilink, IN; Los Angeles Metrolink, CA; Orange County Transit Authority, CA; and South Florida Regional Transportation Authority, FL (Table 23).

The requested information included the addition or loss of a major employer in a service coverage area for each transit agency; factors and conditions external to transit agencies reconfiguration of route structure; significant increase or decrease in fare costs; major transportation infrastructure development (new park and rides, transfer facilities, interchanges); series of high-profile attacks on, or murders of, transit users; rapid growth of the student population at a major university or college; major events (expos, inaugurations, large festivals); an especially cold, snowy winter or hot summer; natural disasters (earthquakes, hurricanes, etc.); strikes by employees; introduction of new transfer policy; a merger of transit systems; and changes in policy (e.g., universities negotiating free rides for students).

The specific events that took place in each of the five transit systems are shown in Table 23.

Table 23. Information of Events and Mode of Service for the Five Transit Systems

Transit System	State	NTD ID	Mode of Service	Event	Date
Fort Wayne Citilink	IN	5044	Motor bus	• Downtown Citiloop service started	2004
				• Snow emergency	02/2007
				• CampusLink offered a free shuttle service between university campuses	2009
Orange County Transit Authority	CA	9036	Motor bus	• Operated 91 express lanes	2003
Broward County	FL	4029	Motor bus	• N/A	N/A

Table 23, Continued

Transit System	State	NTD ID	Mode of Service	Event	Date
Los Angeles Metrolink	CA	9154	Motor bus and rail (heavy and light)	• 91 line opened	2002
				• MTA strike	09/2002 – 10/2002
				• MTA strike	10/2003 – 11/2003
South Florida Regional Transportation Authority	FL	4077	Commuter rail	• New transfer fees for passengers	10/2011

As discussed in the literature review, many factors have been shown to influence transit ridership, including pricing, passenger demand, transit network, and seasonal effects. In this part of the study, monthly transit ridership by mode is used as a dependent variable, while a selected number of explanatory variables are examined as control variables, following the longitudinal analysis approach used in the study by Stover and Bae (2009). Explanatory variables included 1) transit agency operating and financial information collected from the NTD; 2) monthly gasoline prices; 3) aggregated transit ridership of an urbanized area combining all modes available; and 4) agency-specific factors, such as event variables. Table 24 provides a detailed description of variables that were examined in the time-series analysis.

Table 24. Description of Variables for Time-series Analysis

Variable Type	Time Frame	Description	Source
Time	Monthly	Time variable that indicates the time series	Self-created
Ridership	Monthly	Number of passengers served by a specific transport mode provided by a transit agency	NTD
Gasoline price	Monthly	Average of regular gasoline price in the area of a transit agency	U.S. energy information administration
Fare revenue	Annual	Total fare revenue of a specific transport mode provided by a transit agency	NTD
Number of vehicles in maximum operation (voms)	Monthly	Vehicles operated in annual maximum service of a specific transport mode provided a transit agency with do and pt combined	NTD
Vehicle revenue hours (vrh)	Monthly	Vehicles revenue hours of a specific transport mode provided a transit agency with do and pt combined	NTD
Vehicle revenue miles (vrm)	Monthly	Vehicles revenue miles of a specific transport mode provided a transit agency with do and pt combined	NTD
Fleet	Annual	Total number of fleets of a specific transport mode provided a transit agency with standing and seating capacities combined	NTD
Route miles (rm)	Annual	Total route miles	NTD
Bus fare	Annual	Detailed information of bus fare	Transit agencies
Average fare	Annual	Total fare revenue divided by number of passenger trips	Self-created from NTD
Labor size	Monthly	Size of labor force for each area of the transit agencies	Bureau of labor statistics

Table 24, Continued

Variable Type	Time Frame	Description	Source
Unemployment rate	Monthly	Unemployment rate of labor force for each area of the transit agencies	Bureau of labor statistics
Monthly dummies	Monthly	Dummy variables created for months of February to December (January is set as the base)	Self-created
Event dummies	Monthly	Event variables created for each major event that was hypothesized to affect a transport service provided by a transit agency	Self-created
Post event	Monthly	Post event effects of corresponding event variables	Self-created

While average fare is a variable obtained by dividing annual fare revenue by annual ridership for each mode, fare is information obtained directly from transit agencies. The bus fare was recorded annually. Adult fare, student fare, and senior fare information was collected, and only adult fare was incorporated in the model estimation. An event dummy variable is a dummy variable that has a value of 0 before the event and a value of 1 after the event, representing an effect of a specific event on ridership that remains over time. A post-event variable is generated for the time-series analysis, using the function in Equation E-1, which starts with 0 and gradually increases to the value of 1 in twelve months starting from the month of an event. If an estimated coefficient of this post-event variable has a similar value with an opposite sign to a coefficient of an event dummy, a combination of these two represent that the event has an effect on ridership initially, but the effect gradually disappears over 12 months.

$$\text{Post event variable} = 0 \quad \text{if time} < \text{Month\_event}$$
$$= a*(1-e^{(-(\text{time-Month\_event})*(1/12)*(\ln(a)-\ln(a-1)))}) \quad \text{if Month\_event} \le \text{time} \le (\text{Month\_event} + 12)$$

(Equation E-1)

where
time is a time variable in the unit of month
Month_event is the month when the event occurs
a is a coefficient, which is set as 1.1 in this study.

In addition, prior to the regression analysis, all variables except time, monthly dummies, event dummies, and post-event variables, were converted to a log form.

STUDY PROCEDURES

Several tests were performed before running the time-series analysis for each mode-agency-specific regression model. For each model, 1) multicollinearity, 2) heteroskedasticity, 3) serial correlation (autocorrelation), 4) auto-regressive (or lag) selection, and 5) unit root were examined.

Multicollinearity was checked using pairwise correlation between explanatory variables. Highly correlated explanatory variables were not simultaneously included in the same model. In all cases, vehicle revenue hours (VRH) and vehicle revenue miles (VRM) are highly correlated, and only VRH was included in the final specification. *Heteroskedasticity* was checked to see if the error terms have constant variance. When heteroskedasticity is found, it is suggested to use robust regression to obtain more accurate estimates of standard errors, which affects statistical significance of estimated coefficients. *Serial Correlation* is the correlation between a variable and its previous values. In other words, autocorrelation between values of a dependent variable with and without a lag or lags was examined using visual inspection, a Durbin-Watson test, and a Lagrange Multiplier test (Durbin's alternative test in this study). If the Durbin-Watson d-statistic has a value smaller than a critical value, it indicates serial correlation in a dependent variable. A significant Lagrange Multiplier test result is desired to reject the null hypothesis that there is no serial correlation. If serial correlation is detected, it is necessary to select the number of *auto-regressive* (or *lag*) variables of the dependent variable as a way of dealing with serial correlation of the error terms. Too many lags could increase the error in the forecasts, while too few could leave out relevant information. Three information criteria are commonly used to help determine a proper number of lags, and they are: Schwarz's Bayesian information criterion (SBIC), Akaike's information criterion (AIC), and the Hannan and Quinn information criterion (HQIC). These information criteria are reported by the command 'varsoc' in Stata. Specifically, AIC tends to be more accurate with monthly data, HQIC works better for quarterly data on samples over 120, and SBIC works fine with any sample size for quarterly data. A Fuller test is commonly used to test for stationarity. If unit root is detected, one way to deal with stochastic trends (unit root) is by taking the first difference of the variables.

Analysis Results of Time-series Analysis

The model estimation proceeded by subsequently adding/removing explanatory variables based on the R-square and coefficient significance. Variables of interest—gasoline price and event variables—were retained regardless of their statistical significance to examine the statistical significance of their effects on ridership. Monthly dummy variables are treated as control variables as long as they are not highly correlated with any other explanatory variables. If there was no clear time trend detected, time variable was then excluded from the model specifications. In the following regression specifications, the terms AR(1) represents the 1st order autocorrelation), the AR(2) represents the 2nd order autocorrelation in the results tables, and so on.

Results of time-series analysis for all transit agencies are displayed in Table 25. The final models are agency-mode specific and take the following form:

Fort Wayne Citilink (NTDID: 5044) MB

$$\begin{aligned}\ln(\text{mbridership}) &= \beta_0 + \beta_1 \ln(\text{gasprice}) + \beta_2 \text{time} + \gamma_1 \text{newsev1} + \gamma_2 \text{postnewsev1} + \gamma_3 \text{newsev2} \\ &+ \gamma_4 \text{postnewsev2} + \gamma_5 \text{weather} + \gamma_6 \text{postweather} + \beta_3 \ln(\text{vrh}) + \beta_4 \ln(\text{averagefare}) \\ &+ \beta_5 \ln(\text{busfare}) + \beta_6 \ln(\text{rm}) + \gamma_7 \text{feb} + \gamma_8 \text{mar} + \gamma_9 \text{dec} + u + a\end{aligned}$$

Orange County Transit Authority (NTDID: 9036) MB

$$\ln(\text{mbridership}) = \beta_0 + \beta_1 \ln(\text{gasprice}) + \gamma_1 \text{newsev1} + \gamma_2 \text{postnewsev1} + \beta_2 \ln(\text{vrh}) + \beta_3 \ln(\text{averagefare}) + \beta_4 (\text{unemprate}) + \gamma_3 \text{sep} + \gamma_4 \text{oct} + \gamma_5 \text{dec} + u + a$$

Broward County (NTDID: 4029) MB

$$\ln(\text{mbridership}) = \beta_0 + \beta_1 \ln(\text{gasprice}) + \beta_2 \text{time} + \beta_3 \ln(\text{vrh}) + \beta_4 \ln(\text{averagefare}) + \beta_5 \ln(\text{laborsize}) + \beta_6 \ln(\text{rm}) + \beta_7 \ln(\text{voms}) + \gamma_1 \text{mar} + u + a$$

Los Angeles Metrolink (NTDID: 9154) MB and Rail

$$\ln(\text{mbridership}) = \beta_0 + \beta_1 \ln(\text{gasprice}) + \beta_2 \text{time} + \gamma_1 \text{strike1} + \gamma_2 \text{poststrike1} + \gamma_3 \text{strike2} + \gamma_4 \text{poststrike2} + \beta_3 \ln(\text{vrh}) + \beta_4 \ln(\text{averagefare}) + \beta_5 \ln(\text{laborforce}) + \beta_6 \ln(\text{rm}) + \gamma_5 \text{may} + \gamma_6 \text{jun} + \gamma_7 \text{dec} + u + a$$

$$\ln(\text{railridership}) = \beta_0 + \beta_1 \ln(\text{gasprice}) + \beta_2 \text{time} + \gamma_1 \text{strike1} + \gamma_2 \text{poststrike1} + \gamma_3 \text{strike2} + \gamma_4 \text{poststrike2} + \beta_3 \ln(\text{averagefare}) + \beta_3 \ln(\text{vrm}) + \beta_5 \ln(\text{laborforce}) + \gamma_5 \text{oct} + \gamma_6 \text{nov} + \gamma_7 \text{dec} + u + a$$

South Florida Regional Transportation Authority (NTDID: 4077) CR

$$\ln(\text{crridership}) = \beta_0 + \beta_1 \ln(\text{gasprice}) + \beta_2 \text{time} + \gamma_1 \text{newsev1} + \beta_3 \ln(\text{vrh}) + \beta_4 \ln(\text{averagefare}) + \beta_5 \ln(\text{unemprate}) + \beta_6 \ln(\text{fleet}) + \gamma_2 \text{dec} + u + a$$

where

u = unobserved factors,

a = unobserved factors that are constant over time

Table 25 reports multiple results for motor bus and rail services of Los Angeles Metrolink (9154) and commuter rail service of South Florida Regional Transportation Authority (4077), depending on the number of auto-regressive variables examined.

The Durbin-Watson d-statistic of all models reported is below 2, indicating autocorrelation. In addition, the Durbin’s alternative test results of those models are all significant, and the null hypothesis of no serial correlation is rejected.

Table 25. Time-series Results of All Transit Agencies

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Variables	**5044 MB**	**9036 MB**	**4029 MB**	**9154 MB**	**9154 MB**	**9154 Rail**	**9154 Rail**	**4077 CR**	**4077 CR**	**4077 CR**
Regular gasoline price	0.0675***	0.0744***	0.0523***	0.1023***	0.0948***	-0.0062	-0.0093	-0.0470**	-0.0480**	-0.0446**
	(0.0119)	(0.0097)	(0.0163)	(0.0195)	(0.0194)	(0.0340)	(0.0343)	(0.0229)	(0.0225)	(0.0221)
Time	0.0046***		0.0050***	0.0008***	0.0008***	0.0041***	0.0039***	0.0013***	0.0013***	0.0012***
	0.0004		(0.0004)	(0.0003)	(0.0003)	(0.0005)	(0.0006)	(0.0004)	(0.0004)	(0.0004)
Event Variables										
New Service 1 Dummy	0.0364***	-0.0332***						-0.0370*	-0.0346*	-0.0303*
	(0.0063)	(0.0070)						(0.0187)	(0.0185)	(0.0182)
Post New Service 1 Dummy	-0.0221*	0.0146***								
	(0.0126)	(0.0085)								
New Service 2 Dummy	-0.1004***									
	(0.0137)									
Post New Service 2 Dummy	-0.1118***									
	(0.0141)									
Weather Dummy	-0.0863***									
	(0.0178)									
Post Weather Dummy	0.0094									
	(0.0086)									
Strike 1 Dummy				0.0150	0.0106	-0.0726***	-0.0660**			
				(0.0221)	(0.0222)	(0.0238)	(0.0257)			
Post Strike 1 Dummy				-0.0665**	-0.0695**	0.0182	0.0165			
				(0.0272)	(0.0268)	(0.0202)	(0.0197)			
Strike 2 Dummy				-0.0177	0.0109	-0.0932*	-0.0646			
				(0.0340)	(0.0369)	(0.0528)	(0.0546)			
Post Strike 2 Dummy				-0.0183	-0.0471	0.1250**	0.0953			
				(0.0379)	(0.0408)	(0.0593)	(0.0605)			
Autoregressive Variables										
AR(1)	-4.36e-07***	0.0659***			0.0541**		0.0482	0.0720**	0.0350	0.0243

Table 25, Continued

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Variables	**5044 MB**	**9036 MB**	**4029 MB**	**9154 MB**	**9154 MB**	**9154 Rail**	**9154 Rail**	**4077 CR**	**4077 CR**	**4077 CR**
	(1.46e-07)	(0.0246)			(0.0264)		(0.0648)	(0.0306)	(0.0345)	(0.0341)
AR(2)	-3.04e-7**	0.0121							0.0664**	0.0297
	(1.51e-07)	(0.0209)							(0.0317)	(0.0348)
AR(3)	-2.50e-07*	-0.0651*								0.0706**
	(1.47e-07)	(0.0336)								(0.0318)
AR(4)	-5.48e-07***									
	(1.46e-07)									
AR(5)	-4.25e-07**									
	(1.94e-07)									
Control Variables										
Vehicles Revenue Hours	0.0683*	0.6832***	0.2591***	0.6344***	0.6662***			-0.1120***	-0.1093***	-0.1077***
	(0.0345)	(0.0395)	(0.0872)	(0.0566)	(0.0577)			(0.0109)	(0.0109)	(0.0107)
Average Fare	-0.8552***	-0.4234***	-0.5729***	-0.4459***	-0.4175***	-0.4999***	-0.4882***	-0.9689***	-0.9606***	-0.9632***
	(0.0342)	(0.0228)	(0.0472)	(0.0415)	(0.0424)	(0.0703)	(0.0729)	(0.0327)	(0.0324)	(0.0319)
Vehicles Revenue Miles						0.4625***	0.4844***			
						(0.1026)	(0.1080)			
Unemployment Rate		-0.0387***						0.0637***	0.0629***	0.0642***
		(0.0090)						(0.0143)	(0.0142)	(0.0142)
Labor Size			-1.4899***	2.3088***	2.1480***	2.8985***	2.7921***			
			(0.2571)	(0.4391)	(0.4449)	(0.7110)	(0.7311)			
Bus Fare	0.5183***									
	(0.0491)									
Route Miles	0.7637***		-0.0832***	0.4168***	0.3952***					
	(0.2141)		(0.0171)	(0.0659)	(0.0653)					
Number of Fleets								1.1407***	1.1104***	1.0834***
								(0.0717)	(0.0737)	(0.0755)
Vehicles Operated at Maximum Speed			0.2669***							

Table 25, Continued

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Variables	**5044 MB**	**9036 MB**	**4029 MB**	**9154 MB**	**9154 MB**	**9154 Rail**	**9154 Rail**	**4077 CR**	**4077 CR**	**4077 CR**
			(0.0791)							
Feb	0.0156***									
	(0.0046)									
Mar	0.0140**		0.0247***							
	(0.0047)		(0.0064)							
May				0.0205**	0.0191*					
				(0.0099)	(0.0097)					
June				0.0202**	0.0175*					
				(0.0100)	(0.0097)					
Sep		0.0283***								
		(0.0049)								
Oct		0.0120*				-0.0485***	-0.0499***			
		(0.0066)				(0.0178)	(0.0170)			
Nov						-0.0868***	-0.0830***			
						(0.0170)	(0.0167)			
Dec	-0.0079	-0.0432***		-0.0403***	-0.0365***	-0.0880***	-0.0802***	0.0192*	0.0138	0.0168
	(0.0088)	(0.0073)		(0.0100)	(0.0100)	(0.0175)	(0.0183)	(0.0111)	(0.0116)	(0.0110)
Intercept	8.5542***	8.0525***	32.4268***	-29.3705***	-28.1184***	-34.5905***	-34.0292***	10.5134***	10.2241***	10.0265***
	(1.2166)	(0.4435)	(3.3584)	(6.8070)	(6.7751)	(10.7568)	(10.7684)	(0.4181)	(0.4319)	(0.4341)
Observations	114	116	103	119	118	119	118	102	102	102
R-squared	0.9896	0.9795	0.8986	0.9653	0.9671	0.9560	0.9564	0.9778	0.9784	0.9794

Note: Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

DISCUSSION OF TIME-SERIES ANALYSIS RESULTS

Results for variables of interest are discussed first and then results for control variables are briefly discussed. Estimated coefficients for motor bus services provided by the four agencies that offer bus services were found to be positive, ranging from 0.0523 to 0.1023 (Table 25). A closer inspection of the magnitude of coefficients suggest that bus ridership in Los Angeles was more strongly influenced by gasoline price compared to other study areas. However, coefficients for rail services were either statistically insignificant (Los Angeles Metrolink) or negative (South Florida Regional Transportation Authority). These results collectively indicate a large variance that potentially exists in the elasticity of transit ridership to gasoline price among different transit modes and systems.

The time variable was statistically significant with a positive sign except for one model. Although this result implies that ridership would increase in future, controlling for the supply level of transit service and fare level, this is probably not the case.

Table 25 shows mixed results for the effects of event-related variables. For the motor bus service of Fort Wayne Citilink (5044), estimated coefficients indicate that the introduction of the downtown Citiloop service in 2004 had a positive boost effect on overall ridership (0.0364), about 60 percent of which may have dissipated in a year (0.0221). The free shuttle service between university campuses that started in 2009 shows a negative effect on the overall ridership based on the estimated coefficients. In addition, the snow emergency in February 2007 also had a negative effect, which remained even after twelve months. Orange County Transit Authority's (9036) new bus operation on the State Route-91 express lanes shows a small negative impact, controlling for other variables. The estimated coefficients for variables for the first strike in 2002 indicates that the strike had a negative effect on ridership for both bus and rail of Los Angeles Metrolink (9154) and that this negative effect for bus had some time lag to show its full effect. In addition, while the second strike also had a negative effect for rail, ridership recovered after twelve months, as indicated by a positive estimated coefficient for the post-strike 2 dummy. An estimated coefficient of a dummy variable for the new service shows a decrease in ridership for rail service provided by the South Florida Regional Transportation Authority (4077), controlling for other variables.

For motor bus and rail services of Los Angeles Metrolink and commuter rail service of South Florida Regional Transportation Authority, the auto-regressive variables of ridership were found to have an estimated coefficient with a positive sign, ranging from 0.0541 to 0.0720. The magnitude of these coefficients is in the same order of those of gasoline prices. Orange County Transit Authority also has a positive auto-regressive variable of one lag, but the one with three lags has a negative sign. Although auto-regressive variables of up to five lags were included in the model for Fort Wayne Citilink (5044), estimated coefficients were very small and marginal compared to other statistically significant coefficients in the model.

Among the control variables, there is a large variance in a set of explanatory variables that remained in the final specification. Average fare was the only variable that was found statistically significant in all specifications, and it has a negative effect on ridership as expected. Its largest impact was on ridership of rail service provided by the South Florida

Transportation Authority. The elasticity for the rail service by this agency was approximately -0.96, compared to elasticities in the range of -0.8552 to -0.4175 for the other cases.

One of the service supply variables—vehicle revenue hours (VRH) and vehicle revenue miles (VRM)—was found statistically significant in all specifications, as expected. While VRH was a better fit for motor bus services and commuter rail services, VRM was selected for Los Angeles Metrolink rail service. These service supply variables were found to be positively correlated with ridership except for commuter rail of the South Florida Regional Transportation Authority. Elasticity of ridership to VRH varies from -0.1120 to 0.6832. Its positive relationship with bus ridership and its larger magnitude of influence than the other variables are reasonable; the more service provided, the higher the ridership. Surprisingly, VRH had a negative coefficient in the case of rail ridership of South Florida Regional Transportation Authority even after controlling for average fare, employment rate, number of fleets, and a December month dummy.

In all models except motor bus service of Fort Wayne Citilink, one of the two regional economic variables (i.e. unemployment rate and labor size) entered in the models. However, signs of estimated coefficients of these two variables are not consistent across the model specifications. Additional bus fare information obtained directly from transit agencies did not increase the goodness of fit in most cases, being insignificant except for the bus service provided by Fort Wayne Citilink. Average fare showed a negative impact on bus ridership, whereas bus fare had positive effect. Some of the monthly dummy variables, with January as a base month, remained in the final specifications. Overall, February, March, May, June, September, and October had higher bus ridership than January, while December had the lowest bus ridership. July and August had the same level of ridership as January, possibly because of lower student ridership during the summer.

The reason for the mixed results obtained in the longitudinal analysis in this section could be that the service changes included in the analysis were assumed to affect the overall ridershipbut in fact may have affected only part of the transit system, with other factors having a larger effect. The possibility of some econometric problems in this part of analysis is another concern.

ABBREVIATIONS AND ACRONYMS

ACS	American Community Survey
APTA	American Public Transportation Association
CR	Commuter rail
HR	Heavy rail
IV	Instrumental variables
LR	Light rail
NTD	National Transit Database
NY	New York
N/S (n/s)	Not statistically significant
VMT	Vehicle miles traveled
VRH	Vehicle revenue hours
VRM	Vehicle revenue miles
UA	Urbanized areas

ENDNOTES

1. Gasoline price elasticity of transit ridership can be expressed by the equation: e= (dR⁄R)/(dP⁄P) where R and P represent transit ridership and the price of gasoline, respectively. dx represents a change in a variable x.

2. Please refer to Appendix A - Table 17.

3. Bus transit users are less likely to own or have a vehicle available for their trip than are fixed-guideway riders (APTA 2007). Low-income households are more likely to be carless or have inadequate access to private vehicle trips (Pucher and Renne 2003).

4. Please also refer to Appendix A – Table 18.

5. Please also refer to Appendix A – Table 19.

6. Note here that our analysis follows Taylor et al. (2009) by accounting for endogeneity of the supply of services, measured as vehicle revenue hours.

7. Appendix C provides correlations using the pooled data set of all years from 2002 to 2010. Overall, there is not much difference in correlation values between the pooled data sets and the 2007 data set. See Appendix C for more detail.

8. U.S. Department of Transportation Federal Transit Administration, National Transit Database. Ridership of fixed-route (non-demand response) transit in two New York urbanized areas—New York-Newark, NY-NJ-CT and New York, NY-Northeastern NJ—was 3.78 billion, while the national total was 10.25 billion.

9. The analysis was conducted on two data sets with or without New York for all modes. Only the analysis for heavy rail showed significant differences in estimated coefficients.

10. An analysis of the more parsimonious specifications was conducted for all modes except bus by dropping variables that turned out to be statistically insignificant. Results are shown in Appendix D. Estimated coefficients for gasoline prices were very similar between these two sets of models.

11. For commuter rail, a specification without D$3 shows very similar results in estimated coefficients, with a slightly higher coefficient for the interactive term of [Log of GP – Log of $3]*D$3, 0.320, and a statistically insignificant coefficient for monthly gasoline price: -0.032.

12. With a threshold boost effect at the $3 mark and/or a change in elasticity crossing the $3 (or $4) mark, the total effect of gasoline price (GS) increase on transit ridership (R) can be calculated by the following equation. 1 and 2 indicate before and after an increase in gasoline prices:

    (R2 / R1) = (GS2 / GS1)^(estimated coefficient for “Log of GP”)

* (GS2 / 3)^(estimated coefficient for “[Log of GP – Log of $3]*D$3)
* exp(estimated coefficient of D$3)

When no threshold boost effect is found, an estimated coefficient of D$3 is zero, and the last term becomes one. A percent increase in ridership can be obtained by subtracting one from (R2/R1) and multiplying by 100.

13. Transit demand is known to be lower in the summer relative to other seasons because students do not have to use public transit to attend school, college or university. The transit demand is likely to be lower in the summer in areas with a high percentage of students, while it may not vary much in areas with a low percentage of students residing locally. Therefore, Lane (2010) controlled for seasonal effects using dummies for different seasons. However, other studies included monthly dummies. Monthly dummies can capture the effect of factors that affect ridership on a monthly basis and thus provide greater variation than dummies for seasons. Monthly dummies are also able to account for seasonal effects; therefore, it is better to use monthly dummies.

14. The authors used data on the number of linked trips for New Jersey commuter rail (used to commute to and from NYC), unlike most other studies that use unlinked trip data.

BIBLIOGRAPHY

(Also see Annotated Bibliography in Appendix A.)

American Public Transportation Association (APTA): A Profile of Public Transportation Passenger Demographics and Travel Characteristics Reported in On-Board Surveys. Washington, DC, May 2007.

American Public Transportation Agency (APTA), "Public Transportation Protects Americans From Gas Price Volatility." (*Policy Development and Research Paper*, May 2012).

Beresteanu, Arie, and Shanjun Li, "Gasoline Prices, Government Support and the Demand for Hybrid Vehicles in the U.S." *International Economic Review*, Volume 52, Issue 1 (2011): 161–182.

Blanchard, Christopher. *The Impact of Rising Gasoline Prices on US Public Transit Ridership*, Masters Thesis, Duke University; (2009).

Bomberg, Matthew, and Kara M Kockelman, "Traveler Response to the 2005 Gas Price Spike." Proceedings of the 86th Annual Meeting of the Transportation Research Board, Washington DC (2007).

Busse, Meghan R., Christopher R. Knittel, and Florian Zettelmeyer, "Pain at the Pump: The Differential Effect of Gasoline Prices on New and Used Automobile Markets." *NBER Working Paper No. 15590* (2007).

Chen, Cynthia, Don Varley, and Jason Chen, "What Affects Transit Ridership? A Dynamic Analysis involving Multiple Factors, Lags, And Asymmetric Behaviour." *Urban Studies*, Vol. 48, No. 9, July (2010): 1893-1908.

Chu, Xuehao. An Assessment of Public Transportation Markets Using NHTS Data. Prepared for Florida Department of Transportation. March 2012.

Cooper, Michael, "Transit Use Hit Five-decade High in 2008 as Gas Price Rose." *New York Times*, March 9, (2009). http://www.nytimes.com/2009/03/09/us/09transit.html (accessed July 17, 2014).

Currie, Graham, and Justin Phung, "Understanding Links Between Transit Ridership and Auto Gas Prices—US and Australian Evidence." Paper presented at the Transportation Research Board 87th Annual Meeting, Washington, DC, January 13-17, 2008.

Currie, Graham, and Justin Phung, "Transit ridership, auto gas prices, and world events: New drivers of change?" *Transportation Research Record: Journal of the Transportation Research Board* 1992 (-1), (2007): 3-10.

Dormann, Carsten F., Jane Elith, Sven Bacher, Carsten Buchmann, Gudrun Carl, Gabriel Carré, Jaime R. García Marquéz et al. "Collinearity: a review of methods to deal with it and a simulation study evaluating their performance."*Ecography* 36, no. 1 (2013): 027-046.

Florida, Richard, "What Is It Exactly That Makes Big Cities Vote Democratic?" Feb 19, 2013, http://www.theatlanticcities.com/politics/2013/02/what-makes-some-cities-vote-democratic/4598/ (accessed July 08, 2014).

Greene, William H., *Econometric Analysis*. 7th Edition, Upper Saddle River, N.J: Prentice Hall, 2012.

Haire, Ashley, and Randy Machemehl, "Impact of rising fuel prices on US transit ridership." *Transportation Research Record: Journal of the Transportation Research Board* 1992, no. 1 (2007): 11-19.

Kain, John F., and Zvi Liu, "Secrets of success: assessing the large increases in transit ridership achieved by Houston and San Diego transit providers." *Transportation Research A*, 33 (1999): 601–624.

Keyes, Dale L., "Energy for travel: The influence of urban development patterns." *Transportation Research* 16A, 1, (1982): 65-70.

Korkki, Phyllis, "When gas prices lead to roads less traveled." *New York Times*, June 15, 2008. http://www.nytimes.com/2008/06/15/business/15count.html?_r=0 (accessed July 17, 2014).

Lane, Bradley W., "The Relationship Between Recent Gasoline Price Fluctuations and Transit Ridership in Major U.S. Cities." *Journal of Transport Geography*, Vol. 18, No. 2, (2010): 214–225.

Li, Shanjun, Christopher Timmins, and Roger Von Haefen, "How Do Gasoline Prices Affect Fleet Fuel Economy?" *American Economic Journal: Economic Policy* 1, (2009): 113–37.

Litman, Todd, "Transit price elasticities and cross-elasticities." *Journal of Public Transportation*, 7(2), (2004): 37-58.

Linn, Shanjun, and Thomas Klier, "Gasoline Prices and the Demand for New Vehicles: Evidence from Monthly Sales Data." *American Economic Journal: Economic Policy* 2, no. 3 (2010): 134-153.

Maley, Donald W., and Rachel Weinberger, "Rising Gas Prices and Transit Ridership Case Study of Philadelphia." *Transportation Research Record, Journal of the Transportation Research Board* 2139, no. 1 (2009): 183-188.

Mattson, Jeremy, “Effects of Rising Gas Prices on Bus Ridership for Small Urban and Rural Transit Systems.” *Small Urban and Rural Transit Center*, Upper Great Plains Transportation Institute, North Dakota State University, (2008).

Nowak, William P., and Ian Savage, “The cross elasticity between gasoline prices and transit use: Evidence from Chicago.” *Transit Policy*, 29, (2013): 38-45.

Pucher, John and John Renne. “Rural Mobility and Mode Choice: Evidence from the 2001 National Household Transportation Survey,” *Transportation,* 32.3 (2005): 165-186.

Pucher, John and John Renne. “Socioeconomics of Urban Travel: Evidence from the 2001 NHTS,” *Transportation Quarterly*, Vol. 57. No. 3 (2003): 49-77.

Storchmann, Karl, “The impact of fuel taxes on public transport—an empirical assessment for Germany.” *Transport Policy* 8 (1), (2001): 19–28.

Stover, Victor W., and C. H. Christine Bae, “Impact of Gasoline Prices on Transit Ridership in Washington State.” *Transportation Research Record*, Vol. 2217, (2011): 11-18.

Taylor, Brian D., Douglas Miller, Hiroyuki Iseki, and Camille Fink, “Nature and/or Nurture? Analyzing the determinants of transit Ridership across US Urbanized Areas.” *Transportation Research* Part A, 43, (2009): 60-77.

U.S. Census Bureau, 2010-2012 American Community Survey 3-Year Estimates.

Wooldridge, Jeffrey W., *Econometric Analysis of Cross Section and Panel Data*, Second Edition, MIT Press, 2010.

Yanmaz-Tuzel, Ozlem and Kaan Ozbay, “Impacts of Gasoline Prices on New Jersey Transit Ridership.” *Transportation Research Record: Journal of Transportation Research Board*, 2144, no. 1 (2010): 52-61.

ABOUT THE AUTHORS

HIROYUKI ISEKI, PH.D.

Dr. Iseki is an Assistant Professor of Urban Studies and Planning and a research faculty at the National Center for Smart Growth Research and Education of University of Maryland, College Park. His research and teaching interest includes transportation economics and finance, public transit planning and management, travel behavior analysis and modeling, sustainable transportation, regional transportation planning, and applications of GISs to research and practice in public policy and planning. He earned an M.A. and a Ph.D. in Urban Planning from University of California, Los Angeles. For more information, see http://www.arch.umd.edu/ursp/faculty/hiroyuki-iseki and http://www.smartgrowth.umd.edu/hiseki.html.

RUBABA ALI

Rubaba Ali is a Ph.D. candidate at the department of Agricultural and Resource Economics at the University of Maryland, College Park. Her research interests include impact of transport infrastructure on agricultural and non-agricultural sectors and overall welfare in developing countries. She conducts research on energy and transport-related issues in developing and developed countries. She received master’s degrees in Economics and Agricultural and Resource Economics from the University of Maryland, and undergraduate degrees in Mathematics and Economics from Bard College.

PEER REVIEW

San José State University, of the California State University system, and the MTI Board of Trustees have agreed upon a peer review process required for all research published by MTI. The purpose of the review process is to ensure that the results presented are based upon a professionally acceptable research protocol.

Research projects begin with the approval of a scope of work by the sponsoring entities, with in-process reviews by the MTI Research Director and the Research Associated Policy Oversight Committee (RAPOC). Review of the draft research product is conducted by the Research Committee of the Board of Trustees and may include invited critiques from other professionals in the subject field. The review is based on the professional propriety of the research methodology.

SAN JOSÉ STATE
UNIVERSITY

Funded by U.S. Department of
Transportation and California
Department of Transportation