5-1-2018

Improving Demand Modeling in California's Rail Transit System

Rui Liu
San Jose State University

Project 1736 | May 2018

SJSU SAN JOSÉ STATE UNIVERSITY

MTI MINETA TRANSPORTATION INSTITUTE

Improving Demand Modeling in California Rail Transit System

Rui Liu, Ph.D.

MINETA TRANSPORTATION INSTITUTE
LEAD UNIVERSITY OF
Mineta Consortium for Transportation Mobility

Founded in 1991, the Mineta Transportation Institute (MTI), an organized research and training unit in partnership with the Lucas College and Graduate School of Business at San José State University (SJSU), increases mobility for all by improving the safety, efficiency, accessibility, and convenience of our nation's transportation system. Through research, education, workforce development, and technology transfer, we help create a connected world. MTI leads the four-university Mineta Consortium for Transportation Mobility, a Tier 1 University Transportation Center funded by the U.S. Department of Transportation's Office of the Assistant Secretary for Research and Technology (OST-R), the California Department of Transportation (Caltrans), and by private grants and donations.

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: bicycle and pedestrian issues; financing public and private sector transportation improvements; intermodal connectivity and integration; safety and security of transportation systems; sustainability of transportation systems; transportation / land use / environment; and transportation planning and policy development. 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 on TransWeb, the MTI website (http://transweb.sjsu.edu).

Education

The Institute supports education programs for students seeking a career in the development and operation of surface transportation systems. MTI, through San José State University, offers an AACSB-accredited Master of Science in Transportation Management and graduate certificates in Transportation Management, Transportation Security, and High-Speed Rail Management that serve to prepare the nation's transportation managers for the 21st century. With the active assistance of the California Department of Transportation (Caltrans), 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 utilizes a diverse array of dissemination methods and media to ensure research results reach those responsible for managing change. These methods include publication, seminars, workshops, websites, social media, webinars, and other technology transfer mechanisms. Additionally, MTI promotes the availability of completed research to professional organizations and journals and works to integrate the research findings into the graduate education program. MTI's extensive collection of transportation- related publications is integrated into San José State University's world-class Martin Luther King, Jr. Library.

REPORT WP 18-01

IMPROVING DEMAND MODELING IN CALIFORNIA RAIL TRANSIT SYSTEM

Rui Liu, Ph.D.

May 2018

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

1. Report No. WP 18-01	**2. Government Accession No.**	**3. Recipient's Catalog No.**
4. Title and Subtitle Improving Demand Modeling in California Rail Transit System		**5. Report Date** May 2018 **6. Performing Organization Code**
7. Authors Rui Liu, Ph.D.		**8. Performing Organization Report** CA-MTI-1736
9. Performing Organization Name and Address Mineta Transportation Institute College of Business San José State University San José, CA 95192-0219		**10. Work Unit No.** **11. Contract or Grant No.** 69A3551747127
12. Sponsoring Agency Name and Address U.S. Department of Transportation Office of the Assistant Secretary for Research and Technology University Transportation Centers Program 1200 New Jersey Avenue, SE Washington, DC 20590		**13. Type of Report and Period Covered** Final Report **14. Sponsoring Agency Code**
15. Supplemental Notes		

16. Abstract

This paper analyzes urban rail-fare elasticity and compares the results across four California transit systems. A method of Internet search is adopted to collect monthly transit-fare records from 2002 to 2013. This paper contributes towards improving demand modeling for public transit using more precise and monthly data and applies econometric techniques involving autoregressive integrated moving average (ARIMA) and panel data models. Results show that demand for public transit in California is very inelastic. Any ridership promotion policy may have a heterogeneous impact across transit systems.

17. Key Words Public transit; fare elasticity; ARIMA; panel data models	**18. Distribution Statement** No restrictions. This document is available to the public through The National Technical Information Service, Springfield, VA 22161		
19. Security Classif. (of this report) Unclassified	**20. Security Classif. (of this page)** Unclassified	**21. No. of Pages** 21	**22. Price** $15.00

Form DOT F 1700.7 (8-72)

Library of Congress Catalog Card Number:
2018942738

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

052218

ACKNOWLEDGMENTS

The author thanks Mineta Transportation Institute for the Emerging Leaders Seed Grant, two anonymous referees for their valuable comments, and Dr. Matthew Holian (San Jose State University) for his guidance and support.

The authors also thank MTI staff, including Executive Director Karen Philbrick, Ph.D.; Research and Technology Transfer Director Hilary Nixon, Ph.D.; Research Support Assistant Joseph Mercado; Executive Administrative Assistant Jill Carter; and Editor Jan Steckel.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

I. INTRODUCTION

The importance of transit as a means of travel to work has increased substantially over the past few years. According to the *2016 Public Transportation Fact Book* (Neff and Dickens 2017), 6.2 million U.S. workers commuted on public transit in 2005. By 2015 the number of commuters who took public transit had increased to 7.6 million, a 25 percent increase in a decade. Thus, understanding the determinants of the demand for public transit is essential for the development of an efficient U.S. transportation system. In particular, understanding price elasticity of demand is crucial for determining the impact of fare changes on transit ridership.

This elasticity indicator has been estimated by many researchers for different transit systems using annual data and a broad array of econometric techniques (Winston and Maheshri 2007; Blanchard 2009; Chen, Varley and Chen 2011; Litman 2004; Litman 2017). Since fare elasticity is an important input into cost-benefit analysis of rail-transit systems, and monthly data on transit fares are not readily available in the National Transit Database, more precise measures of fare and its elasticity are needed. Therefore, we employ a method of Internet search which enables us to collect the historical monthly fare structure for rail systems. In this way, we are able to determine the precise date at which fares changed. We are also able to use this technique to determine the precise date at which new stations were opened. Although this technique can be scaled up to collect fare data for most systems in the U.S., for now we use this technique to construct a monthly database for four major rail-transit systems in California.

This paper is focused on time-series techniques to analyze the monthly data on California transit ridership and its determinants (fare, service level, etc.). In addition to time-series models, we also estimated panel-data models, as in Winston and Maheshri (2007). Winston and Maheshri (2007) examined 25 U.S. rail systems from 1993-2000. Their panel was wide and short, but the data was from the pre-2000 period. Our panel is not as wide, but it is longer. The main comparative advantage of our panel analysis compared to theirs is that time periods in this study are months, not years, and actual fares were used, not fare revenue per mile. By using more precise and monthly data, the present analysis may result in more accurate fare-elasticity estimates, which would thus be more suitable for drawing out policy recommendations. Value is also added by the present study in that the researcher can make a comparison between econometric techniques using the same data set. Winston and Maheshri (2007) also found that transit-ridership was much more fare elastic than has been found in the previous time-series literature (Small 1992), and it is of interest to see if the estimated elasticities are smaller (in absolute value) in this investigation’s time series analysis than in its panel analysis. Finally, this study uses very recent data; over the last few years in the United States, researchers have documented increased transit usage, reduced driving, and (for the first time in decades) faster population growth in central cities than in suburbs. Thus, there is reason to suppose that estimates produced using older data may no longer be applicable to the current environment.

The study is organized as follows. Section 2 displays the variables and introduces our dataset, while Section 3 provides the details of the adopted models– ARIMA and panel data models. The estimated results from these models are presented in Section 4. Conclusions are discussed in Section 5.

II. DATA

The data is composed of four California urban rail-transit systems that were in operation between 2002:M1 and 2013:M9, generating 564 observations. The systems include San Francisco (SF), Sacramento (SAC), San Jose (SJ) and Los Angeles (LA). Variables adopted in this study are explained as follows.

- UPT: Unlinked Passenger Trips in millions, defined as the number of passengers who board public transportation vehicles.
- VRM: Vehicles Revenue Miles in millions, the miles that vehicles travel while in revenue service. This variable is used to represent transit service level.
- Stations: the number of light rail stations.
- Employment: total non-farm employees in millions.
- Real gasoline price: California gasoline retail price (constant $ per gallon), adjusted for inflation using the Consumer Price Index (CPI) for all urban consumers based on 1982-84.
- Real transit fare: Single ride fare (constant $) for light rail, adjusted for inflation using the Consumer Price Index (CPI) for all urban consumers based on 1982-84.

The monthly fare data are collected by searching cached images of transit agency web pages, using the Internet Archive (archive.org) from 2002 to 2013. The number of stations can also be determined using this technique. Data on UPT and VRM can be retrieved from the National Transit Database. Employment data were obtained from the Bureau of Labor Statistics, whereas monthly gasoline retail price in California can be downloaded from the Energy Information Administration.

Table 1 presents the summary statistics for each variable involved in the study. On average between 2002 and 2013, travelers in these four urban areas paid $1.72 for a single ride, and $3 for gas in real dollars. Passenger trips averaged around 2.34 million, and VRM averaged 0.23 million miles. Tables 2 to 5 display the summary statistics of all variables for each system during the sample period.

Table 1. Summary Statistics for All Transit Systems (2002-2013)

Variable Name	Mean[1]	Standard Deviation	Min	Max
UPT (in millions)	2.34	1.50	0.40	5.62
VRM (in millions)	0.44	0.23	0.11	1.20
Stations	47	19.74	6	80
Employment (in millions)	1.70	1.37	0.81	4.29
Real gasoline price (constant $)	3.02	0.68	1.51	4.48
Real transit fare (constant $)	1.72	0.32	1.17	2.51

Table 2. Summary Statistics for Los Angeles (2002-2013)

Variable Name	Mean	Standard Deviation	Min	Max
UPT (in millions)	3.68	0.85	0.73	5.62
VRM (in millions)	0.77	0.17	0.25	1.20
Stations	64.18	8.18	50	80
Employment (in millions)	4.07	0.11	3.84	4.29
Real gasoline price (constant $)	3.02	0.68	1.51	4.48
Real transit fare (constant $)	1.43	0.11	1.23	1.65

Table 3. Summary Statistics for San Francisco (2002-2013)

Variable Name	Mean	Standard Deviation	Min	Max
UPT (in millions)	3.83	0.39	3	4.97
VRM (in millions)	0.45	0.05	0.31	0.53
Stations	16.34	8.93	6	24
Employment (in millions)	0.99	0.03	0.94	1.08
Real gasoline price (constant $)	3.02	0.68	1.51	4.48
Real transit fare (constant $)	1.54	0.21	1.17	1.92

Table 4. Summary Statistics for San Jose (2002-2013)

Variable Name	Mean	Standard Deviation	Min	Max
UPT (in millions)	0.75	0.17	0.4	1.07
VRM (in millions)	0.23	0.06	0.11	0.3
Stations	56.93	6.20	46	61
Employment (in millions)	0.90	0.03	0.85	0.97
Real gasoline price (constant $)	3.02	0.68	1.51	4.48
Real transit fare (constant $)	1.83	0.12	1.5	2.01

Table 5. Summary Statistics for Sacramento (2002-2013)

Variable Name	Mean	Standard Deviation	Min	Max
UPT (in millions)	1.12	0.22	0.69	1.78
VRM (in millions)	0.30	0.06	0.16	0.37
Stations	43.67	6.23	30	49
Employment (in millions)	0.87	0.03	0.81	0.92

Variable Name	Mean	Standard Deviation	Min	Max
Real gasoline price (constant $)	3.02	0.68	1.51	4.48
Real transit fare (constant $)	2.07	0.28	1.63	2.51

Figures 1 to 4 show the monthly transit ridership and its contributing factors over the sample period for each of the four transit systems in California. (Each variable is rescaled in order to show multiple series in one chart, which also makes comparison easier.) Specifically, as seen in Figure 1, LA transit ridership has increased steadily from 2.3 million in January 2002 to over 5 million in September 2013. During that time, employment has been quite stable and service level (UPT) has more than doubled. Transit fare, adjusted for inflation, has decreased during the study period, except for one fare increase in June 2009. The average real price of gasoline also increased 145 percent over the period.

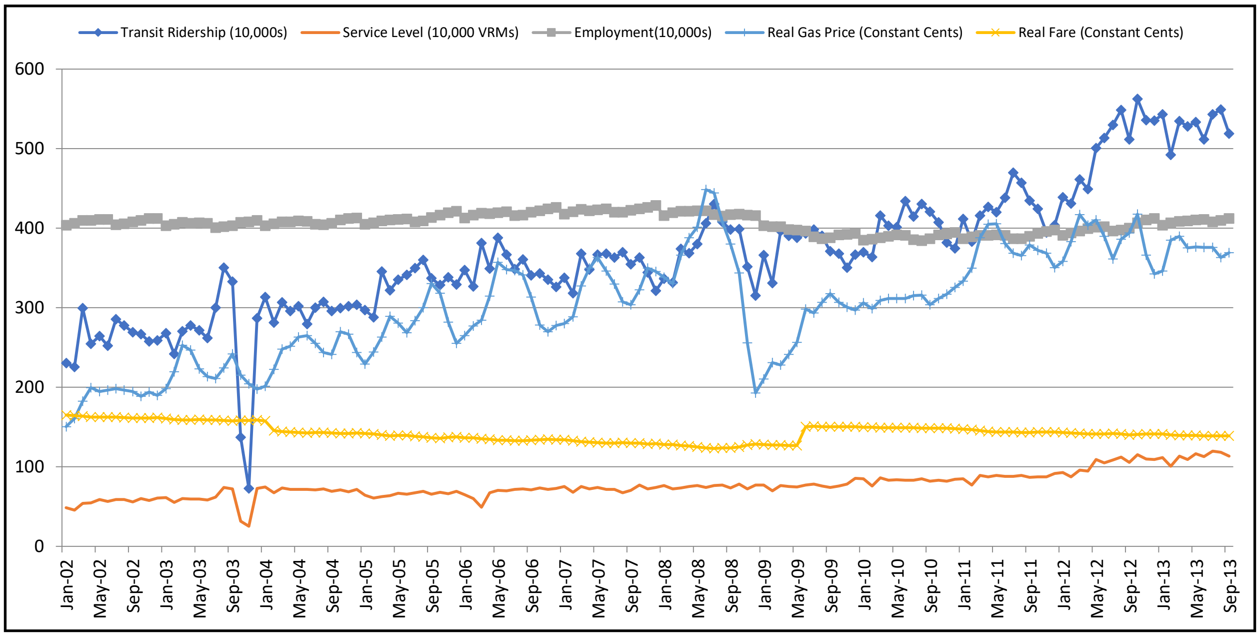

Figure 1. Transit Ridership and All Independent Variables for Los Angeles

Similarly, Figure 2 shows transit ridership and its determinants for San Francisco from 2002 to 2013. During this time, SF transit ridership increased slightly by 0.13 million. Service level has been stable, and unemployment rose during the 2007-2009 recession. Real transit fare increased during the study period, with three sudden fare increases in September 2003, September 2005, and July 2011.

Figure 3 shows transit ridership and its determinants for San Jose. During the study period, SJ transit ridership has increased by more than 0.4 million. At the same time, service level has been quite stable, and employment shrank during the recession phase and reversed in the periods of expansion. Real transit fare has increased during the study period, with four sudden fare increases in July 2002, August 2003, January 2005, and October 2009.

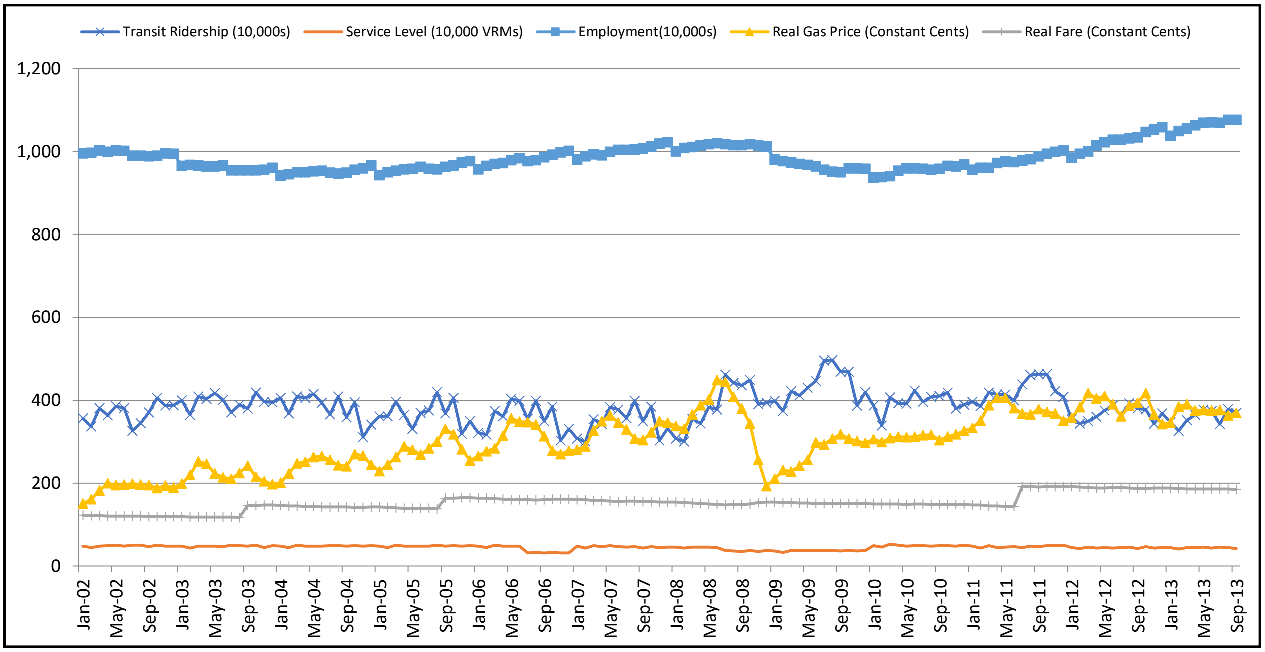

Figure 2. Transit Ridership and All Independent Variables for San Francisco

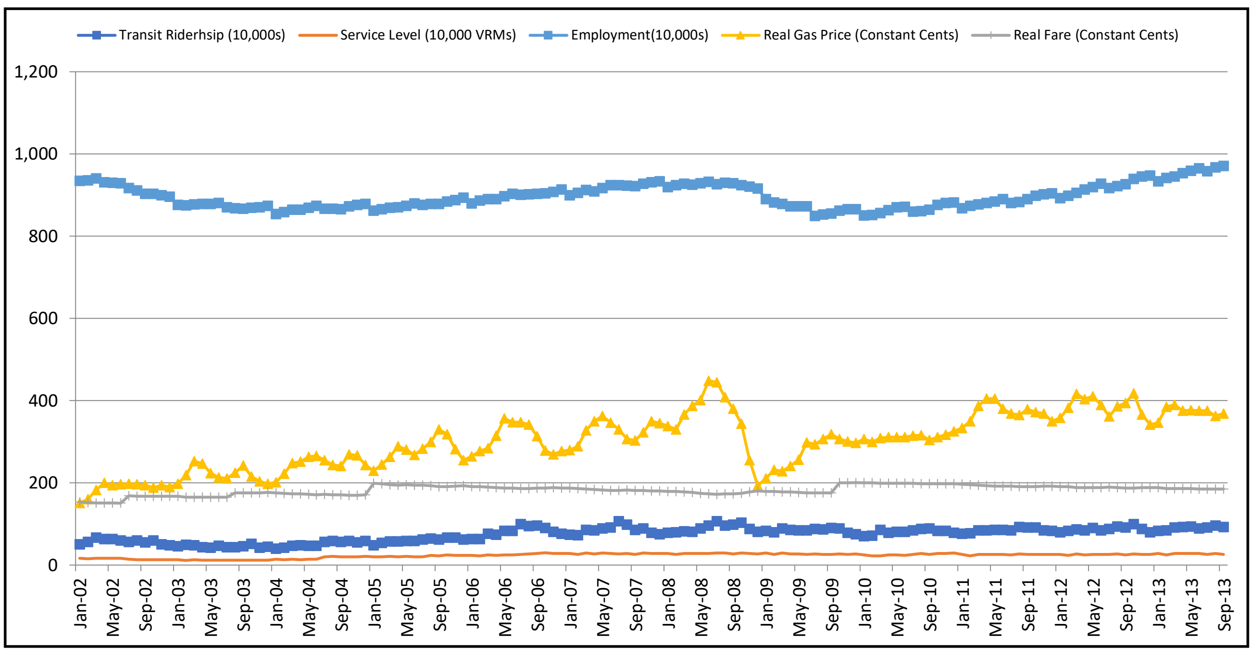

Figure 3. Transit Ridership and All Independent Variables for San Jose

Lastly, Figure 4 shows transit ridership and its contributing factors for Sacramento. From 2002 to 2013, Sacramento transit ridership increased by more than half a million. At the same time, service level and employment showed patterns similar to those found for San Jose. Real transit fare increased during the study period, with four sudden fare increases in November 2005, September 2006, February 2009, and November 2009.

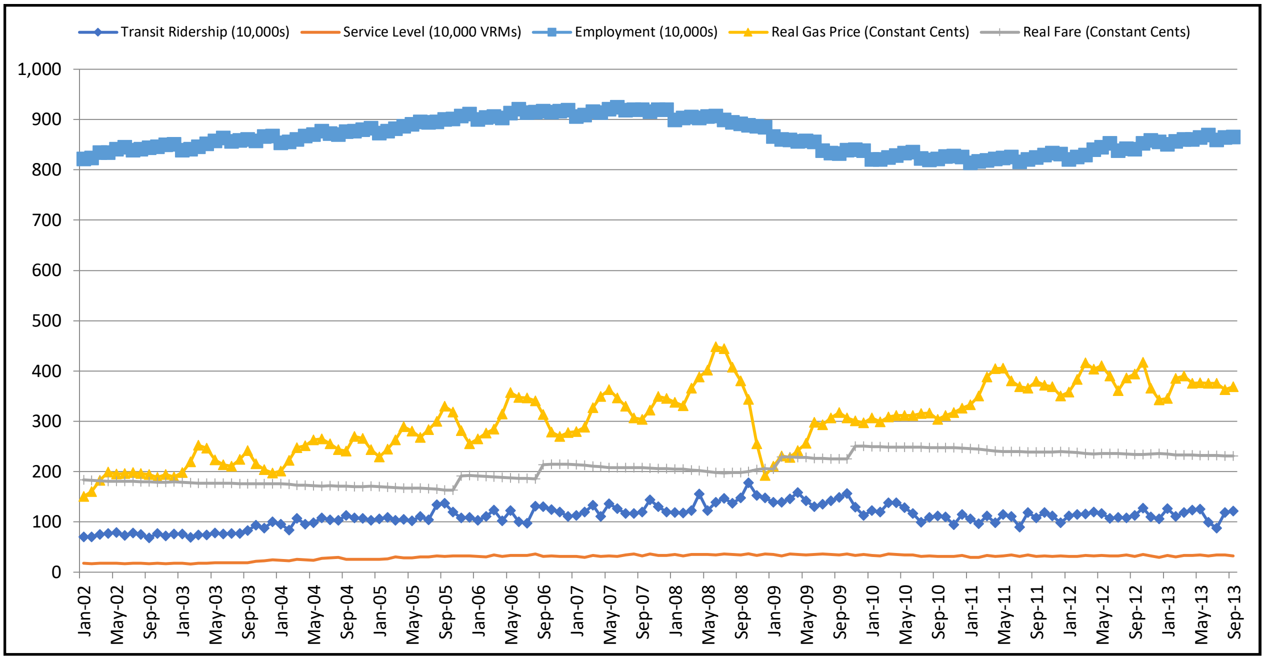

Figure 4. Transit Ridership and All Independent Variables for Sacramento

III. MODELS

The travelers' demand for public transit is specified using UPT. The number of stations is used to represent the transit system network, and the level of employment is used to control for city characteristics. Transit service levels are captured in the form of VRM. Gasoline prices are also included because higher gasoline prices may increase transit ridership, but they may also slow down the overall economy, which might decrease transit ridership.

The data utilized in this study are time series data and therefore require time series analysis. In this section, two models are introduced to estimate fare elasticity of demand for public transit: the ARIMA model and the panel data model. First, the univariate ARIMA model for each transit system in our sample was estimated and the fare elasticities calculated. Next, analysis was carried out based on a panel of fours for transit systems from 2002:M1 to 2013:M9, and the question of whether transit-ridership is more fare elastic than has been found in the ARIMA estimates was examined. In both analyses, natural logarithm of all variables (except Stations) were taken to model elasticity in a convenient fashion.

THE ARIMA MODEL

The author applies the seasonal ARIMA (Autoregressive Integrated Moving Average) model to estimate the impact of various factors on public transit. The seasonal ARIMA model for each system is specified as follows:

$$UPT_t = c + \sum_k \beta^k * X_t^k + u_t$$

where:

- UPT_t is the transit ridership in the form of unlinked passenger trips in period t;
- c is a constant term for the entire sample;
- β^k is the coefficient of the k^{th} explanatory variable;
- X_t^k is the k^{th} explanatory variable in period t;
- u_t is the error term, assumed to follow $ARIMA(p,d,q)(P,D,Q)_{12}$.

This time-series model not only allows for the inclusion of information from the past observations of a series, but also for the inclusion of other relevant information that helps determine the dependent variable. The investigator explicitly models the autocorrelation in the error series by means of a seasonal ARIMA. For example, if u_t follows an ARIMA(1,1,1)$(1,1,1)_{12}$ model, using backshift notation, one can write:

$$(1-\phi_1 B)(1-\Phi_1 B^{12})(1-B)(1-B^{12})u_t = (1+\vartheta_1 B)(1+\Theta_1 B^{12})e_t$$

where e_t is a white noise series.

THE PANEL DATA MODEL

The standard panel model with i = 1, …,n and t = 1, …, T is:

$$UPT_{it} = c + \alpha_i + \sum_k \beta^k * X_{it}^k + u_{it}$$

where:

- UPT_{it} is the transit ridership in the form of unlinked passenger trips for system i in period t;
- c is a constant term for the entire sample;
- α_i is individual specific effect for system i;
- β^k is the coefficient of the k^{th} explanatory variable;
- X_{it}^k is the k^{th} explanatory variable for system i and period t;
- u_{it} is the error term.

There are two common assumptions usually made about the individual specific effect, the random effects assumption and the fixed effects assumption. If α_i is unobserved and correlated with one of the independent variables, one would use a fixed-effect model to solve the problem of omitted variable bias. If α_i is uncorrelated with the independent variables, then a random-effect model would be more efficient than a fixed-effect model.

IV. RESULTS

THE ARIMA MODEL RESULTS

Table 6. ARIMA Model Results

Variable	SF	SJ	SAC	LA
	ARIMA(1, 0, 2) $(1, 0, 1)_{12}$	ARIMA(1,0,0) $(1,0,1)_{12}$	ARIMA(3,0,2)	ARIMA(2, 0, 0) $(1,0,0)_{12}$
Real transit fare (ln)	0.040 (0.109)	-0.463* (0.200)	-0.151 (0.183)	-0.436* (0.238)
Real gasoline price (ln)	-0.034 (0.057)	0.080 (0.066)	-0.032 (0.063)	0.074 (0.073)
Service level (ln(VRM))	-0.067 (0.074)	0.190* (0.082)	0.777* (0.120)	1.095* (0.048)
Employment (ln)	-1.243* (0.445)	1.292* (0.421)	0.359 (0.601)	-0.986* (0.631)
Stations	0.002 (0.002)	0.024* (0.005)	0.002 (0.005)	-0.005* (0.003)
AR1	0.525* (0.141)	0.695* (0.069)	1.722* (0.180)	0.401* (0.081)
AR2			-1.218* (0.228)	0.267* (0.086)
AR3			0.422* (0.107)	
MA1	0.043 (0.127)		-1.315* (0.188)	
MA2	0.351* (0.123)		0.600* (0.155)	
SAR1	0.867* (0.071)	0.816* (0.120)		0.231* (0.091)
SMA1	-0.458* (0.151)	-0.466* (0.208)		
Intercept	1.251* (0.098)	-1.045* (0.347)	1.173* (0.333)	3.380* (0.966)
Log-likelihood	202.59	211.65	159.53	179.29
AIC	-381.18	-403.31	-295.05	-338.58

* Represents statistical significance at the 5% level.
Numbers in parentheses are standard errors.
Dependent variable: logarithm of transit ridership (UPT in millions).

Table 6 shows the results for four California transit systems after applying different seasonal ARIMA models with the same set of independent factors.[1] There are five components of the model, capturing the underlying data-generating process. The components are: the autoregressive (AR) components, the moving average (MA) components, the seasonal autoregressive (SAR) components, the seasonal moving average (SMA) components, and a set of independent variables that help explain transit ridership. The results on autoregressive components show the impact of lagged ridership is particularly great, and the seasonal components indicate seasonal effects are also important for all three geographies except for Sacramento.

Across all four regions, people are not very sensitive to fare changes. Transit fares have a negative and significant effect on ridership in Los Angeles and San Jose, a negative and insignificant effect in Sacramento, and a positive and insignificant effect in San Francisco. The fare elasticities are 0.04 for San Francisco, -0.456 for San Jose, -0.155 for Sacramento, and -0.721 for Los Angeles. The fare elasticity in Los Angeles, for example, indicates that for every 1 percent increase in fare, we would expect the transit ridership to decrease by 0.721 percent. Fare elasticities in San Jose and Los Angeles are consistent with previous studies in which fare elasticities generally fall between -0.20 and -0.90 (McLeod et al. 1991; Litman 2004; Chen, Varley and Chen 2011). However, the elasticities for San Francisco (0.04) and Sacramento (-0.155) are insignificant. The results indicate that there is no immediate effect on transit ridership from fare changes. However, there may exist a lagged relationship between ridership and fare, as evidenced by the significant autoregressive components.

Economic theory predicts that people will drive less when gasoline prices rise. But driving less does not necessarily mean taking public transit more. This is confirmed by our findings that changes in gasoline prices have no contemporaneous effect on transit ridership. Figures 1-4 show that gasoline prices have been quite volatile during the sample period. Prices increased sharply until June 2008, declined, and then rose again until 2013. But the timing of transit’s rise and decline is not amenable to a gasoline price explanation. Thus, it is reasonable to think that gasoline prices have an insignificant impact on transit demand.

One of the most common measures for service level is vehicle revenue miles (VRM). VRM captures the distance that vehicles travel while in revenue service. Hertz (2015) and Harrison (2017), both of whom measure service as VRM, find service decline can lead to falling ridership. On the other hand, reverse causality also exists as more transit riders can potentially increase service (Alam, Nixon and Zhang 2015).

While VRM rise across all three metropolitan areas (SJ, SAC and LA) as in Figures 6-8, they have experienced a significant drop in SF in 2006, 2008 and 2009 (Figure 5). During those periods, service decline has been accompanied by higher transit ridership, making the coefficient on VRM for SF negative. In other periods, expanding service is in concert with increasing ridership. The combined effect of VRM on ridership in SF is therefore inconclusive, thus leading to an insignificant effect of VRM.

Previous findings on service elasticities generally lie between 0.3 and 1.14 (Rose 1986; Litman, 2004; Taylor et al. 2009, Chen, Varley and Chen 2011). Our estimates for SAC and LA are consistent with the range, but the estimate for SJ is slightly lower. The relative low elasticity in SJ may be associated with the specific geographic context, in which the extensive transit service already offered in the area may leave less room for ridership growth (Chen, Varley and Chan 2011).

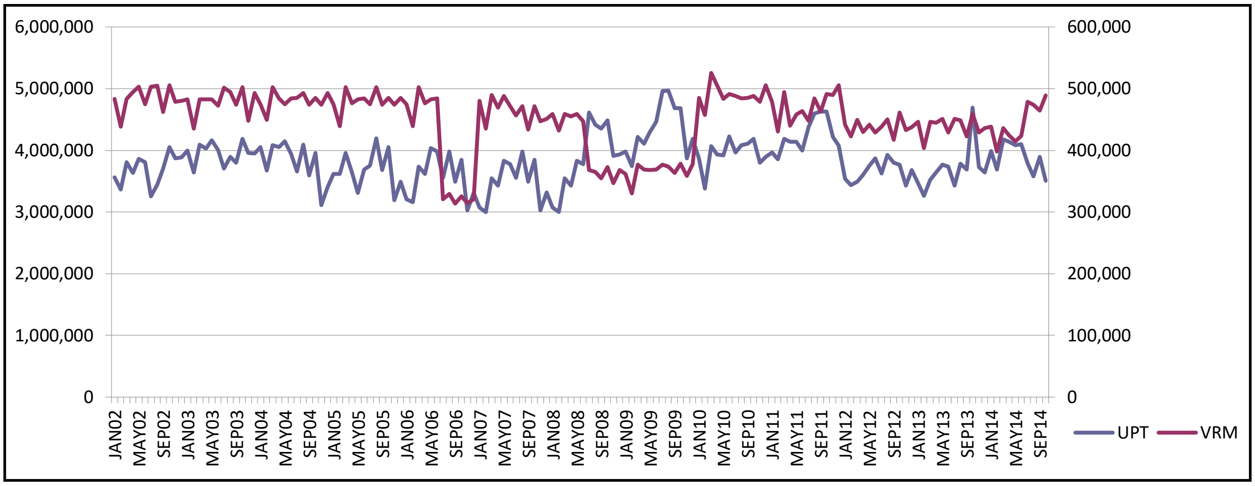

Figure 5. VRM in San Francisco

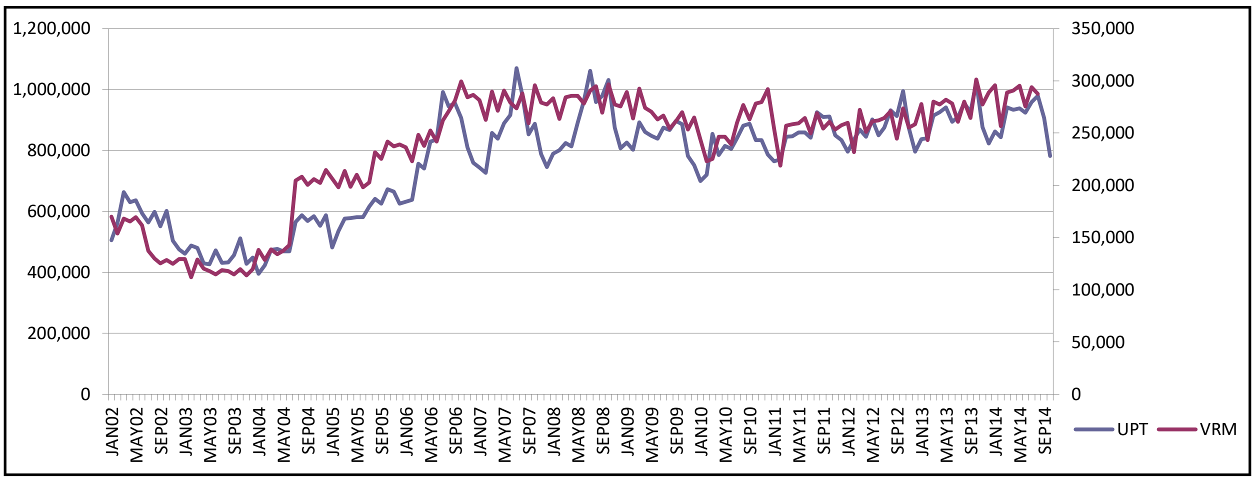

Figure 6. VRM in San Jose

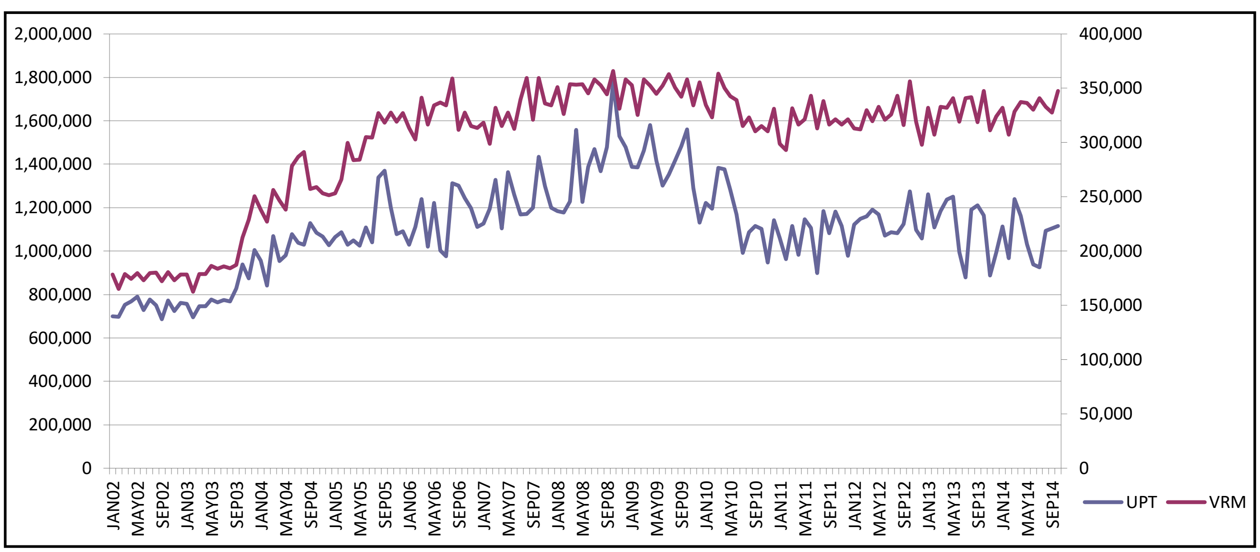

Figure 7. VRM in Sacramento

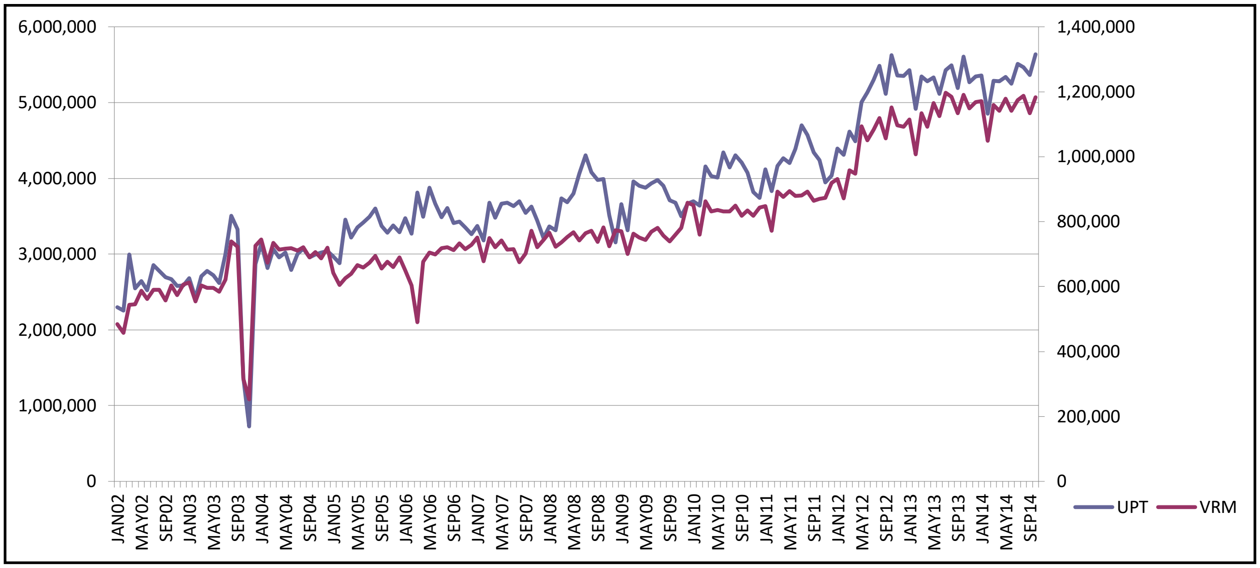

Figure 8. VRM in Los Angeles

Employment is found to be a significant factor in determining transit ridership, in line with the previous findings (McLeod et al. 1991; Taylor et al. 2009; Chen, Varley and Chen 2011). In particular, increased employment reduces ridership in Los Angeles and San Francisco but increases ridership in San Jose and Sacramento. The effect is significant for all systems except for Sacramento.

Previous studies show that employment elasticities are usually positive, within the range of 1.04–1.75 (McLeod et al. 1991; Taylor et al. 2009). Our negative employment elasticities could potentially result from spatial characteristics (i.e. population density) and income effect that were not considered in the analysis. For instance, employment growth arising from population growth may be occurring only in low-density areas which are not served by public transit. In addition, employment growth due to income growth in the area is found to result in less use of public transit (McLeod et al. 1991).

The network variable (number of stations) has a positive effect on transit demand for all geographies except LA. The positive effect is expected, as increase in number of stations would attract more housing, construction and people to move near transit stations, thus leading to higher transit demand. The negative effect for LA is not entirely surprising. For example, an increase in the number of stations in an area where there are already a lot of stations in close proximity will increase travel time for existing riders and may actually discourage large increase in transit ridership. The insignificant effect in San Jose may indicate that San Jose's transit system is saturated in stations, thus leaving less room for increase in transit use.

THE PANEL MODEL RESULTS

Table 7. Panel Data Model Results

Variable	Pooled OLS (1)	Fixed Effect (2)	Random Effect (3)
Real Transit Fare (ln)	-0.475*	-0.307*	-0.302*
	(0.066)	(0.061)	(0.100)
Real Gasoline Price (ln)	0.434*	0.114	0.298*
	(0.081)	(0.061)	(0.070)
Service Level (ln(VRM))	1.073*	0.643*	0.946*
	(0.056)	(0.074)	(0.046)
Employment (ln)	0.205*	-0.014	0.191*
	(0.049)	(0.221)	(0.060)
Stations	-0.018*	0.005*	-0.011*
	(0.001)	(0.002)	(0.001)
Intercept	2.172*		-1.762*
	(0.140)		(0.107)
Entity Fixed Effects			
LA		1.130	
SAC		0.753	
SF		1.776	
SJ		0.409	
Adjusted R^2	0.921	0.741	0.783
Number of Transit Systems (n)	4	4	4
Number of Time Periods (T)	141	141	141
Total Number of Observation (N)	564	564	564

* Represents statistical significance at the 5% level.

Numbers in the parenthesis are Driscoll and Kraay's (1998) robust standard errors for panel models with cross-sectional and serial correlation.

Dependent variable: logarithm of transit ridership (UPT in millions).

Table 7 displays the estimated results based on three panel models: pooled Ordinary Least Squares (OLS), fixed effects and random effects. Column (1) displays the ordinary least squares estimates based on the full sample of 564 observations, ignoring unobserved heterogeneity. The inclusion of "one-way" fixed effects in Column (2) in which dummy variables are specified for each transit system allows unknown cross-sectional effect to be explained through the fixed effects' coefficients and reduces any remaining omitted variable bias (Stock and Watson 2007).

It is of interest to examine whether fixed effects are jointly significant associated with transit ridership. This is done by comparing the pooled-OLS model and the fixed-effect model through an F-statistic. The resulting statistic is significant, with a value of $F(3, 555) = 251.56$, indicating there is substantial variation among transit systems in the sample. Thus, the use of fixed effects is supported in the model estimation.

A comparison of the regression coefficients in columns (2) and (3) shows that fixed- and random-effects methods yield rather similar results for fare elasticity. A selection between these two models concerns the exogeneity of the explanatory variables. Therefore, a Hausman test was employed to compare fixed-effect and random-effect models. The Hausman test is designed to detect violation of the random-effects modeling assumption that the independent variables are uncorrelated with the individual effects. Our finding that is taken as evidence that there is correlation between the covariates and the unit effects. Thus the random-effects model is rejected in favor of the use of the fixed-effect model, and the subsequent analysis is based on the estimates from the fixed-effect model.

According to Column (2), fare, service level and number of stations all show expected signs and are all statistically significant at 5% level. Gasoline price is found to have a positive yet insignificant effect on transit ridership. Employment's impact is negative and statistically insignificant, which is in line with previous finding as in Winston and Maheshri (2007). The association between transit use and employment is not always strong. In summary, we have found that fare reductions, increased service level, and network coverage are associated with increased transit ridership.

In particular, fare elasticity is estimated to be -0.3, which is lower (in absolute value) than the fare elasticities found in ARIMA models for Los Angeles and San Jose. Additionally, the service level elasticity is found to be 0.643, which is lower than those for Sacramento and Los Angeles, but higher than that for San Jose, in comparison to the ARIMA estimates on page 9.

V. CONCLUSION

In order to improve transit demand modeling, the investigator employed a method of Internet search, collecting historical monthly fare records from 2002 to 2013 for four major transit systems in California. By using more precise, more recent, and monthly data, the author's analysis results in more accurate fare elasticity estimates, which would thus be more suitable for drawing out policy recommendations. Based on the data set, the investigator utilizes a seasonal ARIMA model combined with a set of explanatory variables to estimate the fare elasticity for each transit system. In addition, three different panel data models are applied to calculate fare and service level elasticities.

The seasonal ARIMA model results indicate that fare, service level, employment and number of stations all have a significant impact on transit ridership in San Jose and Los Angeles. In San Francisco, higher employment levels are associated with lower ridership. In Sacramento, service expansion is accompanied by higher ridership.

The panel model results show that transit demand in California is in general very inelastic, more so than the fare elasticities found in the ARIMA analysis. We also find that there are significant differences in transit ridership across transit systems as indicated by system-specific fixed effects. Service level and number of stations are found to have a positive and statistically significant impact on California transit ridership. This suggests that an increase in service development could potentially promote transit use.

This study shows that there are significant factors that determine transit demand in California and points out that any ridership promotion policy may have a variation of impact across transit systems. The study sheds some light as to the source of those variations, which might help policymakers draw out system-specific policies that will be more suitable in increasing transit ridership.

ACRONYMS AND ABBREVIATIONS

AIC	Akaike Information Criterion
AR	Autoregressive
ARIMA	Autoregressive Integrated Moving Average
CPI	Consumer Price Index
LA	Los Angeles
MA	Moving Average
OLS	Ordinary Least Squares
SAC	Sacramento
SAR	Seasonal Autoregressive
SF	San Francisco
SJ	San Jose
SMA	Seasonal Moving Average
UPT	Unlinked Passenger Trips
VRM	Vehicles Revenue Miles

ENDNOTES

1. Mean value in Table 1 is calculated as an average of all observations for each variable across four cities over the specified time span.

BIBLIOGRAPHY

Alam, Bhuiyan, Hillary Nixon, and Qiong Zhang. 2015. *Investigating the Determining Factors for Transit Travel Demand by Bus Mode in US Metropolitan Statistical Areas (No. Report 12-30).* San Jose, CA: Mineta Transportation Institute, San Jose State University. http://transweb.sjsu.edu/PDFs/research/1101-transit-bus-demand-factors-in-US-metro-areas.pdf

Blanchard, Christopher. 2009. "The Impact of Rising Gasoline Prices on U.S. Public Transit Ridership." Honors Thesis, Duke University.

Chen, Cynthia, Don Varley, and Jason Chen. 2011. "What Affects Transit Ridership? A Dynamic Analysis involving Multiple Factors, Lags and Asymmetric Behaviour." *Urban Studies*, 48, No. 9: 1893–1908.

Driscoll, John, and Aart Kraay. 1998. "Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data." *Review of Economics and Statistics*, 80: 549–560.

Harrison, David. 2017. "America's Buses Lose Riders, Imperiling Their Future." *Wall Street Journal*, April 12, 2017.

Hertz, Daniel. 2015. "Urban Residents aren't Abandoning Buses; Buses are Abandoning Them." City Observatory. http://cityobservatory.org/urban-residents-arent-abandoning-buses-buses-are-abandoning-them/

Litman, Todd. 2004. "Transit Price Elasticities and Cross-Elasticities." *Journal of Public Transportation*, 7, No. 2: 37–58.

Litman, Todd. 2017. *Evaluating Public Transit Benefits and Costs: Best Practices Guidebook.* Victoria, BC, Canada: Victoria Transport Policy Institute.

McLeod Jr., Malcolm S., Kevin J. Flannelly, Laura Flannelly, and Robert W. Behnke. 1991. "Multivariate Time-Series Model of Transit Ridership Based on Historical, Aggregate Data: The Past, Present and Future of Honolulu." *Transportation Research Record*, 1297: 76–84.

Neff, John and Matthew Dickens. 2017. *2016 Public Transportation Fact Book*. Washington, DC: American Public Transportation Association.

Rose, Geoffrey. 1986. "Transit Passenger Response: Short and Long Term Elasticities Using Time Series Analysis." *Transportation*, 13, No. 2: 131–141.

Small, Kenneth A. 1992. *Urban Transportation Economics*. Chur, Switzerland: Harwood Academic Publishers.

Stock, James H., and Mark W. Watson. 2007. *Introduction to Econometrics*. Boston: Pearson/Addison Wesley.

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

Winston, Clifford, and Vikram Maheshri. 2007. “On the Social Desirability of Urban Rail Systems.” *Journal of Urban Economics*, 62, No. 2: 362–382.

ABOUT THE AUTHOR

RUI LIU, PH.D.

Dr. Rui Liu is an Assistant Professor of Economics at San Jose State University. She is also a faculty member at the Center of Smart Technology, Computing, and Complex Systems at San Jose State University. She completed her Ph.D in Economics in 2013 at the University of California, Irvine (with a specialization in time series econometrics). She has several publications and research papers in the fields of applied econometrics, macroeconometrics, and education economics.

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.