ARE LAND USE PLANNING AND GASOLINE PRICE INCREASE MUTUALLY SUPPORTIVE IN GETTING MORE TRANSIT RIDERS IN THE US URBANIZED AREAS?

BY
YONGSUNG LEE

THESIS
Submitted in partial fulfillment of the requirements
for the degree of Master of Urban Planning in Urban Planning
in the Graduate College of
the University of Illinois at Urbana-Champaign, 2012

Urbana, Illinois

Master's Committee:

Assistant Professor Bumsoo Lee
Assistant Professor Andrew Greenlee

ABSTRACT

In July of 2008, the United States had the highest gasoline price in both nominal and real dollars since the 1950s. As a result, there was a record mass transit ridership increase in the following months. Urban planners have long argued that land use measures (dense, mixed-use, and well-connected developments), transit policies (public subsidy for both operation and infrastructure investment), and a combination of both (transit-oriented development) would attract choice riders to mass transit; however, the rise of motor fuel costs was found to be the greatest factor to influence how people choose their travel modes. For this reason, previous research focused on a statistical relationship between gasoline price fluctuation and transit ridership change, and attempted to quantify that correlation by calculating the price elasticity of gasoline. However, that research ignored the influence of urban form and urban policy on the relationship between gasoline prices and transit ridership.

This thesis focuses on the role of urban form and urban policy in this context and conducts an econometric analysis to see whether there are "synergistic" effects between "urban form and urban policy" and "gasoline price change." A synergistic relationship means that the effect of adopting two policies at the same time yields better outcomes than the effect of individually implementing either of the two. However, quantifying the pure effects of land use measures on transit ridership is beyond the scope of this thesis; instead, this thesis attempts to calculate "interaction effects" of urban form and urban policy and gasoline price change. Based on an analysis of the largest 68 urbanized areas in the US from January 2002 to February 2010, this thesis finds that urbanized areas with either high population density or compact developments (urban form), and extensive urban containment policies (urban policy) are more sensitive to gasoline price changes, meaning that these areas show a greater ridership increase when motor fuel costs rise.

Proponents of land use policies can use the above findings to support compact development and growth management policy. Also, transportation planners who find effective ways to reduce greenhouse gas emissions from automobiles are able to consider not only transit investment but also environmentally sustainable urban form. Because each urban area has different transit systems, urban forms, and urban policies, planners must identify the best ways for their own situations and implement customized policies for the long run. Although land use measures are less effective than a gasoline price change in the short run, long term applications of land use policies will make a difference that pricing policies alone cannot achieve.

Table of Contents

1. Introduction

When we think about people's travel behaviors, it is controversial to argue that land use characteristics highly correlate to the amount of private driving, which is usually measured as Vehicle Miles Traveled (VMT). Some scholars argue that the effect of land use characteristics is too small to predict VMT, because the United States' urban developments are already highly dispersed. For this reason, they assert that land use policy would have a limited impact on the current travel behaviors (Giuliano, 1995).

Giuliano points out that land use measures such as job-housing balance cannot be a critical factor to explain travel behaviors in urban areas. Giuliano cites her previous empirical study which shows that "there is more 'excess' commuting close to the city center, where jobs and housing are balanced, than farther away, where they are not" (1995). Her finding conflicts with the common belief that transportation cost is critical in housing location choice and that modifying the cost by pursuing a better job-housing balance can achieve the goal of lowering individual automobile use. The basic idea underlying her assertion comes from the result of a survey done by the Los Angeles Times that showed that most people prefer living in remote places, and that they are willing to give up accessibility to items such as jobs and shopping places.

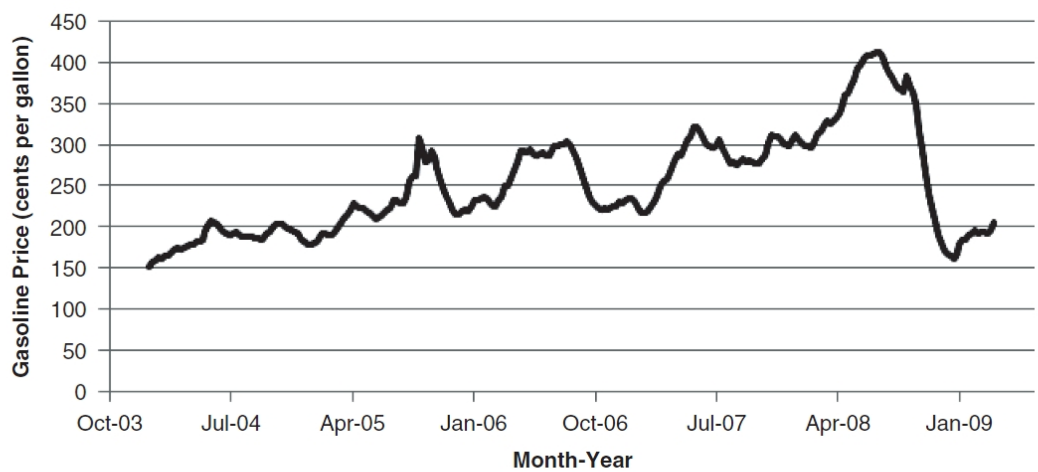
According to her arguments, land use policy is not extensive enough to change the US urban development pattern by, because people can still choose to live in suburbs and drive long distances due to the cheap price of private vehicle travel. In this context, as an alternative to land use policy, she suggests using price policy for reducing automobile use and promote public transit. "If the aim is to reduce environmental damage generated by automobiles, the effective remedy is to directly price and regulate autos and their use, not land use" (Giuliano, 1995).

In contrast, Cervero and Landis believe that land use policy still matters in making the US urban areas less auto-dependent. Cervero and Landis counter Giuliano's arguments stating that, although the relationship between land use and transportation may be weaker than several decades ago when the US metro areas were not sprawled much, there are still much evidence indicating that built environments are affected by travel demand (Cervero & Landis, 1995). Moreover, they mentioned that "synergistic relationships exist and underscores the need to package land use initiatives with other programs like restricted parking" (Cervero et al, 1995), when pointing out Giuliano's incorrect citation of an inappropriate index from the LUTRAQ (Land Use Transportation Air Quality Connection) study of Washington County, Oregon.

Cervero and Landis (1995) introduces an array of research indicating that those who live in job-housing balanced communities have less commuting time and distance, and that rail transit systems have a positive effect on nearby housing prices (people *do* pay premiums for houses with high accessibility). They conclude that, although land use policies alone are not "panacea for today's congestion, air quality, or social problems," they certainly are among the next best options. This is because "the true market-based pricing of transportation" such as "congestion fees and mandatory parking fees" cannot be easily implemented for their political unpopularity (Cevero et al, 1995).

After long debates between price policy supporters and land use initiatives proponents, Boarnet reconciles two seemingly contradicting arguments by providing both theoretical framework and empirical evidence. Theoretically-speaking, both approaches are not substitutes but rather complements. Price policies like the imposition of a Greenhouse Gas (GHG) emission fee will be more effective if the price elasticity of demand for driving is high. People are more willing to reduce driving with the same amount of the GHG fee as before, because the elasticity increases. In the meantime, land use and public transit policies can provide good alternatives to private driving, ending up with increasing the price elasticity of private automobile use (by making public transit cheaper and more attractive). In this context, when two policies are applied at the same time, we can expect higher outcomes in terms of reducing the volume of single drivers, though the specific effect of each policy should be based on empirical studies (Boarnet, 2010).

There are not many empirical studies that show a ”synergistic” relationship between pricing and planning strategies for reducing VMT. Guo, Agrawal, and Dill (2011) clearly show in several cases that both approaches do not substitute but rather complement each other. For example, they used survey data of the 130 households in Portland, OR, who participated either in flat-rate or varying-rate (lowest in peak hours, but highest in off-peak hours) pilot mileage-fee programs for four months. Their regression results present that people in varying-rate program reduced more VMT when they lived in denser and mixed-use communities, and that urban form variables have more explanatory power to VMT variation (in other words, urban form affects VMT more) under a varying-rate mileage-fee program than under a flat-rate one.



[Figure 1] U.S. national weekly regular gasoline price per gallon, January 2004 through April 2009 (Source: *U.S. Retail Gasoline Historical Prices.* Energy Information Administration, U.S. Department of Energy)

Heretofore, the researches which attempted to test the relationship between pricing and land use approaches have been done for VMT reduction, not for public transit demand. This research attempts to apply the above discussions to public transit ridership in the US. In the late 2000, a number of researchers in the field of transportation planning began to focus on the rise of transit ridership in major metro areas in the US, and its relationship to motor fuel price fluctuation. This is because record level gasoline prices seemed to force people to use mass transit for the first time.

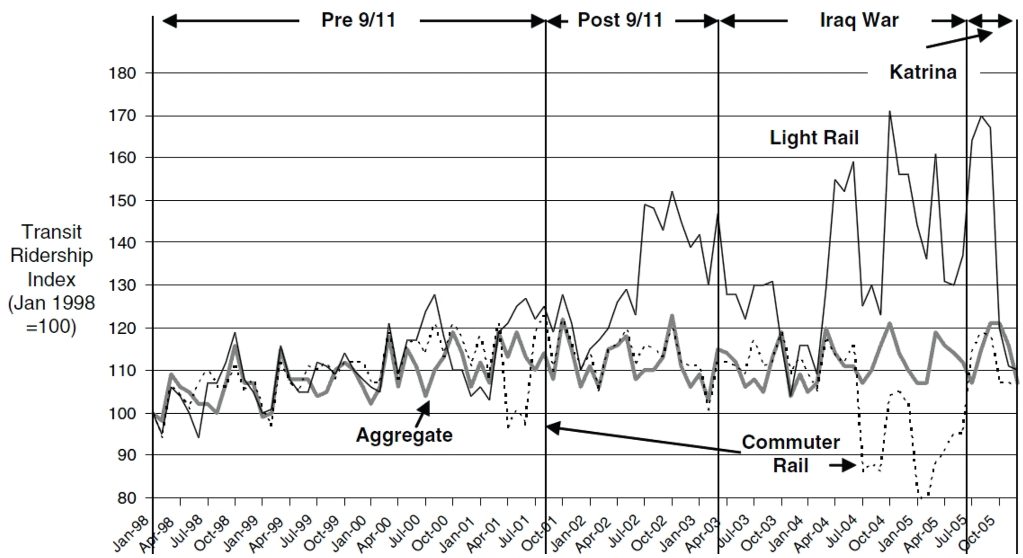
> *"U.S. cities are racing to cope with ever-increasing demand on public transportation as gas prices remain at record levels. High gas prices in recent months have had a considerable impact on commuters using public transportation, statistics show. Even regions that have traditionally resisted giving up cars and have limited access to mass transit are reporting a surge in public transportation use. From trains and trolleys to subways and buses, the growth encompasses all modes of travel, according to the American Public Transportation Association, a Washington D.C.-based industry group." (CNN.com, July 16 2008)*

Urban transportation researchers have two main questions: (1) whether or not gasoline prices statistically correlate to transit ridership, and (2) how much gasoline price fluctuation explains transit ridership change. The scope of their research covers both national and international, specific areas and systems, different transit modes, and even individual household levels (Bhat, Sen, & Elur, 2009; Chen, Varley, & Chane, 2010; Currie & Phung, 2007; Currie & Phung, 2008; Haire & machemehl, 2010; Maley & Weinberger, 2009; Mattson, 2008; Lane, 2009; Stover & Bae, 2011; Yanmaz-Tuzel & Ozbay, 2010). They found that although motor fuel price did have a statistically significant effect on transit ridership, the above two questions were too simple to achieve practical policy implications.

This research assumes (1) that of the price elasticity of gasoline to transit ridership varies in different US urban areas and (2) that the difference mainly comes from different urban form and land use policies of each area. From the perspective of the price policy versus land use initiatives debate, gasoline price fluctuation acts like price policy, while urban form and land use policy is the same as land use initiative in the debate. Though, unlike pricing strategy, we cannot directly control the gasoline price except by levying more tax on it, land use policies such as Compact Development and Growth Management plan can be pursued through local planning efforts. The analysis of this research is expected to reveal whether or not motor fuel price change and land use characteristics are mutually supportive in having more transit riders.

2. Literature Review

National and international analysis - This section will review the previous research about the relationship between gasoline price fluctuation and transit ridership change, and show their shortcomings in terms of policy implication. First of all, there are at least two notable instances of research in a national level analysis. Currie and Phung (2007) explored the cross-modal market effects of auto gasoline price on the demand of transit such as bus, heavy rail, commuter rail, and light rail in the US. They found that the US transit demand was low or medium in its price elasticity of motor fuel compared to European countries and Australia. The aggregate elasticity value for all transit mode is as low as 0.12, implying that the total US transit demand will increase by 1.2% for 10% increase of automobile gasoline price. Their first regression model simply uses monthly gasoline price and monthly dummy variables as explanatory variables, and their second model includes additional time dummies which indicate whether the month is before or after certain world events (9/11, Iraq War, and Hurricane Katrina). Through these methods, they found there are statistical differences in peoples' travel behavior during different time for different transit modes; riders of light rail systems were especially sensitive to changes in gasoline price.



[Figure 2] U.S. transit ridership index, January 1998 to October 2005 (Source: Currie & Phung(2007), Transit Ridership, Auto Gas Prices, and World Events New Drivers of Change? Transportation Research Record: Journal of Transportation Research Board, No.1992 pp3-10)

Another research conducted by Currie and Phung (2008) compared the price elasticity of motor fuel against transit ridership between the US and Australia. In this study, they focused on three Australian cities (Melbourne, Adelaide, and Brisbane) and added home-loan interest rates to their regression model because the rate may be a critical factor that determines the level of financial stress for households and by effect, travel mode as well. They conducted aggregate and disaggregate analyses for different transit modes, different time periods (peak and off-peak), and different travel distances. Interestingly, they linked the different price elasticities of motor fuel for the three cities to urban spatial and economic characteristics: they characterized two cities as “more suburban low-income ridership focus of Melbourne bus service and the strong commuter and CBD access functions of the Brisbane and Adelaide bus network” (Currie & Phung, 2008). However, their explanation was not based on statistical analysis but on simple description about the overall characteristics of cities, for which we cannot find proper policy implication.

Analysis of specific geographic area-Unlike two researches done by Curri and Phung, there are journal articles that focus on specific geographic areas in the US. These studies have more explanatory variables and utilize more detailed analysis to examine the relationship between automobile fuel cost and transit. The additional independent variables are population, economic condition, and level of transit service. Also, other studies attempted to understand not only the relationship between gasoline price and transit ridership but also other aspects of travel behavior by employing advanced statistical methods.

Maley and Weinberger (2009) analyzed a causal relationship between the unprecedented high gasoline price and transit ridership change of Philadelphia, PA. They divided transit services into two categories, Regional Rail and City Transit, and found that Regional Rail demand responds more to gasoline price change. Another point of their research was calculating price elasticity of motor fuel for different price ranges: from $3 to $4, from $4 to $5, and from $5 to $6. Interestingly, the elasticity for both systems increases as the price range increases.

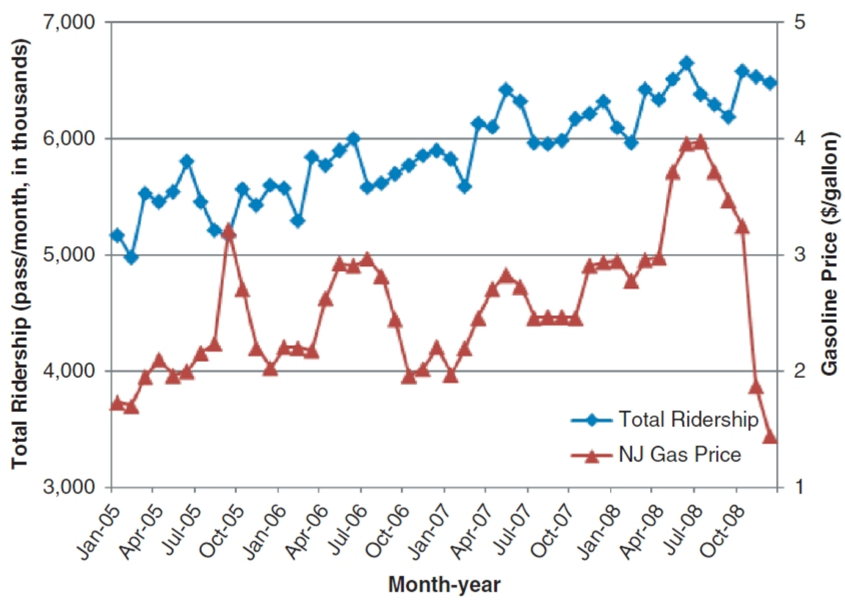
Stover and Bae (2011) showed a typical method when they did the research about eleven counties in Washington State. They added new independent variables such as transit service rate (adult fare price and vehicle revenue hours), economic condition (size of labor force and unemployment rate), and seasonal dummies (winter, spring, and summer). These are what people commonly now use in this field because they are easily obtained online (National Transit Database website and Bureau of Labor Statistics website) and other monthly level data are not available. For each county, they conduct Ordinary Least Squares, and Autoregressive model AR (1) and AR (2), and for eleven counties as a whole, they employed a panel data model with fixed-effect.

Interestingly, their research finds a variation of price elasticity of gasoline for different counties, and describes the different results by introducing different geographic and economic characteristics for each county. For example, four counties have no statistical relationship between gasoline price and transit ridership, and the authors point out their unique travel patterns as a main reason for no correlation.

However, this paper has serious shortcomings. The gasoline price is not available on the county level and, for this reason, four counties in their sample used the monthly price of the Seattle-Tacoma-Bremerton metropolitan area, while the remaining seven counties used the same price of Washington State. This non-differentiated gasoline price weakens the reliability of this research. Moreover, the authors explain different relationships between motor fuel cost and transit ridership of eleven counties by connecting it to their distinctive characteristics, but do not provide any statistical evidence. There should be limitations to such descriptive reasoning to help transportation planners respond with relevant policies.

Yanmaz-Tuzel and Ozbay (2010) analyzed New Jersey transit, “the nation’s largest public transit system by service area.” When estimating the price elasticity of gasoline on transit ridership, they select two times: (1) Hurricane Katrina in September 2005 and (2) the oil price increase in May 2008 to calculate the short-term and long-term impact of major motor fuel changes.

Their analysis is slightly different from the previous one, because it has a longer time period of analysis (from 1980 to 2008, instead of 2004 to 2008 in Stover and Bae(2011)), the addition of a gasoline price lagged variable, and the differentiation of short-term from long-term elasticity. Their advanced regression model reveals that there is a "(two to four months) elapse before travelers respond to gasoline price changes." Moreover, short-term price elasticity of gasoline for two years (one year before and after the big price changes) is found to be larger than medium-term price elasticity for four years) (two years before and after the big price changes).



[Figure 3] New Jersey Transit Ridership versus Gasoline Prices (Source: Yanmaz-Tuzel & Ozbay (2010), Impacts of Gasoline Prices on New Jersey Transit Ridership, Transportation Research Record: Journal of Transportation Research Board, No.2144 pp52-61)

However, the limitations of their research are clear. In regards to short-term and medium-term elasticity, all the data were not available when authors conducted the research, so their estimations were not accurate, i.e. for the spring 2008 event, they used data from May 2007 to December 2008 (1 year and 8 months) for short-term and May 2006 to December 2008 (3 years and 8 months) for long-term. All the necessary data are now available, so additional analysis would need to be done. In addition, in spite of their refined regression model, authors did not introduce new findings into this field. They find the same as what previous researchers say for other areas: there is a statistical relationship between gasoline price and transit ridership, but the size of the relationship is not large. Again, these findings alone cannot give us enough insights for policy improvement. We need to know more about when, where, and how the relationship between gasoline price and transit ridership change.

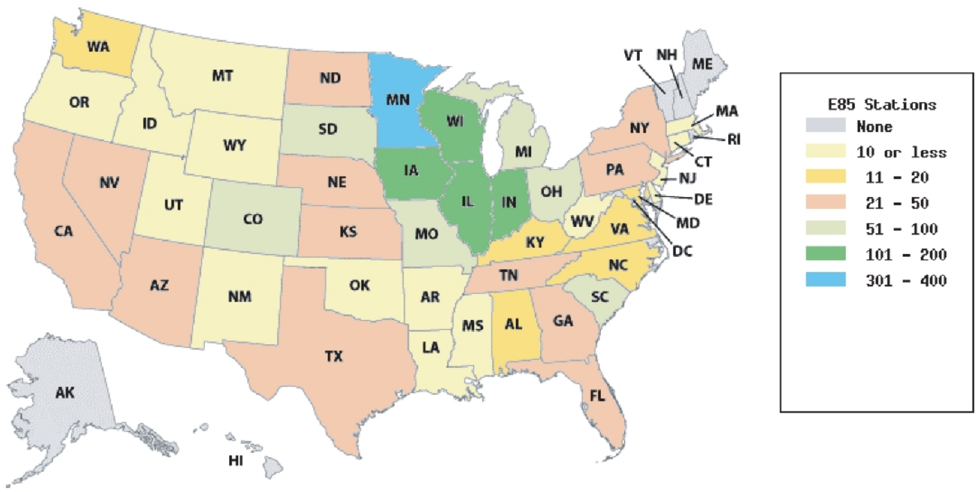
Chen, Varley, and Chen (2010) conducted one of the most in-depth researches in this field. Using New Jersey Transit data from January 1996 to February 2009, they analyzed three aspects about transit ridership change: (1) identifying various factors which can explain transit ridership change, (2) quantifying short-term and long-term effects of these factors on transit ridership, and (3) testing the hypothesis of “asymmetric behaviors in transit ridership in response to rises and falls in gasoline price” (Chen, Varley, & Chen, 2010). They found that long-term effects of gasoline prices, transit services, transit fares, and size of labor force were much higher than short-term effects. Also, they determined that “on the test of equality between a fall and a rise in gasoline price” “at the aggregate level, ridership seems to respond to rises in gasoline price, but not to falls” (Chen et al, 2010). Although they certainly added new findings, they still lacked a spatial dimension. Their analysis will be applied to New Jersey Transit well into the future; however, we do not know whether or not other areas in the US have the same characteristics in terms of the effect of motor fuel cost.

Analysis of multiple-areas- There is another area of research whose focus is not about a single area but about variation among multiple areas. Unfortunately, the analyses here do not provide statistically reliable evidence for what factors cause the variation. Instead, they attempt to provide descriptive explanations by introducing different socio-economic and urban-form characteristics of different areas. Common weakness of these research is their loose connection between a variation in price elasticity of gasoline and its possible causes; for each area, their explanations seem to make sense, but do not always help to understand other areas which are out of their research scope.

Lane (2010) studies the correlation of gasoline price fluctuation to transit ridership change for nine major US cities for monthly data from January 2002 to April 2008. After controlling transit service rates, seasonality, and inherent trending, he found that a small but statistically significant ridership fluctuation came from gasoline price changes.

His samples include Boston, Chicago, Cleveland, Denver, Houston, Los Angeles, Miami, San Francisco, and Seattle (he excluded New York, because NY has very different characteristics from other cities in terms of transit ridership and socio-economic situation). In fact, every month the Energy Information Administration website of the Department of Energy provides gasoline price information for the above cities. Although these are among the biggest in the US, they are not representatives for all US cities. Each and every city in his sample has unique characteristics which we cannot find from other cities. To find consistent patterns throughout the US, we need to have all, or at least "reasonably" selected areas, instead of choosing simply famous cities in an arbitrary way.

Haire and Machemehl (2010) write one of very few research papers that attempts to estimate regional differences in the relationship between fuel price growth and ridership change for U.S. transit systems. They analyzed the 254 transit systems in the U.S., and found that the effects of gasoline price increase on bus ridership varies in ten regions in the US, and that two regions (North Central and Midwest) showed the opposite-direction of change in numbers of bus riders compared to the other eight regions. As a potential explanation for their second finding, they pointed out that the North Central and Midwest regions consume a more corn-based alternative fuel such as ethanol and E85, which might have been used as an alternative for drivers who were "struggling with rising fuel prices" (Haire & Machemehl, 2010).



[Figure 4] Availability of E85 fueling stations (Sources: Haire & Machemehl (2010), Regional and Modal Variability in Effects of Gasoline Prices on U.S. Transit Ridership, Transportation Research Record: Journal of Transportation Research Board, No.2144 pp20-27)

However, when they chose bus systems among their initial 254 samples, they achieved only 82 samples in 10 different regions due to limited data availability. This means that five among ten regions had less than five bus systems individually, whereas only three regions had more than ten systems. For this reason, the regression analysis cannot be easily generalized. Also, they seem to premise that all the bus systems in one region would be similar to one another and different from other regions, but this assumption is too strong to be true. Even in one metro area, we can see huge variations among transit systems due to socio-economic and geographic differences. Thus, the paper cannot be thought to be fully elaborated or refined. Indeed, it shows a big picture of regional variations for the pattern of bus ridership change and their plausible cause, but more research would need to be done to get practical policy implications.

Methodological issues-There is a critical methodological issue in analyzing the relationship between gasoline price fluctuation and transit ridership change. The majority of literature used transit ridership as a dependent variable and gasoline price, transit service rate, and economic condition as explanatory variables. The problem is that-while transit demand (transit ridership) is a result of transit supply (transit service rates)-transit supply is also a function of transit demand. In other words, the relationship between demand and supply is not uni-directional but bi-directional. In this case, if we choose to use traditional regression models, the coefficients and variances for each variable will be biases, and thus cannot be trusted. This is called an endogenous problem in statistics. One of the common ways to deal with endogeneity is to introduce an instrument variable (IV) instead of transit service rates themselves.

Taylor, Miller, Iseki, and Fink (2009) raised this issue and used voting rates for the Democrat in the 2000 presidential election as an IV. They assume that “Democratic-leaning areas are more likely to support public expenditures on transit subsidies” (Taylor, Miller, Iseki, & Fink, 2009). By using predicted value of transit service rates though the IV, they could avoid the endogeneity between dependent and explanatory variables and could get unbiased statistical results.

However, their analysis is not a panel but a cross-section using 2000 data. Thus, although they insert gasoline price in their regression model, they did not check out the price elasticity of gasoline for each city to the city’s transit ridership. They were just trying to understand what factors could explain different transit ridership throughout the 265 US urbanized areas in 2000. However, when it comes to monthly level analysis, which is done to derieve the elasticity, voting rates are not available. Though their research does provide an appropriate insight for this research, additional modification to their model would need to be done.

In sum, there is no research which explores the relationship between automobile fuel price fluctuation and transit ridership change in the US from an inter-area perspective. A few previous researches give insights to inter-area analysis, but they fail to provide strong evidence. As pricing and land use strategies are expected to work together to reduce private driving, gasoline price change and land use initiatives are expected to be mutually supportive in attracting more people to public transit. To test this hypothesis throughout the US, an inter-area approach is critical.

For this reason, this research will analyze correlation of transit ridership, motor fuel price, and urban form and policy conditions for the 68 urbanized areas in the US, whose population is over 500,000 according to the US Census 2000. Analysis using monthly data is essential to derive price elasticity of gasoline price, a panel data approach will be needed to handle multiple areas for the same time period from January 2002 to December 2010, and endogeneity issues between transit supply and demand should be properly dealt with.

The remaining part of this research is organized as followings. In the 3rd chapter, information about data in this research will be provided. The 4th chapter will cover methodological issues and specify statistical models for testing the hypothesis of complementarity of price and land use characteristics. The regression results and interpretation will be explained in the 5th chapter. Finally, policy implications and conclusions will be discussed in the 6th chapter.

3. Data

Research area-This study focuses on the urbanized areas (UAs) whose population are over 500,000 according to the US Census 2000. There are seventy-two UAs in the US which meet this criterion. Among them four UAs are excluded for various reasons: San Juan, PR and Honolulu, HI (due to their different travel behaviors from other UAs), New Orleans, LA (due to its discontinuous travel trend caused by the Hurricane Katrina), and Mission Viejo, CA (there are no transit agencies that report their travel data to American Public Transportation Association). Lastly, sixty-eight UAs are analyzed for their relationship between gasoline price and ridership change, and the complementarity of price and land use characteristics.

US Census The data used in this research come from several sources: the US Census and American Community Survey, Bureau of Labor Statistics, National Transit Database, and two research papers about urban form and policy. The US Census 2000 and American Community Survey 2005 to 2010 provide basic socio-economic data for each analysis unit: population size, number of people who are enrolled in colleges or graduate schools, number of foreign-born people, number of those who are below each year's poverty line, and median household income. Each variable is converted to a logarithmic term and/or percent and then its correlation is tested against transit ridership or other explanatory variables. After these processes, two variables are left in the final statistical models.

lnppl a logarithmic term of population for each UA

lnpchedu a logarithmic term of percent of people in colleges or graduate schools.

lnpden a logarithmic term of population density (population/land area in square miles)

Bureau of Labor Statistics Monthly economic data for each UA or metropolitan area such as unemployment rate (for metro), average price of regular unleaded gasoline (for UA), and consumer price index (for UA) are collected. The following variables are used in regression models.

lnunempr a logarithmic term of monthly unemployment rate for each UA

lngpi a logarithmic term of gasoline price divided by consumer price index $log(\frac{GP_{it}}{CPI_t})$ GP_{it} : gasoline price for UA_i at time t, CPI_t : National Consumer Price Index for one gallon unleaded regular gasoline for time t

National Transit Database (NTD) Transit supply and demand data are downloaded from NTD website. According to the NTD, "all transit properties that are recipients of Urbanized Area Formula Grants from the Federal Transit Administration are required to report. However, transit properties that operate nine or fewer vehicles in peak service are eligible to receive a waiver that exempts them from the monthly reporting." Though monthly supply and demand data are available from January 2002 to April 2012, annual operating fund data are only from 2002 to 2010.

lnupt a logarithmic term of monthly Unlinked Passenger Trip (demand) for UA

lnvrmpt a logarithmic term of monthly Vehicle Revenue Miles (supply) per 1,000 for UA

lnopfp a logarithmic term of annual operating fund per capita for UA

Highway Statistics This gives annual data for road miles, freeway miles, estimation of freeway lane miles, and population. Because the boundary of urban areas in Highway Statistics may be different from the US Census, when dividing miles by population, this research uses population data from the same table in Highway Statistics e-book, not from the US Census. The above data are available from 2002 to 2008, and 2009 and 2010 data are not yet ready for public use.

lnfwylpt a logarithmic term of estimation of freeway lane miles per 1,000 people for UA

Compact Index Compact index in this research comes from Overall Sprawl Score in Ewing, Pendall, and Chen(2002)'s extensive research "Measuring Sprawl and its Impact Volume 1." The score is a result of combining four different indexes for Density, Mixed Use, Centeredness, and Street Connectivity. In fact, the naming of the score is counter-intuitive, because when a metro has compact developments, then its sprawl score is high: i.e. New York Primary Metropolitan Statistical Area (PMSA) has the highest score of 177.8, whereas Riverside-San Bernardino PSMA has the lowest score of 14.2. In its calculating process, population size is also considered, so if a sprawl score is high for a metro, then it means that the metro's sprawl is less severe than it would be based on its population.

lnsprl a logarithmic term of the overall sprawl index from Ewing, et al (2002)

[Table 1] Sprawl Ratings for 2000 in Order of Increasing Sprawl (Ewing et al, 2002)

Urbanized Area Name	Metropolitan Name	Sprawl	Connectivit	Centered	Mixed Use	Density
New York--Newark, NY--NJ--CT Urbaniz	New York, NY PMSA	177.8	154.9	144.6	129.8	242.5
Providence, RI--MA Urbanized Area	Providence-Pawtucket-Woonsocket, RI	153.7	135.9	140.3	140.5	99.1
San Francisco--Oakland, CA Urbanized	San Francisco, CA PMSA	146.8	139.8	128.6	107.3	155.2
Omaha, NE--IA Urbanized Area	Omaha, NE-IA MSA	128.4	104.6	132.3	119.3	96.4
Boston, MA--NH--RI Urbanized Area	Boston-Lawrence-Salem-Lowell-Brockt	126.9	119.1	109.4	124.4	113.6
Portland, OR--WA Urbanized Area	Portland, OR PMSA	126.1	128	121.8	102.3	101.3
Miami, FL Urbanized Area	Miami-Hialeah, FL PMSA	125.7	136.4	92.7	104.7	129.1
Denver--Aurora, CO Urbanized Area	Denver, CO PMSA	125.2	125.7	108.9	115.7	103.7
Albuquerque, NM Urbanized Area	Albuquerque, NM MSA	124.5	117.8	124	103.7	97
Allentown--Bethlehem, PA--NJ Urbanize	Allentown-Bethlehem-Easton, PA-NJ M	124	131	91.7	133.4	86.2
Springfield, MA--CT Urbanized Area	Springfield, MA NECMA	122.5	87.3	148.6	115.7	86.3
Chicago, IL--IN Urbanized Area	Chicago, IL PMSA	121.2	134.9	85.8	115.1	142.9
Buffalo, NY Urbanized Area	Buffalo, NY PMSA	119.1	70.6	135.2	124.7	102.1
Milwaukee, WI Urbanized Area	Milwaukee, WI PMSA	117.3	93.9	117.7	117.9	101.4
El Paso, TX--NM Urbanized Area	El Paso, TX MSA	117.2	102.3	119.5	103.4	100.1
Baltimore, MD Urbanized Area	Baltimore, MD MSA	115.9	105.2	115.6	106.8	104.3
Philadelphia, PA--NJ--DE--MD Urbanize	Philadelphia, PA-NJ PMSA	112.6	113	95.9	119.5	114.7
Phoenix--Mesa, AZ Urbanized Area	Phoenix, AZ MSA	110.9	107.2	92.6	116	106.8
Salt Lake City, UT Urbanized Area	Salt Lake City-Ogden, UT MSA	110.9	117	93.8	103.2	99.5
Austin, TX Urbanized Area	Austin, TX MSA	110.3	94.4	115.8	111.9	89
Fresno, CA Urbanized Area	Fresno, CA MSA	110.3	73	112.6	130.1	93.5
San Jose, CA Urbanized Area	San Jose, CA PMSA	109.7	125.2	93.9	96.6	124.8
Tucson, AZ Urbanized Area	Tucson, AZ MSA	109.1	88	106.4	121.8	90.4
San Antonio, TX Urbanized Area	San Antonio, TX MSA	107.8	103	108.4	100.6	95
Toledo, OH--MI Urbanized Area	Toledo, OH MSA	107.2	77.6	112.2	119.6	91.3
New Haven, CT Urbanized Area	New Haven-Waterbury-Meriden, CT NE	107	86.5	78.9	144.3	91.6
Akron, OH Urbanized Area	Akron, OH PMSA	105.9	84.2	119.5	118.7	86.8
Pittsburgh, PA Urbanized Area	Pittsburgh, PA PMSA	105.9	124.2	104.5	86.8	90.4
Las Vegas, NV Urbanized Area	Las Vegas, NV MSA	104.7	108.8	99.8	80.1	110
Sacramento, CA Urbanized Area	Sacramento, CA MSA	102.6	98.4	87.4	110.9	99.1
San Diego, CA Urbanized Area	San Diego, CA MSA	101.9	106	74.4	105.4	113.4
Los Angeles--Long Beach--Santa Ana, C	Los Angeles-Long Beach, CA PMSA	101.8	123.3	72.4	123.1	151.5
Seattle, WA Urbanized Area	Seattle, WA PMSA	100.9	117.1	98	79.4	103.6
Tulsa, OK Urbanized Area	Tulsa, OK MSA	99.1	96.2	115	88	82.7
Orlando, FL Urbanized Area	Orlando, FL MSA	96.4	120.6	103.5	60.8	93.8
Cincinnati, OH--KY--IN Urbanized Area	Cincinnati, OH-KY-IN PMSA	96	85.4	110.2	95.8	88.8
Minneapolis--St. Paul, MN Urbanized Are	Minneapolis-St. Paul, MN-WI MSA	95.9	87.7	107.8	94.7	94.7
Grand Rapids, MI Urbanized Area	Grand Rapids, MI MSA	95.2	63.7	110.3	115.7	82.7
St. Louis, MO--IL Urbanized Area	St. Louis, MO-IL MSA	94.5	106	76.2	107.4	90.3
Indianapolis, IN Urbanized Area	Indianapolis, IN MSA	93.7	84.5	102.4	96.2	89.3
Houston, TX Urbanized Area	Houston, TX PMSA	93.3	95.6	87	110.1	95.3
Memphis, TN--MS--AR Urbanized Area	Memphis, TN-AR-MS MSA	92.2	76.5	104.2	97	88.9
Cleveland, OH Urbanized Area	Cleveland, OH PMSA	91.8	66.8	100.9	107.4	99.7
Jacksonville, FL Urbanized Area	Jacksonville, FL MSA	91.6	104.6	102.1	72.9	85.6
Kansas City, MO--KS Urbanized Area	Kansas City, MO-KS MSA	91.6	88.8	89	100	90.9
Columbus, OH Urbanized Area	Columbus, OH MSA	91.1	97.2	101.5	76.5	91.5
Washington, DC--VA--MD Urbanized Are	Washington, DC-MD-VA MSA	90.8	98	97.8	78.7	106.9
Birmingham, AL Urbanized Area	Birmingham, AL MSA	88	104	112.5	62.2	77.1
Tampa--St. Petersburg, FL Urbanized A	Tampa-St. Petersburg-Clearwater, FL M	86.3	133.6	51.9	80	93.6
Oklahoma City, OK Urbanized Area	Oklahoma City, OK MSA	85.6	69.1	95.6	101.3	84.5
Hartford, CT Urbanized Area	Hartford-New Britain-Middletown-Bristol	85.2	59.6	84.6	119.4	86.3
Albany, NY Urbanized Area	Albany-Schenectady-Troy, NY MSA	83.3	73.2	98.5	89.3	82.9
Detroit, MI Urbanized Area	Detroit, MI PMSA	79.5	93	63	102.5	97.3
Dallas--Fort Worth--Arlington, TX Urbani	Dallas, TX PMSA	78.3	90.2	81.1	82.6	99.5
Rochester, NY Urbanized Area	Rochester, NY MSA	77.9	37.2	120.7	82.3	91.4
Bridgeport--Stamford, CT--NY Urbanize	Bridgeport-Stamford-Norwalk-Danbury,	68.4	80.7	94.8	137.5	92.5
Atlanta, GA Urbanized Area	Atlanta, GA MSA	57.7	57	82.3	73.7	84.5
Raleigh, NC Urbanized Area	Raleigh-Durham, NC MSA	54.2	80.8	77.2	39.5	76.2
Riverside--San Bernardino, CA Urbanize	Riverside-San Bernardino, CA PMSA	14.2	80.5	41.4	41.5	93.5

Urban Containment Policy Wassmer (2006) analyzed the relationship between different forms of urban containment policies and land consumption of urban areas. He categorized all urban containment policies into four by using two dimensions: strong versus weak, and restrictive versus accommodating. “Based also upon cluster analysis, a “strong” urban containment program facilitated land-use plans that ensured current adequate land supply, offered affordable housing, provided for adequate infrastructure, and promoted land conservation. But a ‘‘weak’’ containment program failed in most of these categories” (Wassmer, 2006) Moreover, a “restrictive” policy means not accommodating projected future growth, while an “accommodating” one supports future growth needs.

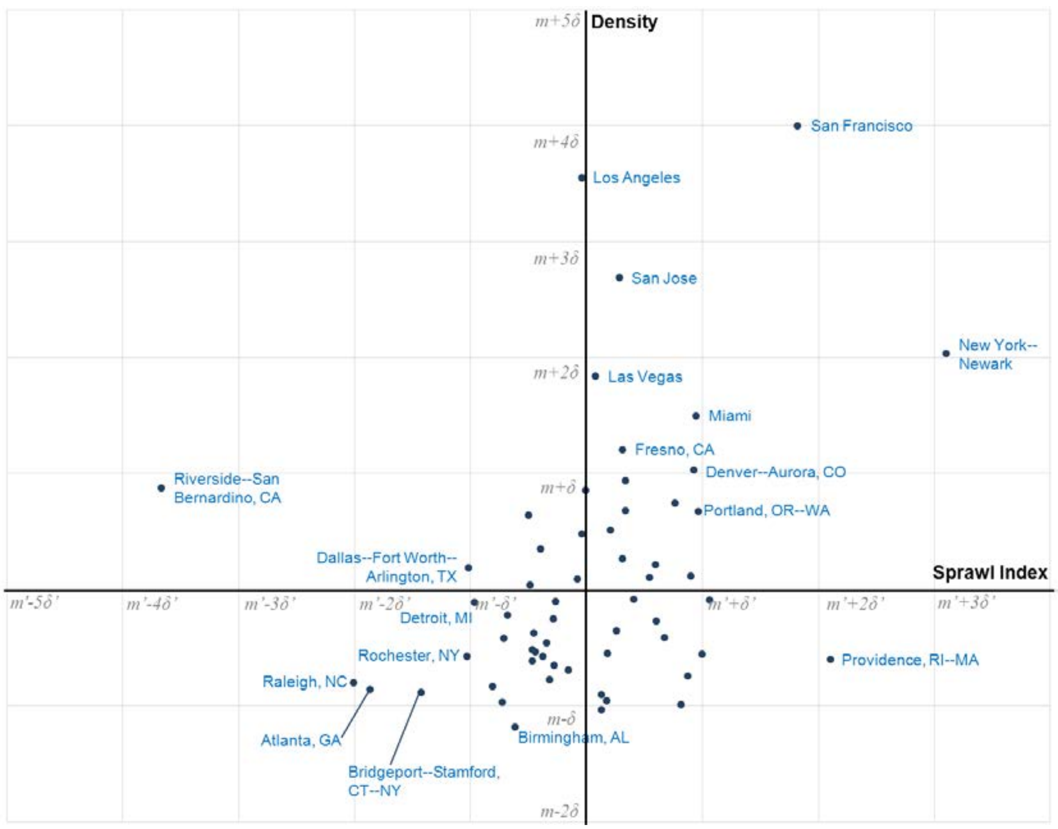
In this context, there is another aspect of these policies: a geographic range of them. When an urban area has a non-local containment policy, the policy is not just applied to some part of the area, but also beyond “just municipal or county jurisdictional boundaries” (Wassmer, 2006). This research employs a dummy variable that shows whether an urbanized area has non-local containment policies regardless of these policies’ characteristics. Public transit systems can cover a number of municipalities, so if a containment policy is only effective in several jurisdictions, its impact on transit ridership would be insignificant.

nlocal dummy variable (1: UA with non-local containment policy, 0: otherwise)

[Table 2] Summary statistics of sample data

Variable	Number	Mean	Std Dev	Minimum	Maximum
lnupt	7,308 [a]	14.6413	1.6254	7.7870	19.7389
lngpi	7,308 [a]	0.1039	0.2505	-0.6301	0.7287
*clngpi**	7,308 [a]	0.0000	0.2505	-0.7340	0.6249
*postpeak**	7,308 [a]	0.2885	0.4531	0	1
*trend**	7,308 [a]	54.6773	31.1431	1	108
lnppl	7,308 [a]	14.1394	0.8133	13.0788	16.7304
pchedu	7,308 [a]	7.1912	1.3428	3.0811	10.9392
lnunempr	7,308 [a]	1.7654	0.3390	0.7885	2.9232
lnfwylpt	7,308 [a]	-0.4493	0.3358	-2.2007	0.2791
lnvrmpt	7,308 [a]	6.9932	0.7858	-15.6459	8.7390
lnpden	7,308 [a]	7.9293	0.3373	7.4021	8.9592
*clnpden**	7,308 [a]	0.0000	0.3373	-0.5273	1.0299
*iclnpden**	7,308 [a]	0.0141	0.0842	-0.4854	0.6318
lnopfp	7,308 [a]	4.4990	0.9005	0.1827	6.7766
*clnopfp**	7,308 [a]	0.0000	0.9005	-4.3163	2.2776
*iclnopfp**	7,308 [a]	0.0384	0.2255	-1.5787	2.0740
lnsprl	6,372 [b]	4.5931	0.3267	2.6532	5.1807
*clnsprl**	6,372 [b]	-0.0121	0.3267	-1.9519	0.5755
*iclnsprl**	6,372 [b]	0.0029	0.0804	-1.1053	1.0643
nlocal	7,308 [a]	0.1921	0.3940	0	1
*inlocal**	7,308 [a]	0.0038	0.1090	-0.6385	0.6049

*: These variables will be explained in the next chapter (c-: centered, ic-: interation with *clngpi*)
a: Concord, CA does not have operating fund information from 2002 to 2004. (68*12*9)-(1*12*3)=7,308
b: Only 59 among 68 Urbanized Areas in the sample have the Overall Sprawl Score. (59*12*9)=6,372



[Figure 5] Sprawl index and population density of 68 urbanized areas (m=mean of density, δ=standard deviation of density, m'=mean of sprawl index, δ'=standard deviation of sprawl index)

4. Methodology and Model Specification

4.1 Methodology

This research basically employs an econometric approach; however, it will not follow the popular approaches in the previous researches, because they have a serious statistical problem and are designed to test the hypothesis of complementarity of gasoline price and land use characteristics. Firstly, the existing research does not deal with the endogenous relationship between supply and demand. Endogeneity can occur in several situations, and one of them is when independent and dependent variable are affected simultaneously. In this research, transit ridership can be determined by transit service rate; however, transit supply can also be a function of numbers of transit riders. In this situation, the Ordinary Least Square analysis will give biased results. (For a detailed discussion, refer to Taylor et al (2009)) As a solution for this problem, this research will adopt two-stage least squares (2SLS), where instrument variables (IV) are used.

Another shortcoming of the literature is that their statistical models are not designed for comparing different geographic units. Most of the existing research focuses on a single area, and even if it has multiple geographic areas in their analysis, it does not consider unobserved time-invariant effects, mainly because their data are not dealt with as a panel. In fact, Stover and Bae (2011) and Mattson (2008) employed fixed-effect models as part of their analyses. However, both researches did not include land use variables in their fixed-effect models, so their analyses could not provide policy implications related to urban form and land use policy. Moreover, they ignored the endogeneity problem between transit supply and demand, and thus, their estimations were biased.

This research chooses the one-way fixed effect model as the best one for several reasons: (1) it considers each urbanized area has specific characteristics, (2) the selection of urbanized areas based on their population makes the random effect model less appropriate, and (3) there is less likely to be area-invariant time-specific effects in transit ridership.

4.2 Model specification

Based on the previous research and availability of relevant data, this research sets up the following statistical models for testing the hypothesis. Before analyzing with the most advanced model, basic models are also employed to get an idea of the statistical limitations. These limitations solved in the advanced models. The four statistical approaches are (1) pooled Ordinary Least Square (OLS), (2) pooled Two-Stage Least Square (2SLS), (3) One-Way Fixed-Effect, and (4) Two-Stage One-Way Fixed Effect models.

(1) $\ln\left(riders_{i,t}\right) = \beta_0 + \ \beta_1 \,\mathrm{cln}\left(\frac{\mathrm{GP}_{i,t}}{\mathrm{CPI}_t}\right) + \beta_2(trend_t) + \beta_3(postpeak_t) + \beta_4 \ln\left(population_{i,t}\right) +$

$\beta_5 \ln\left(percent\ of\ students_{i,t}\right) + \beta_6 \ln\left(unempolyment\ rate_{i,t}\right) +$

$\beta_7 \ln\left(freeway\ lane\ miles\ perthousand\ _{i,t}\right) + \beta_8 \mathrm{cln}\left(operating\ fund\ per\ capita_{i,t}\right) +$

$\beta_9 \mathrm{icln}\left(operating\ fund\ per\ capita_{i,t}\right) + \ \beta_{10} \mathrm{cln}\left(population\ density_{i,t}\right) +$

$\beta_{11} \mathrm{icln}\left(population\ density_{i,t}\right) + \beta_{12} \mathrm{cln}\left(overall\ sprawl\ score_{i,t}\right) +$

$\beta_{13} \mathrm{icln}\left(overall\ sprawl\ score_{i,t}\right) + \beta_{14}(nlocal_i) + \beta_{15}(inlocal_i) + (\ \sum_{j=1}^{4} \gamma_j M_{i,t} +$

$\sum_{j=5}^{12} \gamma_j M_{i,t}\ ) + \epsilon_{i,t}$

(2) First Stage:

$$\ln\left(vehicle\ revenue\ miles\ per\ thousand_{i,t}\right)$$

$$= \beta_0' + \beta_1' \ln\left(population_{i,t}\right) + \ \beta_2' cln\left(operating\ fund\ per\ capita_{i,t}\right)$$

$$+ \ (\sum_{j=1}^{4} \gamma_j' M_{i,t} + \sum_{j=5}^{12} \gamma_j' M_{i,t}\ ) + \epsilon_{i,t}'$$

Second Stage:

$\ln(riders_{i,t}) = \beta_0 + \boldsymbol{\beta_1 \ln(predicted\ vehicle\ revenue\ miles\ per\ thousand_{i,t})} +$

$\beta_2\ \mathrm{cln}\left(\frac{\mathrm{GP}_{i,t}}{\mathrm{CPI}_t}\right) + \beta_3(trend_t) + \beta_4(postpeak_t) + \beta_5 \ln(population_{i,t}) +$

$\beta_6 \ln(percent\ of\ students_{i,t}) + \beta_7 \ln(unempolyment\ rate_{i,t}) +$

$\beta_8 \ln(freeway\ lane\ miles\ per\ thousand\ _{i,t}) + \beta_9 \mathrm{cln}(operating\ fund\ per\ capita_{i,t}) +$

$\beta_{10} \mathrm{icln}(operating\ fund\ per\ capita_{i,t}) + \ \beta_{11} \mathrm{cln}(population\ density_{i,t}) +$

$\beta_{12} \mathrm{icln}(population\ density_{i,t}) + \beta_{13} \mathrm{cln}(overall\ sprawl\ score_{i,t}) +$

$\beta_{14} \mathrm{icln}(overall\ sprawl\ score_{i,t}) + \beta_{15}(nlocal_i) + \beta_{16}(inlocal_i) + (\ \sum_{j=1}^{4} \gamma_j M_{i,t} +$

$\sum_{j=5}^{12} \gamma_j M_{i,t}) + \epsilon_{i,t}$

(3) $\ln(riders_{i,t}) = \beta_0 + \ \beta_1\ \mathrm{cln}\left(\frac{\mathrm{GP}_{i,t}}{\mathrm{CPI}_t}\right) + \beta_2(trend_t) + \beta_3(postpeak_t) + \beta_4 \ln(population_{i,t}) +$

$\beta_5 \ln(percent\ of\ students_{i,t}) + \beta_6 \ln(unempolyment\ rate_{i,t}) +$

$\beta_7 \ln(freeway\ lane\ miles\ perthousand\ _{i,t}) + \beta_8 \mathrm{cln}(operating\ fund\ per\ capita_{i,t}) +$

$\beta_9 \mathrm{icln}(operating\ fund\ per\ capita_{i,t}) + \ \beta_{10} \mathrm{cln}(population\ density_{i,t}) +$

$\beta_{11} \mathrm{icln}(population\ density_{i,t}) + \beta_{12} \mathrm{cln}(overall\ sprawl\ score_{i,t}) +$

$\beta_{13} \mathrm{icln}(overall\ sprawl\ score_{i,t}) + \beta_{14}(nlocal_i) + \beta_{15}(inlocal_i) + (\ \sum_{j=1}^{4} \gamma_j M_{i,t} +$

$\sum_{j=5}^{12} \gamma_j M_{i,t}\ ) + \boldsymbol{\alpha_i} + \epsilon_{i,t}$

(4) First Stage:

$$\ln(vehicle\ revenue\ miles\ per\ thousand_{i,t})$$

$$= \beta'_0 + \beta'_1 \ln(population_{i,t}) + \ \beta'_2 cln(operating\ fund\ per\ capita_{i,t})$$

$$+ \left(\sum_{j=1}^{4} \gamma'_j M_{i,t} + \sum_{j=5}^{12} \gamma'_j M_{i,t} \right) + \boldsymbol{\alpha'_i} + \epsilon'_{i,t}$$

Second Stage:

$$\ln(riders_{i,t}) = \beta_0 + \boldsymbol{\beta_1}\mathbf{ln}(\boldsymbol{predicted\ vehicle\ revenue\ miles\ per\ thousand_{i,t}}) + \beta_2\,\mathrm{cln}\left(\frac{\mathrm{GP}_{i,t}}{\mathrm{CPI}_t}\right) + \beta_3(trend_t) + \beta_4(postpeak_t) + \beta_5\ln(population_{i,t}) + \beta_6\ln(percent\ of\ students_{i,t}) + \beta_7\ln(unempolyment\ rate_{i,t}) + \beta_8\ln(freeway\ lane\ miles\ perthousand\ _{i,t}) + \beta_9\mathrm{cln}(operating\ fund\ per\ capita_{i,t}) + \beta_{10}\mathrm{icln}(operating\ fund\ per\ capita_{i,t}) + \beta_{11}\mathrm{cln}(population\ density_{i,t}) + \beta_{12}\mathrm{icln}(population\ density_{i,t}) + \beta_{13}\mathrm{cln}(overall\ sprawl\ score_{i,t}) + \beta_{14}\mathrm{icln}(overall\ sprawl\ score_{i,t}) + \beta_{15}(nlocal_i) + \beta_{16}(inlocal_i) + \left(\sum_{j=1}^{4}\gamma_j M_{i,t} + \sum_{j=5}^{12}\gamma_j M_{i,t}\right) + \boldsymbol{\alpha_i} + \epsilon_{i,t}$$

Centering of explanatory variables Means of several variables were centered on zero to make regression results easy to interpret. This is done by simply subtracting a mean of each variable from original observations. For example, if the mean of *ln(opfp$_i$)* is *m*, then its centered variable *cln(opfp$_i$)* would be *ln(opfp$_i$)* – *m*. In this way, for all observations, the independent variable is reduced by *m*. The coefficients of both *ln(opfp$_i$)* and *cln(opfp$_i$)* are exactly the same as each other, but their interpretations are different.

Before centering of an independent variable, its coefficient is the amount of a dependent variable change when the independent variable changes by one unit. However, as for the centered independent variable, its coefficient is the amount of the dependent variable change when the independent variable changes by one unit *compared to average*. Because the mean of the centered independent variable is 0 instead of *m*, the coefficient can be easily understood in relation to the explanatory variable's mean. We can see the merit of this centering process, when regression results will be interpreted.

Interaction with gasoline price There are four 'interaction' independent variables in the model (4), which start its name with '*i*': ***i****clnopfp*, ***i****clnpden*, ***i****clnsprl*, and ***i****nlocal*. These variables are made by multiplying *clngpi* (a centered logarithmic term of monthly gasoline price) and each explanatory variable. These variables are included to check whether there are “additional” transit ridership changes caused by the interaction between gasoline price and operating fund, population density, overall sprawl score, and non-local urban containment policy. If these interaction variables are statistically significant and positive, then the urban areas with higher values of the above four explanatory variables will gain “additional” transit riders when gasoline price goes up. The interaction variables are critical to test the complementarity of gasoline price fluctuation and transportation policy and land use initiatives such as compact development and urban growth management plans. However, if we find these interaction variables to be neither statistically significant nor negative, then it is also difficult to conclude that pricing and land use strategies are mutually supportive in travel behaviors of using public transportation.

5. Regression Results

5.1 Pooled Ordinary Least Square

As a first step, this section introduces the results of the simplest regression model, a pooled Ordinary Least Squares. This model treats all the observations of 68 Urbanized Areas from January 2002 to December 2010 as coming from the same dataset. In other words, a pooled OLS ignores both the endogeneity problem and time-invariant area-specific fixed effects for each urbanized area. Thus, standard errors of coefficients are biased and coefficients themselves are not correct due to omitted variables of Urbanized Area dummies.

Employing a simple model delivers a basic understanding about the relationship between transit ridership and explanatory variables and how advanced models differ from simple but wrong ones. Regardless of the appropriateness of the model, if certain variables show similar relationships with transit ridership throughout various models, we can assume that these variables strongly correlate to the number of transit riders and, thus, conclude controlling them will produce better outcomes from the perspective of public transit policy.

The following table shows four different models based on the pooled OLS approach. There are several variables that significantly correlate to transit ridership in all four models.

clngpi If gasoline price goes up by 10%, then transit ridership increases by 1.2% (Model C) to 2.1% (Model B). These numbers do not differ from what other researchers have found in this field. For all the US transit systems, Currie and Phung (2007) found the elasticity of gasoline price to transit ridership ranges from 0.10 to 0.12 from 1998 to 2005 meaning that a 10% gasoline price increase would bring about higher transit ridership by 1 to 1.2%. Chen, Varley, and Chen (2010) mentioned that previous research reported the gasoline price elasticity to range from 0.08 to 0.80, and that most of these numbers fall between 0.10 and 0.30.

trend Interestingly, the coefficients of the trend variable are negative in all four OLS models, showing that as one month passes by, there was a transit ridership loss of 0.005 to 0.006%. This number does not seem to be high enough to be realistic, even if after ten years the decrease is 0.5% to 0.6%, all else being equal. This implies that the recent increase in transit riders is related to other factors, and is not the sign of an overall change in the US travel behaviors.

[Table 3] Pooled OLS estimates of Unlinked Passenger Trip (UPT)

Dep. Variable: *lnupt*	A		B		C		D	
Number of UAs	68		59		68		59	
Time Series Length	108		108		108		108	
R-Square	0.9545		0.9612		0.9546		0.9616	
Explanatory Variable	Estimate (t Value)	Sig.	Estimate (t Value)	Sig.	Estimate (t Value)	Sig.	Estimate (t Value)	Sig.
Intercept	-3.66229 (-28.49)	***	-4.12073 (-34)	***	-3.72856 (-28.74)	***	-4.20139 (-34.4)	***
clngpi	0.148939 (4.77)	***	0.206523 (7.14)	***	0.122005 (3.79)	***	0.178366 (6)	***
trend	-0.00555 (-15.33)	***	-0.00529 (-15.66)	***	-0.00554 (-15.31)	***	-0.00531 (-15.73)	***
Postpeak	0.059471 (2.87)	***	0.048815 (2.53)	**	0.0605 (2.93)	***	0.05281 (2.74)	***
lnppl	1.233969 (173.51)	***	1.259157 (195.2)	***	1.237678 (172.45)	***	1.264731 (194.22)	***
lnpchedu	0.621499 (24.92)	***	0.593137 (23.12)	***	0.63073 (25.21)	***	0.600101 (23.39)	***
lnunempr	-0.06994 (-3.9)	***	-0.00244 (-0.14)		-0.07379 (-4.12)	***	-0.00885 (-0.52)	
lnfwylpt	-0.10014 (-6.95)	***	-0.19303 (-13.23)	***	-0.08477 (-5.61)	***	-0.16969 (-11.17)	***
clnopfp	0.880355 (142.57)	***	0.724473 (112.94)	***	0.877141 (141.01)	***	0.718383 (110.8)	***
iclnopfp	-0.00989 (-0.45)				-0.01195 (-0.55)		-0.05883 (-2.97)	***
clnpden	-0.14507 (-7.96)	***			-0.14726 (-8.09)	***		
iclnpden	0.243509 (4.1)	***			0.225128 (3.77)	***		
clnsprl			0.130856 (10.54)	***			0.146915 (11.6)	***
iclnsprl			0.179615 (3.78)	***			0.243617 (4.85)	***
nlocal					0.044188 (3.93)	***	0.056961 (5.61)	***
inlocal					0.117842 (2.79)	***	0.146029 (3.76)	***
m01	-0.03099 (-1.51)		-0.024 (-1.26)		-0.03123 (-1.53)		-0.02456 (-1.29)	
m02	-0.05315 (-2.61)	***	-0.04835 (-2.55)	**	-0.05344 (-2.63)	***	-0.04899 (-2.6)	***
m03	0.037349 (1.86)	*	0.041819 (2.24)	**	0.037106 (1.85)	*	0.04117 (2.21)	**
m04	0.004916 (0.25)		0.009479 (0.51)		0.004638 (0.23)		0.00879 (0.48)	
m06	-0.04164 (-2.08)	**	-0.04839 (-2.6)	***	-0.04133 (-2.07)	**	-0.04795 (-2.58)	***
m07	-0.04796 (-2.4)	**	-0.05623 (-3.02)	***	-0.0476 (-2.38)	**	-0.05574 (-3.01)	***
m08	0.007665 (0.38)		-0.00036 (-0.02)		0.00787 (0.39)		-0.00025 (-0.01)	
m09	0.054736 (2.74)	***	0.056019 (3.02)	***	0.054761 (2.75)	***	0.055788 (3.01)	***
m10	0.107184 (5.36)	***	0.114271 (6.14)	***	0.106966 (5.36)	***	0.113579 (6.13)	***
m11	0.018348 (0.91)		0.024945 (1.33)		0.017962 (0.89)		0.024044 (1.29)	
m12	-0.01831 (-0.9)		-0.00987 (-0.52)		-0.0188 (-0.92)		-0.01093 (-0.58)	

*** indicates significant at 1 percent level, ** indicates significant at 5 percent level, * indicates significant at 10 percent level

postpeak The national level gasoline price was the highest in June 2008, so the variable was included in the models to test whether there was a difference of people's travel mode choice before and after they experienced the historically high price. All four coefficients of the trend variable are positive, indicating that there were more transit riders after June 2008 by 5% to 6%, all else being equal. Additional research questions including "how long does this different travel behavior last after the peak month" help understand people's travel pattern, but the question is beyond the scope of this research due to the lack of necessary data.

lnppl, lnpchedu, and lnunempr Population size has a positive correlation to unlinked passenger trip, and college and graduate students are found to contribute transit ridership. Interestingly, the effect of monthly unemployment rate is not consistent; for two models, it was statistically significant, but for the other two, it was not. However, its sign is negative in all four models, implying that a poor economic condition results in a low level of travel.

lnfwylpt The estimated freeway lane mile per 1,000 people is found to be "negatively" correlated to the number of transit riders. If an area has higher freeway lane miles per 1,000 people, then the people of the area are less likely to take buses or subways, since congesting is less likely. In this context, making driving expensive, or at lease making its price rational, by both pricing and land use policy would be a key to promote more transit ridership.

As explained before, the first question of this research is whether the gasoline price statistically influences transit ridership in urbanized areas. The second question is whether different characteristics of each urbanized area make a difference in the relationship between gasoline price and transit ridership. If different transit policies, urban forms, and urban policies are found to correlate to transit ridership, and they add to the gasoline price effect, then we can get another strong rationale for supporting appropriate urban forms and policies. In this context, the following variables are selected for answering the second question.

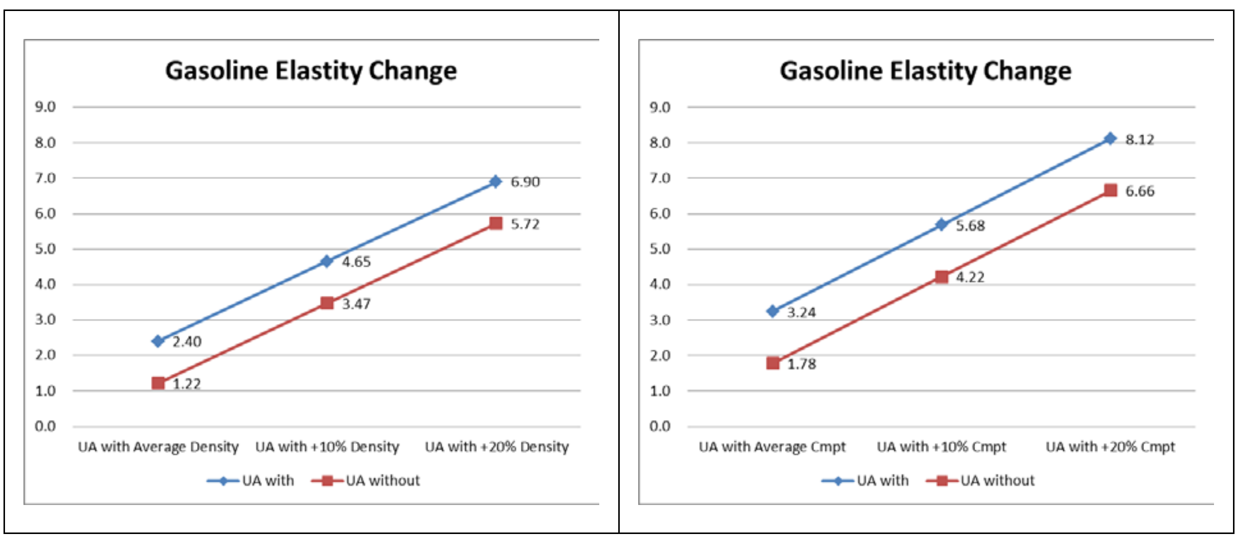
clnopfp and iclnopfp If the operating funds per capita increase by 10%, then there would be 7.2% to 8.8% more transit riders. Although the interaction between operating fund and gasoline price is negative and contrary to the prior expectation, their coefficients are not significant in two of three models. In other words, when an urbanized area has high amount of operating funds, its citizens take public transit more because of the amount of fund.

clnpden and iclnpden Although it is reasonable to assume that people in denser areas would take more public transit, it is not statistically proven yet in the above OLS models. Even if an area has high population density, if it does not have decent transit systems, its density will prevent people from taking more public transit. Of course, urban areas with high density are deemed more likely to have better public transit systems, but this is not always the case. The *iclnpden* variable describes the interaction between population density and gas price change, and it shows positive coefficients in the model A and C of Table 3. This means that if an area is denser than average by 10%, then its transit ridership would increase by 2.3% to 2.4% more that if there was a 10% increase in gasoline price. Thus, the final increase would be 3.5% (Model C: 0.122005 + 0.225128 = 0.347133) to 4.0% (Model A: 0.148939 + 0.243509 = 0.392448), with the same increase of gasoline price.

clnsprl and iclnsprl If the sprawl index of Smart Growth America rises by ten percent, then urbanized areas would enjoy an increase in transit ridership by 1.3% to 1.5%. Considering the elasticity of gasoline is 1.2 to 2.1 in the above models (1.2% to 2.1% increase of ridership, when gasoline goes high by 10%), the effect of urban sprawl (or compact development) should not be underestimated. Moreover, less sprawled urban areas respond more with the public transit system when the gasoline price increases. The coefficients of *iclnsprl* are 0.18 to 0.24, implying more compact urban areas will get more transit riders, when the gasoline price goes up by the same rate. For example, an urbanized area with its sprawl index 10% higher than average will have 3.9% (Model B: 0.206523+ 0.179615= 0.386138) to 4.2% (Model D: 0.178366 + 0.243617= 0.421983) more transit riders, when gasoline price increase by 10%. Instead, the average sprawled areas will gain just 2.1% to 1.8% more riders with the same gasoline price change. This shows how urban form makes a difference in the relationship between gasoline price and transit ridership change.

nlocal and inlocal The dummy variable *nlocal* shows whether or not each urbanized area has urban containment policies that cover both some part of the area and even beyond the area. The positive coefficients of 0.044 to 0.057 means the urban areas with area-wide urban growth policies will have a higher transit ridership by 4.4% to 5.7%, compared to the areas without such policies. The interaction between non-local urban containment policies and gasoline price change is more interesting. The coefficients of *inlocal* are significant at the 1% level, meaning that there would be an "additional" 1.2% to 1.5% ridership increase per 10% increase of gasoline price. In brief, urban areas with growth management policies show higher elasticity of gasoline price to transit ridership.

[Figure 6] Change of Gasoline Price Elasticity (Left: OLS C, Right: OLS D)



The above charts show how gasoline elasticity changes, when population density, urban sprawl index, and urban containment policies vary. All the data points are calculated assuming there is 10%

increase in gasoline price. The gasoline elasticity can be as low as 1.22 and 1.78 when the area has an average density and sprawl index, but does not have non-local containment policies in both charts, the red left-bottom points. However, this range can rise up to 6.90 and 8.12 if the area has 20% more density than average, 20% higher sprawl index than average, and area-wide growth management policies. The previous research only attempted to calculate the lowest numbers in the above charts, while ignoring urban form and urban policy characteristics. The above charts show that, US urban areas can have proper long-term policies that are more sensitive to gasoline prices and inclined-toward-transit This would significantly reduce the levels of carbon dioxide and greenhouse gases originating from the transportation sector.

5.2 Pooled Two Stage Least Square

The second regression model tries to solve the endogeneity problem, by employing Two-Stage Least Square (2SLS) approach. The basic idea of 2SLS is introducing an instrument variable (IV) to avoid problem of simultaneity between independent and dependent variables. In this research, supply variables like VRH, VRM, and VOMS and a demand variable of UPT affect each other simultaneously. Thus, the IV is used to predict VRM at the 1st stage and the predicted VRM is included at the 2nd stage to estimate coefficients of independent variables.

[Table 4] 1st Stage of 2SLS: estimates of Vehicle Revenue Miles (VRM) per 1,000 people

1st Stage Dep: *lnvrmpt*	A		B	
Number of UAs	68		59	
Time Series Length	108		108	
R-Square	0.5181		0.4162	
Explanatory Variable	Estimate (t Value)	Sig.	Estimate (t Value)	Sig.
Intercept	6.178331 (44.95)	***	5.571965 (36.93)	***
lnppl	0.059001 (6.15)	***	0.101712 (9.67)	***
clnopfp	0.59505 (68.67)	***	0.521182 (48.5)	***
m01	-0.01352 (-0.43)		-0.01064 (-0.31)	
m02	-0.07231 (-2.31)	**	-0.07264 (-2.08)	**
m03	0.027336 (0.87)		0.027862 (0.8)	
m04	-0.0012 (-0.04)		-0.00112 (-0.03)	
m06	-0.01319 (-0.42)		-0.01599 (-0.46)	
m07	-0.01458 (-0.47)		-0.01346 (-0.39)	
m08	0.014478 (0.46)		0.014696 (0.42)	
m09	-0.02005 (-0.64)		-0.01937 (-0.55)	
m10	-0.00972 (-0.31)		-0.0117 (-0.34)	
m11	-0.07666 (-2.45)	**	-0.07782 (-2.23)	**
m12	-0.05337 (-1.71)	*	-0.05503 (-1.58)	

*** indicates significant at 1 percent level, ** indicates significant at 5 percent level, * indicates significant at 10 percent level

The above table shows two different 1st stage regression models: Model A uses 68 urbanized areas, but Model B covers only 59 urbanized areas because the 2nd stage of Model B contains the sprawl index, instead of population density, which is available only for 59 urbanized areas. Among three supply (or service) variables, VRM is chosen, since two other variables have problems with inaccurate measurement (VRH) and small variance (VOMS)- characteristics that make the two variables less attractive in regression analysis. VRM is standardized by population (per thousand) and then converted to a logarithmic term (*lnvrmpt*) to make interpretation easy and intuitive.

The first stage OLS shows not a very high level of R-Square, and an instrument variable (*clnopfp:* operating fund per capita, converted to a logarithmic term and then its mean is centered on zero) is significant at the 1% level for both Models A and B. Although population size is expected to show higher coefficients in both models, they are as small as one-fifth to one-tenth of the *clnopfp* coefficients. With a 10% increase in population, VRM rises by 0.6% to 1.0%, while VRM rises by 5.2% to 6.0% with an additional 10% operating fund. This indicates that, when we attempt to explain different levels of public transit services in the US urban areas, those services show higher variances by operating fund than by population size.

Although the coefficients of *clnopfp* are not very different in Model A compared to Model B, those of *lnppl* are not very similar to each other: the coefficient of *lnppl* of Model B is two times as large as Model A. This is of course, because the analysis samples are different, implying that there are not negligible variations among urban areas in terms of the effect of population size on public transit service level.

The results of the 2nd Stage Least Square are shown in the following table. Their R-Squares surpass 0.95, meaning that 95% of variance of *lnupt* can be explained by explanatory variables in the two models. Overall, the sign of coefficients are the same as the pooled Ordinary Least Square model in the previous chapter, so this section will focus mainly on the difference between two approaches.

clngpi, trend, and postpeak Their coefficients are very similar to the pooled OLS results. The elasiticities of *clngpi* were from 0.122 to 0.207 in the pooled Ordinary Least Square models, and are from 0.122 to 0.175. The coefficients of *trend* (-0.005 to -0.006) are very similar to the previous model which were -0.005 to -0.006. As for *postpeak*, estimations are 0.049 to 0.060, which also are almost the same as before. These results suggest that these explanatory variables have a stable effect on transit ridership change.

lnppl, lnpchedu, and lnunempr The effect of population size reduces a little from 1.234 ~ 1.265 in the pooled OLS to 1.125 ~ 1.150 in the Two Stage Least Square. The *lnpchedu* showed 0.593~0. 631 before, and indicates 0.605~0.631. The local economic condition, *lnunempr*, also showed little changes in its coefficients from -0.070~-0.074 to -0.074. The interesting point here is that, once we include the sprawl index in the regression models, the absolute values of *lnunempr* coefficients decrease-almost to zero-and they become insignificant. This is true for both pooled OLS and 2 SLS.

lnfwylpt The *lnfwylpt* has a noticeable change when regression models include (or exclude) the sprawl index as an explanatory variable. Its coefficient is -0.084 without the sprawl variable, and -0.167 with the variable. This is true in the previous regressions, which showed -0.085 ~ -0.100 without the variable and -0.170 ~ -0.193 with the variable. The absolute values increase almost twice after controlling sprawl of urban areas in regression models. Although the reason for the *lnunempr* coefficient change is not clear, the variable *lnfwylpt* seems to correlate to the sprawl index, which has four different factors in its calculation: density, mix-use, centeredness, and street connectivity. (Ewing et al, 2002)

clnpden and iclnpden These two variables show the same pattern as before. The urbanized area which is denser by 10% than the average would lose about 1.47% of their transit riders . The amount lost was between -1.45% to -1.47% in pooled OLS. The regression result that denser areas would lose (instead of gain) transit riders when gasoline price increases is counter-intuitive. However, the interaction between population density and gasoline price change shows a positive number of 0.207, suggesting that, with a 10% increase in gasoline price, the area that is 10% denser than average will gain an additional 2.07% transit riders with a total ridership increase of 3.29% (gasoline elasticity 0.122 + 0.207 = 0.329). Again, based on the Two Stage Least Square models, we can say that denser urban areas are more sensitive to gasoline price changes in terms of public transit use.

[Table 5] 2ndStage of 2SLS: estimates of Unlinked Passenger Trips (UPT)

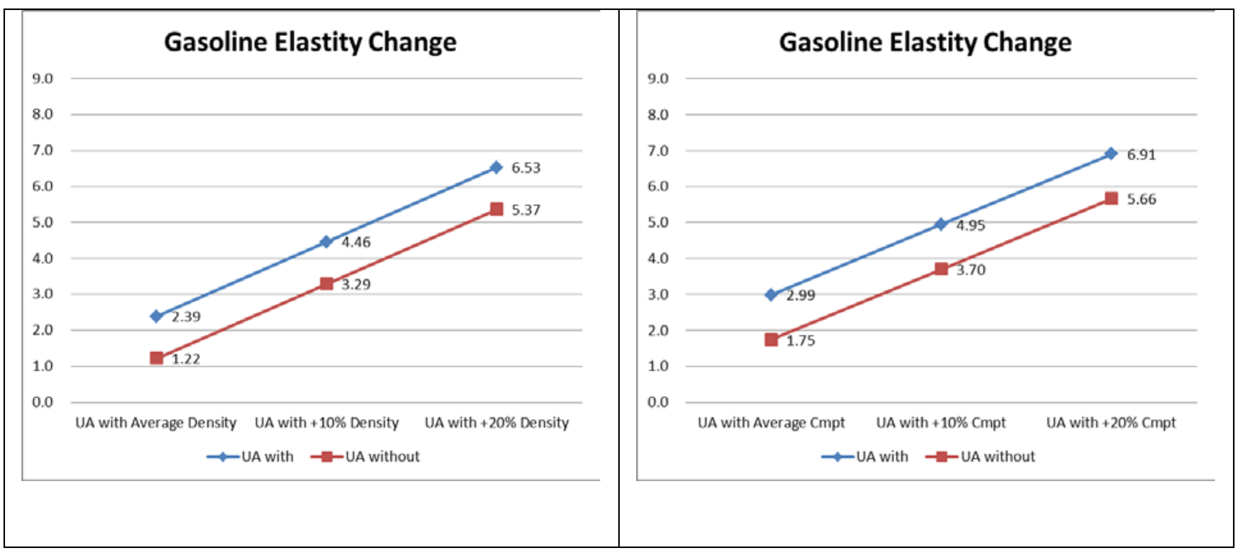
2nd Stage Dep: *lnupt*	A		B	
Number of UAs	68		59	
Time Series Length	108		108	
R-Square	0.9546		0.9615	
Explanatory Variable	Estimate (t Value)	Sig.	Estimate (t Value)	Sig.
Intercept	-12.8372 (-109.15)	***	-11.9017 (-118.71)	***
lnvrmpthat	1.473979 (141.02)	***	1.377677 (110.69)	***
clngpi	0.121906 (3.79)	***	0.174657 (5.88)	***
trend	-0.00553 (-15.3)	***	-0.00525 (-15.58)	***
Postpeak	0.060081 (2.91)	***	0.049044 (2.55)	**
lnppl	1.150763 (154.18)	***	1.125183 (153.42)	***
lnpchedu	0.631133 (25.23)	***	0.604719 (23.6)	***
lnunempr	-0.0738 (-4.12)	***	-0.00675 (-0.4)	
lnfwylpt	-0.08438 (-5.59)	***	-0.1674 (-11.02)	***
clnpden	-0.14666 (-8.07)	***		
iclnpden	0.207304 (4.15)	***		
clnsprl			0.147263 (11.62)	***
iclnsprl			0.195831 (4.11)	***
nlocal	0.044292 (3.94)	***	0.057777 (5.69)	***
inlocal	0.116631 (2.76)	***	0.124346 (3.26)	***
m01	-0.01135 (-0.56)		-0.00977 (-0.51)	
m02	0.053127 (2.61)	***	0.051232 (2.71)	***
m03	-0.00315 (-0.16)		0.003048 (0.16)	
m04	0.006471 (0.33)		0.010669 (0.58)	
m06	-0.02189 (-1.1)		-0.02593 (-1.4)	
m07	-0.0261 (-1.31)		-0.0372 (-2.01)	**
m08	-0.01345 (-0.68)		-0.0203 (-1.1)	
m09	0.084326 (4.23)	***	0.08281 (4.47)	***
m10	0.121329 (6.08)	***	0.13016 (7.02)	***
m11	0.130955 (6.51)	***	0.131612 (7.04)	***
m12	0.059821 (2.93)	***	0.065222 (3.44)	***

*** indicates significant at 1 percent level, ** indicates significant at 5 percent level, * indicates significant at 10 percent level

clnsprl and iclnsprl The coefficients of these two variables are close to the previous ones. *clnsprl*: 0.147 (2SLS) and 0.131 ~ 0.147 (OLS), and *iclnsprl*: 0.196 (2SLS) and 0.180 ~ 0.244 (OLS) These numbers and their statistical significance at the 1% level show urban sprawl (or compact development) *does* have an impact on public transit ridership, when there is a gasoline price change.

nlocal and inlocal The estimations of non-local containment policy variables are almost the same as ones from OLS. *nlocal*: 0.044 ~ 0.058 (2SLS) and 0.044 ~ 0.057 (OLS), and *inlocal*: 0.117 ~ 0.124(2SLS) and 0.118 ~ 0.146(OLS). All of these estimates are significant at the 1% level, showing that urban areas with proper growth management policies have a merit of promoting public transit, not to mention other benefits stemming from less land consumption.

[Figure 7] Change of Gasoline Price Elasticity (Left: 2SLS A, Right: 2SLS B)



The above charts present the elasticity change of gasoline price to transit ridership. Different urban form (sprawl and population density) and existence of urban containment policy do make a meaningful difference in their cities' sensitivity to gasoline price change. The left chart shows how gasoline elasticity changes based on Model A: the right chart gives an idea of the same kind of change based on Model B. According to the above charts, we can understand that urban areas have a huge variation in terms of sensitivity to gasoline price, and that public transit in different areas need to be understood by linking their urban form and policies. These coefficients are slightly smaller than the ones from the previous analysis, and point out that ignoring endogeneity leads to overestimation of coefficients.

5.3 One-Way Fixed-Effect

The fixed-effect model assumes that there would be a time-invariant area-specific effect for all the observations, and that the effects should be dealt with for getting accurate estimates of independent variables. The following is a basic regression model for a fixed-effect approach. In the equation, α_i is the fixed-effect for each urbanized area *i*.

$$y_{it} = x'_{it}\beta + (\boldsymbol{\alpha_i} + \epsilon_{it})$$

In this research, the fixed-effect which is not captured by independent variables would be controlled in the one-way fixed-effect model. However, this research does not employ the two-way fixed-effect model, because area-invariant time-specific fixed-effects are not likely to exist in all the observations. Instead, it is reasonable to assume that people behave differently after they experienced the peak gasoline price in June 2008, so the *postpeak* dummy variable is inserted in all the regression models of this research.

Most of the independent variables appear as smaller coefficients, or as half the previous estimates. The significance and signs of these variables do not greatly differ from previous two models. This means that a considerable amount of variation in UPT can now be explained by area-specific fixed-effect. It is important to note that most of explanatory variables can still predict transit ridership with a statistical significance, though area-specific fixed-effects are excluded from the previous estimates of these variables.

clngpi, trend, and postpeak Gasoline price fluctuation still has a significant impact on transit ridership change at the 1% level, but its estimates decrease from 0.122 ~ 0.207 (OLS) to 0.070 ~ 0.090 (one-way fixed-effect). The effect of *trend* also reduces elasticity from -0.005 ~ -0.006 (OLS) to -0.003 ~ +0.000 (one-way fixed-effect). Though there is one model where *trend* has a positive coefficient, this research focuses more on the model C and D which include population density, urban sprawl, and containment policy. In this context, the estimate of *trend* would be -0.003 which is almost half the previous ones. The OLS models show that there was a slight increase in public transit use after June 2008 with all else being equal. However, two out of four one-way fixed-effect models indicate the dummy variable is not statistically significant.

lnppl, lnpchedu, and lnunempr Population size still matters at the 1% level, but their coefficients are smaller than the ones of the former analyses. When the OLS models consider not a small part of the variation of transit ridership as an effect of population size, the one-way fixed-effect models find that some of the effect is actually coming from area-specific fixed-effects which are not measured by any of the independent variables in the models. The change of a 1 % of those who are enrolled in colleges or graduate schools still positively correlates to the number of transit riders, but the coefficients of *lnpchedu* reduces by two-thirds from 0.593~0.631 to 0.155~0.181. The effect of change of local monthly unemployment rate is significant and negatively correlates to transit ridership.

lnfwylpt, clnopfp, and iclnopfp As opposed to expectation, the change of freeway lane miles per thousand, *lnfwylpt*, positively correlates to transit ridership. With a 10% increase in freeway lane miles per thousand people, there would be 0.66 ~ 0.79% ridership gain. However, the operating fund per capita, *clnopfp*, shows expected estimates of positive 0.532 ~ 0.808 indicating that urban areas with 10% more operating funds than average will enjoy a 5.3% ~ 8.1% transit ridership increase. Unfortunately, the interaction of operating fund and gasoline price does not show a consistent result: coefficients are positive (0.040) in Model C, but negative (-0.038) in Model D. The confusing results of *lnfwylpt* and *iclnopfp* call for additional research.

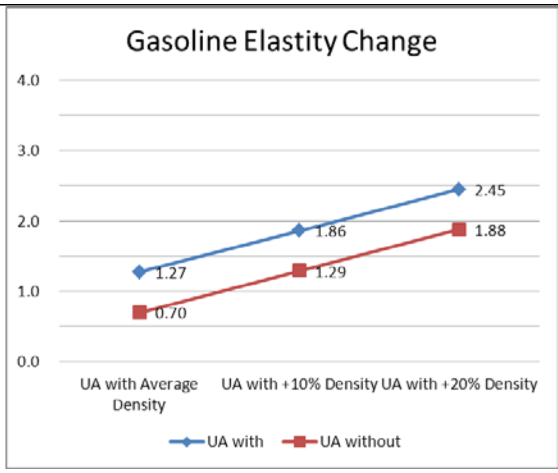
iclnpden, iclnsprl, and inlocal These explanatory variables about urban form and policy present positive and significant estimates of 0.059, 0.090, and 0.051 ~ 0.057 respectively. These are all interactions with gasoline price change, so their positive sign means that urban areas with higher density, less sprawl, and urban containment policies will get additional transit riders than the average area, when gasoline price increases. However, after controlling each urban area's fixed effect, these coefficients become as small as one-half to one-fourth of the results from the OLS model. Like the previous analyses, the following charts show the relationship between gasoline elasticity and urban form/policy conditions of urban areas.

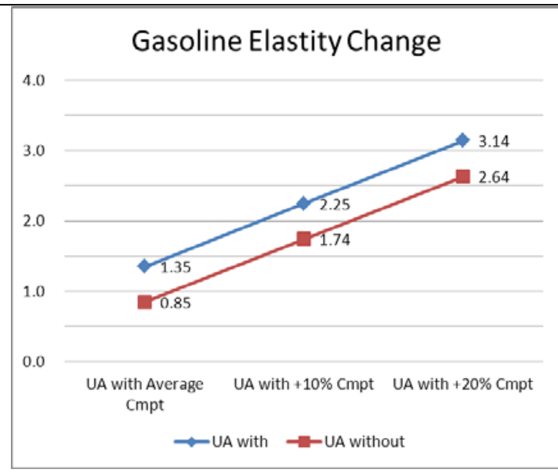
[Table 6] One-Way Fixed-Effect estimates of Unlinked Passenger Trips (UPT)

Dep. Variable: *lnupt*	A		B		C		D	
Number of UAs	68		68		68		59	
Time Series Length	108		108		108		108	
R-Square	0.9878		0.9903		0.9903		0.9944	
Explanatory Variable	Estimate (t Value)	Sig.	Estimate (t Value)	Sig.	Estimate (t Value)	Sig.	Estimate (t Value)	Sig.
Intercept	11.61882 (10.57)	***	9.977872 (10.03)	***	10.51179 (10.46)	***	0.448032 (0.52)	
clngpi	0.08908 (5.13)	***	0.077916 (5.03)	***	0.069987 (4.39)	***	0.08453 (6.86)	***
trend	0.000431 (1.92)	*	-0.00299 (-13.83)	***	-0.003 (-13.88)	***	-0.00278 (-16.43)	***
Postpeak	0.028737 (2.43)	**	0.016912 (1.6)		0.016602 (1.57)		0.056599 (6.97)	***
lnppl	0.391852 (5.48)	***	0.414758 (6.42)	***	0.376964 (5.77)	***	1.061532 (19.13)	***
lnpchedu	-0.02336 (-0.63)		0.154927 (4.58)	***	0.169417 (4.95)	***	0.181343 (6.82)	***
lnunempr	-0.02386 (-1.69)	*	-0.03632 (-2.88)	***	-0.03566 (-2.83)	***	-0.05826 (-5.93)	***
lnfwylpt	0.108971 (3.71)	***	0.079136 (3)	***	0.065745 (2.47)	**	0.068074 (2.85)	***
clnopfp			0.807062 (42.43)	***	0.808386 (42.49)	***	0.531513 (32.21)	***
iclnopfp			0.054006 (6.18)	***	0.037942 (3.6)	***	-0.0384 (-4.89)	***
iclnpden					0.05892 (2.04)	**		
iclnsprl							0.089516 (4.38)	***
inlocal					0.057302 (2.84)	***	0.050882 (3.33)	***
m01	-0.02372 (-2.23)	**	-0.03815 (-4.01)	***	-0.03737 (-3.93)	***	-0.03133 (-4.29)	***
m02	-0.04927 (-4.66)	***	-0.06014 (-6.37)	***	-0.05958 (-6.31)	***	-0.05575 (-7.69)	***
m03	0.040452 (3.88)	***	0.033452 (3.59)	***	0.033665 (3.62)	***	0.037614 (5.27)	***
m04	0.009191 (0.89)		0.005504 (0.6)		0.005502 (0.6)		0.006484 (0.91)	
m06	-0.04726 (-4.54)	***	-0.04151 (-4.47)	***	-0.0415 (-4.47)	***	-0.04536 (-6.37)	***
m07	-0.05992 (-5.77)	***	-0.05062 (-5.46)	***	-0.05054 (-5.45)	***	-0.05558 (-7.81)	***
m08	-0.00883 (-0.85)		0.003649 (0.39)		0.003782 (0.41)		-0.00423 (-0.6)	
m09	0.033428 (3.22)	***	0.049085 (5.3)	***	0.049256 (5.32)	***	0.047957 (6.75)	***
m10	0.079017 (7.59)	***	0.09764 (10.5)	***	0.097896 (10.53)	***	0.098728 (13.83)	***
m11	-0.01995 (-1.9)	*	0.001765 (0.19)		0.002187 (0.23)		0.002154 (0.3)	
m12	-0.06545 (-6.12)	***	-0.04074 (-4.26)	***	-0.04 (-4.18)	***	-0.04027 (-5.48)	***

*** indicates significant at 1 percent level, ** indicates significant at 5 percent level, * indicates significant at 10 percent level

[Figure 8] Change of Gasoline Price Elasticity (Left: OWFE C, Right: OWFE D)





5.4 Two Stage One-Way Fixed-Effect

The final statistical model, the Two Stage One-Way Fixed-Effect model, combines two different approaches for the endogeneity problem and fixed-effect. By comparing this model to simple Two Stage Least Square model, we can expect unbiased estimations of coefficients. In this section, a regression model considers time-invariant fixed-effect for each urban area in both first and second stages. Like the pooled 2SLS model, the first stage uses *clnopfp* as an instrument variable for predicting *lnvrmpt* while and then the second stage estimates elasticity of each independent variable by including a predicted *lnvrmpt* instead of its real value.

[Table 7] 2 Stage One-Way Fixed-Effect estimates of Vehicle Revenue Miles per 1,000 people (VRM)

Dep. Variable: *lnvrmpt*	A		B	
Number of UAs	68		59	
Time Series Length	108		108	
R-Square	0.6194		0.5217	
Explanatory Variable	Estimate (t Value)	Sig.	Estimate (t Value)	Sig.
Intercept	6.159817 (2.45)	**	4.508913 (1.43)	
lnppl	0.077885 (0.47)		0.193682 (0.93)	
clnopfp	0.464836 (11.33)	***	0.381942 (7.49)	***
m01	-0.01352 (-0.48)		-0.01064 (-0.34)	
m02	-0.07231 (-2.59)	***	-0.07264 (-2.29)	**
m03	0.027336 (0.98)		0.027862 (0.88)	
m04	-0.0012 (-0.04)		-0.00112 (-0.04)	
m06	-0.01319 (-0.47)		-0.01599 (-0.5)	
m07	-0.01458 (-0.52)		-0.01346 (-0.42)	
m08	0.014478 (0.52)		0.014696 (0.46)	
m09	-0.02005 (-0.72)		-0.01937 (-0.61)	
m10	-0.00972 (-0.35)		-0.0117 (-0.37)	
m11	-0.07666 (-2.74)	***	-0.07782 (-2.45)	**
m12	-0.05337 (-1.91)	*	-0.05503 (-1.73)	*

*** indicates significant at 1 percent level, ** indicates significant at 5 percent level, * indicates significant at 10 percent level

The above table shows the result of the first stage fixed-effect regression. Although the coefficients of *clnopfp* are lower than the ones from the pooled 2SLS model, the R-Square values are higher than those of the previous model. Interestingly, *lnppl* itself does not have a significant statistical relation to the transit service variable. After controlling each urban area's specific effect, population size loses its explaining power to the logarithmic term of Vehicle Revenue Miles. The following table has the results of two different two-stage one-way fixed-effect models: Model A uses *iclnpden*, while Model B has *ilnsprl* instead. The R-Square values of these models are over 99%, and all the explanatory variables are significant except *postpeak* dummy in Model A.

lnvrmpthat $(\boldsymbol{ln\widehat{vrmpt}})$ As expected, it highly correlates to *lnupt* with coefficients of 1.739 and 1.392. If the estimated value of VRM per thousand people increases by 10%, then the total number of transit riders goes up by 17.39% ~ 13.92% for both models.

lnppl The population size has different effects on transit ridership because it was originally considered in the first stage. Its previous estimates were 0.377 ~ 1.062 in the simple One-Way Fixed-Effect models and now its coefficients are 0.242 ~ 0.792. This means that once the service level of mass transit systems is controlled, the population size has less explanatory power in predicting transit ridership. Although the service level of public transit, population size, and number of transit riders correlates to one another, with the help of the two-stage approach in this research, we find that transit supply (operating fund per capita) can explain transit demand (unlinked passenger trip) better than the number of people in an urban area.

Except the above two variables, there is no difference in coefficients for both simple One-Way Fixed-effect and Two-Stage One-Way Fixed-Effect models. For the same reason, the charts for gasoline elasticity change are the same as the simple One-Way Fixed-Effect model.

[Table 8] 2nd Stage of 2S One-Way Fixed-Effect estimates of Unlinked Passenger Trips (UPT)

Dep. Variable: *lnupt*	A		B	
Number of UAs	68		59	
Time Series Length	108		108	
R-Square	0.9903		0.9944	
Explanatory Variable	Estimate (t Value)	Sig.	Estimate (t Value)	Sig.
Intercept	-0.20061 (-0.19)		-5.8266 (-6.73)	***
lnvrmpthat	1.739077 (42.49)	***	1.391606 (32.21)	***
clngpi	0.069987 (4.39)	***	0.08453 (6.86)	***
trend	-0.003 (-13.88)	***	-0.00278 (-16.43)	***
Postpeak	0.016602 (1.57)		0.056599 (6.97)	***
lnppl	0.241516 (3.69)	***	0.792003 (13.94)	***
lnpchedu	0.169417 (4.95)	***	0.181343 (6.82)	***
lnunempr	-0.03566 (-2.83)	***	-0.05826 (-5.93)	***
lnfwylpt	0.065745 (2.47)	**	0.068074 (2.85)	***
iclnopfp	0.037942 (3.6)	***	-0.0384 (-4.89)	***
iclnpden	0.05892 (2.04)	**		
ilnsprl			0.089516 (4.38)	***
inlocal	0.057302 (2.84)	***	0.050882 (3.33)	***
m01	-0.01385 (-1.46)		-0.01652 (-2.26)	**
m02	0.06617 (6.74)	***	0.045331 (5.8)	***
m03	-0.01387 (-1.48)		-0.00116 (-0.16)	
m04	0.00759 (0.82)		0.008039 (1.13)	
m06	-0.01856 (-2)	**	-0.02311 (-3.23)	***
m07	-0.02518 (-2.71)	***	-0.03684 (-5.15)	***
m08	-0.0214 (-2.31)	**	-0.02469 (-3.47)	***
m09	0.084118 (9.02)	***	0.074906 (10.42)	***
m10	0.114804 (12.32)	***	0.115014 (16.02)	***
m11	0.135497 (13.48)	***	0.110455 (13.61)	***
m12	0.052805 (5.32)	***	0.036314 (4.62)	***

*** indicates significant at 1 percent level, ** indicates significant at 5 percent level, * indicates significant at 10 percent level

6. Conclusion

This research examines the relationship between gasoline price fluctuation and land use characteristics such as population density, compact development, and urban containment policy for their impacts on transit ridership change by employing Two-Stage One-Way Fixed-Effect model from 68 urbanized areas over nine years (January 2002 to December 2010).

The regression results show that urban spatial characteristics and urban policy increase transit ridership when they interact with the rise of motor fuel cost. When population density is included in regressions, the price elasticity of gasoline can be more than doubled if urban areas are denser than average by 10% and have urban grow management plans that also cover wide nearby areas-the elasticity was initially 0.070, but it goes up to 0.186to indicate a 1.86% increase in transit ridership when gasoline price rises 10%. If urban sprawl is considered instead of population density, the elasticity will increase slightly more in an identical situation-the elasticity itself is 0.085, but it would rise to 0.225, meaning a 2.25% ridership gain when motor fuel costs increase 10%. This result would be another rationale as to why we need to support dense and compact urban developments with less land consumption.

Policy implications are as follows. Transportation planners should expand their perspective and be able to link urban form and policy to transit system, because the amount of operating fund per capita alone has a limited role in attracting more riders. If we want to promote public transit systems in the US urban conditions, we had better support less sprawled developments at the same time, i.e. denser, mixed-use, concentrated, and well-connected urban areas. Of course, such land use initiatives work better when combined. For example, increasing density in addition to adopting growth management plans would be more productive than one without the other. Each urban area has different characteristics in terms of urban form and political situation for adopting extensive growth management plans, planners need to adjust findings of this research to match their situations.

REFERENCES

Journal Articles

Boarnet, M. (2010). Planning Climate change, and transportation: Thoughts on Policy analysis, *Transportation Research Part A*, 44, 587-595.

Cervero, R., & Landis, J. (1995). The Transportaion-Land Use Connection Still Matters, *Access*, No.7, 2-10.

Chen, C., Varley, D., & Chen, J. (2010). What affects transit ridership? A dynamic analysis involving multiple factors, lags and asymmetric behavior, *Urban Studies*, 48 (9), 1893-1908.

Currie, G., & Phung, J. (2007). Transit Ridership, Auto Gas Prices, and World Events New Drivers of Change?, *Transportation Research Record: Journal of the Transportation Research Board*, 1992, 3-10.

Currie, G., & Phung, J. (2008). Understanding links between transit ridership and gasoline prices evidence from the United States and Australia, *Transportation Research Record: Journal of the Transportation Research Board*, 2063, 133-142.

Giuliano, G. (1995). The Weakening Transportation-Land Use Connection, *Access*, No.6, 3-11.

Guo, Z., Agrawal, A.W., & Dill, J. (2011). Are Land Use Planning and Congestion Pricing Mutually Supportive?, *Journal of the American Planning Association*, 77:3, 232-250.

Haire, A. R. and Machemehl, R. B. (2010). Regional and Modal Variability in Effects of Gasoline Prices on U.S. Transit Ridership, *Transportation Research Record: Journal of the Transportation Research Board*, 2144, 20-27.

Lane, B. W. (2010). The relationship between recent gasoline price fluctuations and transit ridership in major US cities, *Journal of Transport Geography*, 18, 214-225.

Maley, D. W. & Weinberger, R. (2009). Riding gas price and transit ridership case study of Philadelphia, Pennsylvania, *Transportation Research Record: Journal of the Transportation Research Board*, 2139, 183-188.

Stover, V. W. & Bae, C.-H. C. B. (2011). Impact of Gasoline Prices on Transit Ridership in Washington State, *Transportation Research Record: Journal of the Transportation Research Board*, 2217, 11-18

Taylor, B. D., Miller, D., Iseki, H., & Fink, C. (2009). Nature and/or nurture? Analyzing the determinants of transit ridership across US urbanized areas, *Transportation Research Part A*, 43, 60-77.

Wassmer, R. (2006). The influence of local urban containment policies and statewide growth management on the size of United States urban areas, Journal of Regional Science, 46(1), 25-65.

Yanmaz-Tuzel, O. & Ozbay, K. (2010). Impacts of Gasoline Prices on New Jersey Transit Ridership, *Transportation Research Record: Journal of the Transportation Research Board,* 2144, 52-61.

Electronic Books & Data

Ewing, R., Pendall, R., & Chen, D. (2002). Measuring Sprawl and Its Impact Volume 1. Retrieved from http://www.smartgrowthamerica.org/documents/MeasuringSprawlTechnical.pdf

Mattson, J. (2008). Effects of Rising Gas Prices on Bus Ridership for Small Urban and Rural Transit Systems. Retrieved from http://www.ugpti.org/pubs/pdf/DP201.pdf

Tanneeru, M. (2008, July 18). U.S. cities scrambling to meet rising mass transit demands. CNN. Retrieved from http://www.cnn.com/2008/US/07/16/mass.transit/#cnnSTCText

U.S. Energy Information Administration. (2012). Monthly Retail Gasoline and Diesel Prices (Dollar per Gallon, including Taxes). Retrieved from http://www.eia.gov/dnav/pet/pet_pri_gnd_dcus_nus_m.htm