Measuring Long-Run Price Elasticities in Urban Travel Demand

Donna, Javier D.

The Ohio State University

5 November 2018

Measuring Long-Run Price Elasticities in Urban Travel Demand

Javier D. Donna*

First version: December, 2010.
This version: November 5, 2018.

Abstract

This paper develops a structural model of urban travel to estimate long-run price elasticities. A dynamic discrete choice demand model with switching costs is estimated, using a panel dataset with public market-level data on automobile and public transit use for Chicago. The estimated model shows that long-run own- (automobile) and cross- (transit) price elasticities are more elastic than short-run elasticities, and that elasticity estimates from static and myopic models are downward biased. The estimated model is used to evaluate the response to a gasoline tax. Static and myopic models mismeasure long-run substitution patterns, and could lead to incorrect policy decisions.

JEL Codes: L71, L91, L98.

Keywords: Long-run price elasticities, Dynamic demand travel, Hysteresis.

---

*The Ohio State University, donna.1@osu.edu. I am indebted to my advisors and members of my dissertation committee for helpful discussions, guidance, and support: Rob Porter (committee chair), Meghan Busse, Aviv Nevo, and Florian Zettelmeyer. I especially thank Igal Hendel and Tiago Pires for their many and helpful suggestions, and to Yun Pu for outstanding research assistance. Discussions with John Asker, Jason Blevins, Chris Conlon, Daniel Donna, Jean-Pierre Dubé, Cristian Huse, Ryan Kellogg, Matt Lewis, Charles Manski, James Roberts, Mark Roberts, William Rogerson, Oleksandr Shcherbakov, Alex Torgovitsky, Ija Trapeznikova, Greg Veramendi, Bruce Weinberg, and Matt Weinberg, as well as seminar participants at Aarhus University, Johns Hopkins University, Northwestern University, Northwestern University Transportation Center (NUTC), Universidad Carlos III, and The Ohio State University have greatly benefited this paper. I would also like to express my gratitude to Rob Robinson, chief of the planning section at the Illinois Department of Transportation (IDOT), and Catherine Kannenberg, head of systems performance from the Metra division, for providing me access to the transportation data. A version of this paper was part of my Ph.D. dissertation at Northwestern University. This paper combines two previously unpublished working papers, "Estimating Switching Costs in Urban Travel Demand in Chicago," and "The Effects of Gasoline Prices on Urban Travel Demand in Chicago." I acknowledge financial support from the CSIO at Northwestern University, the allocation of computing time from the Ohio Supercomputer Center (grant PAS1350-2), and financial support from the SBS at The Ohio State University. The analysis in this paper, based on data from the IDOT, are by the author and not by the IDOT. All errors are mine.

1 Introduction

Long-run gasoline price elasticities are a key input to evaluate the impact of transportation policies. Examples include gasoline taxes, investments in capacity for public transportation, and zoning laws. There is a remarkable large number of empirical studies, described in the related literature, investigating gasoline price elasticities. The literature has emphasized the importance of distinguishing between responses at different time horizons. Yet, most of the empirical work estimates short-run elasticities, whereas long-run elasticities are relevant for most policy questions. Furthermore, short-run elasticities may mismeasure long-run effects when switching costs are present as is common in urban travel demand.[1] This paper develops a structural dynamic model of urban travel, estimates the model using public market-level data on automobile and public transit use for Chicago, and uses the estimated model to compute long-run price elasticities and evaluate policy interventions in this market.

The transportation industry plays a central role in modern economies. Private automobile is the main mode of transportation in the U.S. (U.S. Department of Transportation 2011). It has a direct impact on the demand for gasoline, emission of greenhouse gases, and traffic congestion. Public transportation ridership is high in large metropolitan areas. In the U.S., 85 percent of transit agencies reported capacity constraints in 2008, when gasoline prices reached an all-time high (American Public Transportation Association 2008). It shifted the attention of policy makers to a better understanding of how transit ridership responds to gasoline prices. The ability to effectively measure the impact of regulation policies relies on estimates of long-run elasticities of automobile and transit use to the gasoline cost of driving.

In this paper I develop a structural model of urban travel to estimate these long-run elasticities. To do that, I model the demand for modes of transportation using a random coefficient mixed logit dynamic model. For the structural estimation, I build a GMM estimator adapting the procedure by Gowrisankaran and Rysman (2012) and Shcherbakov (2016), who nest a demand system (Berry, Levinsohn, and Pakes 1995) within a dynamic framework (Rust 1987). To compute long-run elasticities, I follow Hendel and Nevo (2006) and evaluate the long-run response to a change in the whole path of gasoline prices, not just the current price. In the model, consumers are rational, forward looking, and have persistent heterogeneous preferences for modes of transportation. A consumer may incur a switching cost if the mode chosen in the current period differs from the one in the previous period. No switching costs are incurred if the consumer chooses the same mode. The switching costs include the costs of seeking information and setting up an alternative mode as emphasized by the transportation literature described below.[2] Switching costs introduce persistence in the choice of the mode of transportation. A traveler changes the mode only if the value of switching exceeds the

[1] I define switching costs below. I discuss their presence in my setting in subsection 2.3.

[2] For public transit, switching costs include the cost of obtaining information about transit services (schedules, routes, transfers, decipherable of timetables), information about bike and/or transit integration, the cost of obtaining a monthly ridership ticket, *etc.* For automobile, switching costs include the cost of obtaining information about routes, parking facilities search, costs for regular use of the car, *etc.*

value of the current mode plus the switching cost. Travelers may make discontinuous choices between modes, based on their valuation of the mode's characteristics, such as the gasoline cost of driving. Thus, market shares may display *hysteresis* or state dependence when the gasoline cost of driving varies, holding other characteristics constant.[3]

There has been a long debate regarding the role of hysteresis and switching costs on price elasticities in travel demand. Studies investigating hysteresis in travel demand date back to Goodwin (1977), Banister (1978), and Blase (1980). One implication is that the response of travel demand may be asymmetric. The response to a price increase in gasoline may be different from an identical price decrease.[4] Another implication is the relative insensitivity of travel demand to gasoline prices due to inertia. Hysteresis in travel demand is typically explained by the presence of switching costs as noted in the survey by Button (2010, p. 95): "[Hysteresis] may be explained in some cases quite simply by the fact that there are costs involved in seeking out information about alternatives and continuing as before is thus a rational response until more major price changes occur." To date, attempts to empirically validate this hysteresis/switching costs framework have focused either on summarizing asymmetries in demand responses to price variations (*e.g.* increasing *vs.* decreasing prices) without specifying a model of conduct, or by modeling the state dependence/inertia without estimating the structural parameters of the dynamic decision process.[5] A limitation in these approaches may arise to account for factors affecting long-run substitution patterns, such as changes in expectations or heterogeneity in consumers' preferences. Another difficulty may arise to evaluate counterfactual policies. For example, under an announced gasoline tax, consumers may internalize in their expectations the permanent price increase, thus affecting long-run elasticities. Evaluating policies aimed to change factors causing the hysteresis requires uncovering the model structural parameters (*e.g.* switching costs). For these reasons, the main question of how long-run price elasticities are affected by switching costs in a dynamic setting where hysteresis is present, has not been directly addressed empirically.

To address this question I combine the structural model with public data obtained for this study from several government agencies, as described in section 2. Public transit ridership data were obtained for the three public transit operators in the Chicago area. Vehicle circulation data are obtained from the Illinois Department of Transportation (IDOT). The IDOT uses automatic traffic recorders (ATRs) to measure the number of vehicles circulating at the ATRs locations. I use the location of the ATRs, to link vehicle counts and public transit ridership. To construct market shares, I use information from the Census Transportation

[3]In the context of travel behavior of this paper, hysteresis is defined as the change in the market share of a mode of transportation that persist after the reversal of the initial variation in the gasoline cost of driving that gave rise to it. See subsection 2.3 for details.

[4]The survey by Dargay (1993) provides evidence about these asymmetries, and raised the question of how price elasticities are affected by the presence of hysteresis.

[5]See, *e.g.*, Golob, van Wissen, and Meurs (1986), Ben-Akiva and Morikawa (1990), Dargay (1993), Cantillo, Ortúzar, and Williams (2007), Maley and Weinberger (2009), Sorrell, Dimitropoulos, and Sommerville (2009), Chen, Varley, and Chen (2011), Cherchi, Meloni, and Ortúzar (2013). See also footnote 15.

Planning Package. I complement the market level data about car and public transit use with information about the relative shares of consumers who do not switch modes of transportation, obtained from a travel survey by Pace, one of the public transit operators in Chicago. The survey was administered to 1,330 randomly chosen households in the Chicago metropolitan area. This additional information is more desegregated than the market level data above. I use the data from this survey to define additional micro moments used for the estimation. I focus on the period 2003 to 2009 for the estimation of the model. This period is characterized by a large variation in the gasoline prices in the U.S., which induced consumers to switch modes of transportation as shown in subsection 2.3. Such variation is useful for the estimation as explained next.

To identify switching costs I rely on instruments, micro moments defined by the shares of consumers who do not switch, and functional form restrictions for the unobserved consumer preferences. First, I rely on instruments to identify the switching costs. The basic insight is to use instruments that shift consumer's previous period decision, holding constant the current period decision. I use the previous period international price of oil, and the average gasoline cost of driving in Chicago during the previous three periods. The exclusion restriction is that gasoline prices follow a martingale process. This is a common assumption made in the literature (see, *e.g.*, Busse, Knittel, and Zettelmeyer 2013 in the context of car purchases).[6] Second, I use the information about the share of consumers who do not switch (micro moments). These additional moments help identify the switching costs and heterogeneity parameters (as in, *e.g.*, Petrin 2002; Berry, Levinsohn, and Pakes 2004; Goeree 2008). Finally, I assume that the consumers' heterogeneous preferences are constant over time. This assumption allows to separate the effect on choices resulting from the heterogeneity in preferences, and the hysteresis due to switching costs. Two features from the empirical setting are also useful for the estimation of the switching costs parameters. First, the frequent and large variation in the gasoline cost of driving during the period under analysis. For example, the large increase in the gasoline price in 2008 induced certain types of consumers to switch to public transit. Second, that the variations in the gasoline cost of driving in Chicago are driven mostly by variation in the cost of international oil prices. Thus, the response of car and public transit use in Chicago is arguably exogenous. See subsection 4.2 for details.

I report three main findings using the estimated model. First, long-run price elasticities with respect to the gasoline cost of driving are more elastic than the short-run elasticities.[7] Long-run own- (for automobile) and cross- (for public transit) price elasticities are 13 and 16 percent higher than short-run elasticities, respectively. Measuring the long-run response is relevant for most applications. Second, static own- and cross-price elasticities underestimate the dynamic ones by 33 and 27 percent, respectively, on average. These biases vary across

[6] See Anderson, Kellogg, Sallee, and Curtin (2011) and Anderson, Kellogg, and Sallee (2013) for detailed discussions. Kellogg (2018) studies the implications for fuel economy policies of gasoline price volatility and uncertainty about its future prices.

[7] See subsection 5.3 for definitions of these elasticities.

markets. It is up to 35 percent for the own- and cross-price elasticities in some markets. Myopic estimates also underestimate the elasticities. The bias is larger than for the static model, 56 and 81 percent for own- and cross-price elasticities, respectively, on average. Finally, I investigate the response to a gasoline tax in a counterfactual analysis. The analysis shows that static and myopic models mismeasure the long-run substitution effects. Static and myopic models do not account for the fact that consumers internalize the permanent increase in the gasoline price due to the announced tax. The bias is larger for the myopic model, because the static model predicts higher probabilities of switching than the myopic model, due to the static model not accounting for switching costs. Overall these results indicate that, when hysteresis is present, elasticities from static and myopic models might be biased estimates of long-run elasticities.

In summary, this paper makes three main contributions. First, it develops a dynamic demand model of urban travel to estimate long-run price elasticities when switching costs may be present. The model can be estimated using aggregate level data, which is typically available from government agencies. Second, it uses a novel dataset obtained from several public sources in Chicago, along with the dynamic demand model, to estimate its structural parameters and obtain long-run price elasticities consistent with the presence of switching costs and hysteresis. Finally, it shows that long-run price elasticities are more elastic than short-run elasticities, and that elasticity estimates from static and myopic models are downward biased, which could lead to incorrect predictions when evaluating public policies.

The rest of the paper is organized as follows. Section 2 describes the industry, the data, and presents stylized facts about the industry. Section 3 presents the dynamic demand model of urban travel. Section 4 discusses identification and estimation of the model. Section 5 presents the estimation results, implications for the estimates of the elasticities, and counterfactuals. Section 6 concludes. Additional data description, robustness analysis, extensions, and details about the estimation and the model are in the appendix.

Related Literature

Variation in crude oil prices after the 1973 oil embargo attracted considerably research attention to model the demand for gasoline, and its relationship to the underlying demand for transportation services and elasticities.[8] Such studies were interested not only in understanding the market response to changes in gasoline prices, but also as policy tools to evaluate the impact of regulations.[9]

[8]See Ben-Akiva and Lerman (1985), Small, Verhoef, and Lindsey (2007), Button (2010), and Small (2013) for comprehensive surveys.

[9]There was also a growing number of studies focused on modeling econometrically urban travel demand, and the automobile industry. The application of discrete choice models to study motor vehicle demand and transportation behavior date back to Burns and Golob (1975) and Lerman and Ben-Akiva (1976). Train (1980) develops a model to explain vehicle ownership and mode choice, and estimate it using survey data. Berry, Levinsohn, and Pakes (1995) studies the estimation of discrete choice models of product differentiation in the automobile industry. See Grigolon, Reynaert, and Verboven (2017) and Miravete, Moral, and Thurk (2018) for recent studies investigating the welfare effects of alternative tax and tariff policies.

There is a vast literature investigating these elasticities. See, *e.g.*, Dahl (1986), Dahl and Sterner (1991), Goodwin (1992), Oum, Waters, and Yong (1992), Espey (1998), Graham and Glaister (2002), Basso and Oum (2007), for thorough surveys of the literature and the interpretation of these elasticities. Although the large number empirical studies investigating these elasticities is astounding, the findings are sensible after proper interpretation. As noted by Dahl and Sterner (1991, p. 203): "by a careful comparison [...] if properly stratified, compared and interpreted, different models and data types do tend to produce a reasonable degree of consistency." There are several recent empirical studies examining gasoline demand elasticities, and the sensitivity of public transportation and vehicle use to gasoline price. Hughes, Knittel, and Sperling (2008) estimate average per capita demand for gasoline in the U.S. for two periods, from 1975 to 1980 and 2001 to 2006, of high gasoline prices. They find that the short-run price elasticities differ considerably across periods. Levin, Lewis, and Wolak (2017) specify a model of gasoline purchase behavior that allows them to identify a measure of the short-run elasticity of gasoline demand from data on gasoline expenditures. They estimate their model using high frequency panel data on gasoline prices and expenditures, and find price elasticities significantly more elastic than estimates from recent studies using more aggregated data. The main difference between these papers and mine, is my focus on long-run elasticities. Litman (2004, 2017) provides comprehensive surveys of studies investigating public transit elasticities. Currie and Phung (2007) studies U.S. transit demand using ridership and fuel data similar to the one in this paper. Haire and Machemehl (1992) also uses ridership and fuel prices to investigate transit response in Atlanta, Dallas, Los Angeles, San Francisco, and Washington, D.C. Mattson (2008) investigates how gasoline prices impact bus ridership for small urban and rural transit systems, while Rose (1986) performs the study using data from Chicago, Maley and Weinberger (2009) from Philadelphia, Yanmaz-Tuzel and Ozbay (2010) from New Jersey, and Stover and Bae (2011) from Washington State.[10] Studies using aggregate, or market, level data have focused in two main sources of variation. First, using time series and panel data, from different cities or countries, researchers have computed the cross-price elasticity of transit ridership with respect to gasoline price. Second, an alternative approach has been to use cross-sectional variation in the data. Several studies have used city disaggregation within United States (*e.g.* Kohn et al. 1999; Taylor, Miller, Iseki, and Fink 2008), or even variation across countries (*e.g.* Wheaton 1982). Individual data from surveys has been used extensively to estimate gasoline and urban travel demand (*e.g.* Archibald and Gillingham 1980, 1981; Hausman and Newey 1995; Schmalensee and Stoker 1999). Horowitz (1980) structurally estimates a model of the demand for multi destination, nonwork travel using data from the Washington D.C. area transportation survey. Manski and Sherman (1980) use a sample from a nationwide rotating

[10]Other studies using aggregate level data include, *e.g.*, Agthe and Billings (1978), Doi and Allen (1986), Storchmann (2001), Chen (2007), Currie and Phung (2008), Mattson (2008), and Currie and Phung (2006, 2008).

consumer panel to estimate household motor vehicle choices.[11] There have also been studies analyzing traffic volumes using data from ATRs similar to one in this paper (*e.g.* Horowitz and Emslie 1978; Hogema 1996; Lingras 2000).[12] While there is a large literature that has studied either gasoline demand or public transit determinants, few studies have used structural models to analyze the long-run impact of gasoline prices on both, public transportation and vehicle, using aggregate data.

There is also a large transportation literature studying hysteresis, inertia, and state dependence resulting from repeated choices (*e.g.* Goodwin 1977; Banister 1978; and Blase 1979, 1980; Clarke, Dix, and Goodwin 1982; Daganzo and Sheffi 1982; Williams and Ortúzar 1982).[13] Dargay (1993) discusses the impact of hysteresis on elasticities, and provides a survey.[14] Hysteresis and nonlinear responses to gasoline prices have also been documented in public transit ridership (*e.g.* Maley and Weinberger 2009; Chen, Varley, and Chen 2011). The literature has emphasized the presence of repeated choices in travel patterns over time (*e.g.* Pendyala, Parashar, and Muthyalagari 2001; Gärling and Axhausen 2003), suggesting inertia in individual choices. A number of papers have developed and estimated dynamic models to conceptualize such inertia.[15] Ben-Akiva and Morikawa (1990) propose and estimate a switching model, where consumers switch if the utility of the new choice is greater than the utility of the current choice plus a threshold reflecting switching costs or an inertia effect. Cantillo, Ortúzar, and Williams (2007) estimate a discrete choice model incorporating randomly distributed inertia thresholds and allowing for serial correlation. Cherchi, Meloni, and Ortúzar (2013) estimate a hybrid choice model combining a joint discrete choice and a latent inertia model. None of these papers estimates the structural parameters characterizing

[11] A further review of urban travel demand elasticities and factors influencing it can be found in Taylor and Fink (2003), Fearnley and Bekken (2005), and more recently in Brons, Nijkamp, Pels, and Rietveld (2008).

[12] This paper is also related to the literature that studies the influence of gasoline on the transportation industry, and the oil industry more generally. Busse, Knittel, and Zettelmeyer (2013) investigate how sensitive are consumers to expected future gasoline costs when they make new car purchases. Sallee, West, and Fan (2016) study whether consumers value fuel economy in the market for used automobiles. See Anderson and Sallee (2016) for a review assessing the efficiency of fuel taxes and efficiency standards mitigating environmental externalities from cars. Asker, Collard-Wexler, and De Loecker (2017) quantify the misallocation attributable to market power in the global oil industry.

[13] Hysteresis or "stickiness" in travel behavior has also been documented in laboratory experiments by Innocenti, Lattarulo, and Pazienza (2013), both for car and public transit use.

[14] There is also a large literature discussing heterogeneities/asymmetries in elasticities arising from the "rebound effect." The rebound effect refers to the increased consumption of energy services due to improvements in energy efficiency that makes those services cheaper. See, *e.g.*, Small and Van Dender (2007) and Frondel and Vance (2013). Sorrell, Dimitropoulos, and Sommerville (2009) survey studies on the rebound effect. West, Hoekstra, Meer, and Puller (2017) study the rebound effect in the context of household driving response induced by the "Cash for Clunkers" program. They find that households who purchase more fuel-efficient vehicles do not drive additional miles after purchase (*i.e.* no rebound effect).

[15] Frameworks in the style of Heckman (1981) have been broadly used in the literature to analyze dynamic choices and/or incorporate inertia (*e.g.* Daganzo and Sheffi 1982; Johnson and Hensher 1982; Hirobata and Kawakami 1990; Morikawa 1994; Srinivasan and Bhargavi 2007; Chatterjee 2011; Yáñez, Cherchi, Ortúzar, and Heydecker 2009; see Golob, Kitamura, and Long 2013 for a survey). In contrast, this paper takes a complementary approach by focusing on structural estimation of the dynamic decision process under the hypothesis that decision and state variables are realizations of a controlled stochastic process as in Rust (1994). See Heckman (1981), Eckstein and Wolpin (1989), and Rust (1994) for detailed discussions.

the dynamic optimization problem of the agents. As emphasized by Cantillo, Ortúzar, and Williams (2007, p. 196): "Although the concept of inertia in travel choice modeling is not new [...], it has remained an important (although unresolved) topic because of its potential bearing on transport policy (*e.g.*, how to reduce car dependency)." This paper contributes to this literature by using structural methods to uncover the primitives of the dynamic decision process, and showing the effects on substitution patterns of policies affecting such primitives, in particular long-run elasticities. The proposed model and estimation techniques are adapted from the industrial organization (IO) literature described below.

Theories of hysteresis were first developed by Baldwin (1989), Dixit (1989), Krugman (1989), and Baldwin and Krugman (1989) in the context of international trade. These theories were used to explain persistence of the trade deficit, and investment under uncertainty (Dixit and Pindyck 1994). Roberts and Tybout (1997) empirically investigate such hysteresis framework studying entry in foreign markets. They estimate a dynamic discrete choice model of exporting behavior to separate profit heterogeneity and sunk entry costs in explaining plants' exporting status. Also following the hysteresis literature, Clerides, Lach, and Tybout (1998) estimate a dynamic discrete choice model to characterize export market participation decisions. Das, Roberts, and Tybout (2007) estimate a dynamic structural model of firms' exporting behavior. Their model features firm level heterogeneity, uncertainty about future profits, and sunk costs to enter foreign markets. Dunne, Roberts, and Samuelson (1988) document patterns of firm entry, growth, and exit in the U.S. manufacturing industries. Aw, Roberts, and Xu (2011) estimate a dynamic structural model of a producer's joint decisions to invest and export using plant-level data for the Taiwanese electronics.[16] Hysteresis has also been extensively studied in unemployment since the work by Blanchard and Summers (1986). The main insight here is that the natural rate is history dependent or, more generally, that the history of unemployment may affect the natural rate. See Blanchard and Katz (1997) for a survey. Dunne, Roberts, and Samuelson (1989) study fluctuations in labor demand arising from the entry, growth, and exit of individual employers.

For the empirical analysis, I estimate a dynamic discrete choice model with switching costs. There is a large literature in IO on dynamic demand estimation (*e.g.* Boizot, Robin, and Visser 2001; Pesendorfer 2002; Hendel and Nevo 2006; Gowrisankaran and Rysman 2012),[17] switching costs and inertia (*e.g.* Klemperer 1987a,b; Shum 2004; Kim 2006; Dubé, Hitsch, and Rossi 2009, 2010; Handel 2013; Honka 2014; Sudhir and Yang 2014; Shcherbakov 2016; Hortaçsu, Madanizadeh, and Puller 2017),[18] and hysteresis in dynamic choice models

[16] Peters, Roberts, and Vuong (2018) extend the model of Aw, Roberts, and Xu (2011) by incorporating product and process innovations. Alfaro, Cuñat, Fadinger, and Liu (2018) estimate a dynamic firm level model in which real depreciations raise the cost of importing intermediates, and where hysteresis arises due to the large magnitude of R&D sunk costs. Costa and Gerard (2018) study the implications of hysteresis for the welfare evaluation of corrective policies, and estimate the long-run impact of a temporary energy saving program in the context of residential electricity use in Brazil.

[17] See Aguirregabiria and Mira (2010) and Aguirregabiria and Nevo (2013) for surveys.

[18] See Luco (2013) and Illanes (2016) for recent applications to pension plan choices. See Klemperer (1995) and Farrell and Klemperer (2007) for surveys.

(*e.g.* Roberts and Tybout 1997; Clerides, Lach, and Tybout 1998; Das, Roberts, and Tybout 2007; Aw, Roberts, and Xu 2011).[19] Yang (2010) shows that switching costs are identified in a dynamic framework with market level data as in this paper. To estimate the econometric model I construct a GMM estimator using an adapted version of the procedure proposed by Gowrisankaran and Rysman (2012) and Shcherbakov (2016), who nest a demand system, in the style of Berry, Levinsohn, and Pakes (1995), within a dynamic framework in the style of Rust (1987). The procedure exploits a population moment condition, that is a product of instrumental variables and a structural error term, to form a nonlinear GMM estimator. One difference in my model relative to Gowrisankaran and Rysman (2012) and Shcherbakov (2016), is that I do not need to use the inclusive value sufficiency assumption to estimate dynamic model. This is because of two features of my empirical setting/industry: there are only two inside modes of transportation (instead of J inside products as it is typically the case in IO), and the other characteristics of the modes of transportation are fixed. Thus, using the vector of prices as a state variable is computationally feasible in my case. The base price coefficient, denoted by α_1 in my model, becomes a "nonlinear" parameter in the estimation algorithm, using the terminology by Nevo (2001). For these reasons, the instruments and identifying assumptions are different from the ones used in these papers, as explained in section 4. The finding that estimates of price sensitivity might be biased in misspecified static or myopic models is consistent with prior work in IO, such as, *e.g.*, Hendel and Nevo (2006), Shcherbakov (2016), Perrone (2017), Donna and Espin-Sanchez (2018).

2 Urban Transportation in Chicago

In this section, I describe the urban transportation industry in Chicago, the dataset, and perform the preliminary analysis describing empirical regularities about the industry.

2.1 Urban Transportation Industry in Chicago

The Chicago metropolitan area is characterized by high levels of access to public transportation and population density. The mean market shares of modes of transportation to work were 70 percent for car (drive alone) and 12 percent for public transit (bus/streetcar and subway/rail) for all the Chicago metropolitan area during the period under analysis, according to the Census Transportation Planning Package.

The Chicago metropolitan area has the second largest public transit share in the U.S. after New York/Northern New Jersey. There are three public transit operators in Chicago and northeastern Illinois: the Chicago Transit Authority (CTA), Metra commuter rail, and Pace suburban bus. These public transit operators are supervised by the Regional Transportation Authority (RTA), which is the regional planning body. The Regional Transportation Asset Management System (RTAMS) provides planning and financial information on the

[19] Given the small number of options in my setting, limited consumer information/awareness about availability of these options does not play a major role. See, *e.g.*, Goeree (2008), Honka (2014), and Hortaçsu, Madanizadeh, and Puller (2017) for applications of models of limited consumer information/awareness.

transportation system in the northeastern Illinois area surrounding Chicago.

The Illinois Department of Transportation (IDOT) collects vehicle circulation information. It uses automatic traffic recorders (ATR) that measure the number of vehicles circulating on specific locations. There are seven major expressways crossing the city of Chicago. Vehicle circulation is available for five of these routes during the period under analysis: Bishop Ford, Dan Ryan, Edens, Eisenhower, and Kennedy. For the estimation of the structural model in section 4, I link vehicle circulation in each of these expressways, to public transit use around the location of the ATRs, as described in next subsection.

There is substantial heterogeneity between types of commuters at a given time and, for certain types, across their choices of modes of transportation over time in response to their relative cost. Some types use persistently the same mode, such as car or public transit. Other types may switch modes over time in response to the gasoline cost of driving (Cambridge Systematics Inc. 2007; Owen, Jane, and Kopp 2007; Long, Lin, and Proussaloglou 2010).

2.2 The Dataset

I combine data from several sources. I use public transit ridership and fare cost of public transit, obtained from the three public transit operators supervised by the RTA. Vehicle circulation counting and vehicle classification are obtained from the IDOT. Gasoline cost of driving and world crude oil are obtained from the U.S. Energy Information Administration. I use the information from the Census Transportation Planning Package to define the size of the market. The shares of consumers who use only car or public transit (*i.e.* consumers who do not switch) are obtained from a travel survey by Pace. For the empirical analysis I focus on the period 2003 to 2009.[20] Below I briefly describe each source. See appendix A for details.

Monthly ridership data and fare cost of public transit were obtained from the corresponding service board of the RTA for the three public transit operators in the area. The data consist of average bus ridership from CTA and Pace, and average rail ridership from CTA. This ridership information is disaggregated by day of the week (weekdays, Saturdays, and Sundays/Holidays), and by bus route or rail station. I also collected Metra rail data, which is only available by branch and is a combined total for weekday, Saturdays, and Sundays/Holidays.

Vehicle circulation is collected by the IDOT using ATR stations. The following characteristics about the road, where the ATR is located, are collected: the functional classification of the roadway by the IDOT, the number of lanes in the roadway, the location of the ATR, an urban area indicator, the county where the ATR is located, and an expressway indicator. These characteristics are constant during the period under analysis. The IDOT maintained

[20] I collected the data for the period June 2000 to October 2009. However, there are some months with "not enough data to pass successful month edits" (*i.e.* months with missing data as reported by the IDOT) during the period 2000 to 2002 for the ATRs defining the routes in the structural model (see in appendix D the definition of the variable: Good months indicator). This feature precludes using years 2000 to 2002 for the structural estimation.

a network of 85 ATR locations throughout the state under this period. There was a subset of 36 ATRs that also collected vehicle classification counts for each of the 13 vehicle types defined in FHWA's guide. See appendix A for details about the ATR district distribution, its distribution by functional class, and the vehicle types classification.

Real gasoline cost of driving and world crude oil are constructed using data from U.S. Energy Information Administration, and the consumer price index (CPI) from the U.S. Bureau of Labor Statistics. For the gasoline cost of driving, I use monthly city average retail price of all grades all formulations gasoline prices for Chicago, and monthly national average of retail price of unleaded regular gasoline. Two series of world crude oil prices are used as instruments for the empirical analysis. The weekly all countries spot price FOB weighted by estimated export volume, and weekly U.S. spot price FOB weighted by estimated import volume. I use the CPI for the area Chicago-Gary-Kenosha to adjust for inflation the gasoline cost of driving in Chicago. I use the U.S. city average CPI to adjust for inflation the gasoline cost of driving in the U.S., and the real world crude oil prices. Almost identical results are obtained using other measures of inflation. To get a sense of the variability of these variables, figure 1 displays the evolution of the monthly gasoline cost of driving in Chicago and in the U.S., as well as crude oil prices for the period. While in some months the series are closer together and in other months they are farther apart, reflecting cross-sectional/regional variation across stations, the series generally track well each other.

I complement the data above with information about the relative shares of consumers who switch, and who do not switch inside modes of transportation. These data are obtained from a cross-sectional detailed household level survey conducted in 2006 by Pace, one of the three public transit operators in Chicago. The survey was administered to 1,330 randomly chosen households in the Chicago metropolitan area. This additional information is more desegregated than the market level data above. It describes the behavior of consumer types who continue using the same inside mode, without switching. The survey asked about households' travel patterns and attitudes towards everyday commuting. Based on the main factors determining their mode of transportation for traveling to work (*e.g.* public transit attitudes, safety, time, *etc.*), the survey respondents were grouped into seven distinct market segments. These segments define households' types with similar travel attitudes, such as households who are auto-dependent, transit-captive, transit-friendly, or households whose commute patterns vary considerably. I use these types to define the shares of consumers who do not switch from car or public transit. See appendix A for details.

I use the data described above to build a dataset of inside modes of transportation used (car and public transit) in each month during the period under analysis, and their characteristics. For vehicle circulation, I observe the number of vehicles that pass through the ATRs located at different locations through the expressways (interstate highways). As mentioned above, there are seven major expressways (routes henceforth), and vehicle circulation is available for five of them. A market is defined as a combination of route and month. I link traffic volume to public transit use in each route, by defining a radius around the location of the ATRs. For each route and month, I compute the public transit use as total weekday ridership (for each CTA rail station, and CTA bus and Pace lines) that lies within the radius.[21] Market shares are defined by dividing vehicle circulation and ridership by the market size. The market size was assumed to be measured by the population that lies within the radius around the ATRs. The radius was chosen to ensure that the market share of the outside mode equals the one from the Census Transportation Planning Package (share of people whose work trip mode is neither auto nor public transit). The outside mode can be conceptualized as including other modes of transportation, such as biking, walking, *etc.* An observation in this dataset represents a market share of car, public transit, or outside mode in a given month and route, the characteristics of the mode, which are constant under the period under analysis except its monetary cost, and the share of the consumers who use only car or public transit. See appendix A for details.

2.3 Preliminary Analysis

2.3.1 Vehicle and Public Transit Substitution

I begin by documenting the response of vehicle circulation and public transit use to the variation in the gasoline cost of driving. Tables A3 and A4 display the results. They display, respectively, OLS regressions of public transit ridership (CTA, Pace, and Metra) and vehicle circulation (Illinois, Chicago area, and by type of vehicle or classification) on the gasoline cost of driving and a set of covariates (observable characteristics for the station, route, or ATR), fixed effects (ATR, station, or route fixed effects), and monthly seasonal effects. Tables A3 and A4 also display two stage least squares (IV henceforth) regressions using the world crude oil prices and the U.S. city average gasoline cost of driving (along with the fixed effects and monthly seasonal effects) as instruments for the gasoline cost of driving in Chicago.

Tables A3 and A4 show the following. (1) Public transit use increases as the gasoline cost of driving increases, and vice versa. The substitution patterns are consistent across different public transports (CTA rail, CTA bus, Pace bus, and Metra rail) and specifications (with and without covariates, and with and without fixed effects), stronger for weekdays relative to Saturdays or Sundays/Holidays, and similar between OLS and IV regressions. (2) Vehicle use decreases as the gasoline cost of driving increases, and vice versa. Again the substitution patterns are robust across specifications, and similar between OLS and IV regressions. The substitution patterns for vehicle use are stronger in urban relative to rural areas, indicating the importance of having access to public transportation. The effects are statistically and economically significant for public transit and vehicle circulation. Overall, the results indicate substantial substitution between public transit and vehicle use in response to variation in the gasoline cost of driving. The substitution patterns are stronger in the Chicago area relative

[21] Bus ridership is only disaggregated by lines (not by stops within the line), so I include a complete bus line when 70 percent or more lies within the radius. I obtained similar results using thresholds of 50 and 90 percent. Results available upon request.

to rural areas. See appendix B for details and additional results.

2.3.2 Heterogeneity and Hysteresis in the Response to Gasoline Cost of Driving

Table 1 investigates heterogeneity in the response to the gasoline cost of driving for CTA rail ridership.[22] Two types of heterogeneity are presented. First, whether the response to a one percent increase in the gasoline cost of driving varies if gasoline prices have been increasing, flat, or declining over the previous months. Second, whether the response to the gasoline cost of driving differs by the absolute level of the gasoline price. On the one hand, table 1 shows that the elasticity estimates are 3 times higher (compared to the baseline estimation in row 5 of in panel A, table A3) when the gasoline cost of driving have been increasing over the previous three months. A similar result is found for high absolute values of the gasoline price. On the other hand, when the gasoline cost of driving have been flat or declining, and also for lower absolute values, the elasticity estimates are considerably lower. Figure 1B shows the evolution of the gasoline cost of driving in Chicago, along with the major oil/gasoline shocks (obtained from Kilian 2010 and Hamilton 2011) during the period under analysis, and the mean monthly CTA rail ridership elasticity by year (top number in figure 1B).[23] The estimates of the elasticities vary considerably in response to the different oil shocks. These patterns indicate that there is substantial heterogeneity in the response to gasoline prices. They also suggest that hysteresis may be present, which I investigate next.

Table 2 tests for the presence of hysteresis. It focuses on a specific shock, the hurricane Katrina, which was by far the largest refining shock (Kilian 2010), affecting gasoline prices in the U.S.[24] In this manner, the initial substitution between car and public transit may not just reflect heterogeneity in consumers' preferences. Table 2 panel A displays vehicle circulation (raw vehicle counts as measured by the corresponding ATRs) and public transit use (raw ridership corresponding to the CTA rail station closest to the corresponding ATR) for the 15 ATRs stations in the Chicago metropolitan area with the largest vehicle counts in the routes used for the structural analysis. Column 1 displays the pre shock levels (the month prior to hurricane Katrina) of vehicle counts and public transit ridership. Columns 2 and 3 display the levels of these variables during the shock (during hurricane Katrina); columns 4 and 5 display the levels after the shock (the month after hurricane Katrina). During the oil shock public transportation ridership increases considerably, 11.3 percent on average. Vehicle use also decreases considerably, 6.1 percent on average. Nevertheless, once the shock disappears and the gasoline cost of driving returns to the pre shock level,

[22] The table focuses on CTA rail for weekdays, the main public transportation used in the routes of the structural estimation. Similar patterns are obtained for CTA and Pace buses, and for vehicle circulation in the Chicago area, consistent with the evidence in subection 2.3.1. See also table 2.

[23] The elasticities reported in figure 1B correspond to the break down of the baseline specification in table A3, panel A, row 1, for different periods of 12 months.

[24] Hurricane Katrina formed in August 23, 2005, and dissipated in August 31, 2005. It was an exogenous shock affecting U.S. crude oil refining capacity in the Gulf of Mexico. At the time, other U.S. refineries were operating close to capacity. Thus, hurricane Katrina caused a major unanticipated reduction of the supply of gasoline in the U.S., producing a large increase in the price of gasoline in the United States.

public transit and vehicle use do not return to their initial values. Public transportation decreases only 3.8 percent on average, and vehicle use increases only 2.1 percent on average. Consumers do not switch back and remain using public transportation for a longer period of time, thus generating an asymmetry over time. This shows evidence of hysteresis in public transportation ridership and vehicle use in the Chicago area. Table 2 panel B repeats the analysis for the most popular ATRs stations located in rural areas without access to public transit. It shows that the decrease in vehicle use during the shock is similar in magnitude as the increase after the shock. There is no hysteresis in rural areas without access to public transit.[25] The presence of hysteresis in the Chicago metropolitan area and absence in rural areas are suggestive of a structural relationship between current and past mode of transportation choices in areas where consumers can substitute car and public transit. In next section I present an econometric model of optimal choice of modes of transportation, where the presence of switching costs can recreate the documented hysteresis.

3 A Model for Urban Modes of Transportation

In this section I present the econometric model. I model the demand for modes of transportation using a random coefficient mixed logit dynamic model. Consumer have idiosyncratic preferences for the different modes of transportation. They incur a switching cost when they choose a mode of transportation different from the one in the previous period.

3.1 Set Up

Assume that there are $r = 1, \dots, R$ routes, each with $i = 1, \dots, I_r$ types of rational, utility-maximizing consumers. Denote by t the discrete time periods measured in months. Consumers have different preferences for different modes of transportation. A market is defined as a period-route, as described in the previous section. For each consumer in each period-route, there are two mutually exclusive inside modes of transportation, indexed by $m_{irt} \in \{c, b\}$, denoting car and public transit in route r and period t. I index with $m_{irt} = 0$ the outside mode of transportation. This is a fictitious mode of transportation that allows consumers not to choose any of the inside modes. In each period-route, each consumer chooses one mode of transportation among car, public transit, or the outside mode. I abstract from consumer heterogeneity in the trip distance (unobserved in my data), and follow the literature assuming that average monthly trip distance is constant for consumers.[26] Denote by $M = \{0, c, b\}$ the set of choices that consumers face each route and period. Consumers have common discount

[25] Similar results to the ones in table 2 are obtained for the other oil shocks reported in figure 1B.

[26] This is due to data limitations (unobserved vehicle-miles and transit-miles traveled) and to keep the model tractable computationally. Note that although the average monthly distance is constant, the monthly trip distance is not necessarily constant over time. The latter is my preferred interpretation. See Ben-Akiva and Lerman (1985), Small, Verhoef, and Lindsey (2007), Button (2010), and the references there, for discussions of this issue using aggregate level data. See the second paragraph in the related literature for reviews of the literature studying the short-run responses of miles traveled to fuel/gasoline prices. For the period under analysis in Chicago, the mean one-way trip distance for work purposes is 7.3 miles (Bricka 2007, Table T-24).

factor $\beta \in (0,1)$, and the horizon is infinite.

The consumer incurs a switching cost in period t if the inside mode of transportation chosen in t is different from the mode of transportation chosen in $t-1$ for a given route r. In subsequent periods, consumers do not incur a switching cost if they choose the same inside mode as in the previous period, or if the switch to the outside mode. Consumers do incur a switching cost, however, if the switch from the outside mode of transportation to one of the inside modes. Let $\phi_m > 0$ with $m \in \{c, b\}$ denote the switching cost to inside mode of transportation m. Switching costs are defined as in Klemperer (1987a,b) and Shcherbakov (2016) as constant across periods and known to consumers.[27]

The preferences of consumer i for inside mode m in route r, period t are represented by:

$$U(p_t, m_{irt}, m_{irt-1}, \epsilon_{imrt}) = \alpha_{0mr} - \alpha_{1i}p_{mt} - \phi_m \mathbf{1}\{m_{irt-1} \neq m_{irt}\} + \tau_r^D + \tau_t^D + \xi_{mrt} + \epsilon_{imrt},$$

where $p_t \equiv (p_{ct}, p_{bt})$ with p_{ct} and p_{bt} denoting, respectively, the gasoline cost of driving and the fare cost of public transit in period t;[28] m_{irt} denotes the mode of transportation chosen by the consumer i in route r in period t; ϵ_{imrt} is an additive *i.i.d.* utility shock to consumer i in route r in period t described below; α_{0mr} is a mode and route specific constant; α_{1i} are individual specific parameters that capture consumers' preferences for price described below; $\phi \equiv (\phi_c, \phi_b)$ is the vector of switching costs parameters for the inside modes; $\mathbf{1}\{\cdot\}$ is an indicator function; τ_r^D, τ_t^D capture the preferences in route r and period t, using fixed dummy variables for route and monthly seasonal effects, respectively; and ξ_{mrt} is the valuation of unobserved, by the econometrician, characteristics of inside mode m_{irt} in route r in period t. In each route r and period t, I normalize the characteristics of the outside mode, $m_{irt} = 0$, such that $p_{0rt} = \xi_{0rt} = 0$ for all (r, t). I model the distribution of consumers' preferences for the gasoline cost of driving and fare cost of public transit as $\alpha_{1i} = \alpha_1 + \Sigma\nu_i$, where α_1 is a parameter that captures the mean price sensitivity, and where the shocks ν_i are constant over time and drawn *i.i.d.* from a distribution with pdf $P_\nu(\nu_i)$ assumed to be a standard normal; Σ is a parameter that governs the distribution of the random coefficients or heterogeneity in consumer preferences for price; ν_i captures unobserved, by the econometrician, individual characteristics, and $P_\nu(\nu_i)$ is a parametric distribution assumed to be a standardized Normal, $\mathcal{N}(0,1)$, for the estimation. Consumer preferences for inside modes of transportation in this setting are captured by preference heterogeneity, switching costs, monetary costs, the set of fixed effects, and unobserved characteristics. Let $\alpha \equiv (\alpha_{0mr}, \alpha_1)$, and let $\theta \equiv (\alpha, \Sigma, \phi)$ be the vector of structural parameters to be estimated. Denote by

[27] This specification assumes that the initial fixed cost of choosing inside mode of transportation j, the start-up cost, is the same as the fixed cost of switching to a different inside mode of transportation, the switching cost. This assumption is made because the initial choice of the mode of transportation is unobserved in the data. In this industry consumers do not switch frequently inside modes of transportation, so the assumption seems reasonable (see, *e.g.*, Button 2010 for a general discussion, and figure 2A for the estimated mean probabilities in the empirical application in this paper).

[28] For the empirical analysis I use, respectively, the gasoline price in Chicago (cents per gallon) and CTA rail fare (cents per trip). Variables definitions are in appendix D.

$\delta_{mrt} \equiv \alpha_{0mr} - \alpha_1 p_{mt} + \xi_{mrt}$, the mean utility for inside mode m, in rout r, in period t. This is the portion of the utility that is constant across types of consumers net of the switching cost. Let $\overline{U}(p_t, m_{irt}, m_{irt-1}) \equiv U(p_t, m_{irt}, m_{irt-1}, \epsilon_{imrt}) - \epsilon_{imrt}$, be the utility function net of the additive utility shock.

3.2 State Variables and Value Function

There are four state variables.

Mode of transportation chosen in the previous period. As noted above, the consumer incurs a switching cost if the inside mode of transportation chosen in the current period is different from the mode chosen in the previous period.

Gasoline cost of driving and fare cost of public transit. Consumers observe the gasoline cost of driving car in the current period, p_{ct}, and form expectations about the future gasoline cost of driving. I model consumers' expectations about the evolution of the future gasoline cost of driving using the following specifications: (i) rational expectations using an autoregressive model of order 1; (ii) rational expectations using a frequency estimator, (iii) perfect foresight, and (iv) myopic expectations.[29] The fare cost of public transit, p_{brt}, is constant across routes and periods during the period under analysis due to regulation. I assume this is common knowledge for the consumers. Denote the pdf of $p_t = (p_{ct}, p_{bt})$ by $f(p_t|p_{t-1})$. I estimate the following specifications for the price transition matrix, $f(p_t|p_{t-1})$:

(i) Rational expectations using an autoregressive model (RE-AR): Under rational expectations using an autoregressive model, an AR(1) process is estimated assuming $p_{c,t+1} = \gamma_0 + \gamma_1 p_{c,t} + \iota_{c,t+1}$, where ι is normally distributed. I estimate γ_0, γ_1, and the standard deviation of ι from the data. Then, the distribution of $p_{c,t+1}$ is generated from $p_{c,t}$, and the estimated parameters.

[29] Predicting changes in crude oil and gasoline prices is difficult (see Hamilton 2009 and Alquist, Kilian, and Vigfusson 2013 for reviews of recent approaches used in the literature). In the context of purchases of automobiles Busse, Knittel, and Zettelmeyer (2013) model consumers' expectations of gasoline prices as following a random walk for real gasoline prices. This has the convenient implication that the current gasoline price is the expected future real gasoline price. Also in the context of demand for automobiles, Anderson, Kellogg, and Sallee (2013) find that average consumer beliefs are indistinguishable from a no-change forecast (*i.e.* that consumers believe that gasoline prices follow a martingale process). They also find deviation from the no-change forecast during the financial crisis of 2008. Anderson, Kellogg, Sallee, and Curtin (2011) find that the forecast accuracy of the Michigan Survey of Consumers predictions is similar to that of a no-change forecast, on average. For the analysis in this paper, I report the results from the specifications above to investigate the role of consumers' expectations on the estimates. An alternative approach would be to allow consumers to be more or less sophisticated in their predictions, than the extreme cases of either rational or myopic expectations, respectively. For example, in principle one could allow consumers to be more sophisticated than the rational expectations assumption, without assuming perfect foresight, by using information on crude oil markets or other macroeconomic variables to make projections into the future gasoline future prices. Another approach would be to use a variable that represents gasoline price expectations directly. These approaches would require incorporating these additional variables, used to model expectations month by month, as state variables, increasing the computational burden substantially. If under alternative approaches, consumers' expectations lie between (iii) and (iv), one can use the current estimates under those expectations as a benchmark.

(ii) Rational expectations using a frequency estimator (RE-FE): Under rational expectations using a frequency estimator, a non parametric frequency estimator is used to estimate the distribution of $p_{c,t+1}$.

(iii) Perfect foresight (PF): Under perfect foresight, it is assumed that there is no uncertainty, and consumers correctly predict the future gasoline cost of driving.

(iv) Myopic expectations (ME): Under myopic expectations, it assumed that consumers have myopia or static expectations in that they assume that the current gasoline cost of driving will prevail forever.

Additive utility shocks. I assume that the additive utility shocks, ϵ_{imrt}, are drawn *i.i.d.* across individuals, modes of transportation, routes, and time, from a standardized Gumbel distribution with pdf $g(\epsilon_{imrt})$.

Value function. The value function is given by:

$$\begin{aligned} V(p_t, m_{irt-1}, \epsilon_{imrt}) = \underset{m_{irt}\in M}{max}\{U(p_t, m_{irt}, m_{irt-1}, \epsilon_{imrt})+ \\ +\beta\ \mathbb{E}[V(p_{t+1}, m_{irt}, \epsilon_{imrt+1}|p_t, m_{irt-1}, \epsilon_{imrt})]\}, \end{aligned} \tag{1}$$

subject to the evolution of the state variables as described above. The expectation is taken over the cost of use of the inside modes, the gasoline cost of driving and the fare cost of public transit, and ϵ_{imrt}.

3.3 Choice Probabilities and Market Shares

The computation of the market shares used for the estimation in section 4 proceeds in three steps. The first step defines a contraction mapping using the conditional value function. The second step uses the conditional value function and the law of total probability to compute the probability of observing mode m_{irt} conditional on p_t, which is observed in the data, and unconditional on the previous choice, m_{irt-1}, which is unobserved. The third step computes the market shares as a function of the parameters and the observed states, by integrating over the distribution of consumer types.

Contraction mapping. The conditional value function of choosing mode m_{irt} in period t, and behaving optimally from period $t+1$ on, net of the utility shock ϵ_{imrt}, denoted by $v(p_t, m_{irt-1}, m_{irt})$, is given by:

$$\begin{aligned} v(p_t, m_{irt-1}, m_{irt}) = \overline{U}(p_t, m_{irt}, m_{irt-1})+ \\ + \beta \int \log\left(\sum_{m'_{t+1}\in M} \exp\left[v(p_{t+1}, m_{irt}, m_{irt+1})\right]\right) f(p_{t+1}|p_t)\mathrm{d}p_{t+1} + \beta\gamma, \end{aligned} \tag{2}$$

where the equality follows by replacing the integrated or *ex ante* value function, $\overline{V}(p_t, m_{irt-1}) =$

$\int V(p_t, m_{irt-1}, \epsilon_{imrt})g(\epsilon_{imrt})\mathrm{d}\epsilon_{imrt}$, by its closed-form expression using the well known properties of the Gumbel distribution; and $\gamma = 0.5772$ is the Euler's constant.

Conditional choice probabilities. The probability of observing the choice m_{irt} conditional on p_t and m_{irt-1}, denoted by $\mathbb{P}(m_{irt}|p_t, m_{irt-1})$, is:

$$\begin{aligned}\mathbb{P}(m_{irt}|p_t, m_{irt-1}) &= \mathbb{P}[v(p_t, m_{irt-1}, m_{irt}) + \epsilon_{imrt} \geq v(p_t, m_{irt-1}, m'_{irt}) + \epsilon_{im'rt}, \forall m'_{irt} \neq m_{irt}],\\ &= \frac{\exp\left[v(p_t, m_{irt-1}, m_{irt})\right]}{\sum_{m'_t \in M} \exp\left[v(s_t, m_{irt-1}, m'_{irt})\right]},\end{aligned} \tag{3}$$

where the second equality follows from the well known properties of the Gumbel distribution.

The probability of observing mode m_{irt} conditional on p_t, denoted by $\mathbb{P}(m_{irt}|p_t)$, is:

$$\mathbb{P}(m_{irt}|p_t) = \sum_{m'_{t-1} \in M} \mathbb{P}(m_{irt}|p_t, m'_{irt-1}) \;\times\; \mathbb{P}(m'_{irt-1}|p_t), \tag{4a}$$

$$= \sum_{m'_{t-1} \in M} \underbrace{\frac{\exp\left[v(p_t, m'_{irt-1}, m_{irt})\right]}{\sum_{\tilde{m}_t \in M} \exp\left[v(s_t, m'_{irt-1}, \tilde{m}_{irt})\right]}}_{\mathbb{P}(m_{irt}|p_t, m'_{irt-1})} \;\times\; \mathbb{P}(m'_{irt-1}|p_t), \tag{4b}$$

where the equality in (4a) follows from the law of total probability; and the equality in (4b) follows from the equation in (3).

Market shares. The market share function for mode m in route r in period t, denoted by $s_{mrt}(\cdot)$, is obtained by integrating over the distribution of consumer types:

$$s_{mrt}(p_t, \delta_{mrt}; \alpha_{0mr}, \Sigma, \phi) = \int \mathbb{P}(m_{irt}|p_t) \; P_\nu(\nu_i)\mathrm{d}\nu_i, \tag{5}$$

where $\mathbb{P}(m_{irt}|p_t)$ is given by the equation in (4b) with the $v(p_t, m_{irt-1}, m_{irt})$ given by the equation in (2); and $P_\nu(\nu_i)$ is the pdf of standardized Normal, $\mathcal{N}(0, 1)$.

4 Estimation and Instruments

I estimate the model by GMM using an adapted version of the procedure proposed by Gowrisankaran and Rysman (2012) and Shcherbakov (2016), as described below. To identify the parameters for the price coefficient, the heterogeneity, and the switching costs, I rely on instruments with the exclusion restrictions discussed below and the micro moments. The discount factor is set at $\beta = 0.99$, similar to Gowrisankaran and Rysman (2012).

4.1 Estimation

The model is estimated by GMM using the following algorithm. First, the outer routine performs the GMM estimation of the structural parameters of the model, $\theta = (\alpha, \Sigma, \phi)$. It relies on the moment condition $\mathbb{E}[Z' \cdot \xi(\theta^*)] = 0$, where Z is a matrix of instruments described in next subsection, $\xi(\theta)$ is a structural error term defined below, and $\theta^* = (\alpha^*, \Sigma^*, \phi^*)$ is the true value of the parameters. Second, within this routine, and given a candidate parameter vector, there is a nested subroutine that outputs $\xi(\theta)$. This is done by finding the value of the mean utility that equates the predicted market shares by the model to the observed market shares in the data. This subroutine takes the conditional value function and choice probabilities from the inner routine. Finally, the inner subroutine is nested within the previous subroutine. It solves the stochastic control problem numerically, given a candidate parameter vector. It outputs the conditional value function and choice probabilities. Below I describe each routine.

1. GMM search. This subroutine searches for the value of θ that minimizes the GMM objective. The GMM estimate is:

$$\hat{\theta} = \arg\min_{\theta} \left[\xi(\theta)' Z W^{-1} Z' \xi(\theta)\right], \tag{6}$$

where W is a positive definite weighting matrix chosen with a standard two step approach described in appendix C. To evaluate the GMM objective, first one needs to solve numerically for $\xi(\theta)$. This is done in the next subroutine.

2. Structural error. For each candidate parameter vector, this subroutine outputs the structural error, $\xi(\theta)$, taking as inputs the conditional value function and choice probabilities computed in the inner subroutine described below. For each candidate parameter vector, I use the equation in (5) with the choice probability in equation (4b) to compute the market shares as a function of the parameters.[30] The error term is defined as the unobserved products' characteristics, and compute it by solving for the mean utility level, δ_{mrt}, that equates:

$$s_{mrt}(p_t, \delta_{mrt}; \alpha_{0mr}, \Sigma, \phi) = S_{mrt}, \tag{7}$$

where $s_{mrt}(\cdot)$ is the market share function given by the equation in (5); and S_{mrt} are the observed market shares obtained from the data. I solve for $\delta_{mrt}(\theta)$ in the system of equations in (7) using Broyden's method for finding roots.[31]

The structural error is then defined as $\xi_{mrt} = \delta_{mrt} - \alpha_{0mr} + \alpha_1 p_{mt}$ for $m \in \{c, b\}$. To solve for $\xi(\theta)$, one needs to compute the market share function, which requires computing

[30] In this step, the main difference relative to Gowrisankaran and Rysman (2012) and Shcherbakov (2016) is that I also compute the shares of the consumers who do not switch predicted by the model as a function of the parameters, $s_{brt|b\tilde{t}}(p_t, \delta_{mrt}; \alpha_{0mr}, \Sigma, \phi)$ and $s_{crt|c\tilde{t}}(p_t, \delta_{mrt}; \alpha_{0mr}, \Sigma, \phi)$, given by the equation in (C.3) in appendix C. Thus, the instruments and identifying assumptions to recover the demand primitives are different (see next subsection). Once all the market share functions are computed, the estimation procedure in this step is similar Gowrisankaran and Rysman (2012) and Shcherbakov (2016).

[31] As in Gowrisankaran and Rysman (2012) an important issue that arises is whether this system of equations have a unique fixed point, which is an avenue of future research. I have experimented with wide variety of different starting values, using random number generators to pick them, and have always obtained the same solution. For robustness, all models have also been estimated using the following two methods to solve for $\delta_{mrt}(\theta)$ in the nonlinear system of equations in (7): a simplex search method, and iterating on it market by market analogously to the contraction mapping used by Berry (1994) and Berry, Levinsohn, and Pakes (1995). Similar results were obtained. See appendix C for details.

the conditional value function and conditional choice probabilities. This is done in the next subroutine.

3. Conditional value function and choice probabilities. For each candidate parameter vector, the conditional choice probabilities are computed using the equation in (4b). To evaluate the expression in (4b), one first need to compute the conditional value function $v(p_t, m_{irt-1}, m_{irt}; \theta)$. This is done by solving numerically the stochastic control problem (contraction mapping henceforth) in (2) for each given candidate parameter vector, as discussed in appendix C. To solve this contraction mapping, one needs to specify the distribution characterizing consumers' expectations about the gasoline cost of driving and fare cost of public transit, denoted by $f(p_t|p_{t-1})$ in (2). In the next section, I report the estimation results using the four specifications for $f(p_t|p_{t-1})$ described in subsection 3.2.

For each candidate parameter vector, the estimation algorithm starts with step 3, which is nested within step 2. These nested routines are repeated interchangeably until both converge. Finally, these steps are nested in the GMM search in step 1.

See appendix C for details about the computation of the weighting matrix in step 1, the procedure to solve for the mean utility in step 2, the contraction mapping in step 3, the initial condition of the consumers and initial market shares, the discretization of p_t, simulation of consumer types, the computation of the shares of consumers who do not switch, and the computation of the standard errors.

4.2 Identification and Instruments

I discuss identification informally. I rely on instruments and the micro moments to identify the price coefficient, the heterogeneity parameters, and the switching costs parameters. Identification requires at least one instrument for price, the heterogeneity parameter (Berry and Haile 2014), and each switching cost parameter. Identification of the model also requires a unique fixed point for the contraction mapping in (2).

The main challenge for identification in my model is to separate the effect resulting from the heterogeneity in consumer preferences, Σ, from the hysteresis or structural state dependence caused by switching costs, ϕ. This is well known issue (*e.g.* see Heckman 1981; Dubé, Hitsch, and Rossi 2010; Sudhir and Yang 2014; and the references there). Consider, for example, a consumer who chose car in the last period and does not switch to public transit in the current period, when the gasoline cost of driving increased. One explanation could be that the consumer has strong idiosyncratic preference for car, low α_i. In this case the consumer may continue using the car for several periods, even if the gasoline cost of driving increases substantially. An alternative explanation could be that the consumer has a relatively high α_i, but the value of continue choosing car exceeds the value of switching to public transit minus its switching cost, $v(p_t, m_{irt-1} = c, m_{irt} = c) > v(p_t, m_{irt-1} = c, m_{irt} = b) - \phi_b$. In this case, the consumer would switch to public transit if the gasoline cost of driving increases substantially.

To identify switching costs I rely on instruments, use micro moments defined by the shares of consumers who do not switch, and use functional form restrictions for the unobserved consumer preferences. First, I rely on instruments to identify ϕ. The basic insight is to use instruments that shift the previous period decision, to identify the switching costs using the current period decisions. Following the intuition in table 2 (Katrina oil shock), consider an exogenous increase in the gasoline cost of driving in t, relative to $t-1$ and $t+1$. Assume that the other characteristics of the modes of transportation do not change in $t-1$, t, and $t+1$. Let $p_{c,t} > p_{c,t-1} = p_{c,t+1}$ (*i.e.* exogenous oil shock in t). If the market share of car in $t+1$ is lower than the market share of car in $t-1$, this provides evidence of switching costs, because the characteristics of the modes of transportations (including the gasoline cost of driving) are the same in $t-1$ and $t+1$. An exogenous change in the gasoline cost of driving in the previous period affects consumer's choice today only if switching costs are present, but has no effect on today's choice if there are no switching costs. Then the resulting dependence between current and past market shares is attributed to the switching costs. This argument relies on the distribution $P_\nu(\nu_i)$ being constant across periods, which is discussed below. For the estimation I use the following exclusion restriction $\mathbb{E}[Z' \cdot \xi_{mrt}] = 0$, and use as instruments the previous period international price of oil, and the average gasoline cost of driving in Chicago during the previous three periods. This approach is similar to the one used by Shcherbakov (2016) to identify the switching cost parameters using data aggregated at the market level.

Second, I use information about share of consumers who did not switch. These data, described in subsection 2.2 and appendix A, helps identify the switching costs and heterogeneity parameters. For example, consider the information from observing the market share of consumers who chose to use car and continue using it during $\tilde{t}$ periods, when the gasoline cost of driving was increasing. Asking the model to reproduce such market shares helps to identify more precisely the heterogeneity in consumer preferences, Σ, because its value has to be such that there sufficient types of consumers, α_i, who have strong idiosyncratic preference for car, low α_i. It also helps identify the switching cost parameters, because the model is to reproduce the observed hysteresis using the information of the consumers who did switch. Similarly, using market shares of consumers who do not switch from public transit, when the gasoline cost of driving decreases, helps to identify the switching costs to car. (See also figure 2B discussed in subsection 5.1.) I use these data to define additional micro moments (no-switching moments, henceforth). The no-switching moments are defined as the difference between the share of consumers who do not switch from car and public transit during the $\tilde{t}$ prior periods obtained from the data, and the model's predictions. These no-switching moments (for car and public transit) are added to the GMM objective in the equation in (6). The additional computational burden does not increase substantially. See appendix C for details about the computation of the shares of consumers who do not switch. In the next section I show that these moments improve the precision of Σ and ϕ.

Finally, I assume that the consumers' preferences for the gasoline cost of driving and the fare cost of public transit, $P_\nu(\nu_i)$, are constant over time and follow a Normal distribution. The critical restriction is the former. It allows to separate the effect on choices resulting from the heterogeneity in preferences, and the hysteresis due to switching costs. It is not possible to generate the hysteresis documented in subsection 2.3 with idiosyncratic preferences that are constant over time.[32] I believe that the constant preferences assumption is justified for urban transit choice in Chicago during the period under analysis, because the overall distribution of demographics of the consumers and the characteristics of the modes of transportation did not change substantially.[33]

Two features from the empirical setting are also useful for identification of the switching costs parameters. First, there are frequent and large variation in the gasoline cost of driving during the period under analysis, as can be seen in figure 1. Consider, for example, the large and rapid increase in the gasoline cost of driving in 2008. Certain types of consumer will switch to public transit. In the following periods, the gasoline cost of driving decreased substantially. So a subset of those types, the ones with higher α_i (weaker preference for car), will continue using public transit on subsequent periods, even after the price returns to the previous pre-shock level. This helps improve the precision of the switching costs estimates, because such hysteresis can only be generated from the switching costs in the model. Second, the variation in the gasoline cost of driving in Chicago is driven mostly by variation in the cost of international oil prices. Thus, the response of vehicle and public transit use in Chicago has arguably a small effect on the global demand for crude oil, reinforcing the argument in the first point. I discuss potential price endogeneity below.

For robustness, I also present estimates of different specifications of the model with and without switching costs, and with and without consumer heterogeneity. A similar approached was used by Dubé, Hitsch, and Rossi (2010).[34]

I use world crude oil and the U.S. city average gasoline cost of driving as instruments for the gasoline cost of driving in Chicago. This is for robustness, to address potential concerns about the gasoline cost of driving in Chicago responding endogenously to car and public

[32] In principle, one could use a more general specification and allow for serial correlation over time in the unobserved consumer preferences, ν_{it}. This would significantly increase in computational burden. More important for my application, with data aggregated at the market level, I believe there is not enough empirical variation to estimate serially correlated preferences with a reasonable level of precision or that these preferences can be estimated using flexible functional forms.

[33] *E.g.* learning as discussed by Dubé, Hitsch, and Rossi (2010) did not play major role during the period under analysis. See Goodwin, Kitamura, and Meurs (1990) for an earlier discussion.

[34] Dubé, Hitsch, and Rossi (2010) also fit models with more flexible specifications for the unobserved heterogeneity, such as a mixture of normals. Using such specifications in my model results in substantial increase in the computational burden. In addition, the increased number of parameters introduces additional identification issues due to the nature of my data, which is aggregated at the market level, while Dubé, Hitsch, and Rossi (2010) use a household panel data.

transit use in Chicago.[35] It helps to identify the parameter α. The identifying assumption is that monthly- route-specific car and public transit demand shocks are uncorrelated with monthly gasoline prices, conditioning on mode of transportation, route, and seasonal fixed effects. The gasoline cost of driving in Chicago is correlated with the world crude oil and the U.S. city average gasoline cost of driving across months, because world crude oil is the main determinant of the marginal cost of driving. They are uncorrelated with month-specific valuations due to the exclusion restriction.

5 Results

5.1 Parameter Estimates

The estimated parameters are present in table 3. The estimates are presented using the following specifications of the model. (1) A static model without switching costs, where $\phi = 0$ and consumers only value the current period utility. Two specifications of this model are presented in panel A, with and without random coefficients. These specifications are equivalent, respectively, to a standard logit or mixed logit model, that has dominated the prior literature. (2) A myopic expectations (ME) model, where consumers assume that the current period gasoline cost of driving will prevail forever. This is implemented using a static model with switching costs, where consumers only value the current period utility, but they incur the switching costs if the current period inside mode of transportation is different from the mode chosen in the previous period. Again two specifications of this model are presented in panel A, with and without random coefficients. (3) A dynamic model, as outlined in section 3 with rational expectations, using an autoregressive model (RE-AR) to model consumers' expectations, and without additional no-switching moments (micro moments). Four specifications of this model are presented in panel B, with and without switching costs, and with and without random coefficients. (4) Same as 3 with additional no-switching moments. This is the preferred specification, and is in panel C.

The estimation algorithm from subsection 4.1 is applied to each model in table 3, with the obvious modifications. For example, the specification with a static model without switching costs and without random coefficients corresponds to standard discrete choice logit model. So step 3 can be skipped because the conditional value function are the current period utility function, and the inversion in step 2 has a closed-form expression.

The estimated parameters are sensible in magnitude and sign. Three conclusions stand out from this table. First, from panel A, one can see how the observed hysteresis is rationalized in a model with myopic expectations, relative to a static model without switching costs. It is useful to start by considering a dynamic model with switching costs and non myopic

[35] One possible source of endogeneity could be, for example, that variations in Chicago economic conditions (correlated with personal income) may be correlated with changes in gasoline prices in Chicago. Hence, gasoline price estimates will capture this cyclical effect on ridership/vehicle use. Other source could be that gasoline local taxes might be correlated with economic cycles, therefore biassing the estimates.

expectations. In such model, a price increase in the current period generates two incentives. On the one hand, because of the switching costs, the consumers are less likely to switch. This is due to the effect of the switching cost on the current period flow utility. On the other hand, the more permanent the consumers expect to be the price increase, the more likely is the consumer to switch. This is due to the effect of the future prices on the continuation value. In panel A, for the myopic expectations model, the consumer believes that the current price increase in the gasoline cost of driving will prevail forever. This makes some consumers quite responsive to the price increase, even if they have to incur a switching cost in the current period. Thus, the myopic expectations model rationalizes the observed hysteresis with a high price sensitivity (due to the role of the second effect) and high switching costs. The static model without switching costs attempts to rationalize the hysteresis with a large estimate for the standard deviation of the random coefficient for price, Σ. However, the hysteresis cannot be reproduced with idiosyncratic preferences that are constant over time. Finally, note that the static model without switching costs and without random coefficients shows a larger value of the GMM objective relative to the one with random coefficients, indicating further difficulties to explain the patterns in the data.

Second, from panel B, one can see the role of switching costs rationalizing the observed hysteresis. The specifications of the models in panel B hold constant the expectations of the consumers, which are modeled as rational expectations using an autoregressive model (RE-AR). As expected, the estimated base price coefficient is smaller in absolute value in the dynamic model with switching costs, relative to the dynamic model without switching costs. Again, the two incentives from the previous paragraph are present. By holding constant the effect of the expectations, one can "isolate" the first effect, that switching costs makes consumers less price sensitive, due to the switching cost in the current period utility. A model that does not account for the switching costs would attribute too much importance to consumer preferences for price, thus biasing the price elasticities as discussed in subsection 5.3. Finally, note that the estimated mean price coefficient from the static model (right of panel A) is smaller (larger in absolute value) than the one from the dynamic model with switching costs and rational expectations. This is because the static model rationalizes when consumers do not switch using only the price coefficient (*e.g.* α_1), while the dynamic model also uses the switching costs (*e.g.* α_1, ϕ_b, and ϕ_c). Thus, a smaller (larger in absolute value) α_1 is needed by the static model without switching costs to rationalize the *same* data.

Third, from panel C, one can see the role of the additional no-switching moments (micro moments). Incorporating the no-switching moments helps to identify the switching costs and heterogeneity parameters. By comparing the specification in panel C, with the first specification in panel B, one can see that the base price coefficient, the standard deviation of the price coefficient, and the switching costs are more precisely estimated. This is similar to Petrin (2002), Berry, Levinsohn, and Pakes (2004), and Goeree (2008).

Figure 2 shows the conditional choice probabilities of the inside modes of transportation, and the switching patterns of a particular consumer type. Figure 2A shows the mean, across consumer types and markets, conditional choice probabilities (left vertical axis) given by the equation in (4) as a function of the observed gasoline cost of driving (right vertical axis), p_t. The estimated conditional choice probabilities respond to the variations in the gasoline cost of driving, consistent with the substitution patterns described in tables A3 and A4. It can also be seen that the large price variation in 2008 (last price hike in the figure) generates large variations in these probabilities. It induced certain types of consumers to switch to public transit. Such variation is useful for identification of the switching costs parameters.

Figure 2B shows the behavior of a particular transit-friendly consumer type.[36] On the left vertical axis, the graph shows using dots, the difference in conditional value function of choosing car today given previous period choice of car, minus the conditional value function of choosing public transit today given previous period choice of public transit: $V_{cc} - V_{bb} \equiv v(p_t, m_{irt-1} = c, m_{irt} = c) - v(p_t, m_{irt-1} = b, m_{irt} = b)$, where the conditional value functions $v(p_t, m_{irt-1}, m_{irt})$ are given by the equation in (2). The vertical dotted lines indicate a switch in the inside mode of transportation by the consumer in such period. The graph also displays horizontal lines with the estimated switching cost parameters, ϕ_c and $-\phi_b$, and the gasoline cost of driving. The shaded area between the switching costs, ϕ_c and $-\phi_b$, defines a band of inaction using the terminology from Dixit (1989, figure 1). Inside this band the consumer does not switch, because although the value of the alternative inside mode is larger than the value of the current mode, the difference is less than the switching cost. The consumer switches from car to public transit (*e.g.* vertical line in 2007m5) if the difference in conditional value functions, $V_{cc} - V_{bb}$, crosses the bottom horizontal line, ϕ_b, from above.[37] Similarly, the consumer switches from public transit to car (*e.g.* last vertical line in 2008m11) if the difference in conditional value functions, $V_{cc} - V_{bb}$, crosses the top horizontal line, ϕ_c, from below. The band of inaction created by the switching costs causes inertia or "stickiness" in consumers' decisions, thus creating hysteresis at the aggregate, or market, level.

5.2 Robustness

The robustness of the model was tested in several ways. First, in table 3 I tested for consistency between the theoretical and estimated predictions for different specifications of the model. Second, also in table 3 and for each specification in (1), (2), and (3) in the previous subsection, I present the estimates with and without random coefficients, where the latter

[36]This is a transit-friendly type from the Pace survey (Cambridge Systematics Inc. 2007). It is a relatively price sensitive consumer type, who switches several times during the period under analysis. Naturally the mean consumer does not switch as frequent.

[37]In the graph, the consumer using car in the current period will only switch to public transit if the value of continue using car minus the value of switching is negative: $v(p_t, m_{irt-1} = c, m_{irt} = c) - v(p_t, m_{irt-1} = c, m_{irt} = b) < 0$. But $v(p_t, m_{irt-1} = c, m_{irt} = c) - v(p_t, m_{irt-1} = b, m_{irt} = b) < -\phi_b \implies v(p_t, m_{irt-1} = c, m_{irt} = c) - v(p_t, m_{irt-1} = c, m_{irt} = b) < 0$. This is because $v(p_t, m_{irt-1} = b, m_{irt} = b) = v(p_t, m_{irt-1} = c, m_{irt} = b) + \phi_b$, *i.e.*, the conditional value of choosing mode $\tilde{m}_{irt}$ by switching from m_{irt-1} to $\tilde{m}_{irt} \neq m_{irt-1}$ is the same as the conditional value of choosing mode $\tilde{m}_{irt}$ given that the consumer chose the same mode in the previous period $\tilde{m}_{irt} = m_{irt-1}$, minus the cost of switching from $m_{irt-1} \neq \tilde{m}_{irt}$ to $\tilde{m}_{irt}$. See page A-17 in appendix C for details.

is implemented by having one representative consumer type with $\alpha_i = \alpha$. Third, I tested for different specifications of the expectations in the model in the appendix, in table A6. I present the estimates using different specifications for the price transition matrix, $f(p_t|p_{t-1})$, representing consumers' expectations: (i) rational expectations using an autoregressive model (RE-AR), where consumers predict the gasoline cost of driving using an autoregressive model of order 1, presented in panel B in table 3; (ii) rational expectations using a frequency estimator (RE-FE), where consumers predict the gasoline cost of driving using a frequency estimator; and the correct cost of public transit, presented in panel A in table A6; (iii) perfect foresight (PF), where consumers correctly predict the future gasoline cost of driving and the fare cost of public transit, presented in panel B in table A6; (iv) Myopic expectations (ME) as described in the model 2 above, presented in panel A in table 3. The estimation algorithm from subsection 4.1 is applied to each specification of the price transition matrix, $f(p_t|p_{t-1})$. Fourth, I tested for different specifications of the additional micro moments in the appendix, in table A7. The table shows the estimates of the dynamic model using different combinations of rational expectations (RE-AR and RE-FE), and additional moments. For the additional moments I use the share of consumers who do not switch for 3, 6, and 12 months periods. Finally, in appendix C are described additional robustness analysis related to the estimation procedure. The estimated parameters did vary sometimes across some of these robustness tests. However, the implications discussed in the next subsections are robust in the cases examined.

5.3 Implications

In this section I present the implications of the estimates, and compare them to static and myopic ones. I define own- and cross-price elasticities, respectively, as car and public transit response to a change in the gasoline cost of driving. To compute the elasticities I adapt the procedure by Hendel and Nevo (2006). I proceed in four steps. First, I simulate the choice probabilities and market shares using the estimated parameters and the observed gasoline cost of driving. Second, I simulate an increase in the gasoline cost of driving by adding a small amount to the observed path of the gasoline price. To evaluate short-run price response, I generate a temporal increase by changing only the current gasoline cost of driving. To evaluate the long-run price response, I generate a permanent increase by changing the whole path of the gasoline cost of driving, not just the current price. Third, I reestimate the price process, and solve for the optimal behavior given the new price process. Fourth, I simulate the new choice probabilities and market shares. For all models, the elasticities are computed by calculating the percent change in the market share in step 4 relative to the initial values in step 1, with a 1 percent change in the gasoline cost of driving. The elasticities are evaluated at each of the observed data points, and then averaged over the observations. Next table reports such average by market, and the weighted average across markets, denoted by mean.

Table 4 presents the average ratio of elasticities computed from the relevant model, relative to the long-run elasticity computed from the dynamic model with switching cots and with no-switching moments (micro moments) from table 3, panel C. Three patterns emerge. First, long-run elasticities from the dynamic model are 13.0 and 15.6 percent higher than the short-run elasticities, respectively, for the mean own- and cross-price elasticities.[38] That is, long-run elasticities are more elastic as expected. The difference between short and long-run elasticities might have important consequences for the evaluation of policies, as previously documented in the literature. Second, the estimates suggest that the static own- and cross-price elasticities are smaller than the long-run elasticities. Static own- and cross-price elasticities underestimate the dynamic ones by 32.7 and 26.5 percent, respectively, on average. The bias varies across routes (markets). The bias is up to 34.5 percent for the own-price elasticity, and up to 34.8 percent for the cross-price elasticity, for routes 2 and 1 respectively. Finally, myopic estimates also underestimate the long-run elasticities. The bias is larger than for the static model, 56.2 and 80.8 percent for own- and cross-price elasticities, respectively.

Finally, I investigate heterogeneity in the estimated elasticities, similar to the analysis in subsection 2.3.2. I find that there is still heterogeneity as a result of the hysteresis, although smaller in magnitude to one reported in subsection 2.3.2.[39] The results indicate that the estimates of the elasticities from the dynamic model tend to be more stable.

5.4 Counterfactuals

I use the estimates from subsection 5.1 to study the implications of two main policy interventions, or counterfactuals: a reduction in switching costs, and a permanent increase in the gasoline cost of driving due to an announced gasoline tax. Below I describe these policies, and results.

Welfare measures. The expected consumer surplus in dollars for consumer type i, denoted by $\mathbb{E}(CS_i)$, is:[40]

$$\mathbb{E}(CS_i) = \frac{1}{\alpha_{1i}}\mathbb{E}[V(p_t, m_{irt-1}, \epsilon_{imrt})], \tag{8}$$

where $\mathbb{E}(\cdot)$ denotes the expectation operator taken over the random shocks ϵ_{imrt}; and the expression $V(p_t, m_{irt-1}, \epsilon_{imrt})$ is given by the equation in (1).

Consumer welfare for type i is defined as the change in the consumer surplus or compensating variation, CV, that results from the policy change. I compute the difference between the consumer surplus before and after the policy change. Two main policy changes are considered below. For the welfare results below I compute the total consumer surplus calculated as the weighted sum of $\mathbb{E}(CS_i)$ using the weights reflecting the number of consumers who

[38] Obtained as follows: $13.0 = (1 - 0.87) \times 100$, and $15.6 = (1 - 0.844) \times 100$. Similarly for the biases reported in the second and third patterns.

[39] Results are available upon request.

[40] This is a partial equilibrium measure from the model in section 3. It does not account for changes in welfare derived from features not accounted by the model (such as the gasoline impact on the environment, road accidents, *etc.*), or externalities due to the changes in the use of the modes of transportation (such as traffic congestion, public transit delays, *etc.*).

face the same representative utilities as the sampled consumer. That is:

$$\mathbb{E}(CV) = \int_{\nu_i} \left[\mathbb{E}(CS_i^1) - \mathbb{E}(CS_i^0)\right] \, \mathrm{d}P_\nu(\nu_i), \tag{9}$$

where $\mathbb{E}(CV)$ denotes the weighted sum across types of consumers of the compensating variation; the superscripts 0 and 1 refer, respectively, to before and after the policy change; and $\mathbb{E}(CS_i)$ is given by the equation in (8).

Counterfactual results. Two policy changes are considered. (1) A policy that would change the switching cost for public transit, without affecting the gasoline cost of driving. (2) Same as the policy in point 1, plus a permanent gasoline tax increase by \$1 per gallon.[41]

Table 6 reports the results from the counterfactuals. Panel A reports the results from policy 1, and panel B from policy 2. The table displays the value (mean per individual, per market, per period) of the compensating variation, $\mathbb{E}(CV)$, given by the equation in (9); the switching costs incurred from the inside modes; and the mean, across consumer types and markets, conditional choice probabilities of each inside mode given by the equation in (4).

Three patters emerge from the counterfactual analysis. First, a moderate decrease in the switching costs to public transit of 20 percent, increases the mean conditional choice probability of public transit by 2 percent ($0.0754 - 0.0563 = 0.0191$), and has a small effect on the conditional choice probability of car. Second, the gasoline tax (without change in the switching cost of public transit) has a large and negative impact of \$33.00 on the mean (per month, per market, and per consumer type) compensating variation, as expected (recall that the use of the gasoline tax revenue is not accounted in the analysis in this section). It decreases the mean probability of switching to car, from public transit or the outside option, by 4 percent ($0.1011 - 0.0624 = 0.0387$), and increases the mean probability of switching to public transit, from car or the outside option, by 3 percent ($0.0563 - 0.087 = -0.0307$). Interestingly, when the gasoline tax is bundled with a decrease in the switching costs to public transit of 20 percent, the negative impact on welfare is only \$11.69 on the mean compensating variation. Of course, there are distributional effects across consumer types. Consumer types who continue using automobile after the policy (gasoline tax plus decrease in the switching costs to public transit of 20 percent) are worse off than the mean consumer. Consumers who switch from car to public transit due to the policy (and those who continue using public transit) are better off than the mean consumer. Third, the analysis in tables 4 and 6 shows that estimates of using static or myopic models tend to mismeasure the effects of substitution between the inside modes of transportation. Consider the case of the gasoline tax without change in the switching cost of public transit. The static model underestimates the substitution from car to public transit by 29.7 percent on average. This is a result of the two effects discussed in subsection 5.1. On the one hand, the static model does not

[41] Li, Linn, and Muehlegger (2014) find that consumers respond to gasoline tax changes. Goulder, Hafstead, Kim, and Long (2018) study the distributional effects of a carbon tax. Knittel and Sandler (2018) study the implications of indirect gasoline taxes on pollution externalities in the California transportation market.

account for switching costs. It predicts higher probabilities of switching to both, car and public transit (0.216 and 0.123, respectively), than the dynamic model (0.062 and 0.107, respectively). On the other hand, the static model does not account either for the effect of consumers expectations on the decisions to switch. As expected in the case of the announced gasoline tax, the latter has a larger effect on the decisions to switch for consumers in the dynamic model. Several consumer types switch to public transit even if they have to incur the switching cost, because they correctly internalize the permanent increase in gasoline prices due to the gasoline tax. Relative to the dynamic model, the myopic model does account for the switching cost, but it does not account for the fact that consumers internalize the permanent increase in the gasoline price due to the tax. The myopic model predicts (for the counterfactual) lower probabilities of switching to public transit (0.067), than the dynamic model (0.107). Thus, it generates a larger underestimate of the substitution from car to public transit than the static model, 68.4 percent on average. These long-run effects, underestimated by static and myopic models, are relevant for public policy. In either case, static and myopic models could lead to incorrect policy decisions. Finally, note that this is a partial analysis framework. It does not account for capacity constraints, uses of the gasoline tax revenues, or deadweight losses. Enriching the model in such dimensions are avenues for future research.

6 Concluding Remarks

Despite the large volume of research on the transportation sector, "[M]uch of the most credible empirical work estimates short-run effects, whereas long-run effects are relevant for policy." (Anderson and Sallee 2016, p. 177). This paper complements such fundamental research, by developing a structural approach to estimate longer-run substitution patterns. In particular, I develop a structural model of urban travel consistent with the presence of hysteresis and switching costs. I estimate the model using a panel dataset with market level data for Chicago, constructed using public data obtained from government agencies. The estimated model is used to compute long-run price demand elasticities, to compare the estimated elasticities among alternative models, and as inputs in a counterfactual analysis, where I simulate policies that reduce switching costs and a gasoline tax.

Three main results are discussed. First, long-run price elasticities, with respect to the gasoline cost of driving, are more elastic than the short-run elasticities. Second, static own- (for automobile) and cross- (for public transit) price elasticities underestimate the dynamic ones. Myopic estimates also underestimate the elasticities. The bias is larger than for the static model. Finally, the counterfactual analysis shows that static and myopic models mismeasure the long-run substitution responses to a gasoline tax. A number of robustness checks are performed, including different specifications of the model, expectations, micro moments, and estimation procedure. The implications are robust to these alternative specifications. Overall, the results indicate that, when hysteresis and switching costs are present, elasticities from static and myopic models might be biased, and could lead to incorrect policy decisions.

References

Agthe, D. E., and R. B. Billings (1978): "The impact of gasoline prices on urban bus ridership," *The Annals of Regional Science*, 12(1), 90–96. (Cited on p. 5.)

Aguirregabiria, V., and P. Mira (2010): "Dynamic discrete choice structural models: A survey," *Journal of Econometrics*, 156(1), 38–67. (Cited on p. 7.)

Aguirregabiria, V., and A. Nevo (2013): "Recent developments in empirical IO: Dynamic demand and dynamic games," *Advances in Economics and Econometrics*, 3, 53–122. (Cited on p. 7.)

Alfaro, L., A. Cuñat, H. Fadinger, and Y. Liu (2018): "The real exchange rate, innovation and productivity: regional heterogeneity, asymmetries and hysteresis," Discussion paper, National Bureau of Economic Research. (Cited on p. 7.)

Alquist, R., L. Kilian, and R. J. Vigfusson (2013): "Forecasting the price of oil," in *Handbook of economic forecasting*, vol. 2, pp. 427–507. Elsevier. (Cited on p. 15.)

American Public Transportation Association (2008): "Rising Fuel Costs: Impacts on Transit Ridership and Agency Operations Survey Results," *American Public Transportation Association. (Accessed on April 25, 2018.)*. (Cited on p. 1.)

Anderson, S. T., R. Kellogg, and J. M. Sallee (2013): "What do consumers believe about future gasoline prices?," *Journal of Environmental Economics and Management*, 66(3), 383–403. (Cited on pp. 3 and 15.)

Anderson, S. T., R. Kellogg, J. M. Sallee, and R. T. Curtin (2011): "Forecasting gasoline prices using consumer surveys," *American Economic Review: Papers & Proceedings*, 101(3), 110–14. (Cited on pp. 3 and 15.)

Anderson, S. T., and J. M. Sallee (2016): "Designing Policies to Make Cars Greener: A Review of the Literature," *Annual Review of Resource Economics*, 8, 157–180. (Cited on pp. 6 and 28.)

Archibald, R., and R. Gillingham (1980): "An analysis of the short-run consumer demand for gasoline using household survey data," *The Review of Economics and Statistics*, pp. 622–628. (Cited on p. 5.)

——— (1981): "A decomposition of the price and income elasticities of the consumer demand for gasoline," *Southern Economic Journal*, pp. 1021–1031. (Cited on p. 5.)

Asker, J., A. Collard-Wexler, and J. De Loecker (2017): "Market Power and Production (Mis) Allocation. A Study of the World Oil Market," Discussion paper, mimeo. (Cited on p. 6.)

Aw, B. Y., M. J. Roberts, and D. Y. Xu (2011): "R&D investment, exporting, and productivity dynamics," *American Economic Review*, 101(4), 1312–44. (Cited on pp. 7 and 8.)

Baldwin, R. (1989): "Sunk-cost hysteresis," . (Cited on p. 7.)

Baldwin, R., and P. Krugman (1989): "Persistent trade effect of large exchange rate changes," *Quarterly Journal of Economics*, 104(4). (Cited on p. 7.)

Banister, D. (1978): "The influence of habit formation on modal choice—a heuristic model," *Transportation*, 7(1), 5–33. (Cited on pp. 2 and 6.)

Basso, L. J., and T. H. Oum (2007): "Automobile fuel demand: a critical assessment of empirical methodologies," *Transport Reviews*, 27(4), 449–484. (Cited on p. 5.)

Ben-Akiva, M., and T. Morikawa (1990): "Estimation of switching models from revealed preferences and stated intentions," *Transportation Research Part A: General*, 24(6), 485–495. (Cited on pp. 2 and 6.)

Ben-Akiva, M. E., and S. R. Lerman (1985): *Discrete choice analysis: theory and application to travel demand*, vol. 9. MIT press. (Cited on pp. 4 and 13.)

Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile prices in market equilibrium," *Econometrica*, pp. 841–890. (Cited on pp. 1, 4, 8, 18, and A-16.)

——— (2004): "Differentiated products demand systems from a combination of micro and macro data: The new car market," *Journal of political Economy*, 112(1), 68–105. (Cited on pp. 3 and 23.)

Berry, S. T. (1994): "Estimating discrete-choice models of product differentiation," *The RAND Journal of Economics*, pp. 242–262. (Cited on pp. 18 and A-16.)

Berry, S. T., and P. A. Haile (2014): "Identification in differentiated products markets using market level data," *Econometrica*, 82(5), 1749–1797. (Cited on p. 19.)

Blanchard, O., and L. F. Katz (1997): "What we know and do not know about the natural rate of unemployment," *Journal of Economic Perspectives*, 11(1), 51–72. (Cited on p. 7.)

Blanchard, O. J., and L. H. Summers (1986): "Hysteresis and the European unemployment problem," *NBER macroeconomics annual*, 1, 15–78. (Cited on p. 7.)

Blase, J. (1979): "Hysteresis and catastrophe theory: A demonstration of habit and thresholds in travel behaviour," in *Proc. 7th PTRC Summer Annual Meeting*. (Cited on p. 6.)

——— (1980): "Hysteresis in travel demand," *Transportation Planning and Technology*, 6(2), 109–115. (Cited on pp. 2 and 6.)

Boizot, C., J.-M. Robin, and M. Visser (2001): "The demand for food products: an analysis of interpurchase times and purchased quantities," *The Economic Journal*, 111(470), 391–419. (Cited on p. 7.)

Bricka, S. (2007): "Chicago Regional Household Travel Inventory Final Report," *Prepared for Chicago Metropolitan Agency for Planning*. (Cited on p. 13.)

Brons, M., P. Nijkamp, E. Pels, and P. Rietveld (2008): "A meta-analysis of the price elasticity of gasoline demand. A SUR approach," *Energy Economics*, 30(5), 2105–2122. (Cited on p. 6.)

Broyden, C. G. (1965): "A class of methods for solving nonlinear simultaneous equations," *Mathematics of computation*, 19(92), 577–593. (Cited on p. A-16.)

Burns, L., and T. Golob (1975): "An Investigation of the Role of Accessibility in Basic Transportation Choice Behavior," *Research Publication GMR-1900, Research Laboratories. General Motors Corp., Warren, Michigan*. (Cited on p. 4.)

Busse, M. R., C. R. Knittel, and F. Zettelmeyer (2013): "Are consumers myopic? Evidence from new and used car purchases," *American Economic Review*, 103(1), 220–56. (Cited on pp. 3, 6, and 15.)

Button, K. (2010): *Transport economics*. Edward Elgar Publishing. (Cited on pp. 2, 4, 13, and 14.)

Cambridge Systematics Inc. (2007): "Pace Market Research Report," Discussion paper, Available at `http://www.rtams.org/reportLibrary/564.pdf`. (Accessed on May 10, 2018.). (Cited on pp. 9, 24, 34, and A-6.)

Cantillo, V., J. d. D. Ortúzar, and H. C. Williams (2007): "Modeling discrete choices in the presence of inertia and serial correlation," *Transportation Science*, 41(2), 195–205. (Cited on pp. 2, 6, and 7.)

Chatterjee, K. (2011): "Modelling the dynamics of bus use in a changing travel environment using panel data," *Transportation*, 38(3), 487–509. (Cited on p. 6.)

Chen, C., D. Varley, and J. Chen (2011): "What affects transit ridership? A dynamic analysis involving multiple factors, lags and asymmetric behaviour," *Urban Studies*, 48(9), 1893–1908. (Cited on pp. 2 and 6.)

Chen, N. (2007): "Modelling Demand for Rail Transport with Dynamic Econometric Approaches," *International Review of Business Research Papers*, 3, 85–96. (Cited on p. 5.)

Cherchi, E., I. Meloni, and J. d. D. Ortúzar (2013): "The latent effect of inertia in the choice of mode," in *13th International Conference on Travel Behavior Research, Toronto, Canada*. (Cited on pp. 2 and 6.)

Clarke, M., M. Dix, and P. Goodwin (1982): "Some issues of dynamics in forecasting travel behaviour—a discussion paper," *Transportation*, 11(2), 153–172. (Cited on p. 6.)

Clerides, S. K., S. Lach, and J. R. Tybout (1998): "Is learning by exporting important? Micro-dynamic evidence from Colombia, Mexico, and Morocco," *The quarterly journal of economics*, 113(3), 903–947. (Cited on pp. 7 and 8.)

Costa, F., and F. Gerard (2018): "Hysteresis and the Welfare Effect of Corrective Policies: Theory and Evidence from an Energy-Saving Program," Discussion paper, National Bureau of Economic Research. (Cited on p. 7.)

Currie, G., and J. Phung (2006): "Exploring the Impacts of Fuel Price Increases on Public Transport Use in Melbourne," *29th Australasian Transport Research Forum. Gold Coast*. (Cited on p. 5.)

——— (2007): "Transit ridership, auto gas prices, and world events: new drivers of change?," *Transportation Research Record: Journal of the Transportation Research Board*, 1(1992), 3–10. (Cited on p. 5.)

——— (2008): "Understanding links between transit ridership and auto gas prices: US and Australian evidence," in *Transportation Research Board 87th Annual MeetingTransportation Research Board*. (Cited on p. 5.)

Daganzo, C. F., and Y. Sheffi (1982): "Multinomial probit with time-series data: unifying state dependence and serial correlation models," *Environment and Planning A*, 14(10), 1377–1388. (Cited on p. 6.)

Dahl, C., and T. Sterner (1991): "Analysing gasoline demand elasticities: a survey," *Energy economics*, 13(3), 203–210. (Cited on p. 5.)

Dahl, C. A. (1986): "Gasoline demand survey," *The Energy Journal*, 7(1), 67–82. (Cited on p. 5.)

Dargay, J. M. (1993): "Demand elasticities: a comment," *Journal of Transport Economics and Policy*, 27(1), 87–90. (Cited on pp. 2 and 6.)

Das, S., M. J. Roberts, and J. R. Tybout (2007): "Market entry costs, producer heterogeneity, and export dynamics," *Econometrica*, 75(3), 837–873. (Cited on pp. 7 and 8.)

Dixit, A. (1989): "Entry and exit decisions under uncertainty," *Journal of political Economy*, 97(3), 620–638. (Cited on pp. 7, 24, and 34.)

Dixit, A. K., and R. S. Pindyck (1994): *Investment under uncertainty*. Princeton university press. (Cited on p. 7.)

Doi, M., and W. B. Allen (1986): "A time series analysis of monthly ridership for an urban rail rapid transit line," *Transportation*, 13(3), 257–269. (Cited on p. 5.)

Donna, J. D., and J. A. Espin-Sanchez (2018): "The Illiquidity of Water Markets: Efficient Institutions for Water Allocation in Southeastern Spain," Discussion paper, The Ohio State University. (Cited on p. 8.)

Dubé, J.-P., G. J. Hitsch, and P. E. Rossi (2009): "Do switching costs make markets less competitive?," *Journal of Marketing research*, 46(4), 435–445. (Cited on p. 7.)

——— (2010): "State dependence and alternative explanations for consumer inertia," *The RAND Journal of Economics*, 41(3), 417–445. (Cited on pp. 7, 19, and 21.)

Dunne, T., M. J. Roberts, and L. Samuelson (1988): "Patterns of firm entry and exit in US manufacturing industries," *The RAND journal of Economics*, pp. 495–515. (Cited on p. 7.)

——— (1989): "The growth and failure of US manufacturing plants," *The Quarterly Journal of Economics*, 104(4), 671–698. (Cited on p. 7.)

Eckstein, Z., and K. I. Wolpin (1989): "The specification and estimation of dynamic stochastic discrete choice models: A survey," *The Journal of Human Resources*, 24(4), 562–598. (Cited on p. 6.)

Espey, M. (1998): "Gasoline demand revisited: an international meta-analysis of elasticities," *Energy Economics*, 20(3), 273–295. (Cited on p. 5.)

Farrell, J., and P. Klemperer (2007): "Coordination and lock-in: Competition with switching costs and network effects," *Handbook of industrial organization*, 3, 1967–2072. (Cited on p. 7.)

Fearnley, N., and J. Bekken (2005): "Short-and long-run demand effects in transport: A literature survey," *Institute of Transport Economics, Norway*. (Cited on p. 6.)

Frondel, M., and C. Vance (2013): "Re-identifying the rebound: what about asymmetry?," *The Energy Journal*, pp. 43–54. (Cited on p. 6.)

Gärling, T., and K. Axhausen (2003): "Introduction: Habitual travel choice," *Transportation*, 30(1), 1–11. (Cited on p. 6.)

Goeree, M. S. (2008): "Limited information and advertising in the US personal computer industry," *Econometrica*, 76(5), 1017–1074. (Cited on pp. 3, 8, and 23.)

Golob, T. F., R. Kitamura, and L. Long (2013): *Panels for transportation planning: methods and applications*. Springer Science & Business Media. (Cited on p. 6.)

Golob, T. F., L. van Wissen, and H. Meurs (1986): "A dynamic analysis of travel demand," *Transportation Research Part A: General*, 20(5), 401–414. (Cited on p. 2.)

Goodwin, P., R. Kitamura, and H. Meurs (1990): "Some Principles of Dynamic Analysis of Travel Behaviour," *in* Developments in dynamic and activity-based approaches to travel analysis, *Edited by Peter Jones*. (Cited on p. 21.)

Goodwin, P. B. (1977): "Habit and hysteresis in mode choice," *Urban studies*, 14(1), 95–98. (Cited on pp. 2 and 6.)

——— (1992): "A review of new demand elasticities with special reference to short and long run effects of price changes," *Journal of transport economics and policy*, pp. 155–169. (Cited on p. 5.)

Goulder, L. H., M. A. C. Hafstead, G. Kim, and X. Long (2018): "Impacts of a Carbon Tax across US Household Income Groups: What Are the Equity-Efficiency Trade-Offs?," Working Paper 25181, NBER. (Cited on p. 27.)

Gowrisankaran, G., and M. Rysman (2012): "Dynamics of consumer demand for new durable goods," *Journal of political Economy*, 120(6), 1173–1219. (Cited on pp. 1, 7, 8, 17, and 18.)

Graham, D. J., and S. Glaister (2002): "The demand for automobile fuel: a survey of elasticities," *Journal of Transport Economics and Policy*, pp. 1–25. (Cited on p. 5.)

Grigolon, L., M. Reynaert, and F. Verboven (2017): "Consumer valuation of fuel costs and tax policy: Evidence from the European car market," *American Economic Journal: Economic Policy*. (Cited on p. 4.)

Haire, A. R., and R. B. Machemehl (1992): "Impact of Rising Fuel Prices on U.S. Transit Ridership," *Transportation Research Record*, pp. 11–19. (Cited on p. 5.)

Hamilton, J. D. (2009): "Understanding Crude Oil Prices," *The Energy Journal*, 30(2). (Cited on p. 15.)

——— (2011): "Historical oil shocks," Discussion paper, National Bureau of Economic Research. Available at: `https://www.nber.org/papers/w16790` (Accessed on September 22, 2012.). (Cited on pp. 12 and 33.)

Handel, B. R. (2013): "Adverse selection and inertia in health insurance markets: When nudging hurts," *American Economic Review*, 103(7), 2643–82. (Cited on p. 7.)

Hansen, L. P. (1982): "Large sample properties of generalized method of moments estimators," *Econometrica*, pp. 1029–1054. (Cited on p. A-19.)

Hausman, J. A., AND W. K. Newey (1995): "Nonparametric estimation of exact consumers surplus and deadweight loss," *Econometrica: Journal of the Econometric Society*, pp. 1445–1476. (Cited on p. 5.)

Heckman, J. J. (1981): *Statistical Models for Discrete Panel Data.* In C.F. Manski and D. McFadden, eds., Structural Analysis of Discrete Data with Econometric Applications. Cambridge, Mass.: MIT Press. (Cited on pp. 6 and 19.)

Hendel, I., AND A. Nevo (2006): "Measuring the implications of sales and consumer inventory behavior," *Econometrica*, 74(6), 1637–1673. (Cited on pp. 1, 7, 8, 25, and A-18.)

Hirobata, Y., AND S. Kawakami (1990): "Modeling disaggregate behavioral modal switching models based on intention data," *Transportation Research Part B: Methodological*, 24(1), 15–25. (Cited on p. 6.)

Hogema, J. (1996): "Effects of Rain on Daily Traffic Volume and on Driving Behaviour," *Report TM-96-B019*. (Cited on p. 6.)

Honka, E. (2014): "Quantifying search and switching costs in the US auto insurance industry," *The RAND Journal of Economics*, 45(4), 847–884. (Cited on pp. 7 and 8.)

Horowitz, J. (1980): "A utility maximizing model of the demand for multi-destination non-work travel," *Transportation Research Part B: Methodological*, 14(4), 369–386. (Cited on p. 5.)

Horowitz, J., AND R. Emslie (1978): "Comparison of measured and forecast traffic volumes on urban interstate highways," *Transportation Research*, 12(1), 29–32. (Cited on p. 6.)

Hortaçsu, A., S. A. Madanizadeh, AND S. L. Puller (2017): "Power to choose? An analysis of consumer inertia in the residential electricity market," *American Economic Journal: Economic Policy*, 9(4), 192–226. (Cited on pp. 7 and 8.)

Hughes, J. E., C. R. Knittel, AND D. Sperling (2008): "Evidence of a Shift in the Short-Run Price Elasticity of Gasoline Demand," *The Energy Journal*, 29(1). (Cited on p. 5.)

Illanes, G. (2016): "Switching Costs in Pension Plan Choice," Discussion paper, Working Paper. (Cited on p. 7.)

Illinois Department of Transportation (2004): "Illinois Traffic Monitoring Program," *Office of Planning and Programming.* (Cited on p. A-4.)

Innocenti, A., P. Lattarulo, AND M. G. Pazienza (2013): "Car stickiness: Heuristics and biases in travel choice," *Transport Policy*, 25, 158–168. (Cited on p. 6.)

Johnson, L., AND D. Hensher (1982): "Application of multinomial probit to a two-period panel data set," *Transportation Research Part A: General*, 16(5-6), 457–464. (Cited on p. 6.)

Kellogg, R. (2018): "Gasoline price uncertainty and the design of fuel economy standards," *Journal of Public Economics*, 160, 14–32. (Cited on p. 3.)

Kilian, L. (2010): "Explaining fluctuations in gasoline prices: a joint model of the global crude oil market and the US retail gasoline market," *The Energy Journal*, pp. 87–112. (Cited on pp. 12 and 33.)

Kim, J. (2006): "Consumers' dynamic switching decisions in the cellular service industry," *NET Institute Working Paper No. 06-24. Available at: `https://ssrn.com/abstract=941222orhttp://dx.doi.org/10.2139/ssrn.941222` (Accessed on November 16, 2017.).* (Cited on p. 7.)

Klemperer, P. (1987a): "The competitiveness of markets with switching costs," *The RAND Journal of Economics*, pp. 138–150. (Cited on pp. 7 and 14.)

——— (1987b): "Markets with consumer switching costs," *The Quarterly Journal of Economics*, 102(2), 375–394. (Cited on pp. 7 and 14.)

——— (1995): "Competition when consumers have switching costs: An overview with applications to industrial organization, macroeconomics, and international trade," *The Review of Economic Studies*, 62(4), 515–539. (Cited on p. 7.)

Knittel, C., AND R. Sandler (2018): "The Welfare Impact of Second-Best Uniform-Pigouvian Taxation: Evidence from Transportation," *American Economic Journal: Economic Policy*, 10(4), 211–42. (Cited on p. 27.)

Kohn, H. M., et al. (1999): *Factors affecting urban transit ridership.* Statistics Canada. (Cited on p. 5.)

Krugman, P. R. (1989): *Exchange Rate Instability.* Cambridge, MA: MIT Press. (Cited on p. 7.)

Lagarias, J. C., J. A. Reeds, M. H. Wright, AND P. E. Wright (1998): "Convergence properties of the Nelder–Mead simplex method in low dimensions," *SIAM Journal on optimization*, 9(1), 112–147. (Cited on p. A-16.)

Lerman, S., AND M. Ben-Akiva (1976): "A Dissagregatge Behavioral Model of Auto Ownership," *Transportation Res. Rec.* (Cited on p. 4.)

Levin, L., M. S. Lewis, AND F. A. Wolak (2017): "High frequency evidence on the demand for gasoline," *American Economic Journal: Economic Policy*, 9(3), 314–47. (Cited on p. 5.)

Li, S., J. Linn, AND E. Muehlegger (2014): "Gasoline taxes and consumer behavior," *American Economic Journal: Economic Policy*, 6(4), 302–42. (Cited on p. 27.)

Lingras, P. (2000): "Traffic Colume Time-Series Analysis According to the Type of Road Use," *Computer-Aided Civil and Infrastructure Engineering*, pp. 365–373. (Cited on p. 6.)

Litman, T. (2004): "Transit price elasticities and cross-elasticities," *Journal of Public Transportation*, 7(2), 3. (Cited on p. 5.)

——— (2017): "Transit price elasticities and cross-elasticities," *Victoria Transport Policy Institute.* (Cited on p. 5.)

Long, L., J. Lin, AND K. Proussaloglou (2010): "Investigating contextual variability in mode choice in Chicago using a hierarchical mixed logit model," *Urban Studies*, 47(11), 2445–2459. (Cited on pp. 9 and A-6.)

Luco, F. (2013): "Switching costs and competition in retirement investment," *The Center for the Study of Industrial Organization-Northwestern University.* (Cited on p. 7.)

Maley, D., AND R. Weinberger (2009): "Rising gas price and transit ridership: case study of Philadelphia, Pennsylvania," *Transportation Research Record: Journal of the Transportation Research Board*, 1(2139), 183–188. (Cited on pp. 2, 5, and 6.)

Manski, C. F., AND L. Sherman (1980): "An empirical analysis of household choice among motor vehicles," *Transportation Research Part A: General*, 14(5-6), 349–366. (Cited on p. 5.)

Mattson, J. W. (2008): *Effects of rising gas prices on bus ridership for small urban and rural transit systems.* Upper Great Plains Transportation Institute, North Dakota State University. (Cited on p. 5.)

Miravete, E. J., M. J. Moral, AND J. Thurk (2018): "Fuel Taxation, Emissions Policy, and Competitive Advantage in the Diffusion of European Diesel Automobiles," *Rand Journal of Economics,* forthcoming. (Cited on p. 4.)

MORIKAWA, T. (1994): “Correcting state dependence and serial correlation in the RP/SP combined estimation method,” *Transportation*, 21(2), 153–165. (Cited on p. 6.)

NEVO, A. (2001): “Measuring market power in the ready-to-eat cereal industry,” *Econometrica*, 69(2), 307–342. (Cited on p. 8.)

NEWEY, W. K., AND D. MCFADDEN (1994): “Large sample estimation and hypothesis testing,” *Handbook of econometrics*, 4, 2111–2245. (Cited on p. A-19.)

OHIO SUPERCOMPUTER CENTER (1987): “Ohio Supercomputer Center,” http://osc.edu/ark:/19495/f5s1ph73. (Cited on pp. and A-18.)

OUM, T. H., W. G. WATERS, AND J.-S. YONG (1992): “Concepts of price elasticities of transport demand and recent empirical estimates: an interpretative survey,” *Journal of Transport Economics and Policy*, pp. 139–154. (Cited on p. 5.)

OWEN, B., T. JANE, AND C. KOPP (2007): “Determined drivers, cautious individuals, and the great middle: Pace’s approach to market segmentation and service planning,” *Transport Chicago Conference, Chicago, IL. Available at: http://www.transportchicago.org/conference-archives.html (Accessed on July 22, 2011.)*. (Cited on pp. 9 and A-6.)

PENDYALA, R., A. PARASHAR, AND G. MUTHYALAGARI (2001): “Measuring day-to day variability in travel characteristics using GPS data,” in *79th annual meeting of the Transportation Research Board*. (Cited on p. 6.)

PERRONE, H. (2017): “Demand for nondurable goods: a shortcut to estimating long-run price elasticities,” *The RAND Journal of Economics*, 48(3), 856–873. (Cited on p. 8.)

PESENDORFER, M. (2002): “Retail sales: A study of pricing behavior in supermarkets,” *The Journal of Business*, 75(1), 33–66. (Cited on p. 7.)

PETERS, B., M. J. ROBERTS, AND V. A. VUONG (2018): “Firm R&D Investment and Export Market Exposure,” Working Paper 25228, National Bureau of Economic Research. (Cited on p. 7.)

PETRIN, A. (2002): “Quantifying the benefits of new products: The case of the minivan,” *Journal of political Economy*, 110(4), 705–729. (Cited on pp. 3 and 23.)

ROBERTS, M. J., AND J. R. TYBOUT (1997): “The decision to export in Colombia: an empirical model of entry with sunk costs,” *The American Economic Review*, pp. 545–564. (Cited on pp. 7 and 8.)

ROSE, G. (1986): “Transit passenger response: short and long term elasticities using time series analysis,” *Transportation*, 13(2), 131–141. (Cited on p. 5.)

RUST, J. (1987): “Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher,” *Econometrica*, pp. 999–1033. (Cited on pp. 1 and 8.)

——— (1994): “Structural estimation of Markov decision processes,” *Handbook of econometrics*, 4, 3081–3143. (Cited on p. 6.)

——— (1996): “Numerical dynamic programming in economics,” *Handbook of computational economics*, 1, 619–729. (Cited on p. A-17.)

SALLEE, J. M., S. E. WEST, AND W. FAN (2016): “Do consumers recognize the value of fuel economy? Evidence from used car prices and gasoline price fluctuations,” *Journal of Public Economics*, 135, 61–73. (Cited on p. 6.)

SCHMALENSEE, R., AND T. M. STOKER (1999): “Household gasoline demand in the United States,” *Econometrica*, 67(3), 645–662. (Cited on p. 5.)

SHCHERBAKOV, O. (2016): “Measuring consumer switching costs in the television industry,” *The RAND Journal of Economics*, 47(2), 366–393. (Cited on pp. 1, 7, 8, 14, 17, 18, 20, and A-17.)

SHUM, M. (2004): “Does advertising overcome brand loyalty? Evidence from the breakfast-cereals market,” *Journal of Economics & Management Strategy*, 13(2), 241–272. (Cited on p. 7.)

SMALL, K. (2013): *Urban transportation economics*, vol. 4. Taylor & Francis. (Cited on p. 4.)

SMALL, K., E. VERHOEF, AND R. LINDSEY (2007): *The economics of urban transportation*. Routledge. (Cited on pp. 4 and 13.)

SMALL, K. A., AND K. VAN DENDER (2007): “Fuel efficiency and motor vehicle travel: the declining rebound effect,” *The Energy Journal*, pp. 25–51. (Cited on p. 6.)

SORRELL, S., J. DIMITROPOULOS, AND M. SOMMERVILLE (2009): “Empirical estimates of the direct rebound effect: A review,” *Energy policy*, 37(4), 1356–1371. (Cited on pp. 2 and 6.)

SRINIVASAN, K. K., AND P. BHARGAVI (2007): “Longer-term changes in mode choice decisions in Chennai: a comparison between cross-sectional and dynamic models,” *Transportation*, 34(3), 355–374. (Cited on p. 6.)

STORCHMANN, K. (2001): “The Impact of Fuel Taxes on Public Transport Ð An Empirical Assessment for Germany,” *Transport Policy*, 8, 19–28. (Cited on p. 5.)

STOVER, V., AND C. C. BAE (2011): “The impact of gasoline prices on transit ridership in Washington State.,” *Transportation Research Board of the National Academies*. (Cited on p. 5.)

SUDHIR, K., AND N. YANG (2014): “Exploiting the choice-consumption mismatch: A new approach to disentangle state dependence and heterogeneity,” *Cowles Foundation Discussion Paper No. 1941*. (Cited on pp. 7 and 19.)

TAYLOR, B., D. MILLER, H. ISEKI, AND C. FINK (2008): “Nature and/or nurture? Analyzing the determinants of transit ridership across US urbanized areas,” *Transportation Research Part A: Policy and Practice*, 43, 60–77. (Cited on p. 5.)

TAYLOR, B. D., AND C. N. FINK (2003): “The factors influencing transit ridership: A review and analysis of the ridership literature,” *UCLA Department of Urban Planning Working Paper, Institute of Transportation Studies*. (Cited on p. 6.)

TRAIN, K. (1980): “A structured logit model of auto ownership and mode choice,” *The Review of Economic Studies*, 47(2), 357–370. (Cited on p. 4.)

U.S. DEPARTMENT OF TRANSPORTATION (2011): “Summary of travel trends: 2009 national household travel survey,” *Technical Report. Available at: https://nhts.ornl.gov/publications.shtml (Accessed on December 28, 2016.)*. (Cited on p. 1.)

WEST, J., M. HOEKSTRA, J. MEER, AND S. L. PULLER (2017): “Vehicle miles (not) traveled: Fuel economy requirements, vehicle characteristics, and household driving,” *Journal of public Economics*, 145, 65–81. (Cited on p. 6.)

WHEATON, W. C. (1982): “The long-run structure of transportation and gasoline demand,” *The Bell Journal of Economics*, pp. 439–454. (Cited on p. 5.)

WILLIAMS, H., AND J. D. D. ORTÚZAR (1982): “Behavioural theories of dispersion and the mis-specification of travel demand models,” *Transportation Research Part B: Methodological*, 16(3), 167–219. (Cited on p. 6.)

YÁÑEZ, M., E. CHERCHI, J.-D. ORTÚZAR, AND B. HEYDECKER (2009): “Inertia and shock effects on mode choice panel data: implications of the Transantiago implementation,” *International Conference on Travel Behaviour Research*. (Cited on p. 6.)

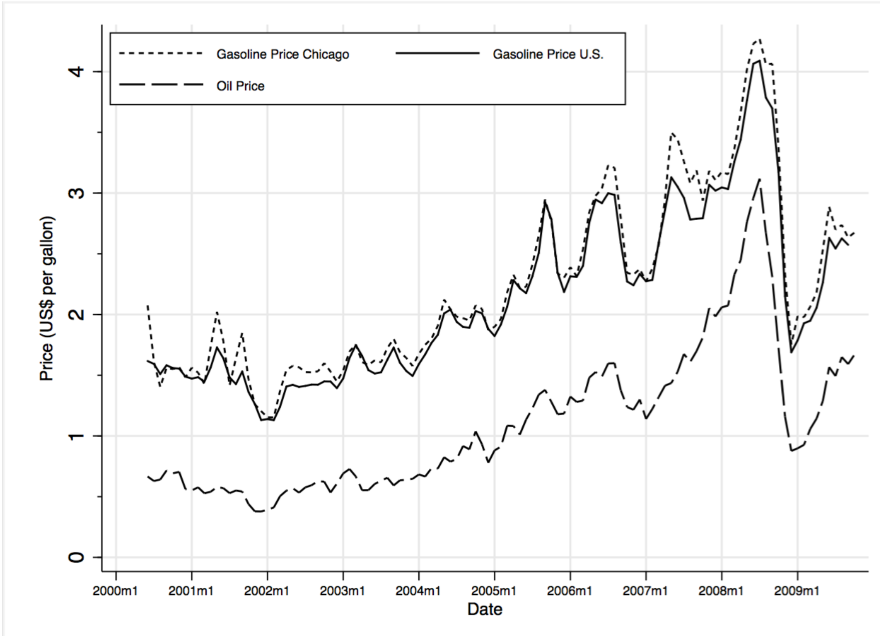
YANG, Y. (2010): “Identification of consumer switching behavior with market level data,” *Unpublished manuscript, UCLA. Available at: http://econyang.bol.ucla.edu/research/Id.Switching.pdf (Accessed on Nov.11 2017.)*. (Cited on p. 8.)

YANMAZ-TUZEL, O., AND K. OZBAY (2010): “Impacts of gasoline prices on New Jersey transit ridership,” *Transportation Research Record: Journal of the Transportation Research Board*, 1(2144), 52–61. (Cited on p. 5.)

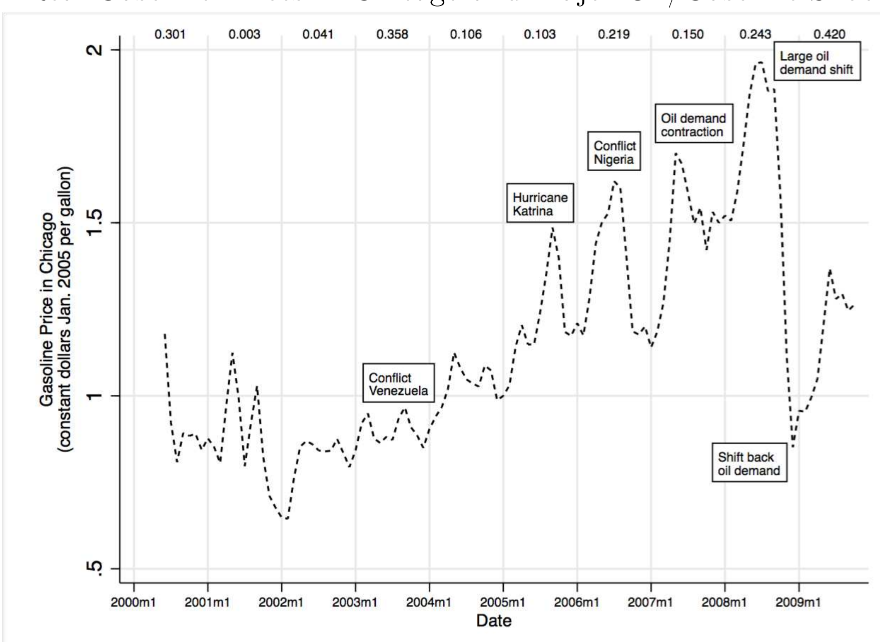
Figures and Tables

Figure 1: Gasoline Prices in U.S. and Oil Prices.

A. Gasoline and Oil Prices.



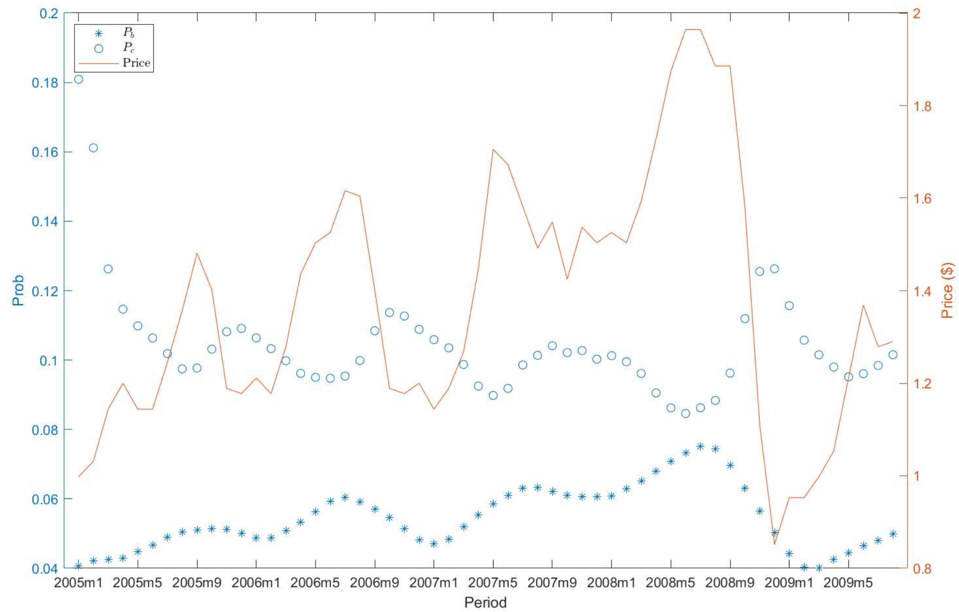
B. Real Gasoline Prices in Chicago and Major Oil/Gasoline Shocks.



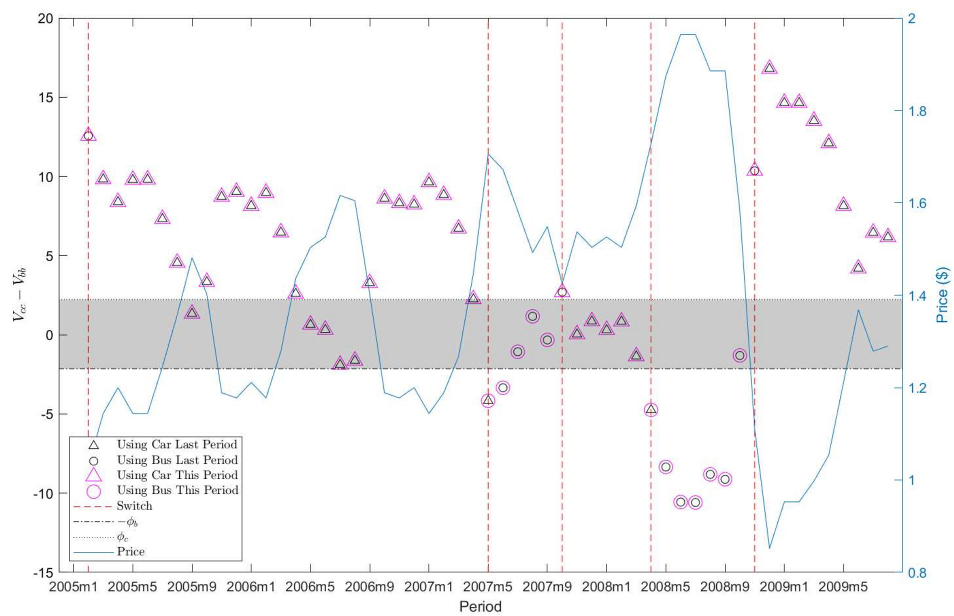
Notes: Figure 1A depicts the monthly evolution of the gasoline prices in Chicago (nominal gasoline price in Chicago in the variables definition appendix) and in the U.S. (nominal gasoline price in the U.S.), and Oil Prices (nominal oil price wtotusa) in dollars per gallon, for the period under analysis in subsection 2.3, June 2000 (the first month with no missing information for the gasoline price in Chicago) to Oct 2009 (the latest month with vehicle counts information provided by the IDOT at the moment of the data collection). Figure 1B depicts the monthly evolution of the gasoline price in Chicago in dollars of January 2005 per gallon (gasoline price), and major oil/gasoline shocks obtained from Kilian (2010) and Hamilton (2011), for the same period as the top figure. The top number in figure 1B shows the mean monthly CTA rail ridership elasticity by year. Each of these numbers are obtained by running a regression for the corresponding year, similar to the baseline specification in table A3, panel A, row 1, which has coefficient for $ln(p_t) = .174$. Similar results are obtained for the other OLS and IV specifications in panel A, and the specifications in panels B, C, and D in table A3. Results are available upon request. Variables definitions are in appendix D.

Figure 2: Estimated model.

A. Conditional Choice Probabilities and Gasoline Prices.



B. Switching Behavior of a Transit-Friendly Consumer.



Notes: Figure 2A shows the mean, across consumer types and markets, conditional choice probabilities (left vertical axis) given by the equation in (4), for public transit P_b and automobile P_c, holding fixed ξ_{mrt}, as a function of the observed gasoline cost of driving (right vertical axis), p_t, for the period 2005-2009. Figure 2B displays the switching behavior of a particular transit-friendly consumer type from the Pace survey (Cambridge Systematics Inc. 2007), holding fixed ξ_{mrt} and ϵ_{imrt}. On the left vertical axis, the graph shows using dots, the difference in conditional value function of choosing car today given previous period choice of car, minus the conditional value function of choosing public transit today given previous period choice of public transit: $V_{cc} - V_{bb} \equiv v(p_t, m_{irt-1} = c, m_{irt} = c) - v(p_t, m_{irt-1} = b, m_{irt} = b)$, where the conditional value functions, $v(p_t, m_{irt-1}, m_{irt})$, are given by the equation in (2). The vertical dotted lines indicate a switch in the inside mode of transportation by the consumer in such period. The graph also displays horizontal lines with the estimated switching cost parameters, ϕ_c and $-\phi_b$, and the gasoline cost of driving. The shaded area between the switching costs, ϕ_c and $-\phi_b$, defines the band of inaction, where the consumer does not switch, analogue to Dixit (1989, figure 1). See page 24 for a description of the switching behavior. See subsection 5.1 for details.

Table 1: CTA rail weekdays ridership and heterogeneity.

Interaction	Estimate	Nmbr. obs. where $\mathbf{1}\{\cdot\} = 1$
Panel A: Price trend in last three months.		
$ln(p_t)\times$		
$\mathbf{1}\{p_t > p_{t-1} > p_{t-2}\}$	.1872	4,532
	(.0274)	
$\mathbf{1}\{p_t < p_{t-1} < p_{t-2}\}$	.0565	3,836
	(.0311)	
$\mathbf{1}\{\text{else}^a\}$	.1896	7,357
	(.0221)	
Nmbr. obs.	15,725	15,725
R^2	0.9234	
Panel B: Absolute level of gasoline price.[b]		
$ln(p_t)\times$		
$\mathbf{1}\{P_t < 150\}$	.0028	1,570
	(.0151)	
$\mathbf{1}\{150 \leq P_t < 250\}$	.1259	9,396
	(.0404)	
$\mathbf{1}\{250 \leq P_t < 350\}$	.2048	3,792
	(.0268)	
$\mathbf{1}\{350 \leq P_t\}$	.2537	967
	(.0143)	
Nmbr. obs.	15,725	15,725
R^2	0.9238	

Notes: Dependent variable is the natural logarithm of CTA rail ridership for weekdays. Independent variable is the natural logarithm of the gasoline cost of driving in real terms, $ln(p_t)$, interacted with the relevant indicators in each panel. Regressions are OLS specifications, and include a constant, monthly seasonal fixed effects, and unit of observation fixed effects, where the unit is the CTA rail station. The regressions reported in this table are the interactions' versions of the baseline specification in table A3, panel A, row 1, which has coefficient for $ln(p_t) = .174$ with standard error .020. Similar results are obtained for the other OLS and IV specifications in panel A, and the specifications in panels B, C, and D in table A3 (ridership using other means of public transportation). Results are available upon request. Robust standard errors clustered at month level are reported in parentheses for all regressions. Variables definitions are in appendix D. In all regressions the period included is June 2000 (the first month with no missing information for the gasoline price in Chicago) to August 2009 (the latest month with ridership information at the moment of the data collection). The number of years included in the regressions is 10. The number of months included is 111. The number of stations included is 143. The notation $\mathbf{1}\{A\}$, denotes an indicator function that equals 1 if condition A is satisfied, and 0 otherwise.

[a]The condition "else" means that none of the previous conditions holds. That is, neither $\mathbf{1}\{p_t > p_{t-1} > p_{t-2}\}$ or $\mathbf{1}\{p_t < p_{t-1} < p_{t-2}\}$ hold.

[b]The absolute level of gasoline price is denoted with uppercase, P_t.

Table 2: Katrina Shock and Hysteresis.

Locations	Pre Shock	Shock	%Δ demand	Post shock	%Δ demand
Panel A: Urban areas with access to public transit.					
ATR urban 1	76,358	71,037	-7.49%	71,443	0.57%
“L” Kedzie	1,454	1,615	9.97%	1,562	-3.28%
ATR urban 2	55,432	51,282	-8.09%	52,438	2.25%
“L” Pulaski	981	1,101	10.90%	1,057	-4.00%
ATR urban 3	51,290	46,147	-11.14%	52,185	13.08%
“L” Oak Park	1,497	1,671	10.41%	1,611	-3.59%
ATR urban 4	115,189	97,817	-17.76%	98,402	0.60%
“L” 47th-Dan Ryan	2,610	3,086	15.42%	3,021	-2.11%
ATR urban 5	89,421	84,356	-6.00%	85,405	1.24%
“L” 63rd-Dan Ryan	2,852	3,265	12.65%	3,191	-2.27%
ATR urban 6	92,600	87,123	-6.29%	88,544	1.63%
“L” Kimball	3,376	3,893	13.28%	3,766	-3.26%
ATR urban 7	79,116	73,308	-7.92%	73,649	0.47%
“L” Cicero	1,454	1,615	9.97%	1,562	-3.28%
ATR urban 8	84,914	81,070	-4.74%	82,557	1.83%
“L” Francisco	1,006	1,128	10.82%	1,098	-2.66%
ATR urban 9	97,282	94,748	-2.67%	95,412	0.70%
“L” Western-Congress	948	1,108	14.44%	1,037	-6.41%
ATR urban 10	102,489	101,031	-1.44%	101,923	0.88%
“L” Western-Douglas	684	773	11.51%	743	-3.88%
ATR urban 11	104,745	101,213	-3.49%	102,801	1.57%
“L” Western-Midway	3,256	3,657	10.97%	3,448	-5.72%
ATR urban 12	111,576	108,061	-3.25%	109,977	1.77%
“L” Western-Ravenswood	3,256	3,657	10.97%	3,448	-5.72%
ATR urban 13	116,210	112,595	-3.21%	114,322	1.53%
“L” Western/Milwaukee	3,702	4,158	10.97%	3,985	-4.16%
ATR urban 14	112,880	107,758	-4.75%	109,052	1.20%
“L”Logan Square	4,872	5,437	10.39%	5,179	-4.75%
ATR urban 15	93,579	90,998	-2.84%	92,392	1.53%
“L” Belmont	4,573	4,929	7.22%	4,874	-1.12%
Panel B: Rural areas without access to public transit.					
ATR rural 1	9,221	8,638	-6.75%	9,287	7.51%
ATR rural 2	32,450	28,794	-12.70%	32,301	12.18%
ATR rural 3	3,572	3,536	-1.02%	3,599	1.75%
ATR rural 4	1,812	1,703	-6.40%	1,803	5.87%
ATR rural 5	19,187	17,557	-9.28%	18,825	7.22%

Notes: The Katrina hurricane formed in August 23, 2005, and dissipated in August 31, 2005. Pre shock refers to the month prior to the hurricane Katrina, August 2005. Shock refers to September 2005. Post shock refers to October 2005. The gasoline prices (real gasoline price in Chicago, in dollars of January 2005) were: pre shock 1.4, shock 1.5, post shock 1.4. Numbers reported in each row correspond to raw ATRs vehicle counts, and raw CTA rail (“L” station name) ridership in each location. Rail locations correspond to the closest CTA rail location to the ATR in the prior row. In panel A, for urban ATRs, the 15 most popular ATR locations (and their closest rail locations) are included. In panel B, for rural ATRs, the 5 most popular ATR locations without access to public transit are included. ATRs numbers are masked. The term %Δ demand, denotes the percentage change in the number of vehicles (ridership) recored by the ATR (“L” station) during the shock and post shock, relative to the previous month. For example, for the first line in panel A, %Δ demand for ATR urban 1 during the shock is computed as $(71{,}037-76{,}358)/71{,}037 \times 100 = -7.49\%$, and post shock as $(71{,}443-71{,}037)/71{,}443 \times 100 = 0.57\%$. Similarly, for each line in the table.

Table 3: Model Estimates.

Panel A: Myopic Expectations and Static Model.

	Estimates	95% C.I.	Estimates	95% C.I.	Estimates	95% C.I.	Estimates	95% C.I.
	Myopic Expectations ME				Static Model without switching costs			
	with random coefficients		without random coefficients		with random coefficients		without random coefficients	
α_1	-0.9	(-1.127, -0.674)	-0.901	(-1.160, -0.641)	-0.629	(-0.773, -0.486)	-0.542	(-0.746, -0.338)
Σ	0.045	(0, 0.164)	–	–	0.772	(0.642, 0.901)	–	–
ϕ_b	4.14	(3.555, 4.725)	4.14	(3.631, 4.649)	–	–	–	–
ϕ_c	4.257	(3.423, 5.091)	4.257	(3.380, 5.134)	–	–	–	–
Value of GMM Objective	80.441		80.443		7.145		38.036	

Panel B: Dynamic Model with Rational Expectations RE-AR.

	Estimates	95% C.I.	Estimates	95% C.I.	Estimates	95% C.I.	Estimates	95% C.I.
	with switching costs				without switching costs			
	with random coefficients		without random coefficients		with random coefficients		without random coefficients	
α_1	-0.462	(-1.057, -0.134)	-0.522	(-1.280, -0.236)	-0.685	(-0.783, -0.587)	-0.587	(-0.677, -0.496)
Σ	0.294	(0, 0.657)	–	–	0.774	(0.628, 0.921)	–	–
ϕ_b	3.098	(2.643, 3.554)	2.306	(1.517, 3.095)	–	–	–	–
ϕ_c	3.353	(2.587, 4.119)	3.143	(2.323, 3.963)	–	–	–	–
Value of GMM Objective	26.268		42.858		6.377		37.533	

Panel C: Dynamic Model with Rational Expectations RE-AR and Additional Moments.

	Estimates	95% C.I.
	with switching costs	
	with random coefficients	
α_1	-0.526	(-0.657, -0.395)
Σ	0.308	(0.040, 0.575)
ϕ_b	2.142	(1.550, 2.734)
ϕ_c	2.224	(1.842, 2.605)
Value of GMM Objective	93.742	

Notes: Estimates of selected parameters from the structural model. See section 2 for details about the data used in the estimation. A description of the model is in section 3. Details about the estimation procedure are in section 4. See subsection 5.1 for details about the specifications of the models in the different panels. The term 95% C.I., denotes 95 percent confidence intervals and are provided in parenthesis. Panel C: The additional moments (no-switching moments) included in this panel correspond to the share of the consumers who did not switch from car and public transit during the $\tilde{t} = 3$ periods prior to t, in route r, denoted respectively by $s_{brt|b\tilde{t}}(p_t, \delta_{mrt}; \alpha_{0mr}, \Sigma, \phi)$ and $s_{crt|c\tilde{t}}(p_t, \delta_{mrt}; \alpha_{0mr}, \Sigma, \phi)$. These additional moments are computed using the conditional value function and choice probabilities in (C.3), implemented in the inner subroutine in step 3 as described in subsection 4.1 in the paper. See appendix C for details.

Table 4: Average ratios of elasticities computed from different models relative to the long-run elasticity computed from the dynamic model.

	Dynamic model short-run	Static model	Myopic model
Car			
Mean	0.870	0.438	0.673
Route:			
1	0.885	0.433	0.664
2	0.867	0.406	0.655
3	0.872	0.453	0.689
4	0.863	0.495	0.662
5	0.864	0.413	0.687
Public Transit			
Mean	0.844	0.192	0.735
Route:			
1	0.854	0.203	0.652
2	0.832	0.156	0.673
3	0.855	0.168	0.728
4	0.856	0.265	0.927
5	0.824	0.160	0.737
Outside option			
Mean	0.903	0.526	0.613
Route:			
1	0.890	0.456	0.655
2	0.900	0.511	0.630
3	0.932	0.517	0.619
4	0.901	0.610	0.545
5	0.901	0.513	0.641

Notes: The table presents the average ratio of elasticities computed from the model indicated in each column, divided by the long-run elasticity computed from the dynamic model. The elasticities for all models are the percent change in market share of mode m with a 1 percent change in the gasoline cost of driving. The dynamic model corresponds to the dynamic model with both switching cots and additional micro moments, from panel C in table 3. The static model corresponds to the static model with neither switching costs nor random coefficients, from the last columns in panel A in table 3. The myopic model corresponds to the myopic model with both switching costs and random coefficients, from the first columns in panel A in table 3. See subsection 5.1 for details about the specifications of the models. See subsection 5.3 for details about the computation of the elasticities. For each mode of transportation, the mean elasticity is computed by averaging the elasticity computed across periods and routes. The routes correspond to the five major expressways, labeled 1 to 5 in the table, in the Chicago area used in the market definition as described in subsection 2.2.

Table 5: Counterfactual Analysis.

Change in switching cost public transit (ϕ_b)	Compensating variation (1)	Switching costs from public transit (2)	Switching costs from car (3)	Probability to car (4)	Probability to public transit (5)
Panel A: Change in Switching Cost to Public Transit, ϕ_b.					
Baseline, without change in ϕ_b	–	0.84	1.46	0.1011	0.0563
Increase ϕ_b by 10%	-7.37	0.80	1.41	0.1005	0.0483
Increase ϕ_b by 20%	-13.67	0.75	1.37	0.1000	0.0412
Decrease ϕ_b by 10%	8.58	0.88	1.52	0.1016	0.0654
Decrease ϕ_b by 20%	18.55	0.90	1.58	0.1022	0.0754
Panel B: Change in Switching Cost to Public Transit, ϕ_b, and gasoline tax of \$1.					
Baseline without gasoline tax, without change in ϕ_b	–	0.84	1.46	0.1011	0.0563
Baseline with gasoline tax and:					
without change in ϕ_b	-33.00	0.97	1.34	0.0624	0.0870
Increase ϕ_b by 10%	-41.66	0.93	1.29	0.0628	0.0779
Increase ϕ_b by 20%	-49.19	0.89	1.25	0.0631	0.0695
Decrease ϕ_b by 10%	-23.06	0.99	1.39	0.0622	0.0968
Decrease ϕ_b by 20%	-11.69	1.01	1.45	0.0619	0.1074

Notes: Counterfactual analysis using the estimates in table 3, panel C. See subsection 5.4 for a discussion about the computation of the welfare measures and counterfactual calculations. Columns 1 to 3 (for compensating variation, and switching costs from public transit and car) are measured in dollars per month, per market, and per consumer type. Switching costs from public transit (car) refers to the mean, across consumer types, markets and months, value in dollars of costs incurred by switching from public transit (car). Probability to public transit (car) is the mean, across consumer types, markets and months, probability of switching to car (from public transit or the outside option). Similarly, Probability to public transit is the mean, across consumer types, probability of switching to public transit (from car or the outside option). These conditional choice probabilities of each inside mode given by the equation in (4). The baseline value of the switching costs, in dollars, are $\phi_b = 2.14$ and with $\phi_c = 2.22$.

Appendix (For Online Publication)

Contents (Appendix)

A Additional Description of the Data

I collected the data for this study from the following sources. (1) From the Chicago Transit Authority (CTA), I obtained CTA monthly ridership for bus and rail. (2) From Pace, I obtained Pace monthly ridership for bus. (3) From Metra, I obtained monthly ridership for rail serving the northeastern Illinois region. (4) From the Illinois Department of Transportation (IDOT), I obtained automatic traffic recorder (ATR) vehicle counts and classification data. (5) From the U.S. Energy Information Administration (EIA), I obtained the gasoline cost of driving and crude oil prices. (6) From the U.S. Bureau of Labor Statistics (BLS), I obtained the consumer price index. (7) From the Census Transportation Planning Package (CTPP), I collected information about the outside mode of transportation. (8) From the Pace 2006 household survey, I obtained information about travel patterns used to construct the share of consumers who do not switch. Summary statistics of selected variables are in table A2. Next I provide details for each of these sources and the data. Next subsections follow closely the relevant descriptions as stated by each source.

A.1 Data Sources

1. Chicago Transit Authority (CTA) ridership. I collected monthly ridership for bus and rail from the CTA (http://www.yourcta.com), the service board of the RTA responsible for the operations and planning for the rapid transit and bus systems serving the City of Chicago and surrounding suburbs. CTA collects bus and rail ridership using the farebox. Each customer boarding a bus or passing through a rail station turnstile is counted as a single rider. Average CTA bus ridership data is available from May 1998 through December 2009. Average CTA rail ridership data is available for April 1998 through December 2009.

2. Pace ridership. I obtained Pace monthly ridership for bus from Pace (http://www.pacebus.com), the service board of the RTA responsible of the operations and planning for the suburban bus system serving the northeastern Illinois region. Pace provides fixed route bus services, along with contracted community dial-a-ride, paratransit services, and sponsored vanpool services. Pace ridership results are primarily collected using the electronic fare collection systems for each passenger boarding a bus. Ridership for Pace's contract bus routes, which do not have electronic fare collections systems, is manually reported. Both datasets are reported in the RTAMS. Average Pace bus ridership data is available for January 1997 through February 2010.

3. Metra ridership. I obtained monthly ridership for Metra's rail from Metra (http://www.metrarail.com), the service board of the RTA responsible of the operations and planning for the commuter rail system serving the northeastern Illinois region. Metra rail summary data is an estimate based on ticket sales (monthly, 10-ride and one way) for a given month. This dataset does not provide a breakdown by day type or time of day. The dataset is by line and branch and is a combined total for weekday, Saturdays and Sundays/Holidays.

Metra station detail data is based on a count of boardings (the number of passengers boarding a train) and alightings (the number of passengers exiting a train) on a given weekday for each station and for each train. Total monthly values (rather than average daily values) for rail ridership data is available for January 2002 through December 2009.

4. IDOT vehicle counts. I obtained vehicle counts directly from the IDOT (http://www.dot.il.gov/), which has responsibility for planning, construction and maintenance of Illinois' transportation network, which encompasses, highways and bridges, airports, public transit, rail freight and passenger systems. According to the IDOT, the Division of Highways and its district offices are responsible for the design, construction, operation, and maintenance of the state highway system and the administration of the state's local roads and streets program. The state highway system of 17,000 miles includes 2,050 miles of interstate roads. This is part of the 138,000-mile network of state, county, municipal, township, and toll roads. This is the third largest system in the nation. The IDOT's conduct an annual monitoring program consisting of volume counting, vehicle classification, and truck weighing.

According to the IDOT, ATRs are located at permanent locations along the roadway where continuous vehicle counts are collected and retrieved throughout the year. IDOT maintained a network of 85 ATR locations in the state throughout the period under analysis. The ATR location consists of induction loops embedded in the pavement for each lane, a cabinet mounted on concrete off the road with the recording unit, modem, battery, and solar panel. ATR sites are polled daily via modem from the central office. A visual review of each station is made to identify missing or irregular daily lane volumes to assure accurate data is being collected from each station. At the beginning of each month the unedited ATR data from the previous month is transferred from the polling PC to the mainframe for storage. The mainframe file containing the hourly ATR data is reviewed and edited before being transferred to a database. Raw (monthly) count data for the period 1995-2009 was collected. Of the 85 ATR sites, 36 of those locations also collect hourly vehicle classification counts for each of 13 vehicle types as defined in FHWA's guide. These locations use two induction loops with a piezo electric cable between the loops embedded in the pavement for each lane. Classification data by hour and vehicle type is polled, reviewed, and transferred to a database. Raw (hourly) classification data for the period 2001-2009 was collected. For the purpose of this study these data was aggregated at the monthly level to be consistent with the frequency for public transit ridership.

ATR distribution and classification data. The ATRs are distributed across the relevant area (*e.g.* urban area, expressway, rural area, *etc.*) to provide adequate representation of the principal roadways. There are seven major interstate highways crossing the city of Chicago. Vehicle circulation is available for five of these expressways: Bishop Ford, Dan Ryan, Edens, Eisenhower, and Kennedy. These five expressways define the routes for the market definition used in this paper. Vehicle circulation for each expressway is recorded at specific permanent locations along the roadway designated to "achieve a statistically valid representation of all roadway systems above township roads and municipal streets" (Illinois Department of Transportation 2004, p. 4).

Table A1: Distribution of ATRs in Illinois by district and functional type.

Panel A: ATR District Distribution.

District	Number of of ATRs	
	Total	With classification data
District One	43	17
District Two	5	2
District Three	3	1
District Four	6	2
District Five	4	2
District Six	6	2
District Seven	6	3
District Eight	9	5
District Nine	3	2
Statewide	85	36

Panel B: ATR distribution by functional class.

District	Number of of ATRs	
	Total	With classification data
Interstate (Urban)	5	1
Interstate (Rural)	14	8
Other Principal Arterial (Urban)	29	14
Other Principal Arterial (Rural)	6	4
Minor Arterial (Rural)	5	3
Major Collector (Rural)	3	1
Minor Arterial (Urban)	16	4
Collector (Urban)	7	1
Satewide	85	36

Source: Illinois Department of Transportation. See Illinois Department of Transportation (2004) for details.

5. U.S. Energy Information Administration (EIA) gasoline cost of driving and crude oil prices. The gasoline cost of driving and crude oil prices were obtained from the EIA (`http://www.eia.doe.gov/`). According to the EIA, every Monday, retail prices for all three grades of gasoline are collected by telephone from a sample of approximately 900 retail gasoline outlets. The prices are published by 5:00 P.M. Monday, except on government holidays, when the data are released on Tuesday, but still represent Monday's price. The reported price includes all taxes, and is the pump price paid by a consumer as of 8:00 A.M. Monday. This price represents the self serve price, except in areas having only full serve. The price data are used to calculate weighted average price estimates at the city, state, regional, and national levels using sales and delivery volume data from other EIA surveys and population estimates from the Bureau of Census. Monthly prices are the average of weekly prices, also reported directly by EIA. Real gasoline prices in Chicago are used to compute the gasoline cost of driving. The EIA also reports crude oil prices, which are used as instruments.

6. U.S. Bureau of Labor Statistics (BLS). Monthly Consumer Price Indexes (CPI) for U.S. city average, and Chicago-Gary-Kenosha area were collected from the BLS (`http://www.bls.gov/`). The BLS produces these indexes through the CPI program, using the prices paid by urban consumers for a representative basket of goods and services. CPIs are used to construct the inflation adjusted variables, as explained in the paper.

7. Census Transportation Planning Package (CTPP). I use the CTPP (`http://ctpp.transportation.org/`) to define the size of the market: the share of people whose work trip mode is neither auto nor public transit, which defines share of the outside mode of transportation, as defined in the next paragraph. I use the "work trip mode share for Chicago" obtained from the municipality disaggregation. This measures the percentage of people for the whole Chicago area whose work trip mode is neither auto nor public transit.

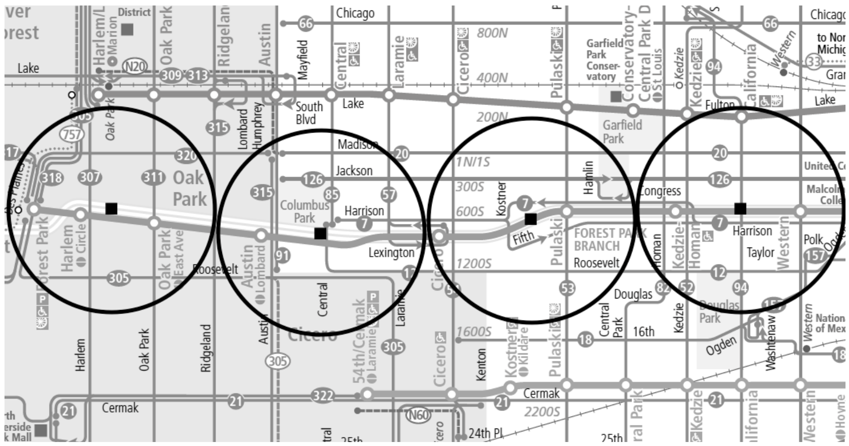
Market Size. A market is defined as a combination of route and month. I link vehicle circulation to public transit use in each route by defining a radius around the location of the ATRs. I analyze spatial information by integrating geographic information about Chicago public transportation and roadway network with ridership data into a geographic information system (GIS). The analysis is performed using ArcGIS, the GIS software produced by ESRI. I merge geospatial information on CTA rail, bus lines, road, and ATR locations with census population by zip codes, along with public transit and vehicle volume data. For each route and month, I compute the public transit use as total ridership for each rail station and bus line that lies within the radius. Figure A1 provides an example for Dwight D. Eisenhower Expressway (Interstate 290). Similarly, car use is computed as the total number of vehicles that circulated in a given route-month according to the ATRs located in the route. Vehicle circulation is captured as follows. For Bishop Ford expressway using ATR stations 4917, 4928, 4946, and 4960. For Dan Ryan express way using ATR stations 4970, 4994, and 5015. For Edens expressway using ATR stations 5045, 5070, and 5091. For Eisenhower expressway and Eisenhower extension using ATR stations 5101, 5107, 5111, 5120, 5139, 5161, 5191, 5196, and 5203. Finally, for Kennedy expressway using ATR stations 5270, 5278, 5298, 5303, 5318, 5345, and 5353.[42] Market shares are defined by dividing vehicle circulation and ridership by the market size. The market size was assumed to be the total census population that lies within the radius around the ATRs. The radius was chosen to ensure that the market share of the outside mode equals the one from the CTPP.[43] I calculate the population that extends over the area defined by the radius using population data disaggregated by zip code

[42]Classification data is not available for these ATRs. So it is not possible to distinguish into categories depending on whether the vehicle carries passengers or commodities. This introduces two empirical challenges. First, that vehicle count for passenger may react differently than total vehicle count to movements in gasoline prices. Second, that vehicle's market shares may need to be adjusted to take into account the fraction of passenger vehicles. In appendix B, I use classification data for other ATRs in Chicago and Illinois, and show that passenger vehicles and total vehicles counts exhibit similar response to the gasoline cost of driving. For details see table A5 described in subsection B.2.2.

[43]For the specifications reported in table 3, I performed the estimation increasing and decreasing the radius definition (market size) by 10 percent, and obtained almost identical results. Results available upon request.

assuming population is distributed homogeneously over the zip code area.

Figure A1: An Example of a Route.



Notes: An example of a *route*, along Dwight D. Eisenhower Expressway (Interstate 290). Black dots denote the location of ATR stations. Circles denote radius around those locations.

8. Pace survey. I use the Pace survey from 2006 (`http://www.pacebus.com`) to compute the relative shares of consumers who switch, and who do not switch inside modes of transportation. The survey was administered between January and April 2006 to 1,330 randomly chosen households in the Chicago metropolitan area. It collected information about: (i) households' observed choices of mode of transportation, (ii) perceptions and attitudes toward transit, and (iii) responses to choices experiment to quantify the tradeoffs that people make in choosing among different travel options. See Cambridge Systematics Inc. (2007) for details about the survey, sampling, and administration procedures. Additional description of the survey is available in Owen, Jane, and Kopp (2007). See Long, Lin, and Proussaloglou (2010) for an application.

This additional information is more desegregated than the market level data above. It describes the behavior of consumer types who, *e.g.*, did not switch modes of transportation. Based on the main factors determining their mode of transportation to travel to work (*e.g.* public transit attitudes, safety, time, *etc.*), the survey respondents were grouped into distinct segments. The shares for these segments and the market size in each route are used to compute the shares of consumers who do not switch the inside modes of transportation.

A.2 Summary Statistics

Table A2: Summary statistics.

Panel A: Transportation data

	Mean	Median	St. dev.	Min.	Max.	Nmbr. obs.
Public transit ridership:[a]						
- CTA rail (143)	3,602.6	2,747	3,011.40	81	20,151	15,725
- CTA bus (167)	6,878.5	4,203	6,968.72	10	37,787	15,646
- Pace bus (297)	486.0	211	756.97	6	7,414	26,090
- Metra (14)	494,029.6	525,129.5	370,672.40	12,263	1,554,401	1,302
Vehicle circulation from ATRs:[b]						
- Illinois (135)	33,200.7	20,793	32,499.21	264	163,895	8,333
- Expressways (28)	86,813.72	87,464.5	19,108.26	43,258	152,754.00	1,300
- Urban areas (101)	37,950.9	23,171	34,030.09	264	163,895	6,399
- Rural areas (27)	14,145.4	9,039	12,779.09	470	85,640	1,743
Classification data:[b]						
- All vehicles Chicago (17)	623,600.1	599,075	491,198.00	3,722	2,389,964	1,301
Passenger (17)	591,386.0	576,195	468,920.70	3,499	2,299,243	1,301
Single unit trucks (17)	14,659.1	10,180	22,943.95	0	309,475	1,233
Multiple units trucks (17)	16,728.3	6,286	16,728.26	0	303,550	1,227

Panel B: Gasoline and oil prices[c]

	Mean	Median	St. dev.	Min.	Max.
Nominal gasoline price in Chicago	2.239	2.047	0.742	1.152	4.271
Nominal gasoline price in the U.S.	2.130	1.949	0.694	1.130	4.090
Nominal oil price (wtotusa)	1.095	0.917	0.593	0.378	3.119
Gasoline price in Chicago in 2005 dollars	1.143	1.073	0.311	0.645	1.964

Panel C: Market shares[d]

	Mean	Median	St. dev.	Min.	Max.	Mean st. dev. across routes[e]	Mean st. dev. across periods[e]
Automobile	0.481	0.502	0.129	0.118	0.801	0.078	0.080
Public transit	0.166	0.189	0.059	0.048	0.239	0.064	0.011

Notes: Summary statistics of selected variables. All variables are at a monthly frequency. See section 2 for details about the data. Variables definitions are in appendix D. Classification data definitions are in appendix E. Period is the same as the one in the preliminary analysis in subsection 2.3, and tables A3, A4, and A5. [a] CTA rail, CTA bus, and Pace bus values refer to monthly average for weekdays, disaggregated by rail station, or bus routes, respectively. Metra rail refers to monthly total combined for weekday, Saturdays, and Sundays/Holidays, disaggregated by branch. The number in parenthesis indicates the number of rail stations, bus routes, or branches, respectively. [b] Vehicle circulation refers to the raw monthly average number of vehicles circulating through the ATRs locations, disaggregated by ATRs. The number in parenthesis indicates the number of ATRs stations included. [c] All prices are in dollars per gallon for the period June 2000 to October 2009. [d] See subsection 2.2 and appendix A for details about the market shares. [e] The mean standard deviation across routes is computed as follows: $\frac{1}{R}\sum_{r=1}^{R} sd_r$, where sd_r is the standard deviation of the market share in route $r = 1, \ldots, R$, where each standard deviation is taken over the set of periods (months) $t = 1, \ldots, T$. The mean standard deviation across periods (months) is computed as follows: $\frac{1}{T}\sum_{t=1}^{T} sd_t$, where sd_t is the standard deviation of the market share in period $t = 1, \ldots, T$, where each standard deviation is taken over the set of routes $r = 1, \ldots, R$.

B Additional Preliminary Analysis

In this section I document the response of vehicle circulation and public transit use to the variation in the gasoline cost of driving. I use variation in public transit ridership and vehicle counts conditional on stations/routes/ATR shocks after taken into account the monthly seasonal effects. The regressions below exploit cross-sectional and time series variation in the data by including subarea disaggregation within Chicago area. Overall the results show that public transportation use (CTA rail, CTA bus, Pace bus, and Metra rail) increases, and vehicle use measured by ATRs decreases as the gasoline cost of driving increases. Single-trailer trucks are substituted by multi-trailer trucks. The effects are statistically and economically significant.

The baseline OLS specification for ridership or traffic volume d_{smy} for station/route/ATR s in month m and year y is:

$$d_{smy} = x_s\alpha - \beta p_{my} + \xi_s + \xi_m + \epsilon_{smy}, \tag{B.1}$$
$$s = 1, \ldots, S, \quad m = 1, \ldots, M, \quad y = 1, \ldots, Y,$$

where x_s is a K-dimensional row vector of observable characteristics for station/route/ATR s which include a vector of ones; p_{my} is the gasoline cost of driving in Chicago in month m and year y, which is the variable of primary interest; ξ_s is a station, route, or ATR fixed effect (depending on the data structure); ξ_m is a monthly seasonal fixed effect; and ϵ_{smy} is a mean-zero stochastic term. Finally, (α, β) are $K + 1$ coefficients. One concern in identifying price estimates to the gasoline cost of driving is that individuals often self select to the location where they live and work. So fixed effects for station, route, and ATR are included to remove the variation in demand shocks that occurs as a result of different rail stations, ATR locations, or bus routes. The reported gasoline effects are identified by variation within station/route/ATR and month.

For robustness I also report results from two stage least squares regressions (IV henceforth) using crude oil prices and U.S. city average gasoline cost of driving as instruments. The top panel in figure 1 shows that the gasoline cost of driving in Chicago and the U.S. city average gasoline cost of driving are closely linked with crude oil prices. U.S. gasoline prices depend on world crude oil prices and refinery margins. The exclusion restriction for the regressions reported below is that world crude oil is not influenced by demand shocks within the same station/route/ATR in Chicago conditional on the seasonal effects. Under this assumption crude oil prices can be used as instruments.[44]

I estimated the reduced-form models above (OLS and IV) using both a linear and a log-linear specification. The latter was implemented by regressing the natural logarithm of the dependent variable on the natural logarithm of the gasoline cost of driving (plus the covariates, fixed effects, and monthly seasonal effects). I present the results using this log-linear specification, so the reported coefficients can be interpreted as elasticities. Similar results are obtained using a linear specification. For each regression specification I report only the estimated coefficient for the gasoline cost of driving. The results from the reduced-form models are presented in the next subsections.

B.1 Public Transit Use

B.1.1 CTA

Panel A in table A3 presents the results for CTA rail ridership. The estimated results are disaggregated for weekdays, Saturdays, and Sundays/Holidays (columns), and 8 specifications (rows) as described in the last two columns. Each reported coefficient corresponds to a different regression of the natural logarithm of CTA rail ridership on the natural logarithm of the gasoline cost of driving, monthly seasonal fixed effects, and rail station specific fixed effects. The estimated coefficient on gasoline price in specification 1 for column weekdays has the expected sign and is statistically significant at the 1 percent level. A one percent increase in gasoline prices in Chicago increases weekday CTA rail ridership by .17 percent. The regression in row 2 adds a station specific monthly time trend. The estimated coefficient, though smaller, has the expected sign and is statistically different from zero. Row 3 adds year fixed effects to the specification in row (1). Finally, row 4 adds both a station specific monthly time trend, and year fixed effects to row 1. Rows 5 to 8 are IV regressions using crude oil prices and city average gasoline cost of driving in the U.S (along with fixed effects and the station month trend) as instruments for the gasoline cost of driving in Chicago. Similar results are obtained. The estimated coefficient on the gasoline cost of driving is positive and significant. The elasticities estimates vary from .04 percent to .17 percent, depending on the specification.

Columns Saturdays and Sunday/Holiday show analogue specifications for the ridership during those days. Similar results are obtained to the ones for weekdays. However, the estimated elasticities are higher, indicating larger substitution patterns as expected. For

[44] One concern to identify gasoline price effects is that ridership or vehicle use may be correlated with station/route/ATR-specific omitted factors, such as some cyclical effect not captured by the included seasonal effects. One possible source of endogeneity, for example, could be that variations in Chicago economic conditions, correlated with personal income, may be correlated with changes in gasoline prices in Chicago. Hence, gasoline price estimates will capture this cyclical effect on ridership/vehicle use. Other source could be that gasoline local taxes might be correlated with economic cycles, thus biassing the estimates. In general, any station/route/ATR-month-specific shock that affects both demand and gasoline prices would produce results that look like the gasoline cost of driving effects. As a robustness check to address this concern, I repeated the analysis adding station/route/ATR-specific monthly time trends by adding the term $m\xi_s$ in (B.1), and obtained similar results, although the level of precision decreases.

example, for Saturdays the estimated elasticity varies from .05 percent to .88 percent. For Sundays/Holidays the elasticity varies from .06 percent to .52 percent.

Panel B in table A3 presents the results for CTA bus ridership using the same specifications as in panel A. The estimated coefficients have the expected positive sign, and are statistically significant. For weekdays the elasticity of CTA bus ridership varies from .08 percent to .1 percent, while for Saturdays and Sundays/Holidays the estimates are also higher and vary from .07 percent to .1 percent, and .139 percent to .145 percent, respectively.

B.1.2 Pace

Panel C in table A3 presents the results for Pace bus ridership using similar specifications as in previous panels. The estimates are similar to the ones for CTA ridership. The estimated coefficient on the gasoline cost of driving is positive and significant. The estimated elasticities vary from .02 percent to .2 percent. For Saturdays and Sundays/Holidays the elasticities are also positive and significant, and higher than the ones for weekdays. For Saturdays the elasticity varies from .11 percent to .27 percent, while for Sundays/Holidays it varies from .13 percent to .29 percent.

B.1.3 Metra

Panel D in table A3 presents the results for Metra rail ridership. Unlike CTA and Pace ridership data, Metra ridership does not provide a breakdown by day type, and is a combined total for weekday, Saturdays, and Sundays/Holidays. It does not provide station disaggregation either, unlike CTA rail ridership dataset, just disaggregation in the 14 branches. This reduces the number of observations considerably, compared to CTA rail. The specification in row 1 includes month of the year seasonal fixed effects, and branch specific fixed effects. Row 2 adds additionally a branch specific monthly time trend to the specification in row 1. The goodness of the fit is relatively high in both regressions, 99.2 percent and 99.7 percent. Most of the variation in the variable of interest is absorbed by the gasoline cost, and the mentioned fixed effects. The reason for this is the way Metra data is disaggregated as explained above. Adding additional year fixed effects leads to weak identification and imprecise estimates of the effect of gasoline effect on Metra ridership, so year fixed effects are not included in the specifications in panel D. The results from the IV specifications are similar to the ones in OLS. Again, the estimated coefficient is positive and significant as expected. The estimated elasticities values vary from .07 percent to .17 percent.

Table A3: Public Transit Response to Gasoline Cost of Driving.

		Weekdays				Saturdays				Sundays/Holidays				Unit specific	Year
		ln(Gas.)	Sd. error	Nmbr. obs.	R^2	*ln*(Gas.)	Sd. error	Nmbr. obs.	R^2	*ln*(Gas.)	Sd. error	Nmbr. obs.	R^2	month trend	fixed effects
Panel A: CTA rail ridership.															
OLS	1	.174	(.020)	15,725	.923	.878	(.036)	15,571	.721	.518	(.066)	15,130	.78	No	No
	2	.066	(.027)	15,725	.945	.266	(.050)	15,571	.910	.126	(.053)	15,130	.891	Yes	No
	3	.037	(.013)	15,725	.926	.074	(.032)	15,571	.731	.091	(.078)	15,130	.789	No	Yes
	4	.035	(.014)	15,725	.946	.048	(.031)	15,571	.913	.055	(.052)	15,130	.894	Yes	Yes
IV	5	.181	(.019)	15,725	.923	.932	(.033)	15,571	.720	.551	(.063)	15,130	.78	No	No
	6	.067	(.027)	15,725	.945	.289	(.053)	15,571	.910	.143	(.046)	15,130	.891	Yes	No
	7	.044	(.015)	15,725	.926	.110	(.037)	15,571	.731	.108	(.083)	15,130	.789	No	Yes
	8	.043	(.015)	15,725	.946	.072	(.034)	15,571	.913	.089	(.056)	15,130	.894	Yes	Yes
Panel B: CTA bus ridership.															
OLS	1	.097	(.018)	15,646	.953	.101	(.018)	10,793	.955	.145	(.025)	9,116	.939	No	No
	2	.060	(.012)	15,646	.972	.017	(.023)	10,793	.968	.042	(.034)	9,116	.952	Yes	No
	3	.089	(.019)	15,646	.954	.074	(.026)	10,793	.956	.148	(.049)	9,116	.94	No	Yes
	4	.077	(.023)	15,646	.972	.072	(.027)	10,793	.969	.139	(.053)	9,116	.952	Yes	Yes
IV	5	.103	(.017)	15,646	.953	.110	(.018)	10,793	.955	.153	(.025)	9,116	.939	No	No
	6	.067	(.013)	15,646	.972	.029	(.023)	10,793	.968	.047	(.035)	9,116	.952	Yes	No
	7	.114	(.018)	15,646	.954	.114	(.030)	10,793	.956	.183	(.054)	9,116	.940	No	Yes
	8	.101	(.020)	15,646	.972	.114	(.028)	10,793	.969	.177	(.058)	9,116	.952	Yes	Yes
Panel C: Pace bus ridership.															
OLS	1	.021	(.015)	26,090	.968	.113	(.020)	10,966	.947	.207	(.032)	4,533	.924	No	No
	2	.227	(.036)	26,090	.981	.269	(.031)	10,966	.971	.291	(.052)	4,533	.961	Yes	No
	3	.074	(.012)	26,090	.969	.101	(.016)	10,966	.949	.135	(.047)	4,533	.926	No	Yes
	4	.078	(.014)	26,090	.982	.110	(.020)	10,966	.973	.130	(.048)	4,533	.962	Yes	Yes
IV	5	.009	(.013)	26,090	.968	.108	(.021)	10,966	.947	.207	(.034)	4,533	.924	No	No
	6	.244	(.040)	26,090	.981	.296	(.033)	10,966	.971	.316	(.058)	4,533	.961	Yes	No
	7	.097	(.016)	26,090	.969	.139	(.026)	10,966	.949	.167	(.039)	4,533	.926	No	Yes
	8	.102	(.013)	26,090	.982	.150	(.024)	10,966	.973	.162	(.040)	4,533	.962	Yes	Yes
		Combined weekdays, Saturdays, and Sundays/Holidays[a]													
Panel D: Metra rail ridership.															
OLS	1	.170	(.017)	1,302	.992									No	No
	2	.071	(.026)	1,302	.997									Yes	No
IV	5	.173	(.016)	1,302	.992									No	No
	6	.058	(.028)	1,302	.997									Yes	No

Notes: All regressions include a constant, monthly seasonal fixed effects, and unit of observation fixed effects, where the unit is the CTA station, CTA bus route, Pace bus route, or Metra branch in panels A, B, C, and D, respectively. Independent variable is, in all regressions, the natural logarithm of the gasoline cost of driving in Chicago in real terms (variable Gasoline Price in the variables definition appendix), *ln*(Gas.). In each panel rows 1 to 4 are OLS regressions, and rows 5 to 8 are the corresponding two stage least squares regressions (labeled "IV") using the natural logarithm of crude oil prices and city average gasoline cost of driving in the U.S (along with fixed effects and the station month trend) as instruments for the gasoline cost of driving in Chicago. Unit specific month trend refers to the unit of analysis month specific time trend in addition to the monthly seasonal fixed effects, where the unit is the CTA station, CTA bus route, Pace bus route, or Metra branch in panels A, B, C, and D, respectively. Robust standard errors clustered at month level are reported in parentheses for all regressions. Variables definitions are in appendix D. In all regressions the period included is June 2000 (the first month with no missing information for the gasoline price in Chicago) to August 2009 (the latest month with ridership information at the moment of the data collection). The number of years included in the regressions is 10. The number of months included is 111. The number of stations, routes, and branches in each panel is reported below. Panel A: All regressions include station specific and month fixed effects. Dependent variable is the natural logarithm of the monthly CTA rail ridership in the days specified. The number of stations included is 143 (weekdays). Panel B: All regressions include bus route specific and month fixed effects. Dependent variable is the natural logarithm of the monthly CTA bus ridership in the days specified. The number of routes included is 167 (weekdays). Panel C: All regressions include bus route specific and month fixed effects. Dependent variable is the natural logarithm of the monthly CTA bus ridership in the days specified. The number of routes included is 297 (weekdays). Panel D: All regressions include branch specific and month fixed effects. Dependent variable is the natural logarithm of the monthly Metra rail ridership for the combined weekdays, Saturdays, and Sundays/Holidays. The number of branches included is 14.

[a] Metra rail ridership does not provide a breakdown by weekday, Saturdays, and Sundays/Holidays. It does not provide station disaggregation either, unlike CTA rail ridership dataset, just disaggregation in the 14 branches. This reduces the number of observations considerably, compared to, *e.g.*, CTA rail. Adding year fixed effects leads to imprecise estimates of the effect of gasoline effect on Metra ridership due to collinearity issues, so year fixed effects are not included in the specifications in panel D. Results with year fixed effects are available upon request. Variables, fixed effects, and time trends definitions are in appendix D.

B.2 Vehicle Circulation

B.2.1 Rural and Urban Areas

Panel A in table A4 displays the results for vehicle counts using the ATRs data. Each reported coefficient corresponds to a different regression of the natural logarithm of the number of vehicles circulating through all ATRs on the natural logarithm of the gasoline cost of driving, monthly seasonal fixed effects, and ATR specific fixed effects. The sample for panel A is the complete network of ATR locations from the IDOT throughout the state of Illinois. Row 1 presents the baseline specification. Traffic volume measured by ATRs decreases when the gasoline cost of driving increases as expected. A one percent increase in the gasoline cost of driving reduces vehicle circulation by .03 percent. The effect is statistically significant at the 1 percent level, and the goodness of the fit is relatively high, 99.1 percent. Rows 2 and 3 add additional covariates (direction, good months, functional classification, and number of lanes), and ATR specific monthly time trend to the specification in row 1. The number of observations decreases considerably because some ATRs do not have information about the additional covariates. The estimated coefficients are qualitatively the same and the absolute value of the elasticity increases. Rows 4 to 6 are the corresponding IV regressions using the same specifications as before. Similar results are obtained.

Panel B in table A4 restricts the sample to ATRs located in expressways. As expected, the absolute value of the elasticity for the gasoline cost of driving is almost four times higher than the one obtained in panel A, for all ATRs. Panels C and D restricts the sample to urban and rural ATRs locations, respectively. Again, the estimated coefficients for the gasoline cost of driving have the expected negative sign, and are statistically significant. The absolute value of the elasticity in urban ATRs is 1.3 times higher than the value obtained for ATRs located in rural areas. This is because urban areas have better access to public transportation services than rural areas.

B.2.2 Classification Data for the Chicago Metropolitan Area

Table A5 presents estimates for the subset of ATRs in the Chicago metropolitan area that collect classification data. This is a subsample of the one in table A4. These locations additionally collect vehicle classification counts for each of 13 vehicle types as defined in FHWAs traffic monitoring guide described in appendix E. Panel A displays the estimates by main classes: all vehicles, only passengers' vehicles, single-unit trucks, and multiple-unit trucks. Row 1 presents the analogue estimates to the ones in panel A in table A4, but for the Chicago area with classification data. The estimated coefficient has the expected sign and is statistically significant. The estimated elasticity for the Chicago area is higher than the value estimated for the whole Illinois area as expected. In row 2 the dependent variable is the natural logarithm of passenger vehicles as defined by the IDOT. The estimated effect is negative but only significant at the 10 percent level. Rows 3 and 4 display the estimates for single-unit trucks and multiple-unit trucks, respectively. The estimated coefficients are strongly negative and significant at the 1 percent level. A one percent increase in the gasoline cost of driving in Chicago reduces single-unit trucks circulation in the Chicago area by 2.07 percent, and multi-unit trucks circulation by 1.97 percent. Similar results are obtained using the IV specifications.

Panel B in table A5 displays the estimates for each of the 13 classes individually. For passenger vehicles (classes 1 to 3) one can observe an increase in motorcycle vehicles (class 1), and a decrease in passenger cars (class 2) and four-tire single unit vehicles, such as pickups or vans (class 3). The estimated coefficient on passenger cars is not statistically different from zero. Panel B shows that the number of buses (class 4) strongly decreases with the gasoline cost of driving consistent with the public transit ridership data. Interestingly, single-trailer trucks circulation (classes 6 to 10) strongly decreases as the gasoline cost of driving increases, while multi-trailer trucks circulation (classes 11 to 13) increases. These results suggest that, as the gasoline cost of driving increase, there is a substitution from single-trailer trucks to multi-trailer trucks, a possible efficiency effect.

Table A4: Vehicle Circulation Response to Gasoline Cost of Driving.

		ln(Gas.)	Sd. error	Nmbr. obs.	R^2	Covariates	ATR specific month trend
Panel A: Vehicles from all ATRs in Illinois.							
OLS	1	-.034	(.008)	8,333	.991	No	No
	2	-.038	(.007)	4,319	.992	Yes	No
	3	-.054	(.026)	4,319	.994	Yes	Yes
IV	4	-.030	(.009)	8,333	.991	No	No
	5	-.037	(.007)	4,319	.992	Yes	No
	6	-.044	(.025)	4,319	.994	Yes	Yes
Panel B: Vehicles from ATRs in Chicago's expressways.[a]							
OLS	1	-.125	(.030)	1,300	.901	No	No
	2	-.615	(.055)	13	.999	Yes	No
IV	4	-.108	(.027)	1,300	.901	No	No
	5	-.615	(.055)	13	.999	Yes	No
Panel C: Vehicles from ATRs in urban areas.							
OLS	1	-.034	(.009)	6,399	.989	No	No
	2	-.043	(.007)	3,292	.991	Yes	No
	3	-.039	(.024)	3,292	.993	Yes	Yes
IV	4	-.030	(.009)	6,399	.989	No	No
	5	-.041	(.007)	3,292	.991	Yes	No
	6	-.032	(.022)	3,292	0.993	Yes	Yes
Panel D: Vehicles from ATRs in rural areas.							
OLS	1	-.027	(.013)	1,743	.992	No	No
	2	-.019	(.012)	1,027	.993	Yes	No
	3	-.115	(.041)	1,027	0.995	Yes	Yes
IV	4	-.021	(.014)	1,743	.992	No	No
	5	-.014	(.013)	1,027	.993	Yes	No
	6	-.108	(.042)	1,027	0.995	Yes	Yes

Notes: All regressions include a constant, monthly seasonal fixed effects, and ATR specific fixed effects. Independent variable is, in all regressions, the natural logarithm of the gasoline cost of driving in Chicago in real terms (variable Gasoline Price in the variables definition appendix), ln(Gas.). In each panel rows 1 to 4 are OLS regressions, and rows 5 to 8 are the corresponding two stage least squares regressions (labeled "IV") using the natural logarithm of crude oil prices and city average gasoline cost of driving in the U.S (along with fixed effects and the station month trend) as instruments for the gasoline cost of driving in Chicago. Covariates include: direction of the route, good months indicator, Illinois functional classification of the road, and number of lanes. The decrease in the number of observations when including the covariates is because covariates' information is not collected by the ATRs in the missing observations. ATR specific monthly trend refers to an ATR month specific trend in addition to the monthly seasonal fixed effects. Robust standard errors clustered at month level are reported in parentheses for all regressions. Variables, fixed effects, and time trends definitions are in appendix D. In all regressions the period included is June 2000 (the first month with no missing information for the gasoline price in Chicago) to July 2009 (the latest month with vehicle counts information provided by the IDOT at the moment of the data collection). Some ATRs has missing observations on specific months. The number of years included in the regressions is 10. The number of months included is 110. The number of ATRs stations in each panel is reported below. Panel A: Dependent variable is the natural logarithm of the number of vehicles circulating through all ATRs in Illinois. The number of ATRs included is 135 (row 1), all ATRs with non missing data provided by the IDOT. Panel B: Dependent variable is the natural logarithm of the number of vehicles circulating through ATRs located in expressways in the Chicago area, as indicated by the IDOT. The number of ATRs included is 28 (row 1). Panel C: Dependent variable is the natural logarithm of the number of vehicles circulating through ATRs located in urban areas, as indicated by the IDOT. The number of ATRs included is 101 (row 1). Panel D: Dependent variable is the natural logarithm of the number of vehicles circulating through ATRs in located in rural areas, as indicated by the IDOT. The number of ATRs included is 27 (row 1).

[a] There is only one ATR station with covariates' information in Chicago's expressways (leaving only 13 observations with no missing data in the regressions with covariates in panel B). Thus regressions in rows 3 and 6 are not included in panel B.

Table A5: Vehicle Circulation Response in Chicago, by Class Type.

	OLS				IV				
	ln(Gas.)	Sd. error	Nmbr. obs.	R^2	ln(Gas.)	Sd. error	Nmbr. obs.	R^2	Mean share[b]
Panel A: Classification data by main classes.									
All vehicles Chicago[a]	-0.170	0.075	1,301	0.696	-0.163	0.074	1,301	0.696	—
- Passenger	-0.157	0.073	1,301	0.690	-0.146	0.073	1,301	0.690	—
- Single unit trucks	-2.065	0.112	1,233	0.500	-2.143	0.106	1,233	0.500	—
- Multiple units trucks	-1.971	0.117	1,227	0.518	-2.077	0.112	1,227	0.518	—
Panel B: Classification data by each class.									
- Class 1	0.656	0.231	1,231	0.587	0.742	0.254	1,231	0.587	.43
- Class 2	-0.043	0.074	1,301	0.701	-0.027	0.075	1,301	0.701	88.32
- Class 3	-2.142	0.141	1,229	0.438	-2.215	0.139	1,229	0.438	11.25
- Class 4	-3.014	0.103	1,227	0.593	-3.138	0.098	1,227	0.592	12.94
- Class 5	-2.548	0.092	1,231	0.434	-2.627	0.087	1,231	0.434	59.03
- Class 6	-0.988	0.130	1,217	0.539	-1.000	0.131	1,217	0.539	25.91
- Class 7	-2.908	0.157	1,172	0.399	-3.069	0.149	1,172	0.398	2.12
- Class 8	-2.353	0.131	1,218	0.511	-2.451	0.121	1,218	0.511	27.62
- Class 9	-1.688	0.149	1,209	0.514	-1.790	0.146	1,209	0.514	64.73
- Class 10	-1.455	0.163	1,169	0.496	-1.534	0.163	1,169	0.496	3.00
- Class 11	0.851	0.146	1,013	0.668	0.863	0.134	1,013	0.668	1.10
- Class 12	4.224	0.218	938	0.625	4.417	0.197	938	0.624	2.29
- Class 13	0.943	0.152	1,110	0.550	1.038	0.150	1,110	0.550	1.26

Notes: Each raw presents a different regression, where the dependent variable is the natural logarithm of the variable indicated in the first column. In each regression, vehicle counts from all ATRs that collect classification data are included. All regressions include a constant, monthly seasonal fixed effects, and ATR specific fixed effects (*i.e.* similar specifications to the ones in row 1 in table A4). Information about covariates is not available for the ATRs that collect classification data. Independent variable is, in all regressions, the natural logarithm of the gasoline cost of driving in Chicago in real terms (variable Gasoline Price in the variables definition appendix), ln(Gas.). The columns labeled "OLS" refer to OLS regressions. The columns labeled "IV" refer to two stage least squares regressions using the natural logarithm of crude oil prices and city average gasoline cost of driving in the U.S (along with the monthly seasonal fixed effects and ATR specific fixed effects) as instruments for the gasoline cost of driving in Chicago. Robust standard errors clustered at month level are reported in parentheses for all regressions. Variables, fixed effects, and time trends definitions are in appendix D. Definitions of each class by the FHWA are provided in appendix E. In all regressions the period included is January 2001 (the first month with vehicle counts classification information provided by the IDOT) to October 2009 (the latest month with vehicle counts classification information provided by the IDOT at the moment of the data collection). Some ATRs has missing observations on specific months. The number of years included in the regressions is 9. The number of months included is 106. The number of ATRs stations in each regression is 17, corresponding to district one with classification data in table A1, panel A.

[a] All vehicles Chicago refers to the raw total number of vehicles recorded by the ATRs. Passenger refers to the raw total number of vehicles in classes 1, 2 and 3, as recorded by the ATRs. Single unit trucks refers to the raw total number of vehicles in classes 4, 5, 6 and 7, as recorded by the ATRs. Multiple units trucks refers to the raw total number of vehicles in classes 8, 9, 10, 11, 12, and 13, as recorded by the ATRs.

[b] Mean share is the mean share of each class in panel B as a percentage of the relevant category in panel A: passenger vehicles, single unit trucks, or multiple units trucks, respectively (depending on which of these 3 categories each class belongs). For example, the category Passenger in panel A includes classes 1, 2 and 3. Thus, in panel B the mean shares of classes 1, 2, and 3 sum 100 percent: $0.43 + 88.32 + 11.25 = 100$. Similarly, for the other two categories in panel A, single and multiple unit trucks.

C Details About the Estimation

In this appendix I provide details about the procedure to solve for the mean utility in step 2, the contraction mapping in step 3, the initial condition of the consumers and initial market shares, the discretization of p_t, simulation of consumer types, computation of the shares of consumers who do not switch, and inference.

Details about step 1. For the weighting matrix, W, a consistent estimate of $\mathbb{E}[Z'\xi(\theta)\xi(\theta)'Z]$ is chosen using the following standard two step approach. First, the estimation algorithm is applied using an initial weighting matrix, W_0, computed by bootstrapping the empirical moments. This outputs a consistent estimate, θ_0, although not efficient. The estimate θ_0 is used to compute a new weighting matrix, $W_1 = \mathbb{E}[Z'\xi(\theta_0)\xi(\theta_0)'Z]$. Second, the estimation algorithm is applied a second time using the weighting matrix from the previous step, W_1. This outputs a consistent estimate, θ_1, that is asymptotically efficient. This procedure requires estimating each model twice with the described estimation algorithm.

Details about step 2. As explained in the text, the structural error term is defined as the unobserved products' characteristics. It is computed by solving for the mean utility level, δ_{mrt}, that equates $s_{mrt}(p_t, \delta_{mrt}; \alpha_{0mr}, \Sigma, \phi) = S_{mrt}$. I solve for $\delta_{mrt}(\theta)$ in the system of equations in (7) using Broyden's method for finding roots (Broyden 1965). I have experimented with a wide variety of different starting values, using random number generators to pick them, and have always obtained the same solution. For robustness, all models have also been estimated using the following two methods to solve for $\delta_{mrt}(\theta)$ in the system of equations in (7). (1) The simplex search method by Lagarias, Reeds, Wright, and Wright (1998). (2) Iterating on it, market by market, analogously to the contraction mapping used by Berry (1994) and Berry, Levinsohn, and Pakes (1995):

$$\delta^{h+1}_{mrt} = \delta^{h}_{mrt} + \ln\left(S_{mrt}\right) - \ln\left[s_{mrt}(p_t, \delta_{mrt}; \alpha_{0mr}, \Sigma, \phi)\right], \tag{C.1}$$

$$m \in \{c, b\}, \quad r = 1, \dots, R, \quad t = 1, \dots, T, \quad h = 0, \dots, H,$$

where H is the smallest integer such that $\|\delta^{h+1}_{mrt} - \delta^{h}_{mrt}\|$ is below a tolerance level of $1e-12$. The value of δ^{H}_{mrt} is taken as the approximation of δ_{mrt}.

Similar results were obtained for the final estimates using these three methods. The simplex and the contraction mapping are substantially more burdensome computationally than Broyden's method. In addition, the contraction mapping did not converge for some values of the parameter space. Although this did not affect the final estimate in my application, it complicates the computation of the hessian, when the objective function cannot be evaluated in neighbor points using the contraction mapping.

To find the expression of $s_{mrt}(\cdot)$, one first need to integrate for the distribution of consumer types in the equation in (5). I approximate this integral by $s_{mrt}(\cdot) = \frac{1}{NS}\sum_{ns=1}^{NS} \mathbb{P}(m_{irt}|p_t)(\nu_{ns})$, where ν_{ns} with $ns = 1, \dots, NS$ are draws from $P_\nu(\nu_{ns}) = \mathcal{N}(0,1)$, a standard normal pdf.

Details about step 3. The conditional value function is found by value function iteration using successive approximations to compute the fixed point of the contraction mapping in (2).[45] For each consumer type, $i = 1, \ldots, I_r$ I solve for $v(p_t, m_{irt-1}, m_{irt})$ by iterating on it as follows:

$$v(p_t, m_{irt-1}, m_{irt})^{(h+1)} = \overline{U}(p_t, m_{irt}, m_{irt-1}) + \\ + \beta \int \log \left(\sum_{m'_{t+1} \in M} \exp \left[v(p_{t+1}, m_{irt}, m_{irt+1})^{(h)} \right] \right) f(p_{t+1}|p_t) \mathrm{d}p_{t+1} + \beta\gamma, \quad \text{(C.2)}$$

$$m_{irt} \in \{0, c, b\}, \quad r = 1, \ldots, R, \quad t = 1, \ldots, T, \quad h = 0, \ldots, H, \quad i = 1, \ldots, I_r,$$

where H is the smallest integer such that $\|v(p_t, m_{irt-1}, m_{irt})^{(h+1)} - v(p_t, m_{irt-1}, m_{irt})^{(h)}\|$ is below a tolerance level of $1e-12$. The value of $v(p_t, m_{irt-1}, m_{irt})^{(H)}$ is taken as the approximation of $v(p_t, m_{irt-1}, m_{irt})$. For the distribution $f(p_t|p_{t-1})$, I use the three specifications described in subsection (3.2), discretizing p_t into $P = 100$ possible values, as described below. Then the integral on the right hand side in (C.2) is approximated by $\frac{1}{P} \sum_{s=1}^{NS_p} \log \left(\sum_{m'_{t+1} \in M} \exp \left[v(p^s_{t+1}, m_{irt}, m_{irt+1})^{(h)} \right] \right) (p^s_{t+1})$, where p^s_{t+1} with $s = 1, \ldots, NS_p$ are draws from $f(p^s_{t+1}|p_{t-1})$, using the specifications in subsection (3.2). The distribution $f(p^s_{t+1}|p_{t-1})$ for the specifications using rational expectations (RE-AR and RE-FE) is estimated outside the estimation algorithm, in a preceding step. For the perfect foresight specification, the value function is solved recursively using the equation in (2) with the actual prices and $v(p_T, m_{irT-1}, m_{irT}) = \overline{U}(p_T, m_{irT}, m_{irT-1})$, and the conditional choice probabilities are solved for all periods applying the contraction mapping to the equation in (4b).

Solving for the conditional value functions in the contraction mapping is computationally burdensome. With 3 choices for the modes of transportation, and assuming we discretize the state vector p_t into $|P|$ values, it involves solving simultaneously a system of $(3 \times 3 \times |P|)$ equations in (2). As noted by Shcherbakov (2016), the problem can be simplified by noting that $v(p_t, m_{irt-1}, m_{irt} = 0) = v(p_t, m_{irt-1} = 0, m_{irt} = 0)$ for all $m_{irt-1} \in M$, because there are no switching costs associated with the outside good; and $v(p_t, m_{irt-1}, \tilde{m}_{irt} \neq m_{irt-1}) = v(p_t, m_{irt-1}, \tilde{m}_{irt} = m_{irt-1}) - \phi_{\tilde{m}_{irt}}$, *i.e.*, the conditional value of choosing mode $\tilde{m}_{irt}$ by switching from m_{irt-1} to $\tilde{m}_{irt} \neq m_{irt-1}$ is the same as the conditional value of choosing mode $\tilde{m}_{irt}$ given that the consumer chose the same mode in the previous period $\tilde{m}_{irt} = m_{irt-1}$, minus the cost of switching from $m_{irt-1} \neq \tilde{m}_{irt}$ to $\tilde{m}_{irt}$. This simplifies the problem from solving a system of $(3 \times 3 \times |P|)$ equations, to one of $(3 \times |P|)$ equations in (2). Thus, when iterating on the contraction mapping in (C.2) I use that $v(p_t, m_{irt-1}, m_{irt} = 0) = v(p_t, m_{irt-1} = 0, m_{irt} = 0)$ for all $m_{irt-1} \in M$, and $v(p_t, m_{irt-1}, \tilde{m}_{irt} \neq m_{irt-1}) = v(p_t, m_{irt-1}, \tilde{m}_{irt} = m_{irt-1}) - \phi_{\tilde{m}_{irt}}$, which simplifies the dimension of the system as explained above.

[45] This is a globally convergent algorithm for computing the fixed point. For a discussion of solution methods see Rust (1996).

Discretization. To perform the estimation, I discretize p_t in $P = 100$ prices. This is done by generating 100 linearly equally spaced points between $\underline{p}$ and $\bar{p}$, where $\underline{p}$ and $\bar{p}$ are the minimum and maximum prices in the grid. The discretized prices are used as state variable in the equation in (2), and to compute the price transition matrix, $f(p_t|p_{t-1})$. In the latter, when a predicted price is different from the discretized prices, the price is replaced by the closest discrete price.

Consumer types. I have access to 28 cores per node and 256 gigabytes of memory per node in the Owens Cluster at the Ohio Supercomputer Center. Thus, I use 28 consumer types in step 3, which corresponds to the maximum number of consumers that I can parallelize per node. For the preferred specification of the model, the estimation was also performed with 83 consumer types, and obtained similar results.[46]

Initial condition. The initial choice of the mode for type i, m_{i0} (initial condition henceforth), is unobserved in the data. I follow Hendel and Nevo (2006), and use the estimated optimal decisions of the consumers to generate an initial distribution as follows: (i) start in 2003 with an *arbitrary* initial condition chosen by a discrete uniform distribution with three possible values; (ii) use the first year in the sample, year 2003, to generate an initial distribution of the initial condition determined by the model conditional on the parameter vector; and (iii) use the remaining years in the data, years 2004 through 2009, and the generated initial condition in (ii) in December 2003 to perform the estimation. The same approach is used for the initial condition needed to compute $\mathbb{P}(m_{irt}|p_t)$ in the equation in (4). Here, for the *arbitrary* initial condition in (i), I use $\mathbb{P}(m_{ir0}|p_t) = S_{mr0}$, the market share of mode m in route r in the first period of data in 2003. I have experimented with different *arbitrary* initial conditions using this procedure, and obtained almost identical results.

Shares of consumers who do not switch. The share of the consumers who did not switch from car and public transit during the $\tilde{t}$ periods prior to t, in route r, denoted respectively by $s_{crt|c\tilde{t}}(\cdot)$ and $s_{brt|b\tilde{t}}(\cdot)$, are computed using the conditional choice probability in (3), and integrating over the distribution of consumers in route r and period t:

$$s_{mrt|m\tilde{t}}(p_t, \delta_{mrt}; \alpha_{0mr}, \Sigma, \phi) = \int \prod_{\hat{t}=t-\tilde{t}+1}^{t} \underbrace{\frac{\exp\left[v(p_{\hat{t}}, m_{ir\hat{t}-1} = m_{ir\hat{t}}, m_{ir\hat{t}})\right]}{\sum_{m'_{\hat{t}} \in M} \exp\left[v(s_{\hat{t}}, m_{ir\hat{t}-1} = m_{ir\hat{t}}, m'_{ir\hat{t}})\right]}}_{\mathbb{P}(m_{ir\hat{t}}|p_{\hat{t}}, m_{ir\hat{t}-1}=m_{ir\hat{t}})} P_\nu(\nu_i)\mathrm{d}\nu_i, \tag{C.3}$$

$$m_{irt} \in \{c, b\}, \quad r = 1, \ldots, R, \quad t = 1, \ldots, T.$$

For the estimation, the conditional value function and choice probabilities in (C.3) are

[46] This number corresponds to the maximum number of consumer types that I can parallelize in the Oakley Cluster at the Ohio Supercomputer Center. I have access to maximum of 7 nodes at Oakley, with 12 cores per node and 48 gigabytes of memory per node, minus 1 core used as the master, leaving a total of $83 = 7 \times 12 - 1$ cores available for the parallelization of the consumer types. Due to the additional computational burden, it is not feasible to perform all the robustness checks with this specification.

computed using the inner subroutine in step 3 as described in subsection 4.1 in the paper. As in the equation in (5), I approximate the integral in (C.3) by:

$$s_{mrt|m\tilde{t}}(\cdot) = \frac{1}{NS}\sum_{ns=1}^{NS}\prod_{\hat{t}=t-\tilde{t}+1}^{t}\mathbb{P}(m_{ir\hat{t}}|p_{\hat{t}}, m_{ir\hat{t}-1} = m_{ir\hat{t}})(\nu_{ns}), \tag{C.4}$$

where ν_{ns} with $ns = 1, \ldots, NS$ are draws from $P_\nu(\nu_{ns}) = \mathcal{N}(0, 1)$, a standard normal pdf. For the estimates in table 3, I use $\tilde{t} = 3$. Table A7 presents robustness analysis for $\tilde{t} = 6$ and $\tilde{t} = 12$. Details about the additional moments are in subsection 2.2. Details about the estimation are in section 4.

Inference. I compute the standard errors for the estimates using the standard procedures (*e.g.* Hansen 1982; Newey and McFadden 1994). I correct the standard errors to account that the simulation draws are the same for all of the observations in a market.

D Definitions of Variables and Fixed Effects

Below are the definitions of the variables and fixed effects used in the empirical analysis.

CTA rail ridership for weekday, Saturdays, or Sundays/Holidays. Monthly average CTA rail ridership for weekdays, Saturdays, or Sundays/Holidays, respectively, disaggregated by rail station. Source: CTA.

CTA bus ridership for weekday, Saturdays, or Sundays/Holidays. Monthly average CTA bus ridership for weekdays, Saturdays, or Sundays/Holidays, respectively, disaggregated by bus route. Source: CTA.

CTA rail fare. CTA standard single rail fare in cents per trip. Source: CTA.

Pace bus ridership for weekday, Saturdays, or Sundays/Holidays. Monthly average Pace bus ridership for weekdays, Saturdays, or Sundays/Holidays, respectively, disaggregated by bus route. Source: Pace.

Metra rail ridership. Monthly Metra rail ridership combined total for weekday, Saturdays, and Sundays/Holidays, disaggregated by branch. Source: Metra.

Vehicle from all ATRs in Illinois. Raw monthly average number of vehicles circulating through each ATR, disaggregated by ATR. ATRs included are all ATRs in the Illinois area. Source: IDOT.

Vehicles from ATRs in Chicago's expressways. Raw monthly average number of vehicles circulating through each ATR, disaggregated by ATR. ATRs included are all ATRs in the Chicago area with expressway indicator equals 1. Source: IDOT.

Vehicles from ATRs in urban areas. Raw monthly average number of vehicles circulating through each ATR, disaggregated by ATR. ATRs included are all ATRs in the Illinois area with Illinois functional classification equals to urban. Source: IDOT.

Vehicles from ATRs in rural areas. Raw monthly average number of vehicles circulating through each ATR, disaggregated by ATR. ATRs included are all ATRs in the Illinois area with Illinois functional classification equals to rural. Source: IDOT.

Class k. Raw monthly average number of vehicles circulating through each ATR. Data are disaggregated by ATR. ATRs included are all ATRs in the Chicago area that collect classification data. Only class k is included, where $k \in 1, \ldots, 13$. Definitions of each class by the FHWA are provided in appendix E. Source: IDOT.

All vehicles Chicago. Raw monthly average number of vehicles circulating through each ATR. Data are disaggregated by ATR. ATRs included are all ATRs in the Chicago area. Source: IDOT.

Passenger. Raw monthly average number of vehicles circulating through each ATR. Data are disaggregated by ATR. ATRs included are all ATRs in the Chicago area that collect classification data. Classes included: 1, 2 and 3. Source: IDOT.

Single unit trucks. Raw monthly average number of vehicles circulating through each ATR. Data are disaggregated by ATR. ATRs included are all ATRs in the Chicago area that collect classification data. Classes included: 4, 5, 6, and 7. Source: IDOT.

Multiple units trucks. Raw monthly average number of vehicles circulating through each ATR. Data are disaggregated by ATR. ATRs included are all ATRs in the Chicago area that collect classification data. Classes included: 8, 9, 10, 11, 12, and 13. Source: IDOT.

Class k. Raw monthly average number of vehicles circulating through each ATR. Data are disaggregated by ATR. ATRs included are all ATRs in the Chicago area that collect classification data. Only class k is included, where $k \in 1, \ldots, 13$. Definitions of each class by the FHWA are provided in appendix E. Source: IDOT.

Direction of the route. Direction of traffic where the ATR is located: E, W, E-W, N, S, N-S. Information about this covariate is collected by certain ATRs. Source: IDOT.

Good months indicator. Number of months that had enough data to pass successful month edits as reported by the IDOT. Source: IDOT.

Illinois functional classification. Illinois functional classification of the roadway where the ATR is located. The functional classes are: Interstate (Urban), Interstate (Rural), Other Principal Arterial (Urban), Other Principal Arterial (Rural), Minor Arterial (Rural), Minor Arterial (Urban), Major Collector (Rural), and Collector (Urban). Information about this covariate is collected by certain ATRs. Source: IDOT.

Number of lanes. Number of lanes in the roadway where the ATR is located. Information about this covariate is available by certain ATRs. Source: IDOT.

Covariates. Covariates include direction of the route, good months indicator, Illinois functional classification of the road, and number of lanes, where these variables are defined as indicated above.

Location. Location of the ATR. Some examples are: I-55 West of Black Lane overpass, IL 143 West of IL 159 (Old Alton Edwardsville Rd.), IL 64 0.9 mile west of il 59, Stevenson Expressway. Source: IDOT.

Urban. Dummy variable equal to one if the ATR is located in an Urban area, and zero otherwise. Source: IDOT.

County. County where the ATR is located. Source: IDOT.

Near city. Nearest city to the ATR. Information about this covariate is available by certain ATRs. Source: IDOT.

Expressway. Dummy variable equal to one if the ATR is located in an Expressway, and zero otherwise. Source: IDOT.

Nominal Gasoline Price in Chicago. Monthly nominal cents per gallon including taxes for Chicago. It corresponds to the monthly Chicago, all grades, all formulations retail gasoline prices. Series ID: MG_TT_C2. Source: U.S. Energy Information Administration.

Nominal Gasoline Price in the U.S. Monthly nominal cents per gallon including taxes for the U.S. It corresponds to the monthly unleaded regular gasoline for U.S. city average retail price. Source: U.S. Energy Information Administration.

Nominal Oil Price (wtotworld). Monthly average oil price in nominal dollars per barrel, measure 1. It refers to the monthly average of the weekly all countries spot price FOB weighted by estimated export volume. Series ID: WTOTWORLD. Source: U.S. Energy Information Administration.

Nominal Oil Price (wtotusa). Monthly average oil price in nominal dollars per barrel, measure 2. It refers to the monthly average of the weekly United States spot price FOB weighted by estimated import volume. Series ID: WTOTUSA. Source: U.S. Energy Information Administration.

CPI U.S. Consumer Price Index for all urban consumers. Area: U.S. city average. Not seasonally adjusted, 1982-84=100. Series ID: CUUS0000SA0. Source: U.S. Bureau of Labor Statistics.

CPI U.S. Gas. Consumer Price Index for gasoline, all types. Area: U.S. city average. Not seasonally adjusted, 1982-84=100. Series ID: CUUS0000SETB01. Source: U.S. Bureau of Labor Statistics. Used for in robustness analysis instead of CPI U.S.

CPI Chicago. Consumer Price Index for all urban consumers. Area: Chicago-Gary-Kenosha, IL-IN-WI. Not seasonally adjusted, 1982-84=100. Series ID: CUUSA207SA0. Source: U.S. Bureau of Labor Statistics.

CPI Chicago Gas. Consumer Price Index for Gasoline, all types. Area: Chicago-Gary-Kenosha, IL-IN-WI. Not seasonally adjusted, 1982-84=100. Series ID: CUURA207SETB01. Source: U.S. Bureau of Labor Statistics. Used for in robustness analysis instead of CPI Chicago.

Gasoline price (Gasoline price in the U.S). Real gasoline price in Chicago (in the U.S.) obtained by dividing the Nominal Gasoline Price in Chicago (in the U.S.) by the CPI Chicago (CPI U.S.).

Oil price, measures 1 and 2. Real crude oil price obtained by dividing Nominal Oil Price measures 1 or 2 by CPI U.S.

Monthly seasonal fixed effects. A set of 11 dummy variables, each corresponding to the month when the variables are registered.

Unit of observation fixed effects. For CTA rail ridership, a set of dummy variables, each corresponding to the CTA station where the ridership is registered. For CTA and Pace bus ridership, a set of dummy variables, each corresponding respectively to the CTA or Pace bus route for which the ridership is registered. For Metra rail ridership, a set of dummy variables, each corresponding to the Metra branch where the ridership is registered.

Unit specific month trend. For CTA rail ridership, a set of monthly time trends, one for each CTA station where the ridership is registered. For CTA and Pace bus ridership, a set of monthly time trends, one respectively for each CTA or Pace bus route for which the ridership is registered. For Metra ridership, a set of monthly time trends, one for each Metra branch where the ridership is registered.

Year fixed effects. A set of years dummy variables, each corresponding to the year when the variables are registered.

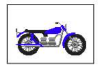
ATR specific fixed effects. A set of dummy variables, each corresponding to ATRs where the vehicle counts are registered.

ATR specific month trend. A set of monthly time trends, one for each ATR where the vehicle counts are registered.

E Federal Highway Administration Vehicle Classifications

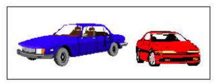
This subsection is obtained directly from the classification by Federal Highway Administration (FHWA). The FHWA classification scheme is separated into categories depending on whether the vehicle carries passengers or commodities. Non-passenger vehicles are subdivided by number of axles and number of units including power, and trailer units.

Class 1 – ***Motorcycles:*** All two or three-wheeled motorized vehicles. This category includes motorcycles, motor scooters, mopeds, motor-powered bicycles, and three-wheel motorcycles.



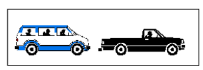
Class 1. Source: FHWA Vehicle Classification Figures.

Class 2 – ***Passenger Cars:*** All sedans, coupes, and station wagons manufactured primarily for the purpose of carrying passengers and including those passenger cars pulling recreational or other light trailers.



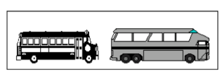
Class 2. Source: FHWA Vehicle Classification Figures.

Class 3 – ***Other Two-Axle, Four-Tire Single Unit Vehicles:*** All two-axle, four-tire, vehicles, other than passenger cars. Included in this classification are pickups, panels, vans, and other vehicles such as campers, motor homes, ambulances, hearses, carryalls, and minibuses. Other two-axle, four-tire single-unit vehicles pulling recreational or other light trailers are included in this classification. Because automatic vehicle classifiers have difficulty distinguishing class 3 from class 2, these two classes may be combined into class 2.



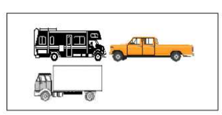
Class 3. Source: FHWA Vehicle Classification Figures.

Class 4 – ***Buses:*** All vehicles manufactured as traditional passenger-carrying buses with two axles and six tires or three or more axles. This category includes only traditional buses (including school buses) functioning as passenger-carrying vehicles. Modified buses should be considered to be a truck and should be appropriately classified.



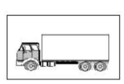
Class 4. Source: FHWA Vehicle Classification Figures.

Class 5 – ***Two-Axle, Six-Tire, Single-Unit Trucks:*** All vehicles on a single frame including trucks, camping and recreational vehicles, motor homes, etc., with two axles and dual rear wheels.



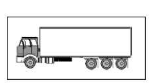
Class 5. Source: FHWA Vehicle Classification Figures.

Class 6 – ***Three-Axle Single-Unit Trucks:*** All vehicles on a single frame including trucks, camping and recreational vehicles, motor homes, etc., with three axles.



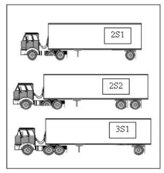
Class 6. Source: FHWA Vehicle Classification Figures.

Class 7 – ***Four or More Axle Single-Unit Trucks:*** All trucks on a single frame with four or more axles.



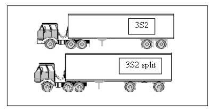
Class 7. Source: FHWA Vehicle Classification Figures.

Class 8 – ***Four or Fewer Axle Single-Trailer Trucks:*** All vehicles with four or fewer axles consisting of two units, one of which is a tractor or straight truck power unit.



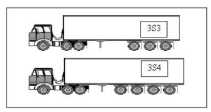
Class 8. Source: FHWA Vehicle Classification Figures.

Class 9 – ***Five-Axle Single-Trailer Trucks:*** All five-axle vehicles consisting of two units, one of which is a tractor or straight truck power unit.



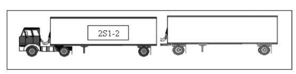
Class 9. Source: FHWA Vehicle Classification Figures.

Class 10 – ***Six or More Axle Single-Trailer Trucks:*** All vehicles with six or more axles consisting of two units, one of which is a tractor or straight truck power unit.



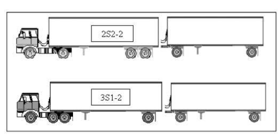
Class 10. Source: FHWA Vehicle Classification Figures.

Class 11 – ***Five or fewer Axle Multi-Trailer Trucks:*** All vehicles with five or fewer axles consisting of three or more units, one of which is a tractor or straight truck power unit.



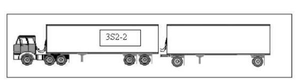
Class 11. Source: FHWA Vehicle Classification Figures.

Class 12 – ***Six-Axle Multi-Trailer Trucks:*** All six-axle vehicles consisting of three or more units, one of which is a tractor or straight truck power unit.



Class 12. Source: FHWA Vehicle Classification Figures.

Class 13 – ***Seven or More Axle Multi-Trailer Trucks:*** All vehicles with seven or more axles consisting of three or more units, one of which is a tractor or straight truck power unit.



Class 13. Source: FHWA Vehicle Classification Figures.

F Robustness of Estimates

F.1 Alternative Specifications of Expectations

Table A6: Robustness. Structural Estimates using Alternative Specifications of Expectations.

Panel A: Dynamic Model with Rational Expectations RE-FE.

	with switching costs				without switching costs			
	with random coefficients		without random coefficients		with random coefficients		without random coefficients	
	Estimates	95% C.I.	Estimates	95% C.I.	Estimates	95% C.I.	Estimates	95% C.I.
α_1	-0.634	(-1.249, -0.019)	-0.552	(-1.077, -0.027)	-0.685	(-0.786, -0.585)	-0.587	(-0.679, -0.494)
Σ	0.163	(0, 0.395)	–	–	0.774	(0.624, 0.924)	–	–
ϕ_b	4.468	(3.854, 5.082)	2.065	(1.295, 2.836)	–	–	–	–
ϕ_c	1.716	(0.887, 2.545)	3.289	(2.456, 4.122)	–	–	–	–
Value of GMM Objective	48.237		50.71		6.377		37.533	

Panel B: Dynamic Model with Perfect Foresight PF.

	with switching costs			
	with random coefficients		without random coefficients	
	Estimates	95% C.I.	Estimates	95% C.I.
α_1	-0.589	(-1.114, -0.064)	-0.583	(-0.706, -0.460)
Σ	0.0065	(0, 0.515)	–	–
ϕ_b	0.785	(0.344, 1.226)	0.173	(0.035, 0.311)
ϕ_c	1.209	(0.605, 1.813)	1.719	(1.314, 2.123)
Value of GMM Objective	147.829		163.542	

Notes: Estimates of selected parameters from the structural model. See section 2 for details about the data used in the estimation. Descriptions of the models are in section 3. Details about the estimation procedure are in section 4. See subsection 5.1 for details about the specifications of the models in the different panels. The term 95% C.I., denotes 95 percent confidence intervals and are provided in parenthesis. The specifications in this table do not include additional moments with the share of the consumers who did not switch from car and public transit.

F.2 Alternative Specifications of Micro Moments

Table A7: Robustness. Structural Estimates using Alternative Specifications of Micro Moments and Expectations.

	Estimates	95% C.I.	Estimates	95% C.I.	Estimates	95% C.I.
Panel A: Dynamic Model with Rational Expectations RE-AR and Additional Moments.						
	$\tilde{t} = 3$ months		$\tilde{t} = 6$ months		$\tilde{t} = 12$ months	
α_1	-0.526	(-0.657, -0.395)	-0.492	(-1.553, -0.568)	-0.458	(-1.001, -0.085)
Σ	0.308	(0.041, 0.576)	0.194	(0, 0.574)	0.275	(0, 0.687)
ϕ_b	2.142	(1.550, 2.734)	2.912	(1.053, 4.771)	3.075	(1.787, 4.362)
ϕ_c	2.224	(1.842, 2.605)	2.731	(1.471, 3.991)	3.142	(2.186, 4.098)
Value of GMM Objective	93.742		47.421		29.949	
Panel B: Dynamic Model with Rational Expectations RE-FE and Additional Moments.						
	$\tilde{t} = 3$ months		$\tilde{t} = 6$ months		$\tilde{t} = 12$ months	
α_1	-0.628	(-0.753, -0.503)	-0.591	(-0.717, -0.465)	-0.521	(-0.791, -0.439)
Σ	0.179	(0, 0.332)	0.193	(0, 0.343)	0.207	(0, 403)
ϕ_b	2.577	(2.114, 3.040)	2.97	(2.438, 3.501)	3.932	(1.811, 4.287)
ϕ_c	2.048	(1.623, 2.473)	2.467	(1.842, 3.091)	3.781	(2.209, 4.104)
Value of GMM Objective	102.281		54.889		28.024	

Notes: Estimates of selected parameters from the structural model. See section 2 for details about the data used in the estimation. A description of the model is in section 3. Details about the estimation procedure are in section 4. See subsection 5.1 for details about the specifications of the models in the different panels. The term 95% C.I., denotes 95 percent confidence intervals and are provided in parenthesis. The additional moments included in this panel correspond to the share of the consumers who did not switch from car and public transit during the $\tilde{t} \in \{3, 6, 12\}$ periods prior to t, in route r, denoted respectively by $s_{brt|b\tilde{t}}(p_t, \delta_{mrt}; \alpha_{0mr}, \Sigma, \phi)$ and $s_{crt|c\tilde{t}}(p_t, \delta_{mrt}; \alpha_{0mr}, \Sigma, \phi)$. These additional moments (no-switching moments) are computed using the conditional value function and choice probabilities in (C.3), implemented in the inner subroutine in step 3 as described in subsection 4.1 in the paper. See appendix C for details.