Measuring Long-Run Gasoline Price Elasticities in Urban Travel Demand

Donna, Javier D.

University of Florida

10 May 2021

Measuring Long-Run Gasoline Price Elasticities in Urban Travel Demand

Javier D. Donna*

May 10, 2021.

Abstract

I develop a structural model of urban travel to estimate long-run gasoline price elasticities. I model the demand for transportation services using a dynamic discrete-choice model with switching costs and estimate it using a panel dataset with public market-level data on automobile and public transit use in Chicago. Long-run own- (automobile) and cross- (transit) price elasticities are substantially more elastic than short-run elasticities. Elasticity estimates from static and myopic models are downward biased. I use the estimated model to evaluate the response to several counterfactual policies. A gasoline tax is less regressive after accounting for the long-run substitution behavior.

JEL Codes: H22, H25, L43, L71, L91, L98.

Keywords: Long-run gasoline price elasticities, Dynamic demand, Switching Costs, Hysteresis, Consumer Inertia, Gasoline Tax Incidence.

---

*University of Florida, jdonna@ufl.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. Thank you to Daniel A. Donna for helping me access the transportation data to start the project and to Doris M. D. Grillo González for her continuous encouragement to finish this long-run project. I especially thank Igal Hendel, Tiago Pires, Yun Pu, and Greg Veramendi for their many and helpful suggestions. I thank Niu Chiao and Yun Pu for outstanding research assistance. I also thank the editor, Marc Rysman, and three anonymous referees for insightful comments. Discussions with John Asker, Jason Blevins, Jean-Pierre Dubé, Cristian Huse, Ryan Kellogg, Matt Lewis, Cynthia Lin Lawell, Charles Manski, James Roberts, Mark Roberts, Bill Rogerson, Oleksandr Shcherbakov, Will Strange, Jurre Thiel, André Trindade, and Matt Weinberg, as well as seminar participants at the International Industrial Organization Conference (2019), Consumer Search and Switching Costs Workshop (UCLA Anderson School of Management, Marketing Department, 2019), Barcelona GSE Summer Forum (Applied Industrial Organization, 2019), Stanford Institute for Theoretical Economics (SITE, Session 3: Empirical Implementation Models of Strategic Interaction and Dynamic Behavior Workshop, 2019), Aarhus University, Johns Hopkins University, Kentucky University, Northwestern University, Northwestern University Transportation Center (NUTC), Penn State University, Queen's University, Ryerson University, the Ohio State University, Universidad Carlos III, University of Florida, University of Illinois at Urbana-Champaign, University of Toronto (Rotman School of Management) have benefited this article. I would also like to express my gratitude to Rob Robinson, chief of the planning section at the Illinois Department of Transportation (IDOT), and to Catherine Kannenberg, head of systems performance from the Metra division. A version of this article was part of my Ph.D. dissertation at Northwestern University. This article 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 and the allocation of computing time from the Ohio Supercomputer Center (grant PAS1350-2). The analyses in this article, based on data from IDOT, are by the author and not by IDOT. First version: December 2010.

1 Introduction

Long-run gasoline price elasticities are a key input to evaluate the impact of transportation policies such as the incidence of gasoline taxes, investments in public transportation capacity, and zoning laws. There is a remarkably large number of empirical studies 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] In this article I develop a structural dynamic model of urban travel, estimate the model using public market-level data on automobile and public transit use in Chicago, and use the estimated model to compute long-run price elasticities and evaluate policy interventions in this market.

The transportation sector plays a central role in modern economies. Private automobiles are the main mode of transportation in the U.S. (U.S. Department of Transportation 2011). They have a direct impact on the demand for gasoline and traffic congestion. They are the largest source of greenhouse gases. 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). This development shifted the attention of policymakers toward obtaining a better understanding of how transit ridership responds to gasoline prices. The ability to effectively measure the impact of regulatory policies relies on estimates of long-run elasticities of automobile and transit use to the gasoline cost of driving.

I develop a structural model of urban travel to estimate these long-run elasticities. 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 by nesting the demand system within a dynamic framework. Long-run elasticities are obtained by using the estimated model to compute changes in market shares over longer periods of time for a permanent gasoline price change. I compute short-run elasticities separately for a temporary and a permanent gasoline price increase. To compute the elasticities for a permanent price increase, I evaluate the short- and long-run responses to a change in the whole path of gasoline prices, not just the current price. I do this to illustrate the effect of consumers' expectations on elasticities when considering permanent price changes (*e.g.*, an announced gasoline tax) relative to shorter run changes (*e.g.*, a gas/oil shock). To compute the elasticities, I use the estimated model to simulate changes in the market shares over shorter and longer periods, as indicated in Section 5.

In the model, consumers are rational, forward-looking, and have persistent heterogeneous preferences over modes of transportation. A consumer may incur a switching cost if the mode chosen in the current period differs from that chosen 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

[1] I define switching costs below. I discuss their presence in my setting in Section 2.

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

There is a longstanding debate regarding the role of hysteresis and switching costs in determining 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 given gasoline price increase 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 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 without specifying a model of conduct (*e.g.*, increasing *vs.* decreasing prices), or by modeling the state dependence/inertia without estimating the structural parameters of the dynamic decision process.[5] These approaches may suffer from limitations when attempting to account for factors affecting long-run substitution patterns, such as changes in expectations or heterogeneity in consumers' preferences. Another difficulty may arise in evaluating counterfactual policies. For example, under an announced gasoline tax increase, consumers may internalize the permanent price increase in their expectations, thus affecting long-run elasticities and the distributional impacts of the gasoline tax increase. Evaluating policies designed to alter factors causing hysteresis requires uncovering the model's 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. Public transit ridership data were obtained for the three public transit operators in the Chicago area. Vehicle circulation data were obtained from the Illinois Department of Transportation (IDOT). IDOT uses automatic traffic recorders (ATRs) to measure the number of vehicles

[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 automobiles, switching costs include the cost of obtaining information about routes, searching for parking facilities, the costs of regular car use, *etc.*

[3] In the context of travel behavior in this article, hysteresis is defined as the change in the market share of a mode of transportation that persists after the reversal of the initial variation in the gasoline cost of driving that gave rise to it. See Section 2 for details.

[4] The survey by Dargay (1993) provides evidence about these asymmetries. It 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 footnote 16.

circulating at ATR 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 Planning Package. I complement the market-level data on car and public transit use with a cross-sectional 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 selected households in the Chicago metropolitan area. It asked about households' travel patterns, income, and attitudes towards everyday commuting. I use the data from this survey to define additional micro moments used for the estimation and to investigate the distributional consequences of a gasoline tax increase. I focus on the period from 2003 to 2009 for the estimation of the model. This period is characterized by substantial variation in the gasoline prices in the U.S. It induced several consumer types to switch modes of transportation. Such variation is useful for the estimation of the model.

To identify switching costs I rely on instruments 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 a consumer's decision in the previous period while holding the decision in the current period constant. I use the previous period's 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 assumption is common in the literature (see, *e.g.*, Busse, Knittel, and Zettelmeyer 2013 in the context of car purchases).[6] Second, I assume that the consumers' heterogeneous preferences are constant over time. This assumption allows me to separate the effect on choices resulting from the heterogeneity in preferences and the hysteresis due to the market friction generated by the switching costs.

Three features from the empirical setting are useful for the estimation of the switching cost and heterogeneity parameters. The first is the use of information about the share of consumers who do not switch (micro moments). These additional moments provide a more precise estimation of the switching costs and heterogeneity parameters (as in, *e.g.*, Petrin 2002; Berry, Levinsohn, and Pakes 2004; and Goeree 2008). The second is the frequent and substantial variation in the gasoline cost of driving during the period under analysis. For example, the large increase in gasoline prices in 2008 induced certain types of consumers to switch to public transit. The third is 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.

I report three main findings from the estimated model. First, long-run price elasticities with respect to the gasoline cost of driving are substantially more elastic than the short-run elasticities.[7] Long-run own- (for automobile) and cross- (for public transit) price elasticities are, respectively, 61 and 67 (55 and 61) percent higher in absolute value than short-run temporal (permanent) elasticities. The gasoline price elasticities for urban travel are inelastic in the short run but elastic in the long run. Measuring the long-run response is relevant for most applications. Second, static own- and cross-price elasticities underestimate

[6]See Anderson, Kellogg, Sallee, and Curtin (2011) and Anderson, Kellogg, and Sallee (2013) for detailed discussions. Kellogg (2018) studies the implications of gasoline price volatility and uncertainty over future prices for fuel economy policies. Lin and Prince (2013) study how gasoline price volatility impacts the gasoline price elasticity.

[7]See Section 5 for definitions of these elasticities.

long-run elasticities by 65 and 70 percent, respectively.[8] The myopic model also underestimates the long-run elasticities. The bias for the myopic model is larger than for the static model, 81 and 93 percent for own- and cross-price elasticities, respectively. Finally, I investigate the response to a gasoline tax increase 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 increase. The bias is larger for the myopic model and for low-income consumers. The former result is because the static model predicts higher probabilities of switching than the myopic model due to the static model not accounting for the switching costs. The latter result is because low-income consumers are more likely to switch from a car to public transit in response to the gasoline tax than middle- and high-income consumers. That is, there is substantial long-run heterogeneity among consumers with different incomes: low-income consumers are more price sensitive in the long run and therefore more responsive to a gasoline tax increase. Overall, these results indicate that when hysteresis is present, elasticities derived from static and myopic models might be biased estimates of long-run elasticities. The gasoline tax is still regressive, but it is less regressive after the gasoline tax increase due to the dynamics and the long-run substitution behavior.

In summary, this article 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 are 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 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 substantially more elastic than short-run elasticities and that the gasoline tax is less regressive after accounting for the longer-run substitution behavior.

The rest of the article is organized as follows. Section 2 describes the industry and the data, and presents stylized facts about the former. 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 presents the concluding remarks. 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.[9] Such studies were interested not only in understanding the market response to changes

---

[8] The static model refers to a model without switching costs, where consumers only value the current period utility. The myopic model refers to a model with switching costs, where consumers face dynamic incentives and have myopic expectations (they believe that the current gasoline price will prevail forever). See Section 5 for details.

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

in gasoline prices but also as policy tools to evaluate the impact of regulations.[10]

There is 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 of 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 prices. Hughes, Knittel, and Sperling (2008) estimate average per capita demand for gasoline in the U.S. for two periods of high gasoline prices, from 1975 to 1980 and 2001 to 2006. 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 articles 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 article. 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. 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.[11] 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 the U.S. (*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). Tyndall (2017) uses the reduction in public transit access in New York City due to Hurricane Sandy to investigate the effect of public transportation on employment outcomes. Horowitz

[10] 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. As regards alternative tax policies, 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. See Grigolon, Reynaert, and Verboven (2017) and Miravete, Moral, and Thurk (2018) for recent studies investigating the welfare effects of alternative tariff policies.

[11] 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).

(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 consumer panel to estimate household motor vehicle choices.[12] There have also been studies analyzing traffic volumes using data from ATRs similar to the data in this article (*e.g.*, Horowitz and Emslie 1978; Hogema 1996; Lingras 2000).[13] 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 use, using aggregate-level 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).[14] Dargay (1993) discusses the impact of hysteresis on elasticities and provides a survey.[15] 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 articles have developed and estimated dynamic models to conceptualize such inertia.[16] 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 articles estimates the structural parameters characterizing the dynamic optimization problem of the agents. As emphasized by Cantillo, Ortúzar,

[12] 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).

[13] This article 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. Severen and Van Benthem (2019) document that commuters in the U.S. who experience a substantial shock to the price of gasoline while coming of driving age are less likely to drive to work and more likely to use public transit later in life. Asker, Collard-Wexler, and De Loecker (2017) quantify the misallocation attributable to market power in the global oil industry.

[14] 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.

[15] 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).

[16] Frameworks in the style of Heckman (1981) have been broadly used in the literature to analyze dynamic choices and/or to 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 article 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 discussions.

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 article 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 next.

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; Donna and Espin-Sanchez 2021),[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; MacKay and Remer 2019),[18] and hysteresis in dynamic choice models (*e.g.*, Roberts and Tybout 1997; Clerides, Lach, and Tybout 1998; Das, Roberts, and Tybout 2007; Aw, Roberts, and Xu 2011).[19,20] 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 difference 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 in my setting are different from the ones used in these articles, 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 Hendel and Nevo (2006), Perrone (2017), and Donna and Espin-Sanchez (2021).

2 Urban Transportation in Chicago

I describe the urban transportation sector in Chicago and the dataset, and perform the preliminary analysis describing empirical regularities about the sector.

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

[18] See Luco (2019) for a recent application to pension plan choice. See Klemperer (1995) and Farrell and Klemperer (2007) for surveys.

[19] 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). Hysteresis has also been extensively studied in unemployment since the work by Blanchard and Summers (1986).

[20] 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.

2.1 Urban Transportation Sector 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 cars (drive alone) and 12 percent for public transit (bus/streetcar and subway/rail) for the entire 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 transportation system in the northeastern Illinois area surrounding Chicago.

The Illinois Department of Transportation (IDOT) collects vehicle circulation information. It uses automatic traffic recorders (ATRs) that measure the number of vehicles circulating in specific locations. Seven major expressways cross 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. To estimate the structural model in Section 4, I link vehicle circulation on each of these expressways to public transit use around the locations of the ATRs, as described in the 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 a 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 the fare cost of public transit obtained from the three public transit operators supervised by the RTA. Vehicle circulation counts and vehicle classification were obtained from IDOT. The gasoline cost of driving and world crude oil prices were 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 a car or public transit (*i.e.*, consumers who do not switch) were obtained from a travel survey by Pace. For the estimation of the structural model, I focus on the period from 2003 to 2009.[21] During this period, ride-sharing was not present in Chicago. Below, I briefly describe each source. Summary statistics of selected variables are in Table 1. See Appendix B for details.

Monthly ridership data and the 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

[21] 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 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..

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 weekdays, Saturdays, and Sundays/Holidays.

Vehicle circulation is collected by IDOT using ATR stations. The following characteristics about the road where an ATR is located are collected: the functional classification of the roadway by 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. IDOT maintained a network of 85 ATR locations throughout the state during 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 B for details on the ATR district distribution, its distribution by functional class, and the vehicle type classification.

The real gasoline cost of driving and real-world crude oil prices are constructed using price data from the U.S. Energy Information Administration. The consumer price index (CPI) is obtained from the U.S. Bureau of Labor Statistics. For the gasoline cost of driving, I use the monthly city average retail price of all grades of all formulations of gasoline prices for Chicago and the monthly national average of the retail price of unleaded regular gasoline. I use two series of world crude oil prices as instruments in the empirical analysis: the weekly all countries spot price FOB weighted by estimated export volume and the 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 world crude oil prices. Almost identical results are obtained when using other measures of inflation. To obtain 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 the series are closer together in some months and farther apart in others, reflecting cross-sectional/regional variation across stations, the series generally track each other well.

I complement the data above with a cross-sectional 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 selected households in the Chicago metropolitan area. I use it to describe the behavior of consumer types who continue using the same inside mode, without switching. The survey provides information about the joint distribution of households' travel patterns, income, and attitudes towards everyday commuting, among other issues. 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 household types with similar travel attitudes, such as households who are auto-dependent, transit-captive, transit-friendly, or households whose commuting patterns vary considerably. I use these types to define the shares of consumers who do not switch from cars or public transit and to investigate the distributional consequences of the gasoline tax increase. See Appendix B for details.

I use the data described above to build a dataset of the inside modes of transportation used (cars 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 on the expressways (interstate highways). As mentioned above, there are seven major expressways (henceforth, routes). 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 on each route by defining a radius around the locations of the ATRs. Figure 2 provides an example for Dwight D. Eisenhower Expressway (Interstate 290). For each route and month, I compute public transit use as total weekday ridership (for each CTA rail station, CTA bus line, and Pace bus line) that lies within the radius.[22] 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 that from the Census Transportation Planning Package (the share of people whose work trip mode is neither automobile nor public transit). The outside mode can be conceptualized as including other modes of transportation, such as biking or walking. It may also include choices to change where the consumer lives/works or to change the route, holding the market size constant, which may also contribute to a more elastic price elasticity.[23] When estimating the structural model, I seasonally adjust the data using the X-13ARIMA-SEATS seasonal adjustment software from the U.S. Census Bureau (U.S. Census Bureau 2017) using its automatic model selection procedure.[24]

An observation in this dataset represents the market share of cars, public transit, or the outside mode in a given month and route, the characteristics of the mode, and the share of consumers who use only cars or public transit. See Appendix B 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 2 and 3 display the results. They display, respectively, the results of 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. The full OLS specification for ridership or traffic volume in station/route/ATR s in month m and year y is:

$$\ln(d_{smy}) = x_s\alpha + \beta \ln(p_{cmy}) + \tau_s + \tau_m + \tau_{ms} + \tau_y + \epsilon_{smy},$$

$$s = 1, \ldots, S, \quad m = 1, \ldots, M, \quad y = 1, \ldots, Y,$$

[22]Bus ridership is only disaggregated by line (not by stops on the line). I include a complete bus line when 70 percent or more of it lies within the radius. I obtained similar results using thresholds of 50 and 90 percent.

[23]As explained below, the model holds the market size constant. Over a period of, *e.g.*, five years, the stock of cars that a family owns can be altered; over a period of, *e.g.*, a decade, the stock of housing can be altered. Consumers' decisions to change where they live/work or to change their cars, that may affect the market size, are held constant in the model. In that sense, this article makes a distinction between a short-run elasticity and a longer-run elasticity thus defined.

[24]Similar results were obtained by seasonally adjusting the data by running regressions on month-of-year dummy variables. I also performed a robustness analysis using the month of the year as an additional state variable instead of using seasonally adjusted data (monthly seasonal fixed effects for static models) with 40 Halton sequence consumer draws and obtained similar results. Tables A7, A8, and A9 in the appendix display the results. Including the month of the year as a state variable enormously increases the computational burden, thus making it unfeasible to estimate all specifications using the importance sampling procedure described in the appendix.

where $\ln(d_{smy})$ denotes the natural logarithm of the ridership or vehicle counts in station/route/ATR s in month m in year y; x_s is a K-dimensional row vector of observable characteristics for station/route/ATR s which include a vector of ones; $\ln(p_{cmy})$ is the natural logarithm of the real gasoline cost of driving in Chicago in month m and year y, which is the variable of primary interest; τ_s are station, route, or ATR fixed effects (depending on the data structure); τ_m are monthly seasonal fixed effects; τ_{ms} are station-, route-, or ATR-specific monthly trends; τ_y are year fixed effects; 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. 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. Tables 2 and 3 also display two-stage least squares (henceforth, IV) 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 2 and 3 show the following results. (1) Public transit use increases as the gasoline cost of driving increases and vice versa. The substitution patterns are consistent across different public transportation types (CTA rail, CTA bus, Pace bus, and Metra rail) and specifications (with and without covariates, and with and without fixed effects), stronger on weekdays than on 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.[25] 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 than in 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 a 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 than in rural areas. See Appendix C for details and additional results.

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

Table 4 investigates heterogeneity in the response to the gasoline cost of driving for CTA rail ridership. I distinguish between the response to rising and falling prices, and when the absolute gasoline price increases above certain thresholds.[26,27] I present two types of heterogeneity results. The first concerns 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. The second concerns whether the response to the gasoline cost of driving differs by the absolute level of the gasoline price. Panels A and B in Table 4 show results from the following OLS specifications, respectively:

[25] Although the number of individuals in the car is unobserved, if consumers respond to an increase in the gasoline price by carpooling, the number of vehicles would decrease, which is captured by the vehicle count variable.

[26] See Dargay (1993) for a discussion.

[27] The table focuses on CTA rail for weekdays, the main public transportation used in the routes considered in the structural estimation. Similar patterns are obtained for CTA and Pace buses, and for vehicle circulation in the Chicago area. This evidence is consistent with the one in Section 2. See also Table 5.

$$\ln(d_{smy}) = x_s\alpha + \beta \ln(p_{cmy}) \times [\mathbf{1}\{p_{cmy} > p_{cm-1y} > p_{cm-2y}\} + \mathbf{1}\{p_{cm} < p_{cm-1y} < p_{cm-2y}\} + \mathbf{1}\{\text{else}\}] + \tau_s + \tau_m + \epsilon_{smy},$$

$$\ln(d_{smy}) = x_s\alpha + \beta \ln(p_{cmy}) \times [\mathbf{1}\{P_{cmy} < 150\} + \mathbf{1}\{150 \leq P_{cmy} < 250\} + \mathbf{1}\{250 \leq P_{cmy} < 350\} + \mathbf{1}\{350 \leq P_{cmy}\}] + \tau_s + \tau_m + \epsilon_{smy},$$

$$s = 1, \ldots, S, \quad m = 1, \ldots, M, \quad y = 1, \ldots, Y,$$

where $\ln(d_{smy})$ denotes the natural logarithm of the ridership or vehicle counts in station/route/ATR s in month t; $\ln(p_{cmy})$ denotes the natural logarithm of the real gasoline cost of driving in month m in year y; the notation $\mathbf{1}\{\text{A}\}$ denotes an indicator function that equals 1 if condition A is satisfied, and 0 otherwise; the condition *else* means that neither $\mathbf{1}\{p_{cmy} > p_{cm-1y} > p_{cm-2y}\}$ or $\mathbf{1}\{p_{c,m} < p_{cm-1y} < p_{cm-2y}\}$ hold; and the absolute level of gasoline price is denoted with uppercase, P_{cmy}.

On the one hand, Table 4 shows that the elasticity estimates are three times higher (than the baseline estimation in row 5 of Panel A, Table 2) when the gasoline cost of driving has 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 has been flat or declining and 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 (the top number in Figure 1B).[28] 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 5 tests for the presence of hysteresis. It focuses on a specific shock, Hurricane Katrina, which was by far the largest refining shock affecting gasoline prices in the U.S (Kilian 2010).[29] In this manner, the initial substitution between cars and public transit may not merely reflect heterogeneity in consumers' preferences. Table 5, 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 ATR stations in the Chicago metropolitan area with the largest vehicle counts on the routes used in the structural analysis. Column 1 displays the pre-shock levels (the month before 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, by 11.22 percent (weighted mean). Vehicle use also decreases considerably, by 5.86 percent (weighted mean). Nevertheless, once the shock disappears and the gasoline cost of driving returns to the pre-shock level, public transit and vehicle use do not return to their initial values. Public transportation decreases by only 3.44 percent, and vehicle use increases by only 1.69 percent. Consumers do not switch back and continue using public transportation for a longer period of time, thus

[28] The elasticities reported in Figure 1B correspond to the breakdown of the baseline specification in Table 2, Panel A, row 1, for different 12-month periods.

[29] Hurricane Katrina formed on August 23, 2005, and dissipated on 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 in the supply of gasoline in the U.S., producing a large increase in the price of gasoline in the U.S.

generating asymmetry over time. This shows evidence of hysteresis in public transportation ridership and vehicle use in the Chicago area. Table 5, Panel B repeats the analysis for the most popular ATR 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 to the increase after the shock. There is no hysteresis in rural areas without access to public transit.[30,31] The presence of hysteresis in the Chicago metropolitan area and its absence in rural areas are suggestive of a structural relationship between current and past modes of transportation choices in areas where consumers can substitute between cars and public transit. I present next an econometric model of the optimal choice of modes of transportation, where the presence of a market friction, switching costs, can recreate the documented hysteresis.

3 A Model for Urban Modes of Transportation

I model the demand for modes of transportation using a random coefficient mixed logit dynamic model. Consumers have persistent idiosyncratic preferences over the different modes of transportation. They incur a switching cost when they choose a mode of transportation different from that chosen in the previous period.

3.1 Set Up

Assume that there are $r = 1, \ldots, R$ routes, each with a continuum of rational, utility-maximizing consumers indexed by i. Denote by t the discrete-time periods measured in months. Consumers differ in their preferences over modes of transportation. A market is defined as a period-route as described in the previous section. In each period-route, each consumer chooses one mode of transportation among car, public transit, or the outside mode. Let $M = \{c, b, 0\}$ be the set of choices that consumers face in each period-route denoting car, public transit, and the outside mode of transportation, respectively. Denote by m_{irt} the mode of transportation chosen by consumer i 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.[32] Consumers have common discount 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 that 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 they switch to the outside mode. Consumers do incur a switching cost, however, if they 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.

[30] Hysteresis may also be driven by better relocation opportunities in urban areas. Note, however, that the hysteresis documented here focuses on the response to exogenous (from the perspective of the consumer) shocks, such as Hurricane Katrina, that affected short-run gasoline prices. During these shocks, consumers' housing decisions (*e.g.*, the decision to rent/buy a house/apartment closer to the workplace) are arguably fixed. Thus, I believe that these factors are not the main force driving the documented hysteresis.

[31] Similar results to those in Table 5 are obtained for the other oil shocks reported in Figure 1B. Table A3 in the appendix shows the gasoline price elasticities during Hurricane Katrina; see also the discussion there. Table A4 repeats Table 5 adjusting for seasonality. Similar evidence of hysteresis in urban areas is observed, with no hysteresis in rural areas without access to public transit.

[32] Following the Census Transportation Planning Package, I define the outside option as the share of people whose work trip mode is neither auto nor public transit. See Section 2 for details.

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

$$U_{imrt} = \alpha_{0mr} - \alpha_{1i}p_{mt} - \phi_m \mathbf{1}\{m_{irt-1} \neq m_{irt}\} + \xi_{mrt} + \epsilon_{imrt}, \tag{1}$$

where p_{mt} with $m \in \{c, b\}$ denote, respectively, the real gasoline cost of driving and the fare cost of public transit in period t;[33] m_{irt} denotes the mode of transportation chosen by the consumer i in route r in period t; ξ_{mrt} is the valuation of unobserved, by the econometrician, characteristics of inside mode m_{irt} in route r in period t;[34] 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, described below, that capture consumers' preferences over price; ϕ_m with $m \in \{c, b\}$ denotes the switching-cost parameter to inside mode m; and $\mathbf{1}\{\cdot\}$ is an indicator function.

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 over the gasoline cost of driving and fare cost of public transit as $\alpha_{1i} = \alpha_1 + \sigma\nu_i$. The parameter α_1 captures the mean price sensitivity. The parameter σ governs the distribution of the random coefficients (heterogeneity in consumer preferences over price), ν_i, which are drawn from the distribution $\mathcal{P}_\nu(\nu_i)$ specified below. The preferences of consumers for the inside modes of transportation in this setting are captured by their idiosyncratic preferences, switching costs, monetary costs, and unobserved characteristics. Let $\phi \equiv (\phi_c, \phi_b)$ and $\alpha \equiv (\alpha_{0mr}, \alpha_1)$. Let $\theta \equiv (\alpha, \sigma, \phi)$ be the vector of structural parameters to be estimated. Denote by $\delta_{mrt} = \alpha_{0mr} - \alpha_1 p_{mt} + \xi_{mrt}$, the mean utility for inside mode m in route r in period t; that is, the portion of the utility that is constant across types of consumers net of the switching cost. Using α_{1i} and δ_{mrt} rewrite the flow utility in (1) as:

$$U_{imrt} = \delta_{mrt} - \phi_m \mathbf{1}\{m_{irt-1} \neq m_{irt}\} - \sigma\nu_i p_{mt} + \epsilon_{imrt}. \tag{2}$$

The flow utility in (2) is expressed as a sum of three terms. The first term, δ_{mrt}, is the mean utility common to all consumers. The second term, $-\phi_m \mathbf{1}\{\cdot\}$, captures the effect of the switching costs. The last two terms, $-\sigma\nu_i p_{mt} + \epsilon_{imrt}$, represent an idiosyncratic deviation from the mean utility net of the switching cost, capturing the effects of the random coefficients and additive shocks. Finally, let $\overline{U}_{imrt} \equiv U_{imrt} - \epsilon_{imrt}$ be the utility function net of the additive utility shock.

I abstract from consumer heterogeneity in the trip distance (unobserved vehicle-miles and transit-miles traveled in my data) and assume that average monthly trip distance is constant for consumers. Although the average monthly distance is constant, the monthly trip distance is not necessarily constant over time. The latter is my preferred interpretation.[35] I abstract from changes in the decisions of consumers to change cars and where they live/work that may affect the market size. One can conceptualize the

[33] 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.

[34] Unobserved characteristics can include, for example, the impact of unobserved factors affecting the quality of the mode or systematic shocks to demand. In principle, some components of the unobserved characteristics can be captured by dummy variables as in static discrete-choice models. For example, one can model $\xi_{mrt} = \xi_m + \xi_r + \xi_t + \Delta\xi_{mrt}$ and capture ξ_m, ξ_r and ξ_t by mode-, route-, and month-specific dummy variables. In practice, introducing these variables as state variables increases the computational burden substantially.

[35] See Ben-Akiva and Lerman (1985), Small, Verhoef, and Lindsey (2007), Button (2010), and the references therein for discussions of this issue using aggregate-level data. See 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).

outside option as including choices to live in the neighborhood of the workplace or to continue using the same model of transportation and changing the route holding constant the market size. These factors contribute to a more elastic long-run demand. Three comments are worth mentioning. First, that the hysteresis documented in the previous section focuses on the response to exogenous shocks (from the perspective of the consumer), such as Hurricane Katrina. These shocks affect short-run gasoline prices. During these shocks, consumers' housing/work/car decisions are arguably fixed. Second, I focus on the consumers who work/live/commute around the five vehicle routes described in the article (instead of focusing on the overall Chicago area). Because these routes extend along rail lines (see, *e.g.*, Figure 2), they have relatively better access to public transit. Thus, even if some consumers were to change where they work/live/commute, the share/distribution of consumers who would continue using these routes/rail lines is relatively stable for the analysis considered. Finally, holding constant the share/distribution of consumers who would continue using these routes/rail lines seems reasonable to compute the 12-month gasoline price elasticities in this article. For these reasons holding constant the market size due to decisions to change the car and where consumers live/work seems reasonable in my context.

Switching costs are defined as in Klemperer (1987a,b) as constant across periods and known to consumers. This specification assumes that the *start-up cost* (initial fixed cost of choosing the inside mode for the first time) is the same as the *switching cost* (fixed cost of subsequently switching to the inside mode after the first time). This simplification is due to the nature of the data and the industry. First, because my data is aggregated at the market level, I do not directly observe the decisions of consumers to switch. When a consumer switches, I cannot tell how many times the consumer switched in the past. Second, in the setting of urban travel in my article, consumers do not switch frequently inside modes of transportation. This feature makes it difficult to obtain precise estimates of both start-up and switching costs (or, more generally, switching-cost parameters that vary over time) using aggregate-level data.[36,37]

3.2 State Variables and Value Function

Consumer i has the following state variables.

Gasoline cost of driving and fare cost of public transit. Consumers observe the gasoline cost of driving 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) boundedly rational expectations using an autoregressive model of order 1; (ii) boundedly rational expectations using a frequency estimator;[38] (iii) perfect foresight; and (iv) myopic expectations. The fare cost of public transit, p_{bt}, is constant across routes and exhibits little variation during the period under analysis due to regulation; I assume consumers know p_{bt}. I comment on these assumptions below. Denote by $\vec{p}_t = (p_{ct}, p_{bt})$. Let us call $\mathcal{P}_{p_c}(p_{ct}|p_{ct-1})$ and $f_{p_c}(p_{ct}|p_{ct-1})$ the cdf and pdf

[36] One can still estimate different start-up and switching costs by making assumptions regarding the initial condition. Table A10 in the appendix show estimates of the main specification of the model allowing for start-up costs to differ from subsequent switching costs. For the initial condition, I use the same procedure described in Appendix A. I obtained similar results to the ones reported in Table 6. The switching costs are less precisely estimated in Table A10 than in Table 6, Panel C due to the second feature mentioned above.

[37] See, *e.g.*, Button (2010) for a general discussion. Figure 3A displays the estimated mean probabilities in the empirical application in this article.

[38] Specifications (i) and (ii) are *boundedly rational* in that they approximate consumers' beliefs about gasoline price as a function of its own lagged values and disregarding other potentially relevant information. See below for a discussion.

of p_{ct}, respectively. I estimate the model using the following specifications for the transition of p_{ct}:

(i) Boundedly rational expectations using an autoregressive model (BRE-AR): Under boundedly rational expectations using an autoregressive model, an AR(1) process is estimated assuming $p_{ct+1} = \gamma_0 + \gamma_1 p_{ct} + \iota_{ct+1}$, 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_{ct+1} is generated from p_{ct} and the estimated parameters.

(ii) Boundedly rational expectations with discretized prices using a frequency estimator (BRE-FE): Under boundedly rational expectations using a frequency estimator, a first-order Markov process is estimated for the conditional distribution, $f_{p_c}(p_{ct}|p_{ct-1})$, using a frequency estimator.

(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.[39]

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

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.

Mean utility levels. The mean utility of inside mode m is given by $\delta_{mrt} = \alpha_{0mr} - \alpha_1 p_{mt} + \xi_{mrt}$. I make two assumptions for the estimation. First, I assume that consumers have myopic or static expectations regarding ξ_{mrt}; that is, consumers believe that the current value of ξ_{mrt} will prevail in the future (no-change forecast for its future values). Second, I assume that prices, $\vec{p}_t$, are exogenous. I comment on these assumptions below. Denote by $\vec{\delta}_{rt} = (\delta_{crt}, \delta_{brt})$ the vector of mean utilities for the inside modes.[40]

Additive utility shock. 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 $g(\epsilon_{imrt})$.

Idiosyncratic preference for prices. Consumer i time-invariant idiosyncratic preference for prices is denoted by the random variable ν_i. I assume that consumers' types are drawn *i.i.d.* across individuals from a distribution $\mathcal{P}_\nu(\nu_i)$, assumed to be a standardized normal, $\mathcal{N}(0,1)$, for the estimation.

The assumptions on the gasoline cost of driving and mean utility levels merit further discussion. For gasoline prices, I test different assumptions regarding consumers' beliefs about the price process. I do this for two reasons. The first is to investigate the role of consumers' expectations on the estimates. The second is because predicting changes in gasoline and crude oil 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 specification has the convenient implication that the current gasoline price is the expected future real gasoline price.

[39] In the perfect foresight model, the decision in the last period, T, is static. Thus, the conditional value function at T is just the flow utility. The conditional value function at period t is computed by backwards recursion, similarly to finite-horizon models.

[40] It is equivalent to use $\vec{\xi}_{rt}$ instead of $\vec{\delta}_{rt}$ as a state variable because, conditional on prices, there is a one-to-one mapping between these two variables under the assumptions in the article.

Additionally, 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.*, consumers believe that gasoline prices follow a martingale process). They also find deviations from the no-change forecast during the financial crisis of 2008. Anderson, Kellogg, Sallee, and Curtin (2011) find that the forecast accuracy of the predictions of the Michigan Survey of Consumers is similar to that of a no-change forecast, on average. An alternative approach to the one I follow above 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 suggested by the boundedly rational expectations assumption, without assuming perfect foresight, by using the information on crude oil markets or other macroeconomic variables to make projections regarding future gasoline prices. Another approach would be to use a variable that directly represents gasoline price expectations. These approaches would require incorporating these additional variables, used to model expectations month by month, as state variables, thereby increasing the computational burden. If under alternative approaches, consumers' expectations lie between (iii) and (iv), one can use the current estimates under those expectations as a benchmark.

For the mean utility levels, I make two assumptions. First, the assumption on expectations is that consumers believe that the other characteristics of the modes will not change; that is, consumers cannot perfectly forecast the mean utility.[41] The validity of this assumption is testable. It can be examined empirically by comparing the realized value and assumed expectation of the consumers. I find that it is reasonable in the context of urban travel analyzed in this article because there were no major disruptions in the quality of the modes (*e.g.*, route/station construction, rail remodeling, closures). Models incorporating inertia in travel choice typically do not account for the role of consumers' expectations about future characteristics, including price. Given my focus on price elasticities, I model prices (and the mean utility) as state variables and test various specifications that differ in consumer expectations of prices, as discussed in the previous paragraph. An alternative approach would be to specify only the mean utility as a state variable (instead of prices and the mean utility) and test specifications that differ in its expectations and evolution (*e.g.*, inclusive-value sufficiency).[42] The assumption that prices are conditionally exogenous seems reasonable in the context of urban travel; for example, it is satisfied if gasoline prices follow a martingale, a standard assumption in the literature as discussed above. These assumptions about the mean utility levels simplify the computational burden of solving the contraction mapping of the consumer defined in equation (4) below. The expectations assumption can be relaxed to a certain extent. The conditional independence assumption is crucial. It makes it possible to obtain consistent estimates of the price coefficient, α_1, even if consumers are incorrect in their expectations about the other characteristics.

Value function. The value function for consumer type i is given by:

[41] This assumption rules out, *e.g.*, potential network effects, whereby a consumer may take into consideration how the aggregate behavior of other consumers may affect the quality of the modes. With additional information about traffic congestion for both inside modes one may incorporate network effects into the analysis, subject to the complications mentioned in this paragraph.

[42] See, *e.g.*, Hendel and Nevo (2006), Schiraldi (2011), Gowrisankaran and Rysman (2012), Melnikov (2013), and Shcherbakov (2016) for different versions of the inclusive-value sufficiency assumption.

$$V_i(\vec{p}_t, m_{irt-1}, \vec{\delta}_{rt}, \epsilon_{imrt}, \nu_i; \theta) = \underset{m_{irt} \in M}{max} \{\delta_{mrt} - \phi_m \mathbf{1}\{m_{irt-1} \neq m_{irt}\} - \sigma \nu_i p_{mt} + \epsilon_{imrt} + \\ + \beta \, \mathbb{E}[V_i(\vec{p}_{t+1}, m_{irt}, \vec{\delta}_{rt+1}, \epsilon_{imrt+1}, \nu_i; \theta | \vec{p}_t, m_{irt-1}, \vec{\delta}_{rt}, \epsilon_{imrt})]\}, \tag{3}$$

subject to the evolution of the state variables as described above.

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 $\vec{p}_t$ and unconditional on the previous choice, m_{irt-1}. 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. Following Rust (1987, 1994), 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 $v_i(\vec{p}_t, m_{irt-1}, m_{irt}, \vec{\delta}_{rt}, \nu_i; \theta)$, is given by:

$$v_i(\vec{p}_t, m_{irt-1}, m_{irt}, \vec{\delta}_{rt}, \nu_i; \theta) = \delta_{mrt} - \phi_m \mathbf{1}\{m_{irt-1} \neq m_{irt}\} - \sigma \nu_i p_{mt} + \\ + \beta \int \log \left(\sum_{m_{irt+1} \in M} \exp \left[v_i(\vec{p}_{t+1}, m_{irt}, m_{irt+1}, \vec{\delta}_{rt+1}, \nu_i; \theta) \right] \right) \mathrm{d}\mathcal{P}(\vec{p}_{t+1}, \vec{\delta}_{rt+1} | \vec{p}_t, \vec{\delta}_{rt}) + \beta\gamma, \tag{4}$$

where $\mathcal{P}(\cdot)$ denotes the distribution function; the equality follows from the independence of the ϵ_{imrt} shocks replacing $\int V_i(p_{t+1}, m_{irt}, \delta_{mrt+1}, \epsilon_{imrt+1}, \nu_i; \theta) g(\epsilon_{imrt+1}) \mathrm{d}\epsilon_{imrt+1}$, by its closed-form expression using the well-known properties of the Gumbel distribution; and $\gamma = 0.577$ is the Euler's constant.

Using the assumptions regarding the independence of prices and expectations:

$$v_i(\vec{p}_t, m_{irt-1}, m_{irt}, \vec{\delta}_{rt}, \nu_i; \theta) = \delta_{mrt} - \phi_m \mathbf{1}\{m_{irt-1} \neq m_{irt}\} - \sigma \nu_i p_{mt} + \\ + \beta \int \log \left(\sum_{m_{irt+1} \in M} \exp \left[v_i(\vec{p}_{t+1}, m_{irt}, m_{irt+1}, \vec{\delta}_{rt+1}, \nu_i; \theta) \right] \right) f_{p_c}(p_{ct+1} | p_{ct}) \mathrm{d}p_{ct+1} + \beta\gamma, \tag{5}$$

where the equality follows because the mean utility levels in period $t+1$ are a known deterministic function of the given prices.

From equations (4) and (5) it can be seen the role of the assumption about the independence of prices. The expectations' assumption about the mean utility can be relaxed using, *e.g.*, different specifications for $\mathcal{P}_\delta(\cdot)$, analogously as I do for the expectations of the gasoline cost of driving. I test for different specifications of price expectations in this article given my focus on price elasticities.

Conditional choice probabilities. The probability of observing the choice m_{irt} for type ν_i conditional on $\vec{p}_t$, $\vec{\delta}_{rt}$, and m_{irt-1}, denoted by $\mathbb{P}_i(m_{irt}, \nu_i | \vec{p}_t, m_{irt-1}, \vec{\delta}_{rt}; \theta)$, is:

$$\begin{aligned}
\mathbb{P}_i(m_{irt}, \nu_i | \vec{p}_t, m_{irt-1}, \vec{\delta}_{rt}; \theta) = \mathbb{P}[&v_i(\vec{p}_t, m_{irt-1}, m_{irt}, \vec{\delta}_{rt}, \nu_i; \theta) + \epsilon_{imrt} \geq \\
&v_i(\vec{p}_t, m_{irt-1}, m'_{irt}, \vec{\delta}_{rt}, \nu_i; \theta) + \epsilon_{im'rt}, \forall m'_{irt} \neq m_{irt}], \\
= &\frac{\exp\left[v_i(\vec{p}_t, m_{irt-1}, m_{irt}, \vec{\delta}_{rt}, \nu_i; \theta)\right]}{\sum_{m'_{irt} \in M} \exp\left[v_i(\vec{p}_t, m_{irt-1}, m'_{irt}, \vec{\delta}_{rt}, \nu_i; \theta)\right]},
\end{aligned} \tag{6}$$

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

Using the law of total probability, write $\mathbb{P}_i(m_{irt}, \nu_i | \vec{p}_t, \vec{\delta}_{rt}; \theta)$ as:

$$\mathbb{P}_i(m_{irt}, \nu_i | \vec{p}_t, \vec{\delta}_{rt}; \theta) = \sum_{m'_{irt-1} \in M} \mathbb{P}_i(m_{irt}, \nu_i | \vec{p}_t, m'_{irt-1}, \vec{\delta}_{rt}; \theta) \times \mathbb{P}_i(m'_{irt-1}, \nu_i | \vec{p}_t, \vec{\delta}_{rt}; \theta), \tag{7a}$$

$$= \sum_{m'_{irt-1} \in M} \underbrace{\frac{\exp\left[v_i(\vec{p}_t, m'_{irt-1}, m_{irt}, \vec{\delta}_{rt}, \nu_i; \theta)\right]}{\sum_{\tilde{m}_{irt} \in M} \exp\left[v_i(s_t, m'_{irt-1}, \tilde{m}_{irt}, \vec{\delta}_{rt}, \nu_i; \theta)\right]}}_{\mathbb{P}_i(m_{irt}, \nu_i | \vec{p}_t, m'_{irt-1}, \vec{\delta}_{rt}; \theta)} \times \mathbb{P}_i(m'_{irt-1}, \nu_i | \vec{p}_t, \vec{\delta}_{rt}; \theta), \tag{7b}$$

where the equality in (7b) follows from equation (6).

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

$$s_{mrt}(\vec{p}_t, \vec{\delta}_{rt}; \theta) = \int \mathbb{P}_i(m_{irt}, \nu_i | \vec{p}_t, \vec{\delta}_{rt}; \theta) \ P_\nu(\nu_i) \mathrm{d}\nu_i, \tag{8}$$

where $\mathbb{P}_i(m_{irt}, \nu_i | \vec{p}_t, \vec{\delta}_{rt}; \theta)$ is given by equation (7b) with the $v_i(\vec{p}_t, m_{irt-1}, m_{irt}, \vec{\delta}_{rt}, \nu_i; \theta)$ given by equation (5); and $P_\nu(\nu_i)$ is the pdf of a 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. I set the discount factor at $\beta = 0.99$.

4.1 Estimation

I estimate the model 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 search 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 as inputs 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:

$$\hat{\theta} = \arg\min_{\theta} \left[\xi(\theta)' Z W^{-1} Z' \xi(\theta) \right], \tag{9}$$

where $\xi(\theta)$ is the vector of structural errors; and W is a positive definite weighting matrix chosen with a standard two-step approach described in Appendix A. To evaluate the GMM objective one needs to first solve numerically for $\xi(\theta)$, which 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 equation (8) with the choice probability in equation (7b) to compute the market shares as a function of the parameters. I define the error term as the unobserved products' characteristics and compute it by solving for the mean utility level, $\vec{\delta}_{rt}$, that equates:

$$s_{mrt}(\vec{p}_t, \vec{\delta}_{rt}; \theta) = S_{mrt}, \tag{10}$$

where $s_{mrt}(\cdot)$ is the market share function given by equation (8); and S_{mrt} are the observed market shares obtained from the data. I solve for $\delta_{mrt}(\theta)$ in the system of equations in (10) using Broyden's method for finding roots.[43]

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 the conditional value function and conditional choice probabilities. These computations are performed in the next subroutine.

3. Conditional value function and choice probabilities. For each candidate parameter vector, the conditional choice probabilities are computed using equation (7b). To evaluate the expression in (7b), one first needs to compute the conditional value function, $v_i(\vec{p}_t, m_{irt-1}, m_{irt}, \vec{\delta}_{rt}, \nu_i; \theta)$. The computation is done by solving numerically the stochastic control problem (henceforth, contraction mapping) in (5) for each given candidate parameter vector, as discussed in Appendix A. To solve this contraction mapping, one needs to specify the distribution characterizing consumers' expectations about the gasoline cost of driving, denoted by $f_{p_c}(p_{ct}|p_{ct-1})$ in (5). In the next section, I report the estimation results using the four specifications for $f_{p_c}(p_{ct}|p_{ct-1})$ described in Section 3.

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 A for details about the computation of the weighting matrix in step 1, the procedure

[43] As in Gowrisankaran and Rysman (2012) an important issue that arises is whether this system of equations has a unique fixed point. 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 nonlinear system of equations (10): (i) a simplex search method and (ii) 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 A for details.

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 $\vec{p}_t$, 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 rely on instruments to identify the price coefficient, the heterogeneity parameters, and the switching-cost parameters. Identification requires at least one instrument for price, the heterogeneity parameter, and each switching-cost parameter. Identification of the model also requires a unique fixed point for the contraction mapping in (4).

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 the switching costs, ϕ. This issue is well known (*e.g.*, see Heckman 1981; Dubé, Hitsch, and Rossi 2010; Sudhir and Yang 2014; and the references therein). Consider, for example, a consumer who chose a 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 a strong idiosyncratic preference for a car, a low α_i.[44] 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, $v_i(\vec{p}_t, m_{irt-1} = c, m_{irt} = c, \vec{\delta}_{rt}, \nu_i; \theta) > v_i(\vec{p}_t, m_{irt-1} = c, m_{irt} = b, \vec{\delta}_{rt}, \nu_i; \theta) - \phi_b$. 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 and 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 consumer's decision in the previous period to identify the switching costs using the decisions in the current period. Following the intuition in Table 5 (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). A market share of cars in $t+1$ lower than the market share of cars in $t-1$ 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 $\mathbb{E}[Z' \cdot \xi(\theta)] = 0$, and use as instruments the international price of oil in the previous period and the average gasoline cost of driving in Chicago in the previous three periods.

I assume that the consumers' preferences over 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. This assumption allows me 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

[44] In my model a random coefficient on the gasoline cost of driving is similar to a random coefficient on a car dummy variable.

documented in section 2 with idiosyncratic preferences that are constant over time and without switching costs.[45] 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.[46]

In terms of empirical variation, three features from the empirical setting are useful for the estimation of the switching cost and heterogeneity parameters. The first is the use of information about the share of consumers who did not switch. These data, described in Section 2 and Appendix B, provide a more precise estimation of the switching costs and heterogeneity parameters. For example, consider the information from observing the market share of consumers who chose to use a 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 estimate more precisely the heterogeneity in consumer preferences, σ, because its value has to be such that there are sufficient types of consumers, α_i, who have a strong idiosyncratic preference for car; that is, a low α_i. It also provides a more precise estimation of 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, provides a more precise estimation of the switching costs to the car. (See also Figure 3B discussed in Section 5.) I use these data to define additional micro moments (henceforth, no-switching moments). The no-switching moments are defined as the difference between the share of consumers who do not switch from the 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 equation (9). The additional computational burden does not increase substantially. See Appendix A for details about the computation of the shares of consumers who do not switch. In the next section, I show that the no-switching moments improve the precision of the estimates of σ and ϕ.

The second is the frequent and substantial variation in the gasoline cost of driving during the period under analysis. Consider, for example, the large and rapid increase in gasoline prices in 2008 as depicted in Figure 1. Certain types of consumers switch to public transit. In the following periods, the gasoline cost of driving decreased substantially. A subset of those types, the ones with higher α_i (weaker preference for car), continue using public transit on subsequent periods, even after the price returns to the previous pre-shock level. This feature helps improve the precision of the switching costs estimates because such hysteresis can only be generated from the switching costs in the model.

The third is 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 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 next.

I use world crude oil price as instruments for the gasoline cost of driving in Chicago to address potential

[45]In principle, one could use a more general specification and/or allow for serial correlation over time in the unobserved consumer preferences, ν_{it}. Such specifications would increase the computational burden. For my application, with data aggregated at the market level, additional assumptions may be needed to estimate serially correlated preferences.

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

concerns about the gasoline cost of driving in Chicago responding endogenously to the car and public transit use in Chicago.[47] The identifying assumption is that monthly- route-specific car and public transit demand shocks are uncorrelated with monthly gasoline prices conditional on the mode of transportation, route, and seasonal fixed effects. The gasoline cost of driving in Chicago is correlated with the world crude oil price across months because world crude oil price is the main determinant of the marginal cost of gasoline. They are uncorrelated with month-specific valuations due to the exclusion restriction.

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 approach was used by Dubé, Hitsch, and Rossi (2010).[48]

5 Results

5.1 Parameter Estimates

I present the estimated parameters in Table 6 using several specifications of the model. The first is a static model without switching costs, where $\phi = 0$ and consumers value only the current period utility. I present two specifications of this model in Panel A, with and without random coefficients. These specifications are equivalent to a standard and mixed logit model, respectively, which have dominated the prior literature. The second is a model with myopic expectations (ME), where consumers face switching costs but do not disregard dynamic incentives. I implement this model using the dynamic model in Section 3 with switching costs and myopic expectations. I also present two specifications of this model in Panel A, with and without random coefficients. The third is a dynamic model, as outlined in Section 3, with boundedly rational expectations using an autoregressive model for consumers' expectations (BRE-AR) and without additional no-switching moments (micro moments). For this model, I present four specifications in Panel B, with and without switching costs, and with and without random coefficients. The fourth is the same as the third except that it has additional no-switching moments. This specification is my preferred specification and is in Panel C.

I apply the estimation algorithm from Section 4 to each model in Table 6 with obvious modifications. For example, the specification with a static model without switching costs and without random coefficients corresponds to a standard discrete-choice logit model. Step 3 can therefore be skipped because the conditional value function is 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 expectations. In such a model, a price increase

[47] 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 capture this cyclical effect on ridership/vehicle use. Another source could be that gasoline local taxes might be correlated with economic cycles, therefore biassing the estimates.

[48] Dubé, Hitsch, and Rossi (2010) also fit models with more flexible specifications for the unobserved heterogeneity, such as a mixture of normals. Using such specification increases the number of parameters to be estimated. In my setting, it introduces challenges due to the nature of my data, which is aggregated at the market level. Dubé, Hitsch, and Rossi (2010) use household panel data.

in the current period generates two incentives. On the one hand, because of switching costs, consumers are less likely to switch. This feature 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 the price increase to be, the more likely they are to switch. This feature is due to the effect of future prices on the continuation value. In Panel A, for the model with myopic expectations, consumers believe that the current price increase in the gasoline cost of driving will prevail forever. It 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 the price, σ.

Second, from Panel B, one can see the role of switching costs for rationalizing the observed hysteresis. The specifications of the models in Panel B hold consumer expectations constant. They are modeled as boundedly rational expectations using an autoregressive model (BRE-AR). As expected, the estimated base price coefficient is smaller in the dynamic model with switching costs than in the dynamic model without switching costs. Again, the two incentives from the previous paragraph are present. By holding constant the effect of expectations, one can "isolate" the first effect, that switching costs make 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 Section 5. Finally, the estimated mean price coefficient from the static model (right of Panel A) is larger than that from the dynamic model with switching costs and boundedly rational expectations. It is because the static model rationalizes consumers not switching using the price coefficient (α_1) while the dynamic model also uses the switching costs (ϕ_b and ϕ_c). Thus, a larger value of 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 obtain more precise estimates of the switching-cost 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 result is similar to Petrin (2002), Berry, Levinsohn, and Pakes (2004), and Goeree (2008).

Figure 3 shows the conditional choice probabilities of the inside modes of transportation and the switching patterns of a particular consumer type. Figure 3A shows the mean, across consumer types and markets, conditional choice probabilities (left vertical axis) given by equation (7) as a function of the observed gasoline cost of driving (right vertical axis), p_{ct}. The estimated conditional choice probabilities respond to the variations in the gasoline cost of driving, consistent with the substitution patterns described in Tables 2 and 3. It can also be seen that the large price variation in 2008 (the 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 to obtain precise estimates of the switching-cost parameters.

Figure 3B shows the behavior of a particular transit-friendly consumer type.[49] Let $v_i^{jr} \equiv v_i(\vec{p}_t, m_{irt-1} = j, m_{irt} = r, \vec{\delta}_{rt}, \nu_i; \theta)$, where the superscripts j and r index the mode of transportation of consumer i in $t-1$ and t, respectively, and the conditional value function $v_i(\vec{p}_t, m_{irt-1}, m_{irt}, \vec{\delta}_{rt}, \nu_i; \theta)$ is given by equation (5). The figure shows on the left vertical axis using dots the difference in the conditional value function of choosing car transport today given the choice to travel by car in the previous period, minus the conditional value function of choosing public transit today given a previous-period choice of public transit for this particular transit-friendly consumer type: $v_i^{cc} - v_i^{bb}$. The vertical dotted lines indicate a switch in the inside mode of transportation by the consumer in that period. The graph also displays the gasoline cost of driving on the right vertical axis using a solid line and the estimated switching-cost parameters, ϕ_c and $-\phi_b$, using horizontal lines. 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 exceeds the value of the current mode, the difference is less than the switching cost.[50] The consumer switches from a car to public transit if the difference in conditional value functions, $v_i^{cc} - v_i^{bb}$, crosses the bottom horizontal line, $-\phi_b$, from above (*e.g.*, vertical line in 2008m4).[51] Similarly, the consumer switches from public transit to a car if the difference in conditional value functions, $v_i^{cc} = v_i^{bb}$, crosses the top horizontal line, ϕ_c, from below (*e.g.*, last vertical line in 2008m10).[52] 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. When the switching costs are zero, $\phi_c = \phi_b = 0$, the band of inaction vanishes and the consumer switches when the difference in conditional value functions, $v_i^{cc} = v_i^{bb}$, crosses the horizontal line at zero. That is, the asymmetry created by the switching costs disappears. The hysteresis therefore vanishes. The graph shows that the response to gasoline prices is more complex than that suggested by static or dynamic models without switching costs.

5.2 Robustness

I tested the robustness of the model in several ways. First, in Table 6 I tested for consistency between the theoretical and estimated predictions of different specifications of the model. Second, also in Table 6, for each specification in the previous subsection I present the estimates with and without random coefficients, where the latter is implemented by having one representative consumer-type with $\alpha_i = \alpha$. Third, I tested for different specifications of the expectations of the agents in the appendix in Table A5. There, I present the estimates using different specifications for the price transition matrix, $f_{p_c}(p_{ct}|p_{ct-1})$, representing consumers' expectations: (i) boundedly rational expectations using an autoregressive model

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

[50] When start-up costs differ from the switching costs incurred in subsequent periods, as in Table A10, the band of inaction in Figure 3B decrease over time if the consumer switches multiple times to the same inside mode.

[51] In the graph, the consumer using a car in the current period switches to public transit only if the value of continue using the car minus the value of switching is negative: $v_i^{cc} - v_i^{cb} < 0$. However, $v_i^{cc} - v_i^{bb} < -\phi_b \implies v_i^{cc} - v_i^{cb} < 0$ because $v_i^{bb} = v_i^{cb} + \phi_b$. That is, the conditional value of choosing mode $\tilde{m}_{irt}$ by switching from m_{irt-1} (with $\tilde{m}_{irt} \neq m_{irt-1}$) 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} to $\tilde{m}_{irt}$ (with $m_{irt-1} \neq \tilde{m}_{irt}$). See Appendix A for details.

[52] Figure A1 in the appendix displays the behavior of a different transit-friendly consumer type who switches more often using the estimated model with month-of-the-year as an additional state variable.

(BRE-AR), where consumers predict the gasoline cost of driving using an autoregressive model of order 1, presented in Panel B in Table 6; (ii) boundedly rational expectations using a frequency estimator (BRE-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 A5; (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 A5; and (iv) myopic expectations (ME) as described in the second model above, presented in Panel A in Table 6. The estimation algorithm from Section 4 is applied to each specification of the price transition matrix, $f_{p_c}(p_{ct}|p_{ct-1})$. Fourth, I tested for different specifications of the additional micro moments in the appendix, in Table A6. The table shows the estimates of the dynamic model using different combinations of boundedly rational expectations (BRE-AR and BRE-FE) and additional moments. For the additional moments, I use the share of consumers who do not switch for 3-, 6-, and 12-month periods. Fifth, I performed a robustness analysis of all the previous specifications using the month of the year as an additional state variable instead of using seasonally adjusted data (monthly seasonal fixed effects for static models) in Tables A7, A8, and A9. Sixth, in Table A10 in the appendix, I performed a robustness analysis allowing for initial start-up costs to differ from the switching costs incurred in subsequent periods, as discussed in Section 3. Seventh, I tested for robustness during the Great Recession. The econometric model I presented presumes the absence of income effects when consumers are making their choice of the mode of transportation. To assess the validity of this premise, I estimated separate models for subsamples including and not including the months during the Great Recession. The robustness analysis presented in Tables A11 and A12 reveals remarkable consistency in the estimates and ratios of elasticities discussed in the next subsection.[53] Finally, in Appendix A, I describe additional robustness analyses related to the estimation procedure. The estimated parameters sometimes varied across some of these robustness tests. However, the implications discussed in the next subsections are robust in the cases examined.

5.3 Elasticities and Implications

I present the implications of the estimates and compare them to static and myopic estimates. I define own- and cross-price elasticities as the car and public transit response to a change in the gasoline cost of driving, respectively. I define short- and long-run elasticities as the change in the market share over 1 and 12 months, respectively. I compute short-run elasticities separately for a temporal and a permanent gasoline price increase (henceforth, short-run temporal and short-run permanent elasticities). I compute long-run elasticities for a permanent gasoline price increase (henceforth, long-run elasticities).[54] Below I describe the procedure to compute these elasticities.

To compute the elasticities, 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 a change in the gasoline cost of driving by adding a small amount to the observed path of the gasoline price. To evaluate the temporal price response, I change only the current gasoline cost of driving.

[53] Although the Great Recession provided an important source of variation in gasoline prices, the implications discussed in the article do not seem to be driven by the Great Recession nor the resulting income effects. I believe that the main reason for this result is that the assumptions discussed in Section 3 seem to be satisfied using the definitions of the elasticities in the empirical setting studied.

[54] In principle, one can also compute long-run elasticities for a temporal gasoline price increase. In my setting, however, this elasticity is very small in absolute value: in the long run, consumers respond very little to a temporal gasoline price increase today because they are forward-looking.

To evaluate the permanent price response, I change 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. To evaluate the short-run price response, I consider the change in market share over 1 month. To evaluate the long-run price response, I consider the change in market share over 12 months. For all models, the elasticities are computed by calculating the percent change in market share in step 4 relative to the initial values in step 1, with a one 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. The next table reports such an average by market and the average across markets, denoted by the mean.

Table 7 presents the average ratio of elasticities computed from different models relative to the long-run elasticity computed from the dynamic model. Three patterns emerge. First, the long-run elasticities from the dynamic model are substantially larger in absolute value than the short-run elasticities. Long-run own- (for automobile) and cross- (for public transit) price elasticities are, respectively, 61 and 67 (55 and 61) percent higher in absolute value than short-run temporal (permanent) elasticities.[55] That is, the long-run elasticities are substantially more elastic than short-run elasticities. The difference between the short- and long-run elasticities has important consequences for the evaluation of policies, as previously noted in the literature and investigated in the next subsections. Second, the estimates show that myopic own- and cross-price elasticities are smaller in absolute value than long-run elasticities. Static own- and cross-price elasticities underestimate the long-run elasticities by 65 and 70 percent, respectively. The myopic model also underestimates the long-run elasticities. The bias is larger for the myopic model than for the static model, 81 and 93 percent for own- and cross-price elasticities, respectively. The latter result is because the static model predicts higher probabilities of switching than the myopic model due to the static model not accounting for the switching costs. The bias varies across routes (markets). For the myopic model, the bias is up to 83 percent for own-price elasticity and up to 95 percent for cross-price elasticity for route 5. Finally, there is substantial heterogeneity in the consumers' long-run response to the gasoline price. Lower-income consumers are more likely to switch in response to a permanent increase in gasoline prices than middle- and high-income consumers. This effect has implications for the distributional consequences of a gasoline tax discussed below.

Table 8 displays the mean own- and cross-price elasticities from reduced-form and structural models together with the ratios relative to the long-run elasticity. Long-run own- (automobile) and cross- (transit) price elasticities are substantially more elastic than short-run elasticities. The magnitudes of the elasticities are sensible. The gasoline price elasticities for urban travel are inelastic in the short run but elastic in the long run. After proper stratification, the reduced-form elasticities are similar to the ones reported in the literature. The reduced-form elasticities are comparable to the ones from the static model. Reduced-form own- and cross-price elasticities underestimate the long-run elasticities by 75 and 78 percent, respectively. Overall, Table 8 shows that there is a reasonable degree of consistency in the elasticity estimates of the different models after proper interpretation.[56]

[55]These are obtained as follows: $60.9 = (1 - 0.391) \times 100$ and $66.7 = (1 - 0.333) \times 100$. The approach is similar for the other numbers reported.

[56]See Appendix C.3 and Table A3 for a discussion of the elasticities during Hurricane Katrina.

I also investigate the heterogeneity in the estimated elasticities with respect to whether gasoline prices have been increasing or decreasing in previous months and with respect to the absolute level of the gasoline price similar to the analysis in Section 2. I find that there is still heterogeneity as a result of the hysteresis, albeit smaller in magnitude than that reported in Section 2. 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 Section 5 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 increase.[57] 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:[58]

$$\mathbb{E}\left[CS_i(\nu_i)\right] = \frac{1}{\alpha_{1i}}\mathbb{E}[V_i(\vec{p}_t, m_{irt-1}, \vec{\delta}_{rt}, \epsilon_{imrt}, \nu_i; \theta)], \tag{11}$$

where $\mathbb{E}[\cdot]$ denotes the expectation operator taken over the random shocks ϵ_{imrt}; and the expression $V_i(\vec{p}_t, m_{irt-1}, \vec{\delta}_{rt}, \epsilon_{imrt}, \nu_i; \theta)$ is given by equation (3).

Consumer welfare for type i is defined as the change in 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. For the welfare results below, I compute the total consumer surplus calculated as the weighted sum of $\mathbb{E}\left[CS_i(\nu_i)\right]$ using the weights reflecting the number of consumers who face the same representative utilities as the sampled consumer. That is:

$$\mathbb{E}(CV) = \int_{\nu_i} \left[\mathbb{E}\left[CS_i^1(\nu_i)\right] - \mathbb{E}\left[CS_i^0(\nu_i)\right]\right] \mathrm{d}P_\nu(\nu_i), \tag{12}$$

where $\mathbb{E}(CV)$ denotes the weighted sum across types of consumers of the compensating variation; the superscripts $k = 0$ and $k = 1$ refer, respectively, to before and after the policy change; and $\mathbb{E}\left[CS_i^k(\nu_i)\right]$ is given by equation (11).

Counterfactual results. I consider two policy changes. First, a policy that would change the switching cost for public transit without affecting the gasoline cost of driving. Second, a policy identical to that in point 1, except for the addition of a permanent gasoline tax increase of one dollar per gallon.

Table 9 reports the results from the counterfactuals. Panel A reports the results from the first policy. Panel B reports the results from the second policy. The table displays the value (mean per individual, per market, per period) of the compensating variation, $\mathbb{E}(CV)$, given by equation (12); 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 equation (7).

Three patterns 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

[57] In practice, the reduction in switching costs to public transit can be implemented, *e.g.*, using advertising campaigns or customized marketing along the lines of the public transportation experiments in Abou-Zeid and Ben-Akiva (2012).

[58] The consumer surplus is obtained 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 impact of gasoline on the environment, road accidents, *etc.*) or externalities due to changes in the use of the modes of transportation (such as traffic congestion, public transit delays, *etc.*). The counterfactual analysis does not account for capacity constraints, uses of gasoline tax revenues, or deadweight losses.

by 2 percent ($0.072 - 0.057 = 0.015$) and has a negligible effect on the conditional choice probability of cars.

Second, the gasoline tax increase (without a change in the switching cost of public transit) has a large and negative impact of 31 dollars on the mean (per month, per market, and per consumer type) compensating variation. This negative impact is expected because the use of the gasoline tax revenue is not considered in the analysis in this section. It decreases the mean probability of switching to a car, from public transit or the outside option, by 4 percent ($0.098 - 0.061 = 0.037$) and increases the mean probability of switching to public transit, from a car or the outside option, by 3 percent ($0.057 - 0.083 = -0.026$). Interestingly, when the gasoline tax increase is bundled with a decrease in the switching costs to public transit of 20 percent, the negative impact on welfare is only 11 dollars at the mean compensating variation (again, without considering the use of the gasoline tax revenue). Of course, there are distributional effects across consumer types. Consumer types who continue using automobiles after the policy (the gasoline tax plus a decrease in the switching costs to public transit of 20 percent) are worse off than the mean consumer. Consumers who switch from cars 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 7 and 9 shows that estimates of the static or myopic models tend to mismeasure the effects of substitution between the inside modes of transportation. Consider the case of the gasoline tax increase without a change in the switching cost of public transit. The static model underestimates the substitution patterns by 68 percent on average. It is a result of the two effects discussed in Section 5. On the one hand, the static model does not account for switching costs. It predicts higher probabilities of switching to both cars and public transit (0.237 and 0.134, respectively) than the dynamic model (0.061 and 0.102, respectively). On the other hand, the static model does not account for the effect of consumer's expectations on the decision to switch. In the case of the announced gasoline tax increase, the latter has a larger effect on the decisions to switch for consumers in the dynamic model. That is, 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 also predicts (for the counterfactual) lower probabilities of switching to public transit than the dynamic model. It underestimates the substitution patterns by 87 percent on average. The bias is larger for the myopic model than for the static model because the static model predicts higher probabilities of switching (than the myopic model) due to the static model not accounting for the switching costs, as noted before. The long-run effects, underestimated by the static and myopic models, are relevant for public policy. In either case, static and myopic models could lead to incorrect policy decisions.

5.5 The Distributional Consequences of a Gasoline Tax Increase

There is a longstanding discussion about whether a gasoline tax is regressive.[59] Allegations of the regressivity of the gasoline tax have been used as a major objection against proposals to increase gasoline taxes. The bottom three deciles spend five times as much of their income on gasoline as the top three deciles in the application studied, as described below. One might therefore be concerned that a gasoline tax increase would be highly regressive. Next, I study the distributional implications of a gasoline tax increase while accounting for dynamics and longer-run substitution behavior.

I use the micro data from the travel survey by Pace and the estimated model. I proceed in three steps. First, I compute the annual gasoline expenditure as a share of the consumer's income before the gasoline tax increase. I compute the moments described below using their empirical analogs, where I use the market shares and the joint distribution of reported income and mode of transportation (for work trips) from the travel survey. That is, I use the survey data as the factual choice of the mode of transportation chosen by each surveyed consumer (*i.e.*, the mode of transportation chosen by consumer i before the gasoline tax increase).[60] Second, I compute the annual gasoline expenditure as a share of the consumer's income after the gasoline tax increase.[61] Here, I use the estimated model to simulate the behavior of each simulated consumer after a permanent gasoline tax increase of one dollar. That is, I use the simulation as the counterfactual choice of the mode of transportation chosen by each simulated consumer (*i.e.*, the mode of transportation chosen by consumer i after the gasoline tax increase). I perform this step twice, using the dynamic and static models. Finally, I compute the mean gasoline expenditure as a share of income by income decile. To do so, I arrange the consumers in order of income and divide them into ten groups of equal size. The first income decile contains the lowest income consumers; the last one contains the highest income consumers. The relative expenditures reflect the relative amounts spent on the tax.

The mean annual gasoline expenditure of consumer i is:

$$\text{Mean Annual Gasoline Expenditure}_i = \frac{1}{T}\sum_{t=1}^{T} \frac{\mathbf{1}\{m_{irt} = c\} \times \mathrm{d} \times p_{ct}}{\mathrm{e}} \times 12,$$

where t index months, T is total number of months, $\mathbf{1}\{\cdot\}$ is an indicator function, $m_{irt} = c$ denotes that consumer i chose car as the mode of transportation in route r and period t, d denotes the mean monthly distance traveled by car (constant across consumers), p_{ct} is the monthly gasoline price per gallon in constant dollars as defined in Appendix D, e denotes the mean fuel economy of a car (constant across consumers). In words, the mean annual gasoline expenditure is one over the total number of months times the sum of the total number of months of car use by the consumer times the mean monthly distance

[59] In an influential work, Poterba (1991) discusses using annual expenditure as a more reliable measure of consumer well-being than annual income and argues that gasoline taxes are less regressive under the former. Poterba (1989) shows that household expenditures on gasoline, alcohol, and tobacco as a share of total consumption (a proxy for lifetime income) are more equally distributed than expenditures as a share of annual income and that excise taxes on those goods are less regressive in the long run. See Bento, Goulder, Jacobsen, and Von Haefen (2009) for a more recent discussion about the impacts of increased gasoline taxes.

[60] The predictions by the dynamic and static models before the gasoline tax increase are similar to the empirical analogs from the data. I use the latter for the main specification of the counterfactual to avoid reporting a ratio of ratios in Figure 5, which is less intuitive than the reported ratio. See footnote 64.

[61] By *after* I mean the long-run response (12-month response in the application analyzed).

traveled by car times the price of gasoline per gallon divided by the mean fuel economy of a car times 12. For the analysis, I use d = 750 miles per month (25 miles per day) and e = 23 miles per gallon.[62]

Figure 4A depicts the situation before the gasoline tax increase. It displays gasoline expenditures as a share of income for consumers in different deciles of the pretax income distribution. The figure shows that low-income consumers have higher expenditure-to-income ratios than middle- and high-income consumers. The bottom three deciles spend on average five times as much of their income on gasoline as the top three deciles ($\frac{12.628+5.898+4.102}{1.878+1.448+1.200} = 5$). This evidence is typically used to argue that gasoline taxes are regressive.

Figure 4B shows the situation after the gasoline tax increase. The figure depicts gasoline expenditures as a share of income for consumers in different deciles of the pretax income distribution predicted by the static and dynamic models after the tax increase.[63] Three conclusions emerge.[64] First, compare the predictions from the static and dynamic models after the tax increase (Figure 4B). The static model predicts larger gasoline expenditure shares than the dynamic model for all deciles of the income distribution. In other words, the ratio of gasoline expenditure from the static model relative to the dynamic model is above one for all deciles of the income distribution. It is a consequence of the elasticity estimates from the static model being smaller in absolute value than the long-run elasticities from the dynamic model, as documented in Table 7.

Second, compare the gasoline expenditure shares before and after the gasoline tax increase (Figures 4A *vs.* 4B). The static model predicts that the gasoline expenditure shares increase after the gasoline tax increase. The opposite is true when using the long-run response from the dynamic model: gasoline expenditure shares decrease after the tax increase. It is a consequence of the elasticity estimates from the static model being inelastic and the long-run elasticities from the dynamic model being elastic, as documented in Table 8.

Third, compare the ratio of gasoline expenditures computed from the static model relative to that computed from the dynamic model across deciles of the income distribution. Figure 5 shows that the ratio is larger for low-income than middle- and high-income consumers. Consumers' long-run response to the gasoline tax increase varies considerably across income deciles. That is, there is substantial long-run heterogeneity among consumers with different incomes. Low-income consumers are more price sensitive in the long run and therefore more responsive to a gasoline tax increase than middle- and high-income consumers. Low-income consumers are thus more likely to switch from cars to public transit in response to the gasoline tax than middle- and high-income consumers. This long-run heterogeneity is not captured

[62]These values were obtained from IDOT and represent the mean respective values during the period under analysis. The results reported in this subsection are robust to using other values of d and e. These values shift up or down the whole distribution in Figures 4 and 5; they do not affect the distributional impact analyzed in this subsection because d and e are constant across consumer types. The results in this subsection assume that there is no substantial variation in the parameters d and e across consumers in different income deciles or after the tax increase. I make these assumptions because I do not have information about the joint distribution of income, driving distance, and fuel efficiency and because the mean driving distances and fuel efficiencies are held constant in the model in Section 3. In addition, for the analysis in this subsection, I focus on the regressive nature of the gasoline tax increase; see the previous subsection for the welfare analysis. Finally, as before, the analysis considered is a partial analysis framework. It assumes that consumer income is unaffected by changes in the gasoline tax burden or indexed transfer payments and it abstracts from the effect of the gasoline tax on efficiency costs and fuel economy.

[63]See Table A13 in the appendix for details.

[64]Similar conclusions are obtained using the predictions by the static and dynamic models before the gasoline tax increase in step 1. See Figure A2 in the appendix.

by the static model. The static model predicts that consumers in the bottom three deciles spend on average 61 percent more of their income on gasoline than the dynamic model; for the top three deciles, the difference is 39 percent ($\frac{1.854+1.559+1.425}{3} = 1.612$ and $\frac{1.367+1.411+1.394}{3} = 1.391$). After the gasoline tax increase, the bottom three deciles spend on average 4 times as much of their income on gasoline as the top three deciles in the long run; the number is 4.9 when using the static model after the tax increase (and 5 before the tax increase, as indicated above). The gasoline tax is still regressive, but it is less regressive after the gasoline tax increase due to the dynamics and the long-run substitution behavior. In other words, the (gasoline tax) increase makes the gasoline tax less regressive.

6 Concluding Remarks

Despite the remarkably large number of empirical studies investigating gasoline price elasticities in urban travel demand, much of the work has focused on short-run effects, whereas long-run effects are typically relevant for policy analysis. This article complements such fundamental research by developing a structural approach to estimate long-run substitution patterns. I build a structural model of urban travel consistent with the presence of hysteresis in the context of the market friction generated by switching costs. I estimate the model using a panel dataset with market-level data from Chicago, constructed using public data obtained from government agencies. I use the estimated model to compute long-run gasoline price 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 or carbon tax increase.

I discuss two main results. First, long-run gasoline price elasticities are substantially more elastic than short-run elasticities. The gasoline price elasticities are inelastic in the short run but elastic in the long run. Second, static own- (for automobile) and cross- (for public transit) price elasticities underestimate the dynamic elasticities. Myopic estimates also underestimate the elasticities. The bias is larger for the myopic model than for the static model.

I performed a number of robustness checks, including different specifications of the model, consumers' expectations, micro moments, how seasonal effects are incorporated, the role of the Great Depression, and the estimation procedures. The implications are robust to these alternative specifications. Overall, I find robust evidence that when hysteresis and switching costs are present, gasoline price elasticities from static and myopic models might be biased and could lead to incorrect policy decisions.

I find substantial long-run heterogeneity among consumers in different income deciles. This feature helps to explain why static and myopic models mismeasure long-run substitution responses to a gasoline tax increase. Low-income consumers are more likely to switch from car to public transit than middle- and high-income consumers. From a longer-run perspective, a gasoline tax is therefore substantially less regressive than predicted by static or short-run analyses. The illustrative application that I presented in this article suggests that when studying the incidence of gasoline taxes, these biases may be substantial.

References

Abou-Zeid, M., and M. Ben-Akiva (2012): "Travel mode switching: Comparison of findings from two public transportation experiments," *Transport Policy*, 24, 48–59. 28

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. 5

Aguirregabiria, V., and P. Mira (2010): "Dynamic discrete choice structural models: A survey," *Journal of Econometrics*, 156(1), 38–67. 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. 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. 16

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.)*. 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. 3, 17

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. 3, 17

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. 6

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. 5

——— (1981): "A decomposition of the price and income elasticities of the consumer demand for gasoline," *Southern Economic Journal*, pp. 1021–1031. 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. 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. 7

Baldwin, R. (1989): "Sunk-cost hysteresis," *National Bureau of Economic Research Cambridge, Mass., USA*. 7

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

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

Basso, L. J., and T. H. Oum (2007): "Automobile fuel demand: a critical assessment of empirical methodologies," *Transport Reviews*, 27(4), 449–484. 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. 2, 6

Ben-Akiva, M. E., and S. R. Lerman (1985): *Discrete choice analysis: theory and application to travel demand*, vol. 9. MIT press. 4, 14

Bento, A. M., L. H. Goulder, M. R. Jacobsen, and R. H. Von Haefen (2009): "Distributional and efficiency impacts of increased US gasoline taxes," *American Economic Review*, 99(3), 667–99. 30

Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile prices in market equilibrium," *Econometrica*, pp. 841–890. 5, 7, 20

——— (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. 3, 24

Berry, S. T. (1994): "Estimating discrete-choice models of product differentiation," *The RAND Journal of Economics*, pp. 242–262. 20, 36

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

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

——— (1980): "Hysteresis in travel demand," *Transportation Planning and Technology*, 6(2), 109–115. 2, 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. 7

Bricka, S. (2007): "Chicago Regional Household Travel Inventory Final Report," *Prepared for Chicago Metropolitan Agency for Planning*. 14

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. 6

Broyden, C. G. (1965): "A class of methods for solving nonlinear simultaneous equations," *Mathematics of computation*, 19(92), 577–593. 36

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*. 5

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. 3, 6, 16

Button, K. (2010): *Transport economics*. Edward Elgar Publishing. 2, 4, 14, 15

Cambridge Systematics Inc. (2007): "Pace Market Research Report," Discussion paper, Available at `http://www.rtams.org/reportLibrary/564.pdf`. (Accessed on May 10, 2018.). 8, 25, 41, A-5, A-21

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. 2, 6

Chatterjee, K. (2011): "Modelling the dynamics of bus use in a changing travel environment using panel data," *Transportation*, 38(3), 487–509. 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. 2, 6

Chen, N. (2007): "Modelling Demand for Rail Transport with Dynamic Econometric Approaches," *International Review of Business Research Papers*, 3, 85–96. 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*. 2, 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. 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. 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*. 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. 5

——— (2008): "Understanding links between transit ridership and auto gas prices: US and Australian evidence," in *Transportation Research Board 87th Annual MeetingTransportation Research Board*. 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. 6
Dahl, C., and T. Sterner (1991): "Analysing gasoline demand elasticities: a survey," *Energy economics*, 13(3), 203–210. 5
Dahl, C. A. (1986): "Gasoline demand survey," *The Energy Journal*, 7(1), 67–82. 5
Dargay, J. M. (1993): "Demand elasticities: a comment," *Journal of Transport Economics and Policy*, 27(1), 87–90. 2, 6, 11
Das, S., M. J. Roberts, and J. R. Tybout (2007): "Market entry costs, producer heterogeneity, and export dynamics," *Econometrica*, 75(3), 837–873. 7
Dixit, A. (1989): "Entry and exit decisions under uncertainty," *Journal of political Economy*, 97(3), 620–638. 7, 25, 41, A-21
Dixit, A. K., and R. S. Pindyck (1994): *Investment under uncertainty*. Princeton university press. 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. 5
Donna, J. D., and J. A. Espin-Sanchez (2021): "The Illiquidity of Water Markets," Discussion paper, Available at SSRN `http://dx.doi.org/10.2139/ssrn.2667654`. 7, 38
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. 7
——— (2010): "State dependence and alternative explanations for consumer inertia," *The RAND Journal of Economics*, 41(3), 417–445. 7, 21, 22, 23
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. 6
Espey, M. (1998): "Gasoline demand revisited: an international meta-analysis of elasticities," *Energy Economics*, 20(3), 273–295. 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. 7
Fearnley, N., and J. Bekken (2005): "Short-and long-run demand effects in transport: A literature survey," *Institute of Transport Economics, Norway*. 6
Frondel, M., and C. Vance (2013): "Re-identifying the rebound: what about asymmetry?," *The Energy Journal*, pp. 43–54. 6
Gärling, T., and K. Axhausen (2003): "Introduction: Habitual travel choice," *Transportation*, 30(1), 1–11. 6
Goeree, M. S. (2008): "Limited information and advertising in the US personal computer industry," *Econometrica*, 76(5), 1017–1074. 3, 7, 24
Golob, T. F., R. Kitamura, and L. Long (2013): *Panels for transportation planning: methods and applications*. Springer Science & Business Media. 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. 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*. 22
Goodwin, P. B. (1977): "Habit and hysteresis in mode choice," *Urban studies*, 14(1), 95–98. 2, 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. 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. 5
Gowrisankaran, G., and M. Rysman (2012): "Dynamics of consumer demand for new durable goods," *Journal of political Economy*, 120(6), 1173–1219. 7, 17, 19, 20
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. 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*. 5
Haire, A. R., and R. B. Machemehl (1992): "Impact of Rising Fuel Prices on U.S. Transit Ridership," *Transportation Research Record*, pp. 11–19. 5
Hamilton, J. D. (2009): "Understanding Crude Oil Prices," *The Energy Journal*, 30(2). 16
——— (2011): "Historical oil shocks," Discussion paper, National Bureau of Economic Research. Available at: `https://www.nber.org/papers/w16790` (Accessed on September 22, 2012.). 12, 39
Handel, B. R. (2013): "Adverse selection and inertia in health insurance markets: When nudging hurts," *American Economic Review*, 103(7), 2643–82. 7
Hansen, L. P. (1982): "Large sample properties of generalized method of moments estimators," *Econometrica*, pp. 1029–1054. 38
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. 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. 6, 21
Hendel, I., and A. Nevo (2006): "Measuring the implications of sales and consumer inventory behavior," *Econometrica*, 74(6), 1637–1673. 7, 17, 38
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. 6
Hogema, J. (1996): "Effects of Rain on Daily Traffic Volume and on Driving Behaviour," *Report TM-96-B019*. 6
Honka, E. (2014): "Quantifying search and switching costs in the US auto insurance industry," *The RAND Journal of Economics*, 45(4), 847–884. 7
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. 5
Horowitz, J., and R. Emslie (1978): "Comparison of measured and forecast traffic volumes on urban interstate highways," *Transportation Research*, 12(1), 29–32. 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. 7
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). 5
Illinois Department of Transportation (2004): "Illinois Traffic Monitoring Program," *Office of Planning and Programming*. A-3, A-4
Innocenti, A., P. Lattarulo, and M. G. Pazienza (2013): "Car stickiness: Heuristics and biases in travel choice," *Transport Policy*, 25, 158–168. 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. 6

Kellogg, R. (2018): "Gasoline price uncertainty and the design of fuel economy standards," *Journal of Public Economics*, 160, 14–32. 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. 12, 39

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.).* 7

Klemperer, P. (1987a): "The competitiveness of markets with switching costs," *The RAND Journal of Economics*, pp. 138–150. 7, 15

——— (1987b): "Markets with consumer switching costs," *The Quarterly Journal of Economics*, 102(2), 375–394. 7, 15

——— (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. 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. 5

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

Krugman, P. R. (1989): *Exchange Rate Instability*. Cambridge, MA: MIT Press. 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. 36

Lerman, S., and M. Ben-Akiva (1976): "A Dissagregatge Behavioral Model of Auto Ownership," *Transportation Res. Rec.* 5

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. 5

Li, S., J. Linn, and E. Muehlegger (2014): "Gasoline taxes and consumer behavior," *American Economic Journal: Economic Policy*, 6(4), 302–42. 5

Lin, C.-Y. C., and L. Prince (2013): "Gasoline price volatility and the elasticity of demand for gasoline," *Energy Economics*, 38, 111–117. 3

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

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

——— (2017): "Transit price elasticities and cross-elasticities," *Victoria Transport Policy Institute. Available at: https://www.vtpi.org/tranelas.pdf (Accessed on January 6, 2018.).* 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. 8, A-5

Luco, F. (2019): "Switching costs and competition in retirement investment," *American Economic Journal: Microeconomics*, 11(2), 26–54. 7

MacKay, A., and M. Remer (2019): "Consumer Inertia and Market Power," *Available at: http://dx.doi.org/10.2139/ssrn.3380390 (Accessed on July 30, 2019.).* 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. 2, 5, 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. 6

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. 5

Melnikov, O. (2013): "Demand for differentiated durable products: The case of the US computer printer market," *Economic Inquiry*, 51(2), 1277–1298. 17

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. 5

Morikawa, T. (1994): "Correcting state dependence and serial correlation in the RP/SP combined estimation method," *Transportation*, 21(2), 153–165. 6

Nevo, A. (2001): "Measuring market power in the ready-to-eat cereal industry," *Econometrica*, 69(2), 307–342. 7

Newey, W. K., and D. McFadden (1994): "Large sample estimation and hypothesis testing," *Handbook of econometrics*, 4, 2111–2245. 38

Ohio Supercomputer Center (1987): "Ohio Supercomputer Center," http://osc.edu/ark:/19495/f5s1ph73.

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. 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.).* 8, A-5

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*. 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. 7

Pesendorfer, M. (2002): "Retail sales: A study of pricing behavior in supermarkets," *The Journal of Business*, 75(1), 33–66. 7

Petrin, A. (2002): "Quantifying the benefits of new products: The case of the minivan," *Journal of political Economy*, 110(4), 705–729. 3, 24

Poterba, J. M. (1989): "Lifetime Incidence and the Distributional Burden of Excise Taxes," *The American Economic Review*, 79(2), 325–330. 30

——— (1991): "Is the gasoline tax regressive?," *Tax Policy and the Economy*, 5, 145–164. 30

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. 7

Rose, G. (1986): "Transit passenger response: short and long term elasticities using time series analysis," *Transportation*, 13(2), 131–141. 5

Rust, J. (1987): "Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher," *Econometrica*, pp. 999–1033. 7, 18

——— (1994): "Structural estimation of Markov decision processes," *Handbook of econometrics*, 4, 3081–3143. 6, 18

——— (1996): "Numerical dynamic programming in economics," *Handbook of computational economics*, 1, 619–729. 37

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. 6

Schiraldi, P. (2011): "Automobile replacement: a dynamic structural approach," *The RAND journal of economics*, 42(2), 266–291. 17

Schmalensee, R., and T. M. Stoker (1999): "Household gasoline demand in the United States," *Econometrica*, 67(3), 645–662. 5

Severen, C., and A. Van Benthem (2019): "Formative Experiences and the Price of Gasoline," Discussion paper, National Bureau of Economic Research. 6

Shcherbakov, O. (2016): "Measuring consumer switching costs in the television industry," *The RAND Journal of Economics*, 47(2), 366–393. 7, 17, 19

Shum, M. (2004): "Does advertising overcome brand loyalty? Evidence from the breakfast-cereals market," *Journal of Economics & Management*

Strategy, 13(2), 241–272. 7
SMALL, K. (2013): *Urban transportation economics*, vol. 4. Taylor & Francis. 4
SMALL, K., E. VERHOEF, AND R. LINDSEY (2007): *The economics of urban transportation*. Routledge. 4, 14
SMALL, K. A., AND K. VAN DENDER (2007): "Fuel efficiency and motor vehicle travel: the declining rebound effect," *The Energy Journal*, pp. 25–51. 6
SORRELL, S., J. DIMITROPOULOS, AND M. SOMMERVILLE (2009): "Empirical estimates of the direct rebound effect: A review," *Energy policy*, 37(4), 1356–1371. 2, 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. 6
STORCHMANN, K. (2001): "The Impact of Fuel Taxes on Public Transport Ð An Empirical Assessment for Germany," *Transport Policy*, 8, 19–28. 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*. 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. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2409115 (Accessed on February 14, 2018.)*. 7, 21
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. 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*. 6
TRAIN, K. (1980): "A structured logit model of auto ownership and mode choice," *The Review of Economic Studies*, 47(2), 357–370. 5
TYNDALL, J. (2017): "Waiting for the R train: Public transportation and employment," *Urban Studies*, 54(2), 520–537. 5
U.S. CENSUS BUREAU (2017): "X-13ARIMA-SEATS Reference Manual, Version 1.1," *Available at: https://www.census.gov/ts/x13as/docX13AS.pdf (Accessed on Jul. 22 2019.)*. 10, A-11
U.S. DEPARTMENT OF TRANSPORTATION (2011): "Summary of travel trends: 2009 national household travel survey," *Technical Report No. FHWA- PL-11-022. Available at: https://nhts.ornl.gov/publications.shtml (Accessed on December 28, 2016.)*. 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. 6
WHEATON, W. C. (1982): "The long-run structure of transportation and gasoline demand," *The Bell Journal of Economics*, pp. 439–454. 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. 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*. 6
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. 5

Appendix

A 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 $\vec{p}_t$, 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 approach 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 approach 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, $\vec{\delta}_{rt}(\theta)$, that equates $s_{mrt}(\vec{p}_t, \vec{\delta}_{rt}; \theta) = S_{mrt}$. I solve for $\vec{\delta}_{rt}(\theta)$ in the system of equations in (10) 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 $\vec{\delta}_{rt}(\theta)$ in the system of equations in (10). (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):

$$\delta_{mrt}^{h+1} = \delta_{mrt}^{h} + \ln(S_{mrt}) - \ln[s_{mrt}(\vec{p}_t, \vec{\delta}_{rt}; \theta)], \tag{13}$$

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

where H is the smallest integer such that $\|\delta_{mrt}^{h+1} - \delta_{mrt}^{h}\|$ is below a tolerance level of $1e-12$. The value of δ_{mrt}^{H} 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 equation (8). I approximate this integral by $s_{mrt}(\cdot) = \sum_{i=1}^{I} \mathbb{P}_i(m_{irt}, \nu_i|\vec{p}_t, \vec{\delta}_{rt}; \theta)w_i$, where ν_i with $i = 1, \ldots, I$ are draws obtained using the importance sampling procedure described below, and w_i are the corresponding weights.

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 (5).[65] For each consumer type, $i = 1, \ldots, I_r$, I solve for $v_i(\vec{p}_t, m_{irt-1}, m_{irt}, \vec{\delta}_{rt}, \nu_i; \theta)$ by iterating on it as follows:

$$\begin{aligned} v_i(\vec{p}_t, m_{irt-1}, m_{irt}, \vec{\delta}_{rt}, \nu_i; \theta)^{(h+1)} &= \delta_{mrt} - \phi_m \mathbf{1}\{m_{irt-1} \neq m_{irt}\} - \sigma \nu_i p_{mt} + \\ &+ \beta \int \log\left(\sum_{m_{irt+1} \in M} \exp\left[v_i(\vec{p}_{t+1}, m_{irt}, m_{irt+1}, \vec{\delta}_{rt+1}, \nu_i; \theta)^{(h)} \right] \right) f_{p_c}(p_{ct+1}|p_{ct}) \mathrm{d}p_{ct+1} + \beta\gamma, \quad (14) \\ m_{irt} \in \{0, c, b\}, \quad r = 1, \ldots, R, \quad & t = 1, \ldots, T, \quad h = 0, \ldots, H, \quad i = 1, \ldots, I_r, \end{aligned}$$

where H is the smallest integer such that $\|v_i(\vec{p}_t, m_{irt-1}, m_{irt}, \vec{\delta}_{rt}, \nu_i; \theta)^{(h+1)} - v_i(\vec{p}_t, m_{irt-1}, m_{irt}, \vec{\delta}_{rt}, \nu_i; \theta)^{(h)}\|$ is below a tolerance level of $1e{-}12$. The value of $v_i(\vec{p}_t, m_{irt-1}, m_{irt}, \vec{\delta}_{rt}, \nu_i; \theta)^{(H)}$ is taken as the approximation of $v_i(\vec{p}_t, m_{irt-1}, m_{irt}, \vec{\delta}_{rt}, \nu_i; \theta)$. For the distribution $f_{p_c}(p_{ct+1}|p_{ct})$, I use the specifications described in Section 3, discretizing $\vec{p}_{ct}$ into $P_c = 100$ possible values, as described below. Then I approximate the integral on the right-hand side in (14) by:

$$\frac{1}{S_p} \sum_{s=1}^{S_p} \log\left(\sum_{m'_{t+1} \in M} \exp\left[v_i(p_{bt+1}, p^s_{ct+1}, m_{irt}, m_{irt+1}, \vec{\delta}_{rt+1}, \nu_i; \theta)^{(H)} \right] \right),$$

where p^s_{ct+1} with $s = 1, \ldots, NS_p$ are draws from $f_{p_c}(p^s_{ct+1}|p_{ct})$, using the specifications in Section 3. For the specifications using boundedly rational expectations (BRE-AR and BRE-FE) the distribution $f_{p_c}(p^s_{ct+1}|p_{ct})$ is estimated outside the estimation algorithm, in a preceding step. For the perfect foresight specification, the value function is solved recursively using equation (5) with the actual prices and $v_i(\vec{p}_T, m_{irT-1}, m_{irT}, \vec{\delta}_{rT}, \nu_i; \theta) = \delta_{mrT} - \phi_m \mathbf{1}\{m_{irT-1} \neq m_{irT}\} - \sigma\nu_i p_{mT}$, and the conditional choice probabilities are solved for all periods applying the contraction mapping to equation (7b).

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 $\vec{p}_t$ into $|P|$ values, it involves solving simultaneously a system of $(3 \times 3 \times |P|)$ equations in (5). The problem can be simplified by noting that $v_i(\vec{p}_t, m_{irt-1}, m_{irt} = 0, \vec{\delta}_{rt}, \nu_i; \theta) = v_i(\vec{p}_t, m_{irt-1} = 0, m_{irt} = 0, \vec{\delta}_{rt}, \nu_i; \theta)$ for all $m_{irt-1} \in M$, because there are no switching costs associated with the outside good; and $v_i(\vec{p}_t, m_{irt-1}, \tilde{m}_{irt} \neq m_{irt-1}, \vec{\delta}_{rt}, \nu_i; \theta) = v_i(\vec{p}_t, m_{irt-1}, \tilde{m}_{irt} = m_{irt-1}, \vec{\delta}_{rt}, \nu_i; \theta) - \phi_{\tilde{m}_{irt}}$; that is, the conditional value of choosing mode $\tilde{m}_{irt}$ by switching from m_{irt-1} to $\tilde{m}_{irt}$ (with $\tilde{m}_{irt} \neq m_{irt-1}$) 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} to $\tilde{m}_{irt}$ (with $m_{irt-1} \neq \tilde{m}_{irt}$). The problem is simplified from solving a system of $(3 \times 3 \times |P|)$ equations to one of $(3 \times |P|)$ equations in (5). Thus, when iterating on the contraction mapping in (14), I use that $v_i(\vec{p}_t, m_{irt-1}, m_{irt} = 0, \vec{\delta}_{rt}, \nu_i; \theta) = v_i(\vec{p}_t, m_{irt-1} = 0, m_{irt} = 0, \vec{\delta}_{rt}, \nu_i; \theta)$ for all $m_{irt-1} \in M$, and $v_i(\vec{p}_t, m_{irt-1}, \tilde{m}_{irt} \neq m_{irt-1}, \vec{\delta}_{rt}, \nu_i; \theta) = v_i(\vec{p}_t, m_{irt-1}, \tilde{m}_{irt} = m_{irt-1}, \vec{\delta}_{rt}, \nu_i; \theta) - \phi_{\tilde{m}_{irt}}$, which simplifies the dimension of the system as explained above.

Discretization. To perform the estimation, I discretize p_{ct} in $P_c = 100$ prices. This discretization 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 equation (4) and to compute the price transition matrix, $f_{p_c}(p_{ct+1}|p_{ct})$. In the latter, when a predicted price is different from the discretized prices, the price is replaced by the closest discrete price. For p_{bt}, I use the values of the CTA rail fare collected from the data.

Importance sampling procedure.[66] The following subsection describes the procedure I use to obtain S_1 draws *via* importance sampling. (1) Draw ν_i, $i = 1, \ldots, S_1$ draws *via* a standard Halton sequence and obtain a preliminary estimate of the parameter vector, θ_1, using the algorithm in Section 4. (2) Obtain $S_2 = 1000 \times S_1$ draws from ν_1. Call the vector with those draws $\tilde{\nu}$. For each $i = 1, \ldots, S_2$, compute the probability of ever choosing the inside mode at θ_1, denoted by P_i using, *e.g.*, the average choice probability: $P_i = \frac{1}{T}\sum_{t=1}^{T}(1 - s_{i0t})$, where s_{i0t} denotes the share of i choosing the outside option at t. Let $p_i = P_i / (\sum_{i=1}^{S_2} P_i)$ denote the normalized probabilities that sum to 1. (3) Let ρ be a vector with $S_1 + 1$ evenly

[65] This procedure is a globally convergent algorithm for computing the fixed point. See Rust (1996) for details.

[66] I thank Marc Rysman for suggesting this procedure. The description of the importance sampling procedure follows closely the one in the note "A method for implementing importance sampling in BLP type models," by Marc Rysman.

spaced elements. Denote by ρ_k the element k of ρ with $k = 1, \ldots, S_1$. (4) Create a vector of size S_1 with the importance sampling draws, ν^*. The element k of ν^* is defined as $\nu_k^* = \tilde{\nu}_n$, where n is the highest integer such that: $\sum_{i=1}^{n-1} p_i < \rho_k$. Let P^* be the S_1 vector of associated normalized probabilities, such that if element k of ν^* equals element n of $\tilde{\nu}$, then element k of P^* equals p_n. (5) Let w^* be the S_1 vector of weights, where each element is given by: $w_k^* = \frac{1/P_k^*}{\sum_{i=1}^{S_1}(1/P_i^*)}$. (6) Compute the market shares using the importance sampling draws and weights: $s_{mrt}(\cdot) = \sum_{i=1}^{S_1} \mathbb{P}_i(m_{irt}, \nu_i | \vec{p}_t, \vec{\delta}_{rt}; \theta) w_i$, where ν_i with $i = 1, \ldots, S_1$ are the importance sampling draws and and w_i are the corresponding weights. The number of draws S_2 needs to be high enough so that no draw n from $\tilde{\nu}$ is drawn more than once.

Consumer draws. I use 100 importance sampling draws in step 3. For the preferred specification of the model, I also performed the estimation using 150 importance sampling consumer draws and obtained similar results. I perform the computation of the conditional value functions in step 3 separately for each consumer draw ν_i, which are parallelized in the estimation routine.

Initial condition. The initial choice of the mode for type i, m_{i0} (henceforth, initial condition), is unobserved in the data. I follow Hendel and Nevo (2006) and Donna and Espin-Sanchez (2021), 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, the 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 $\mathbb{P}(m_{irt}|\vec{p}_t)$ in equation (7). Here, for the *arbitrary* initial condition in (i), I use $\mathbb{P}(m_{ir0}|\vec{p}_t) = S_{mr0}$, 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 the car and public transit during the $\tilde{t}$ periods before 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 (6), and integrating over the distribution of consumers in route r and period t:

$$s_{mrt|m\tilde{t}}(\vec{p}_t, \vec{\delta}_{rt}; \theta) = \int \prod_{\hat{t}=t-\tilde{t}+1}^{t} \underbrace{\frac{\exp\left[v_i(\vec{p}_{\hat{t}}, m_{ir\hat{t}-1} = m_{ir\hat{t}}, m_{ir\hat{t}}, \vec{\delta}_{r\hat{t}}, \nu_i; \theta)\right]}{\sum_{m'_{\hat{t}} \in M} \exp\left[v_i(\vec{p}_{\hat{t}}, m_{ir\hat{t}-1} = m_{ir\hat{t}}, m'_{ir\hat{t}}, \vec{\delta}_{r\hat{t}}, \nu_i; \theta)\right]}}_{\mathbb{P}_i(m_{ir\hat{t}}, \nu_i | \vec{p}_{\hat{t}}, m_{ir\hat{t}-1} = m_{ir\hat{t}}, \vec{\delta}_{r\hat{t}}; \theta)} P_\nu(\nu_i) \mathrm{d}\nu_i, \tag{15}$$

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

For the estimation, the conditional value function and choice probabilities in (15) are computed using the inner subroutine in step 3 as described in Section 4 in the article. As in equation (8), I approximate the integral in (15) by:

$$s_{mrt|m\tilde{t}}(\vec{p}_t, \vec{\delta}_{rt}; \theta) = \sum_{i=1}^{I} \prod_{\hat{t}=t-\tilde{t}+1}^{t} \mathbb{P}_i(m_{ir\hat{t}}, \nu_i | \vec{p}_{\hat{t}}, m_{ir\hat{t}-1} = m_{ir\hat{t}}, \vec{\delta}_{r\hat{t}}; \theta) w_i, \tag{16}$$

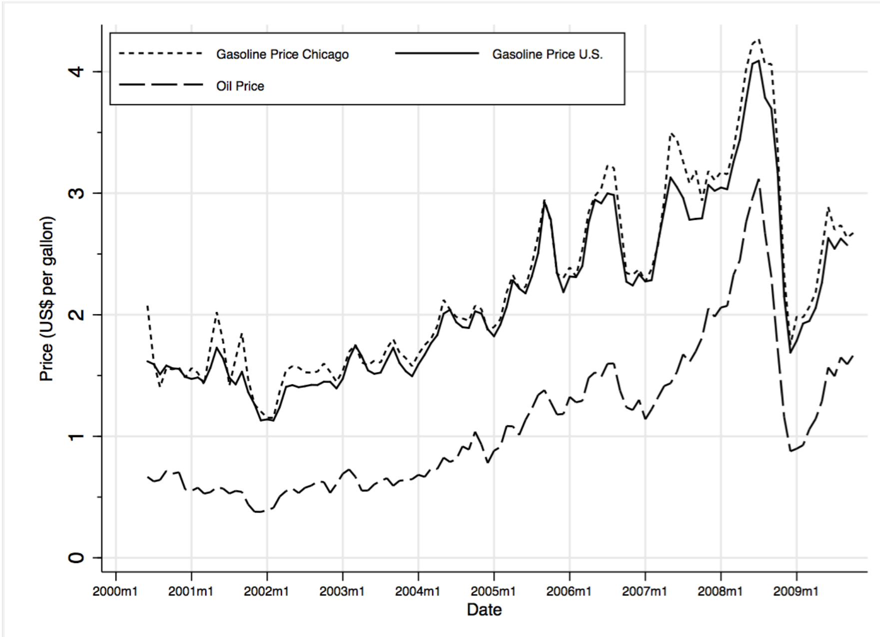
where ν_i with $i = 1, \ldots, I$, are draws obtained using the importance sampling procedure described above, and w_i are the corresponding weights. For the estimates in Table 6, I use $\tilde{t} = 3$. Table A6 presents robustness analysis for $\tilde{t} = 6$ and $\tilde{t} = 12$. Details about the additional moments are in Section 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.

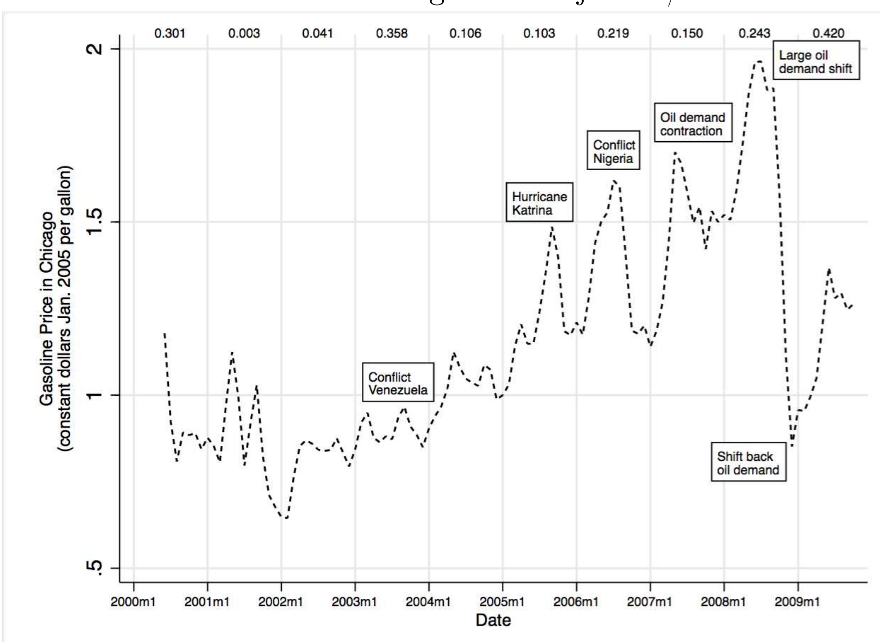
Figures and Tables

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

A. Gasoline and Oil Prices.

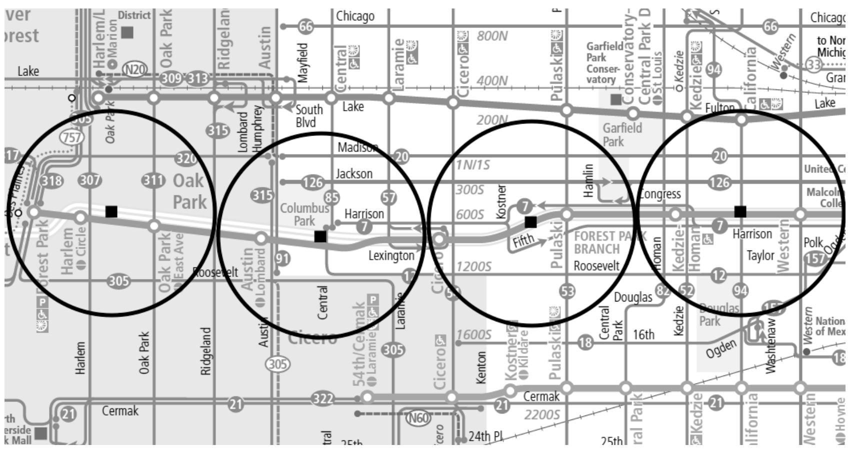


B. Real Gasoline Prices in Chicago and Major Oil/Gasoline Shocks.



Notes: Figure 1A depicts the monthly evolution of the gasoline prices in Chicago (nominal gasoline price in Chicago in the variables definition appendix) and 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 Section 2, 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 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 2, Panel A, row 1, which has coefficient for $ln(\vec{p}_t) equal to .174$. 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 2. Variables definitions are in Appendix D.

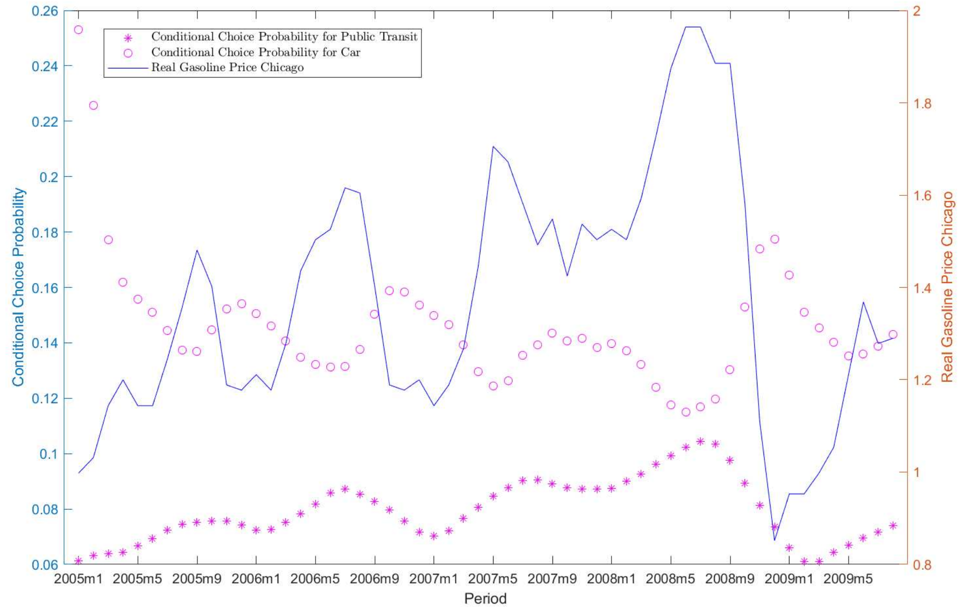
Figure 2: An Example of a Route.



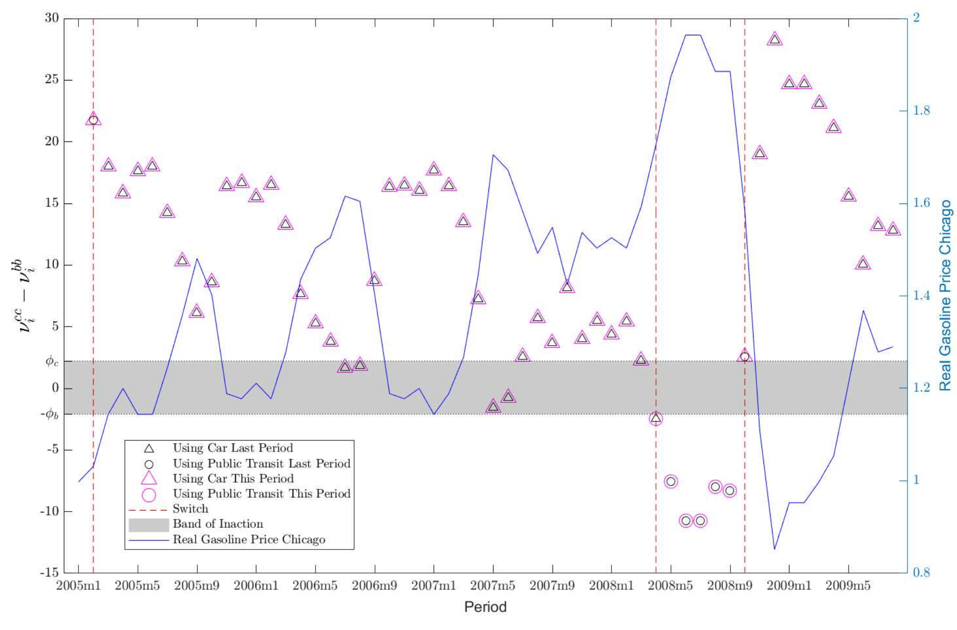
Notes: An example of a *route*, along Dwight D. Eisenhower Expressway (Interstate 290). Black dots denote the location of ATR stations. Circles denote radius around those locations.

Figure 3: Estimated model.

A. Conditional Choice Probabilities and Gasoline Prices.



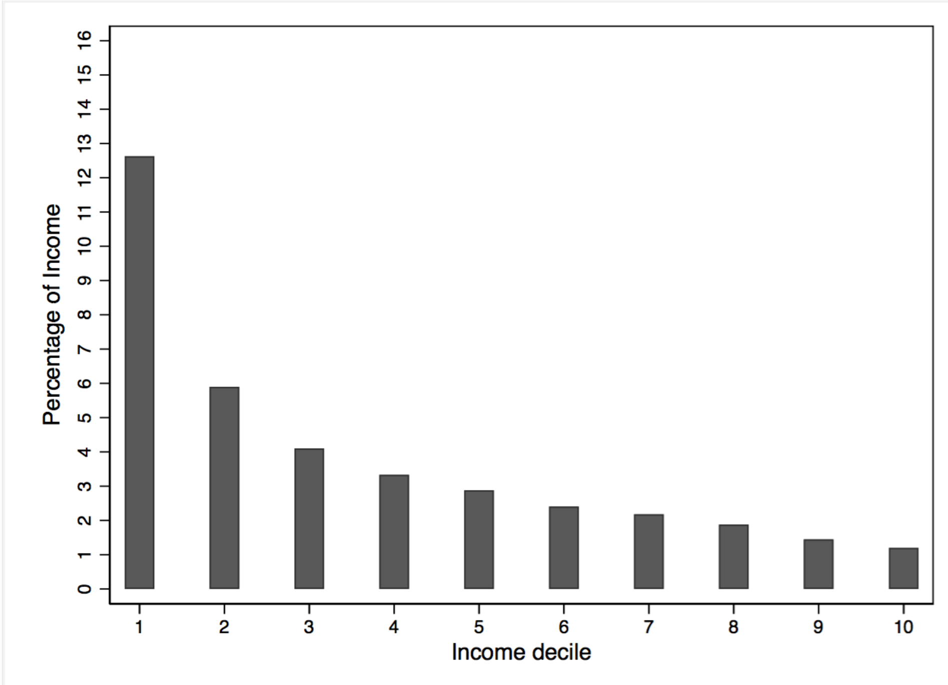
B. Switching Behavior of a Transit-Friendly Consumer.



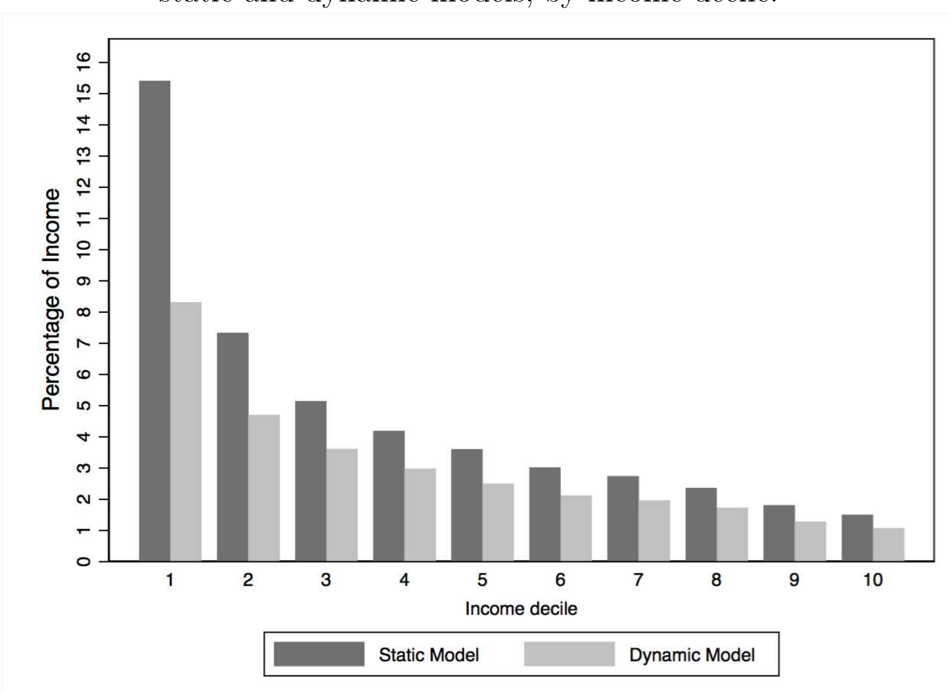
Notes: Figure 3A shows the mean, across consumer types and markets, conditional choice probabilities (left vertical axis) given by equation (7), for public transit and car, holding fixed ξ_{mrt}, as a function of the observed gasoline cost of driving (right vertical axis), $\vec{p}_t$, for the period 2005-2009. Figure 3B 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}. The figure shows on the left vertical axis using dots the difference in the conditional value function of choosing car transport today given the choice to travel by car in the previous period, minus the conditional value function of choosing public transit today given a previous-period choice of public transit for this particular transit-friendly consumer type: $v_i^{cc} - v_i^{bb}$, with $v_i^{jr} \equiv v_i(\vec{p}_t, m_{irt-1} = j, m_{irt} = r, \vec{\delta}_{rt}, \nu_i; \theta)$, where the superscripts j and r index, respectively, the mode of transportation of consumer i in $t-1$ and t, and the conditional value function $v_i(\vec{p}_t, m_{irt-1}, m_{irt}, \vec{\delta}_{rt}, \nu_i; \theta)$ is given by equation (5). The vertical dotted lines indicate a switch in the inside mode of transportation by the consumer in that period. The graph also displays the gasoline cost of driving on the right vertical axis using a solid line and the estimated switching-cost parameters, ϕ_c and $-\phi_b$, using horizontal lines. 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). Both figures use the estimates from the dynamic model with both switching costs and additional micro moments, from Panel C in Table 6. See Section 5 for a description of the switching behavior.

Figure 4: Distributional Consequences of a Gasoline Tax Increase.

A. Before Gasoline Tax Increase:
Gasoline expenditure as a share of income, by income decile.



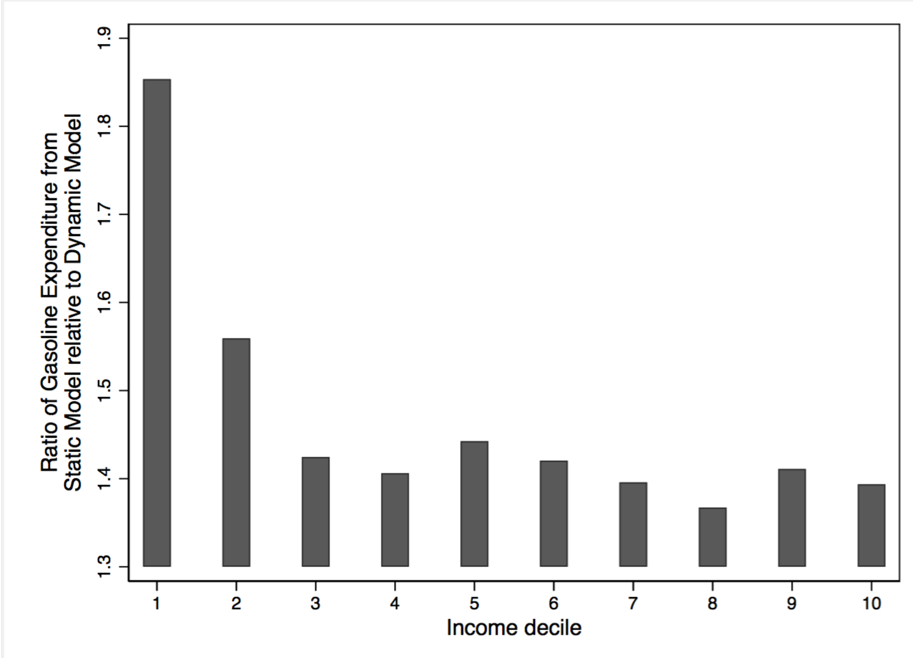
B. After Gasoline Tax Increase:
Gasoline expenditure as a share of income computed from the static and dynamic models, by income decile.



Notes: Figure 4A displays the gasoline expenditures as a share of income before the gasoline tax increase for consumers in different deciles of the pretax income distribution. Figure 4B displays the gasoline expenditures as a share of income after the gasoline tax increase predicted by the static and dynamic models for consumers in different deciles of the pretax income distribution. See Section 5 and Appendix G for details.

Figure 5: Distributional Consequences of a Gasoline Tax Increase.

After Gasoline Tax Increase:
Ratio of gasoline expenditure computed from the static model relative to gasoline expenditure computed from the dynamic model, by income decile.



Notes: The figure displays the ratio of gasoline expenditure computed from the static model relative to gasoline expenditure computed from the dynamic model after the gasoline tax increase for consumers in different deciles of the pretax income distribution. See Section 5 and column 4 in Table A13 for details.

Table 1: 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-unit 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 Section 2, and Tables 2, 3, and A2. [a] CTA rail, CTA bus, and Pace bus values refer to the monthly average for weekdays, disaggregated by rail station, or bus routes, respectively. Metra rail refers to the monthly total combined for weekdays, 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 Section 2 and Appendix B 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 $r = 1, \ldots, R$.

Table 2: 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: The dependent variable is the natural logarithm of CTA rail ridership, at the CTA station level.															
OLS	1	0.174	(0.020)	15,725	0.923	0.878	(0.036)	15,571	0.721	0.518	(0.066)	15,130	0.780	No	No
	2	0.174	(0.020)	15,725	0.923	0.878	(0.036)	15,571	0.723	0.517	(0.066)	15,130	0.787	Yes	No
	3	0.037	(0.013)	15,725	0.926	0.074	(0.032)	15,571	0.731	0.091	(0.078)	15,130	0.790	No	Yes
	4	0.035	(0.014)	15,725	0.946	0.074	(0.033)	15,571	0.733	0.098	(0.082)	15,130	0.796	Yes	Yes
IV	5	0.197	(0.017)	15,725	0.923	0.984	(0.027)	15,571	0.720	0.586	(0.061)	15,130	0.780	No	No
	6	0.198	(0.017)	15,725	0.924	0.984	(0.027)	15,571	0.723	0.584	(0.061)	15,130	0.787	Yes	No
	7	0.027	(0.015)	15,725	0.926	0.030	(0.032)	15,571	0.731	0.052	(0.078)	15,130	0.789	No	Yes
	8	0.024	(0.016)	15,725	0.926	0.031	(0.032)	15,571	0.733	0.061	(0.079)	15,130	0.796	Yes	Yes
Panel B: The dependent variable is the natural logarithm of CTA bus ridership, at the CTA bus route level.															
OLS	1	0.097	(0.018)	15,646	0.953	0.101	(0.018)	10,793	0.956	0.145	(0.025)	9,116	0.939	No	No
	2	0.097	(0.018)	15,646	0.954	0.095	(0.019)	10,793	0.959	0.142	(0.027)	9,116	0.944	Yes	No
	3	0.089	(0.019)	15,646	0.954	0.074	(0.026)	10,793	0.956	0.148	(0.049)	9,116	0.940	No	Yes
	4	0.092	(0.029)	15,646	0.955	0.079	(0.023)	10,793	0.959	0.167	(0.050)	9,116	0.945	Yes	Yes
IV	5	0.120	(0.018)	15,646	0.953	0.126	(0.020)	10,793	0.955	0.176	(0.026)	9,116	0.939	No	No
	6	0.119	(0.018)	15,646	0.954	0.119	(0.021)	10,793	0.959	0.172	(0.028)	9,116	0.944	Yes	No
	7	0.121	(0.024)	15,646	0.954	0.112	(0.035)	10,793	0.956	0.204	(0.060)	9,116	0.940	No	Yes
	8	0.123	(0.024)	15,646	0.955	0.116	(0.032)	10,793	0.959	0.224	(0.061)	9,116	0.945	Yes	Yes
Panel C: The dependent variable is the natural logarithm of Pace bus ridership, at the Pace bus route level.															
OLS	1	0.021	(0.015)	26,090	0.968	0.113	(0.020)	10,966	0.947	0.207	(0.032)	4,533	0.924	No	No
	2	0.021	(0.016)	26,090	0.968	0.112	(0.020)	10,966	0.948	0.207	(0.032)	4,533	0.926	Yes	No
	3	0.074	(0.012)	26,090	0.969	0.101	(0.016)	10,966	0.949	0.135	(0.047)	4,533	0.926	No	Yes
	4	0.070	(0.014)	26,090	0.970	0.101	(0.018)	10,966	0.973	0.134	(0.045)	4,533	0.928	Yes	Yes
IV	5	0.003	(0.013)	26,090	0.968	0.106	(0.020)	10,966	0.947	0.207	(0.032)	4,533	0.924	No	No
	6	0.003	(0.013)	26,090	0.968	0.106	(0.021)	10,966	0.948	0.207	(0.032)	4,533	0.926	Yes	No
	7	0.086	(0.015)	26,090	0.969	0.104	(0.017)	10,966	0.949	0.167	(0.037)	4,533	0.926	No	Yes
	8	0.081	(0.017)	26,090	0.970	0.106	(0.018)	10,966	0.950	0.167	(0.036)	4,533	0.928	Yes	Yes
		Combined weekdays, Saturdays, and Sundays/Holidays[a]													
Panel D: The dependent variable is the natural logarithm of Metra rail ridership, at the Metra branch level.															
OLS	1	0.170	(0.017)	1,302	0.992									No	No
	2	0.170	(0.016)	1,302	0.993									Yes	No
IV	5	0.186	(0.016)	1,302	0.992									No	No
	6	0.186	(0.016)	1,302	0.992									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. The 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, wtotworld and wtotusa, (along with fixed effects and the station month trend) as instruments for the gasoline cost of driving in Chicago. Similar results are obtained using only crude oil prices, wtotworld, (along with fixed effects and the station month trend) as instruments. 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 the month level are reported in parentheses for all regressions. Similar results are obtained clustering standard errors at the unit level and at the unit-month levels. 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. The 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. The 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. The 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. The 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 the CTA rail ridership dataset, just disaggregation in the 14 branches. It 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 prices on Metra ridership due to collinearity issues, so year fixed effects are not included in the specifications in Panel D. Variables, fixed effects, and time trends definitions are in Appendix D.

Table 3: Vehicle Circulation Response to Gasoline Cost of Driving.

The dependent variable: Natural logarithm of vehicles at each ATR.

		ln(Gas.)	Sd. error	Nmbr. obs.	R^2	Covariates	ATR-specific month trend
Panel A: Vehicles from all ATRs in Illinois.							
	1	-0.034	(0.008)	8,333	0.991	No	No
OLS	2	-0.038	(0.007)	4,319	0.992	Yes	No
	3	-0.054	(0.026)	4,319	0.994	Yes	Yes
	4	-0.035	(0.007)	8,333	0.991	No	No
IV	5	-0.037	(0.007)	4,319	0.992	Yes	No
	6	-0.046	(0.025)	4,319	0.994	Yes	Yes
Panel B: Vehicles from ATRs in Chicago's expressways.[a]							
	1	-0.125	(0.030)	1,300	0.901	No	No
OLS	2	-0.615	(0.055)	1,138	0.999	Yes	No
	4	-0.144	(0.022)	1,300	0.901	No	No
IV	5	-0.615	(0.055)	1,138	0.999	Yes	No
Panel C: Vehicles from ATRs in urban areas.							
	1	-0.034	(0.009)	6,399	0.989	No	No
OLS	2	-0.043	(0.007)	3,292	0.991	Yes	No
	3	-0.039	(0.024)	3,292	0.993	Yes	Yes
	4	-0.035	(0.008)	6,399	0.989	No	No
IV	5	-0.041	(0.007)	3,292	0.991	Yes	No
	6	-0.028	(0.021)	3,292	0.993	Yes	Yes
Panel D: Vehicles from ATRs in rural areas.							
	1	-0.027	(0.013)	1,743	0.992	No	No
OLS	2	-0.019	(0.012)	1,027	0.993	Yes	No
	3	-0.115	(0.041)	1,027	0.995	Yes	Yes
	4	-0.022	(0.011)	1,743	0.992	No	No
IV	5	-0.011	(0.012)	1,027	0.993	Yes	No
	6	-0.111	(0.044)	1,027	0.995	Yes	Yes

Notes: All regressions include a constant, monthly seasonal fixed effects, and ATR-specific fixed effects. The 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, wtotworld and wtotusa, (along with fixed effects and the station month trend) as instruments for the gasoline cost of driving in Chicago. Similar results are obtained using only crude oil prices, wtotworld, (along with fixed effects and the station month trend) as instruments. Covariates include: direction of the route, good months indicator, Illinois functional classification of the road, and the 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 the month level are reported in parentheses for all regressions. Similar results are obtained clustering standard errors at the ATR level and at the ATR-month levels. 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 IDOT at the moment of the data collection). Some ATRs have 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: The 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 IDOT. Panel B: The dependent variable is the natural logarithm of the number of vehicles circulating through ATRs located in expressways in the Chicago area, as indicated by IDOT. The number of ATRs included is 28 (row 1). Panel C: The dependent variable is the natural logarithm of the number of vehicles circulating through ATRs located in urban areas, as indicated by IDOT. The number of ATRs included is 101 (row 1). Panel D: The dependent variable is the natural logarithm of the number of vehicles circulating through ATRs located in rural areas, as indicated by IDOT. The number of ATRs included is 27 (row 1).

[a] Regressions in rows 3 and 6 are not included in Panel B because of a collinearity issue due to having only few ATR stations with covariates' information in Chicago's expressways.

Table 4: CTA rail weekdays ridership and heterogeneity.

Interaction	Estimate	Nmbr. obs. where $\mathbf{1}\{\cdot\} = 1$
Panel A: Price trend in last three months.		
$ln(p_{ct})\times$		
$\mathbf{1}\{p_{ct} > p_{ct-1} > p_{ct-2}\}$	.1872	4,532
	(.0274)	
$\mathbf{1}\{p_{ct} < p_{ct-1} < p_{ct-2}\}$	.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]		
$ln(p_{ct})\times$		
$\mathbf{1}\{P_{ct} < 150\}$	.0028	1,570
	(.0151)	
$\mathbf{1}\{150 \leq P_{ct} < 250\}$	.1259	9,396
	(.0404)	
$\mathbf{1}\{250 \leq P_{ct} < 350\}$	.2048	3,792
	(.0268)	
$\mathbf{1}\{350 \leq P_{ct}\}$	.2537	967
	(.0143)	
Nmbr. obs.	15,725	15,725
R^2	0.9238	

Notes: The dependent variable is the natural logarithm of CTA rail ridership for weekdays. The independent variable is the natural logarithm of the gasoline cost of driving in real terms, $ln(p_{ct})$, interacted with the relevant indicators in each panel. Regressions are OLS specifications and include 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 2, Panel A, row 1, which has coefficient for $ln(p_{ct})$ equal to .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 2 (ridership using other means of public transportation). Robust standard errors clustered at the 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 $\mathbf{1}\{\text{A}\}$, 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 $\mathbf{1}\{p_{ct} > p_{ct-1} > p_{ct-2}\}$ or $\mathbf{1}\{p_{ct} < p_{ct-1} < p_{ct-2}\}$ hold.

[b]The absolute level of gasoline price is denoted with uppercase, P_{ct}.

Table 5: 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	55,432	51,282	-8.09%	52,438	2.20%
CTA Rail 1	981	1,101	10.90%	1,057	-4.16%
ATR urban 2	51,290	46,147	-11.14%	52,185	11.57%
CTA Rail 2	1,497	1,671	10.41%	1,611	-3.72%
ATR urban 3	115,189	97,817	-17.76%	98,402	0.59%
CTA Rail 3	2,610	3,086	15.42%	3,021	-2.15%
ATR urban 4	89,421	84,356	-6.00%	85,405	1.23%
CTA Rail 4	2,852	3,265	12.65%	3,191	-2.32%
ATR urban 5	92,600	87,123	-6.29%	88,544	1.60%
CTA Rail 5	3,376	3,893	13.28%	3,766	-3.37%
ATR urban 6	79,116	73,308	-7.92%	73,649	0.46%
CTA Rail 6	1,223	1,470	16.80%	1,455	-1.03%
ATR urban 7	84,914	81,070	-4.74%	82,557	1.80%
CTA Rail 7	1,006	1,128	10.82%	1,098	-2.73%
ATR urban 8	97,282	94,748	-2.67%	95,412	0.70%
CTA Rail 8	948	1,108	14.44%	1,037	-6.85%
ATR urban 9	102,489	101,031	-1.44%	101,923	0.88%
CTA Rail 9	684	773	11.51%	743	-4.04%
ATR urban 10	104,745	101,213	-3.49%	102,801	1.54%
CTA Rail 10	3,256	3,657	10.97%	3,448	-6.06%
ATR urban 11	111,576	108,061	-3.25%	109,977	1.74%
CTA Rail 11	3,108	3,449	9.89%	3,313	-4.11%
ATR urban 12	116,210	112,595	-3.21%	114,322	1.51%
CTA Rail 12	3,702	4,158	10.97%	3,985	-4.34%
ATR urban 13	112,880	107,758	-4.75%	109,052	1.19%
CTA Rail 13	4,872	5,437	10.39%	5,179	-4.98%
ATR urban 14	76,358	71,037	-7.49%	71,443	0.57%
CTA Rail 14	1,454	1,615	9.97%	1,562	-3.39%
ATR urban 15	93,579	90,998	-2.84%	92,392	1.51%
‘CTA Rail 15	4,573	4,929	7.22%	4,874	-1.13%
Panel B: Rural areas without access to public transit.					
ATR rural 1	9,221	8,638	-6.75%	9,287	6.99%
ATR rural 2	32,450	28,794	-12.70%	32,301	10.86%
ATR rural 3	3,572	3,536	-1.02%	3,599	1.75%
ATR rural 4	1,812	1,703	-6.40%	1,803	5.55%
ATR rural 5	19,187	17,557	-9.28%	18,825	6.74%

Notes: The Katrina hurricane formed on August 23, 2005 and dissipated on August 31, 2005. Pre-shock refers to the month before 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 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, 5 popular ATR locations without access to public transit are included. The term %Δ demand, denotes the percentage change in the number of vehicles (ridership) recorded by the ATR (CTA rail 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 $(51{,}282-55{,}432)/51{,}282 \times 100 = -8.09\%$ and post-shock as $(52{,}438-51{,}282)/52{,}438 \times 100 = 2.20\%$. Similarly, for each line in the table.

Table 6: 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	0.911	(0.693, 1.129)	0.890	(0.610, 1.170)	0.627	(0.489, 0.765)	0.618	(0.335, 0.900)
σ	0.039	(0.000, 0.121)	–	–	0.723	(0.604, 0.842)	–	–
ϕ_b	4.114	(3.541, 4.688)	4.138	(3.659, 4.617)	–	–	–	–
ϕ_c	4.350	(3.423, 5.277)	4.392	(3.440, 5.345)	–	–	–	–
Value of GMM Objective	81.976		81.987		255.674		258.260	

Panel B: Dynamic Model with Boundedly Rational Expectations BRE-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	0.528	(0.283, 0.773)	0.523	(0.270, 0.776)	0.952	(0.816, 1.088)	0.947	(0.794, 1.101)
σ	0.390	(0.094, 0.686)	–	–	0.682	(0.562, 0.802)	–	–
ϕ_b	3.302	(2.783, 3.820)	2.306	(1.428, 3.183)	–	–	–	–
ϕ_c	3.240	(2.468, 4.012)	3.144	(2.404, 3.885)	–	–	–	–
Value of GMM Objective	25.183		41.284		221.276		227.671	

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

	Estimates	95% C.I.
	with switching costs	
	with random coefficients	
α_1	0.610	(0.484, 0.736)
σ	0.342	(0.139, 0.545)
ϕ_b	2.089	(1.830, 2.348)
ϕ_c	2.219	(1.808, 2.631)
Value of GMM Objective	92.255	

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 Section 5 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 the car and public transit during the $\tilde{t} = 3$ periods before t, in route r, denoted respectively by $s_{brt|b\tilde{t}}(\vec{p}_t, \vec{\delta}_{rt}; \theta)$ and $s_{crt|c\tilde{t}}(\vec{p}_t, \vec{\delta}_{rt}; \theta)$. These additional moments are computed using the conditional value function and choice probabilities in (15), implemented in the inner subroutine in step 3 as described in Section 4 in the article. See Appendix A for details.

Table 7: Average Ratios of Elasticities Computed from Different Models Relative to the Long-Run Elasticity Computed from the Dynamic Model.

	Dynamic model short-run temporal	Dynamic model short-run permanent	Static model	Myopic model
Car				
Mean	0.391	0.449	0.346	0.190
Route:				
1	0.435	0.492	0.394	0.215
2	0.379	0.436	0.333	0.174
3	0.396	0.453	0.370	0.204
4	0.380	0.442	0.314	0.200
5	0.376	0.434	0.330	0.167
Public Transit				
Mean	0.333	0.394	0.304	0.074
Route:				
1	0.386	0.451	0.338	0.096
2	0.315	0.377	0.293	0.063
3	0.350	0.408	0.343	0.071
4	0.324	0.384	0.290	0.079
5	0.293	0.355	0.271	0.055
Outside option				
Mean	0.536	0.592	0.413	0.298
Route:				
1	0.475	0.535	0.423	0.245
2	0.517	0.574	0.401	0.273
3	0.709	0.759	0.539	0.377
4	0.525	0.580	0.360	0.342
5	0.541	0.599	0.414	0.278

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 terms *temporal* and *permanent* refer to the computation of the elasticities for a temporary and a permanent gasoline price increase, respectively, as described in the text. The long-run elasticity from the dynamic model is computed for a permanent gasoline price increase as indicated in the text. 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 costs and additional micro moments, from Panel C in Table 6. The static model corresponds to the static model with neither switching costs nor random coefficients, from the last columns in Panel A in Table 6. The myopic model corresponds to the myopic model with both switching costs and random coefficients, from the first columns in Panel A in Table 6. See Section 5 for details about the specifications of the models. See Section 5 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 Section 2.

Table 8: Elasticities Computed from Different Models.

		Elasticity Gasoline Price	Ratio Relative to Dynamic Long-Run
Car			
Reduced-form[a]	IV, month and unit fixed effects	-0.314	0.248
Structural[b]	Myopic	-0.240	0.190
	Static	-0.438	0.346
	Dynamic Short-Run	-0.494	0.391
	Dynamic Long-Run	-1.264	1.000
Public Transit			
Reduced-form[c]	IV, month and unit fixed effects	0.262	0.219
Structural[b]	Myopic	0.088	0.074
	Static	0.364	0.304
	Dynamic Short-Run	0.398	0.333
	Dynamic Long-Run	1.197	1.000

Notes: Same sample used in all cases for reduced-form and structural models. In all cases, the period included is January 2003 to December 2009 which corresponds to the data used for the estimation of the structural model (see also footnote 21). The estimated coefficients from all the models are statistically different from zero at the 1% level.

[a] Reduced-form model, car. The 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 sample of ATR stations corresponds to the same sample of ATR stations used for the estimation of the structural model. The independent variable is 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 regression includes constant, monthly seasonal fixed effects, and ATR-specific fixed effects, and it corresponds to two-stage least squares regressions (labeled "IV") using the natural logarithm of crude oil prices, wtotworld and wtotusa, (along with fixed effects) as instruments for the gasoline cost of driving in Chicago. Similar results are obtained using only crude oil prices, wtotworld, (along with fixed effects) as instruments. The elasticity reported in the table corresponds to the estimated coefficient of ln(Gas.), which corresponds to the baseline specification in Table 3, Panel B, row 4, using the sample specified there.

[b] The elasticities for the structural models correspond to the levels of the ratio of elasticities reported in Table 7. The long-run elasticity from the dynamic model is computed for a permanent gasoline price increase as indicated in the text. The short-run elasticity from the dynamic model is computed for a temporal gasoline price increase. 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; mean values across routes are reported. The dynamic model corresponds to the dynamic model with both switching costs and additional micro moments, from Panel C in Table 6. The static model corresponds to the static model with neither switching costs nor random coefficients, from the last columns in Panel A in Table 6. The myopic model corresponds to the myopic model with both switching costs and random coefficients, from the first columns in Panel A in Table 6. See Section 5 for details about the specifications of the models. See Section 5 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 Table 7, in the Chicago area used in the market definition as described in Section 2.

[c] Reduced-form model, public transit. The dependent variable is the natural logarithm of CTA rail ridership, at the CTA station level. The sample of CTA stations corresponds to the same sample of CTA stations used for the estimation of the structural model. The independent variable is 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 regression includes constant, monthly seasonal fixed effects, and unit of observation fixed effects (where the unit is the CTA station), and it corresponds to two-stage least squares regressions using the natural logarithm of crude oil prices, wtotworld and wtotusa (along with fixed effects) as instruments for the gasoline cost of driving in Chicago. Similar results are obtained using only crude oil prices, wtotworld, (along with fixed effects) as instruments. The elasticity reported in the table corresponds to the estimated coefficient of ln(Gas.). The regression reported corresponds to the baseline specification in Table 2, Panel A, row 5, using the sample specified there.

Table 9: 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 public transit (4)	Probability to car (5)
Panel A: Change in Switching Cost to Public Transit, ϕ_b.					
Baseline, without change in ϕ_b	–	0.836	1.484	0.057	0.098
Increase ϕ_b by 10%	-6.141	0.796	1.417	0.049	0.098
Increase ϕ_b by 20%	-12.547	0.746	1.384	0.041	0.099
Decrease ϕ_b by 10%	7.922	0.876	1.531	0.067	0.098
Decrease ϕ_b by 20%	17.633	0.895	1.575	0.072	0.098
Panel B: Change in Switching Cost to Public Transit, ϕ_b, and permanent gasoline tax increase of 1 dollar.					
Baseline without gasoline tax, without change in ϕ_b	–	0.836	1.484	0.057	0.098
Baseline with gasoline tax and:					
without change in ϕ_b	-30.792	0.965	1.329	0.083	0.061
Increase ϕ_b by 10%	-37.232	0.925	1.301	0.074	0.062
Increase ϕ_b by 20%	-48.139	0.885	1.236	0.066	0.062
Decrease ϕ_b by 10%	-22.530	0.985	1.375	0.092	0.062
Decrease ϕ_b by 20%	-10.817	1.004	1.438	0.102	0.061

Notes: Counterfactual analysis using the estimates in Table 6, Panel C. See Section 5 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 equation (7). The baseline value of the switching costs are $\phi_b = 2.09$ and with $\phi_c = 2.22$.

Appendix (For Online Publication)

Contents (Appendix)

B 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's monthly ridership for the 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 1. I provide details for each of these sources and the data. The next subsections follow closely the relevant descriptions as stated by each source.

B.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 weekdays, 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 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 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 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 system is part of the 138,000-mile network of state, county, municipal, township, and toll roads. This system is the third-largest system in the nation. IDOT conducts an annual monitoring program consisting of volume counting, vehicle classification, and truck weighing.

According to 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 the hour and vehicle type is polled, reviewed, and transferred to a database. Raw (hourly) classification data for the period 2001-2009 was collected. For this study, these data were aggregated at the monthly level to be consistent with the frequency of 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. Seven major interstate highways cross 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 article. 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).

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

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.

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 article.

7. Census Transportation Planning Package (CTPP). I use the Census Transportation Planning Package (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 the 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. It 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 locations of the ATRs. I analyze spatial information by integrating geographic information about Chicago's 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 each bus line that lies within the radius. Figure 2 in the article 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 four ATR stations. For Dan Ryan expressway using three ATR stations. For Edens expressway using three ATR stations. For Eisenhower expressway and Eisenhower extension using nine ATR stations. Finally, for Kennedy expressway using seven ATR stations.[67] 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.[68] I calculate the population that extends over the area defined by the radius using population data disaggregated by zip code assuming population is distributed homogeneously over the zip code area.

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 selected 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 disaggregated than the market-level data above. It describes the behavior of consumer types who, *e.g.*, did not switch modes of transportation and their income. 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.

[67] Classification data is not available for these ATRs. It is therefore not possible to distinguish into categories depending on whether the vehicle carries passengers or commodities. This feature introduces two empirical challenges. First, that passenger-vehicle counts may react differently than total-vehicle counts 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 C, I use classification data for other ATRs in Chicago and Illinois and show that passenger-vehicle and total-vehicle counts exhibit a similar response to the gasoline cost of driving. For details see Table A2 described in Appendix C.

[68] For the specifications reported in Table 6, I performed the estimation increasing and decreasing the radius definition (market size) by 10 percent and obtained similar results.

C 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 taking into account the monthly seasonal effects. The regressions below exploit cross-sectional and time-series variation in the data by including subarea disaggregation within the 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.

For robustness, I also report results from two-stage least squares regressions (henceforth, IV) 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 price 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.[69]

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. 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.

C.1 Public Transit Use

C.1.1 CTA

Panel A in Table 2 in the article 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

[69] 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. Another 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.

effects to the specification in row (1). Finally, row 4 adds both a station-specific monthly time trends 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 elasticity estimates vary from .04 percent to .17 percent, depending on the specification.

Columns Saturdays and Sunday/Holiday show analog 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 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 2 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.

C.1.2 Pace

Panel C in Table 2 in the article 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.

C.1.3 Metra

Panel D in Table 2 in the article 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 weekdays, Saturdays, and Sundays/Holidays. It does not provide station disaggregation either, unlike the CTA rail ridership dataset, just disaggregation in the 14 branches. It 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 trends 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 elasticity values vary from .07 percent to .17 percent.

C.2 Vehicle Circulation

C.2.1 Rural and Urban Areas

Panel A in Table 3 in the article 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 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 the number of lanes), and ATR-specific monthly time trends 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 3 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 restrict 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. It is because urban areas have better access to public transportation services than rural areas.

C.2.2 Classification Data for the Chicago Metropolitan Area

Table A2 presents estimates for the subset of ATRs in the Chicago metropolitan area that collect classification data. It is a subsample of the one in Table 3. 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 analog estimates to the ones in Panel A in Table 3, 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 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 A2 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 truck circulation (classes 6 to 10) strongly decreases as the gasoline cost of driving increases, while multi-trailer truck circulation (classes 11 to 13) increases. These results suggest that, as the gasoline cost of driving increases, there is a substitution from single-trailer trucks to multi-trailer trucks, a possible efficiency effect.

Table A2: Vehicle Circulation Response in Chicago, by Class Type.

The dependent variable: Natural logarithm of vehicles at each ATR.

	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	-0.194	(0.071)	1,301	0.696	—
- Passenger	-0.157	(0.073)	1,301	0.690	-0.173	(0.068)	1,301	0.690	—
- Single-unit trucks	-2.065	(0.112)	1,233	0.500	-2.299	(0.103)	1,233	0.499	—
- Multiple-units trucks	-1.971	(0.117)	1,227	0.518	-2.192	(0.110)	1,227	0.517	—
Panel B: Classification data by each class.									
- Class 1	0.656	(0.231)	1,231	0.587	0.778	(0.254)	1,231	0.587	.43
- Class 2	-0.043	(0.074)	1,301	0.701	-0.051	(0.068)	1,301	0.701	88.32
- Class 3	-2.142	(0.141)	1,229	0.438	-2.294	(0.131)	1,229	0.437	11.25
- Class 4	-3.014	(0.103)	1,227	0.593	-3.291	(0.093)	1,227	0.591	12.94
- Class 5	-2.548	(0.092)	1,231	0.434	-2.734	(0.084)	1,231	0.433	59.03
- Class 6	-0.988	(0.130)	1,217	0.539	-1.105	(0.137)	1,217	0.538	25.91
- Class 7	-2.908	(0.157)	1,172	0.399	-3.215	(0.153)	1,172	0.397	2.12
- Class 8	-2.353	(0.131)	1,218	0.511	-2.529	(0.119)	1,218	0.511	27.62
- Class 9	-1.688	(0.149)	1,209	0.514	-1.860	(0.149)	1,209	0.514	64.73
- Class 10	-1.455	(0.163)	1,169	0.496	-1.625	(0.138)	1,169	0.496	3.00
- Class 11	0.851	(0.146)	1,013	0.668	0.971	(0.143)	1,013	0.667	1.10
- Class 12	4.224	(0.218)	938	0.625	4.674	(0.221)	938	0.623	2.29
- Class 13	0.947	(0.152)	1,110	0.551	1.157	(0.164)	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 3 in the article). Information about covariates is not available for the ATRs that collect classification data. The 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, wtotworld and wtotusa, (along with fixed effects and the station month trend) as instruments for the gasoline cost of driving in Chicago. Similar results are obtained using only crude oil prices, wtotworld, (along with fixed effects and the station month trend) as instruments. Robust standard errors clustered at the month level are reported in parentheses for all regressions. Similar results are obtained clustering standard errors at the ATR level and at the ATR-month levels. 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 IDOT) to October 2009 (the latest month with vehicle counts classification information provided by IDOT at the moment of the data collection). Some ATRs have 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 refer to the raw total number of vehicles recorded by the ATRs. Passenger refers to the raw total number of vehicles in classes 1, 21, 2, and 3 and 3, as recorded by the ATRs. Single-unit trucks refer to the raw total number of vehicles in classes 4, 5, 6 and 7, as recorded by the ATRs. Multiple-unit trucks refer 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-unit 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.3 Robustness Analysis: Katrina Shock and Hysteresis

Elasticity during Katrina. For the stations considered, the gasoline price elasticities during Hurricane Katrina calculated using Table A4 are sensible after proper interpretation. Consider Panel A, for urban areas. For vehicle circulation (public transit use) the pre-shock-weighted mean change in vehicle counts (public transit ridership) during the Katrina shock is -3.83 (2.33) percent; the own- (cross-) price elasticity is -0.57 (0.35) during the Katrina shock in urban areas in the stations considered. For vehicle circulation (public transit use) the pre-shock-weighted mean change in demand post-Katrina shock is 0.71 (-1.66) percent; the own- (cross-) price elasticity is -0.10 (0.23) post-Katrina shock in urban areas in the stations considered. The post-shock elasticities are similar in magnitude to the reduced-form elasticities in the Table A3 that show average own- (cross-) price elasticities of -0.11 in Panel B (0.22 in Panel A) for the IV specifications; the elasticities during the shock are higher consistent with the evidence displayed in Table 4. Similarly for rural areas. Rural areas display more inelastic elasticities for vehicle circulation, also consistent with Table 3.

Table A3: Katrina Shock Regressions.

The dependent variable: Natural logarithm at each CTA station or ATR.

	ln(Gas.)	St. error	Nmbr. obs.	R^2	Months fixed effects	Unit fixed effects
Panel A: Public Transit in Urban Areas						
OLS	0.187	(0.013)	1,657	0.972	Yes	Yes
IV	0.215	(0.011)	1,657	0.972	Yes	Yes
Panel B: Vehicles in Urban Areas						
OLS	-0.092	(0.029)	682	0.932	Yes	Yes
IV	-0.109	(0.020)	682	0.932	Yes	Yes
Panel C: Vehicles in Rural Areas						
OLS	-0.038	(0.020)	416	0.996	Yes	Yes
IV	-0.039	(0.018)	416	0.996	Yes	Yes

Notes: All regressions include a constant, monthly seasonal fixed effects, and unit of observation fixed effects, where the unit is the CTA station or ATR station in Panels A or B/C, respectively. The 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 row labeled IV corresponds to two-stage least squares regressions using the natural logarithm of crude oil prices, wtotworld and wtotusa, (along with fixed effects and the station month trend) as instruments for the gasoline cost of driving in Chicago. Similar results are obtained using only crude oil prices, wtotworld, (along with fixed effects and the station month trend) as instruments. Robust standard errors clustered at the month level are reported in parentheses for all regressions. Similar results are obtained clustering standard errors at the unit level and at the unit-month levels. Panel A: The dependent variable is the natural logarithm of the monthly CTA rail ridership on weekdays. The number of stations included is 15, corresponding to the stations in Tables 5 and A4, Panel A. 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). Panel B: The dependent variable is the natural logarithm of the number of vehicles circulating through ATRs located in urban areas, as indicated by IDOT. The number of stations included is 15, corresponding to the stations in Tables 5 and A4, Panel A. 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 IDOT at the moment of the data collection). Some ATRs have missing observations on specific months. Panel C: The dependent variable is the natural logarithm of the number of vehicles circulating through ATRs located in rural areas, as indicated by IDOT. The number of stations included is 5, corresponding to the stations in Tables 5 and A4, Panel B. 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 IDOT at the moment of the data collection). Some ATRs have missing observations on specific months. Variables and fixed effects definitions are in Appendix D.

Table A4: Katrina Shock and Hysteresis using Seasonally Adjusted Data.

Locations	Pre-shock	Shock	%Δ demand	Post-shock	%Δ demand
Panel A: Urban areas with access to public transit.					
ATR urban 1	55,914.86	53,614.78	-4.29%	53,852.53	0.44%
CTA rail 1	994.76	1,006.26	1.14%	997.55	-0.87%
ATR urban 2	51,291.86	49,264.85	-4.11%	50,145.48	1.76%
CTA rail 2	1,494.92	1,532.05	2.42%	1,508.71	-1.55%
ATR urban 3	120,910.50	106,503.00	-13.53%	108,109.50	1.49%
CTA rail 3	2,823.64	2,886.11	2.16%	2,873.74	-0.43%
ATR urban 4	85,967.49	83,519.67	-2.93%	84,160.41	0.76%
CTA rail 4	3,034.68	3,092.06	1.86%	3,078.51	-0.44%
ATR urban 5	91,908.78	88,580.63	-3.76%	89,703.38	1.25%
CTA rail 5	3,530.08	3,583.54	1.49%	3,550.62	-0.93%
ATR urban 6	85,878.30	77,969.72	-10.14%	78,809.42	1.07%
CTA rail 6	1,296.38	1,347.57	3.80%	1,323.01	-1.86%
ATR urban 7	85,554.96	81,546.45	-4.92%	82,518.82	1.18%
CTA rail 7	1,031.36	1,047.74	1.56%	1,042.33	-0.52%
ATR urban 8	97,633.47	96,478.84	-1.20%	96,488.55	0.01%
CTA rail 8	1,016.67	1,058.65	3.97%	1,029.62	-2.82%
ATR urban 9	102,295.70	100,717.70	-1.57%	100,972.60	0.25%
CTA rail 9	689.03	706.63	2.49%	697.43	-1.32%
ATR urban 10	104,600.60	102,790.50	-1.76%	103,346.70	0.54%
CTA rail 10	3,305.55	3,369.44	1.90%	3,343.31	-0.78%
ATR urban 11	111,748.10	110,158.30	-1.44%	110,315.30	0.14%
CTA rail 11	3,172.65	3,210.51	1.18%	3,186.06	-0.77%
ATR urban 12	114,580.50	113,159.70	-1.26%	114,527.80	1.19%
CTA rail 12	3,790.39	3,914.67	3.17%	3,724.57	-5.10%
ATR urban 13	110,736.20	109,110.80	-1.49%	109,316.00	0.19%
CTA rail 13	4,929.71	5,191.54	5.04%	4,995.91	-3.92%
ATR urban 14	77,488.83	74,902.98	-3.45%	75,173.70	0.36%
CTA rail 14	1,491.57	1,507.50	1.06%	1,501.84	-0.38%
ATR urban 15	91,310.96	89,926.51	-1.54%	90,221.84	0.33%
CTA rail 15	4,587.30	4,628.98	0.90%	4,610.34	-0.40%
Panel B: Rural areas without access to public transit.					
ATR rural 1	9,501.27	9,247.19	-2.67%	9,527.98	3.04%
ATR rural 2	33,957.63	33,370.36	-1.73%	33,669.36	0.90%
ATR rural 3	3,637.23	3,580.87	-1.55%	3,648.93	1.87%
ATR rural 4	1,906.39	1,840.50	-3.46%	1,917.69	4.19%
ATR rural 5	20,853.93	20,558.63	-1.42%	20,627.59	0.34%

Notes: The Katrina hurricane formed on August 23, 2005 and dissipated on August 31, 2005. Pre-shock refers to the month before 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 the seasonally adjusted ATRs vehicle counts, and seasonally adjusted CTA rail ridership in each location. Data were seasonally adjusted using the X-13ARIMA-SEATS seasonal adjustment software (using its automatic model selection procedure) produced, distributed, and maintained by the U.S. Census Bureau (U.S. Census Bureau 2017). 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, 5 popular ATR locations without access to public transit are included. The term %Δ demand denotes the percentage change in the number of vehicles (ridership) recorded 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 $(53{,}614.78-55{,}914.86)/53{,}614.78 \times 100 = -4.29\%$, and post-shock as $(53{,}852.53-53{,}614.78)/53{,}852.53 \times 100 = 0.44\%$. Similarly, for each line in the table.

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 weekdays, 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 weekdays, 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 weekdays, 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 weekdays, 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-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: 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 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 the 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 U.S. 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.

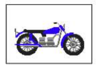
ATR-specific fixed effects. A set of dummy variables, each corresponding to ATRs where the vehicle counts are registered.

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

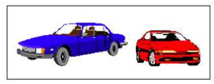
This subsection was obtained directly from the classification by Federal Highway Administration (FHWA). The FHWA classification scheme is separated into categories depending on whether the vehicle carries passengers or commodities. Non-passenger vehicles are subdivided by the number of axles and number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



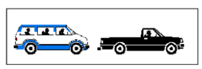
Class 1. Source: FHWA Vehicle Classification Figures.

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



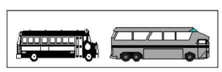
Class 2. Source: FHWA Vehicle Classification Figures.

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.



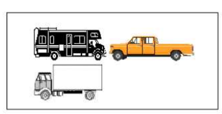
Class 3. Source: FHWA Vehicle Classification Figures.

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.



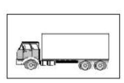
Class 4. Source: FHWA Vehicle Classification Figures.

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.



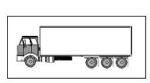
Class 5. Source: FHWA Vehicle Classification Figures.

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.



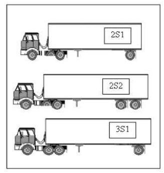
Class 6. Source: FHWA Vehicle Classification Figures.

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



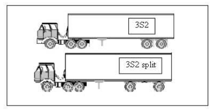
Class 7. Source: FHWA Vehicle Classification Figures.

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.



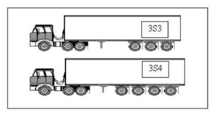
Class 8. Source: FHWA Vehicle Classification Figures.

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.



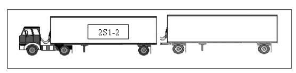
Class 9. Source: FHWA Vehicle Classification Figures.

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.



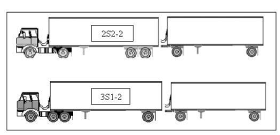
Class 10. Source: FHWA Vehicle Classification Figures.

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.



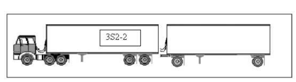
Class 11. Source: FHWA Vehicle Classification Figures.

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 Analysis: Structural Estimates

F.1 Alternative Specifications of Expectations

Table A5: Robustness. Structural Estimates using Alternative Specifications of Expectations.

	Estimates	95% C.I.	Estimates	95% C.I.	Estimates	95% C.I.	Estimates	95% C.I.
Panel A: Dynamic Model with Boundedly Rational Expectations BRE-FE.								
	with switching costs				without switching costs			
	with random coefficients		without random coefficients		with random coefficients		without random coefficients	
α_1	0.566	(0.114, 1.019)	0.545	(0.193, 0.897)	0.647	(0.555, 0.740)	0.633	(0.536, 0.731)
σ	0.133	(0.005, 0.262)	–	–	0.682	(0.564, 0.800)	–	–
ϕ_b	2.794	(2.432, 3.156)	2.177	(1.357, 2.878)	–	–	–	–
ϕ_c	3.053	(1.319, 4.786)	3.162	(2.330, 3.994)	–	–	–	–
Value of GMM Objective	37.680		48.777		221.276		227.671	
Panel B: Dynamic Model with Perfect Foresight PF.								
	with switching costs							
	with random coefficients		without random coefficients					
α_1	0.621	(0.033, 1.275)	0.614	(0.496, 0.732)				
σ	0.197	(0.004, 0.390)	–	–				
ϕ_b	0.726	(0.232, 1.220)	0.780	(0.272, 1.287)				
ϕ_c	1.326	(0.617, 2.034)	1.223	(0.900, 1.545)				
Value of GMM Objective	138.527		143.253					

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 Section 5 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 the car and public transit.

F.2 Alternative Specifications of Micro Moments

Table A6: 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 Boundedly Rational Expectations BRE-AR and Additional Moments.						
	$\tilde{t} = 3$ months		$\tilde{t} = 6$ months		$\tilde{t} = 12$ months	
α_1	0.610	(0.484, 0.736)	0.543	(0.455, 0.632)	0.531	(0.228, 0.833)
σ	0.342	(0.139, 0.545)	0.261	(0.000, 0.644)	0.364	(0.048, 0.680)
ϕ_b	2.089	(1.830, 2.348)	2.953	(1.296, 4.609)	3.098	(1.863, 4.333)
ϕ_c	2.219	(1.808, 2.631)	2.753	(1.205, 4.301)	3.210	(2.074, 4.346)
Value of GMM Objective	92.255		46.226		28.066	
Panel B: Dynamic Model with Boundedly Rational Expectations BRE-FE and Additional Moments.						
	$\tilde{t} = 3$ months		$\tilde{t} = 6$ months		$\tilde{t} = 12$ months	
α_1	0.731	(0.571, 0.891)	0.660	(0.517, 0.802)	0.592	(0.494, 0.690)
σ	0.283	(0.088, 0.478)	0.268	(0.067, 0.469)	0.218	(0.084, 0.352)
ϕ_b	2.448	(2.041, 2.854)	2.745	(2.261, 3.229)	1.707	(1.547, 1.867)
ϕ_c	2.041	(1.648, 2.434)	2.643	(2.045, 3.242)	1.265	(1.152, 1.378)
Value of GMM Objective	99.452		53.352		27.190	

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 Section 5 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 the car and public transit during the $\tilde{t} \in \{3, 6, 12\}$ periods before t, in route r, denoted respectively by $s_{brt|b\tilde{t}}(\vec{p}_t, \vec{\delta}_{rt}; \theta)$ and $s_{crt|c\tilde{t}}(\vec{p}_t, \vec{\delta}_{rt}; \theta)$. These additional moments (no-switching moments) are computed using the conditional value function and choice probabilities in (15), implemented in the inner subroutine in step 3 as described in Section 4 in the article. See Appendix A for details.

F.3 Month of the Year as Additional State Variable

F.3.1 Model Estimates

Table A7: Robustness Using Month of the Year as Additional State Variable. Model Estimates.

Panel A: Myopic Expectations and Static Model.

	Myopic Expectations ME				Static Model 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	0.906	(0.679, 1.134)	0.901	(0.641, 1.160)	0.629	(0.486, 0.773)	0.542	(0.338, 0.746)
σ	0.045	(0.000, 0.164)	–	–	0.772	(0.642, 0.901)	–	–
ϕ_b	4.140	(3.555, 4.725)	4.141	(3.632, 4.650)	–	–	–	–
ϕ_c	4.257	(3.423, 5.091)	4.248	(3.373, 5.123)	–	–	–	–
Value of GMM Objective	80.441		80.443		7.145		38.036	

Panel B: Dynamic Model with Boundedly Rational Expectations BRE-AR.

	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	0.462	(0.134, 0.790)	0.522	(0.236, 0.808)	0.685	(0.587, 0.783)	0.587	(0.496, 0.677)
σ	0.294	(0.000, 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 Boundedly Rational Expectations BRE-AR and Additional Moments.

	with switching costs	
	with random coefficients	
	Estimates	95% C.I.
α_1	0.526	(0.395, 0.657)
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. The data used for the estimation is not seasonally adjusted. The model is the same as in Section 3 adding "month of the year" as state variable; static models include monthly seasonal fixed effects. Details about the estimation procedure are in Section 4. See Section 5 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 the car and public transit during the $\tilde{t} = 3$ periods before t, in route r, denoted respectively by $s_{brt|b\tilde{t}}(\vec{p}_t, \vec{\delta}_{rt}; \theta)$ and $s_{crt|c\tilde{t}}(\vec{p}_t, \vec{\delta}_{rt}; \theta)$. These additional moments are computed using the conditional value function and choice probabilities in (15), implemented in the inner subroutine in step 3 as described in Section 4 in the article. See Appendix A for details.

F.3.2 Alternative Specifications of Expectations

Table A8: Robustness Using Month of the Year as Additional State Variable. Structural Estimates using Alternative Specifications of Expectations.

	Estimates	95% C.I.	Estimates	95% C.I.	Estimates	95% C.I.	Estimates	95% C.I.
Panel A: Dynamic Model with Boundedly Rational Expectations BRE-FE.								
	with switching costs				without switching costs			
	with random coefficients		without random coefficients		with random coefficients		without random coefficients	
α_1	0.634	(0.019, 1.249)	0.552	(0.027, 1.077)	0.685	(0.585, 0.786)	0.587	(0.494, 0.679)
σ	0.163	(0.000, 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					
α_1	0.589	(0.064, 1.114)	0.583	(0.460, 0.706)				
σ	0.007	(0.000, 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. The data used for the estimation is not seasonally adjusted. The model is the same as in Section 3 adding "month of the year" as state variable; static models include monthly seasonal fixed effects. Details about the estimation procedure are in Section 4. See Section 5 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 the car and public transit.

F.3.3 Alternative Specifications of Micro Moments

Table A9: Robustness Using Month of the Year as Additional State Variable. Structural Estimates using Alternative Specifications of Micro Moments and Expectations.

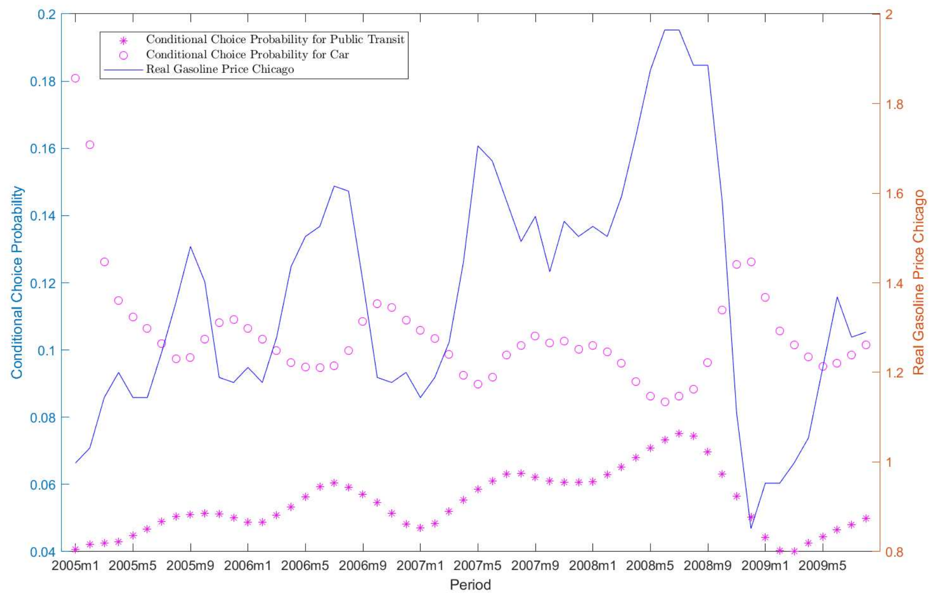
	Estimates	95% C.I.	Estimates	95% C.I.	Estimates	95% C.I.
Panel A: Dynamic Model with Boundedly Rational Expectations BRE-AR and Additional Moments.						
	$\tilde{t} = 3$ months		$\tilde{t} = 6$ months		$\tilde{t} = 12$ months	
α_1	0.526	(0.395, 0.657)	0.492	(0.416, 0.568)	0.458	(0.085, 0.831)
σ	0.308	(0.041, 0.576)	0.194	(0.000, 0.574)	0.275	(0.000, 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 Boundedly Rational Expectations BRE-FE and Additional Moments.						
	$\tilde{t} = 3$ months		$\tilde{t} = 6$ months		$\tilde{t} = 12$ months	
α_1	0.628	(0.503, 0.753)	0.591	(0.465, 0.717)	0.521	(0.439, 0.603)
σ	0.179	(0.000, 0.332)	0.193	(0.000, 0.343)	0.207	(0.000, 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	(3.458, 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. The data used for the estimation is not seasonally adjusted. The model is the same as in Section 3 adding "month of the year" as state variable; static models include monthly seasonal fixed effects. Details about the estimation procedure are in Section 4. See Section 5 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 the car and public transit during the $\tilde{t} \in \{3, 6, 12\}$ periods before t, in route r, denoted respectively by $s_{brt|b\tilde{t}}(\vec{p}_t, \vec{\delta}_{rt}; \theta)$ and $s_{crt|c\tilde{t}}(\vec{p}_t, \vec{\delta}_{rt}; \theta)$. These additional moments (no-switching moments) are computed using the conditional value function and choice probabilities in (15), implemented in the inner subroutine in step 3 as described in Section 4 in the article. See Appendix A for details.

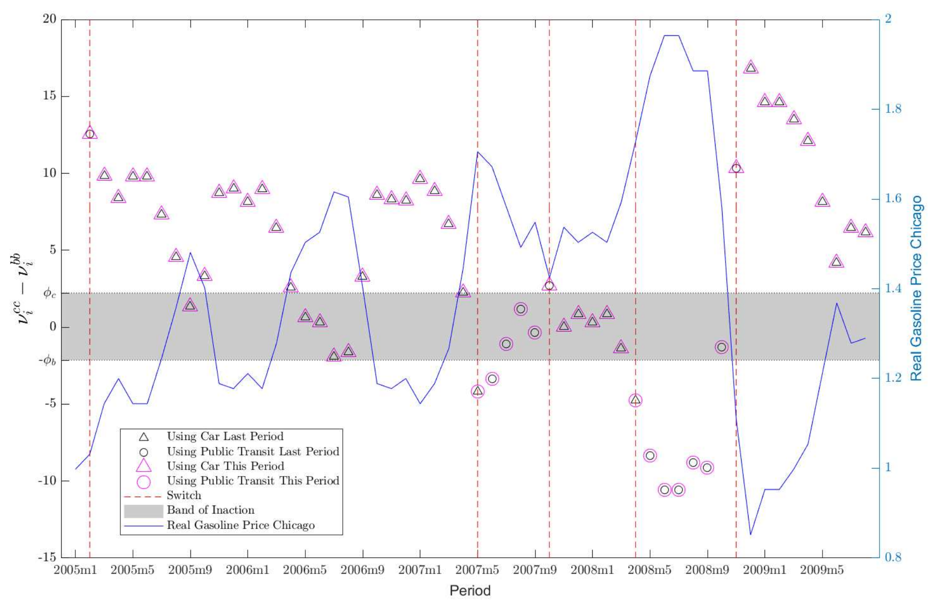
F.3.4 Figures

Figure A1: Robustness Using Month of the Year as Additional State Variable. Estimated model.

A. Conditional Choice Probabilities and Gasoline Prices.



B. Switching Behavior of a Transit-Friendly Consumer.



Notes: Figure A1A shows the mean, across consumer types and markets, conditional choice probabilities (left vertical axis) given by equation (7), for public transit and car, holding fixed ξ_{mrt}, as a function of the observed gasoline cost of driving (right vertical axis), $\vec{p}_t$, for the period 2005-2009. Figure A1B 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}. The figure shows on the left vertical axis using dots the difference in the conditional value function of choosing car transport today given the choice to travel by car in the previous period, minus the conditional value function of choosing public transit today given a previous-period choice of public transit for this particular transit-friendly consumer type: $v_i^{cc} - v_i^{bb}$, with $v_i^{jr} \equiv v_i(\vec{p}_t, m_{irt-1} = j, m_{irt} = r, \vec{\delta}_{rt}, \nu_i; \theta)$, where the superscripts j and r index, respectively, the mode of transportation of consumer i in $t-1$ and t, and the conditional value function $v_i(\vec{p}_t, m_{irt-1}, m_{irt}, \vec{\delta}_{rt}, \nu_i; \theta)$ is given by equation (5). The vertical dotted lines indicate a switch in the inside mode of transportation by the consumer in this period. The graph also displays the gasoline cost of driving on the right vertical axis using a solid line and the estimated switching-cost parameters, ϕ_c and $-\phi_b$, using horizontal lines. 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). Both figures use the estimates from the dynamic model with both switching costs and additional micro moments, from Panel C in Table A7. See Section 5 for a description of the switching behavior.

F.4 Start-Up Costs and Switching Costs

Table A10: Structural Estimates with Start-Up Costs and Switching Costs.

	Dynamic Model with Boundedly Rational Expectations BRE-AR and Additional Moments. with switching costs with random coefficients	
	Estimates	95% C.I.
α_1	0.582	(0.460, 0.704)
σ	0.329	(0.085, 0.573)
ϕ_b^1	3.800	(2.622, 4.978)
ϕ_c^1	3.800	(2.622, 4.978)
ϕ_b^2	0.515	(0, 1.362)
ϕ_c^2	0.515	(0, 1.362)
Value of GMM Objective	94.620	

Notes: Estimates of selected parameters from the main specification of the structural model. Same specification as in Table 6, Panel C, with the following modifications. The specification allows that the *start-up cost* (initial fixed cost of choosing the inside mode for the first time) denoted by ϕ_b^1 and ϕ_c^1 to differ from the *switching cost* (fixed cost of subsequently switching to the inside mode after the first time) denoted by ϕ_b^2 and ϕ_c^2. For the estimation, it is assumed that the start-up (switching) costs are the same for the inside modes of transportation, $\phi_b^1 = \phi_c^1$ ($\phi_b^2 = \phi_c^2$) for the reasons discussed in Section 3. However, the start-up and switching costs are specific to the mode of transportation; that is, the start-up cost of inside mode m is incurred the first time that the consumer switches to m, even if the consumer previously incurred the start-up cost of the other inside mode $\tilde{m} \neq m$. No constraints are imposed in terms of the relative magnitude of the start-up and switching costs. See Section 3 for details. 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 Section 5 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 (no-switching moments) included in this panel correspond to the share of the consumers who did not switch from the car and public transit during the $\tilde{t} = 3$ periods before t, in route r, denoted respectively by $s_{brt|b\tilde{t}}(\vec{p}_t, \vec{\delta}_{rt}; \theta)$ and $s_{crt|c\tilde{t}}(\vec{p}_t, \vec{\delta}_{rt}; \theta)$. These additional moments are computed using the conditional value function and choice probabilities in (15), implemented in the inner subroutine in step 3 as described in Section 4 in the article. See Appendix A for details.

F.5 Great Recession

The Great Recession began in December 2007 and ended in June 2009 for a total of 18 months according to the definition by the National Bureau of Economic Research (NBER).[70] I performed a robustness analysis to all the specifications in Table 6 and the elasticities reported in Table 7. To do that, I estimated all the models for a subsample not including the months during the Great Recession, where one suspects that income effects might vary—over time or by household size, location, or other characteristics. Table A11 reports the estimation results for the main specification of the structural model (similar results obtained for the other models). Table A12, which is similar to Table 7, reports the ratios of elasticities. The robustness analysis reveals remarkable consistency in the estimates and ratios of elasticities discussed in the article. Although some parameters are less precisely estimated, the implications discussed in section 5 are the same.

F.5.1 Model Estimates

Table A11: Robustness Great Recession. Model Estimates.

Dynamic Model with Boundedly Rational Expectations BRE-AR and Additional Moments.
with switching costs
with random coefficients

	Estimates	95% C.I.
α_1	0.836	(0.535, 1.137)
σ	0.403	(0.000, 0.809)
ϕ_b	1.936	(0.708, 3.164)
ϕ_c	1.195	(0.030, 2.360)
Value of GMM Objective	41.593	

Notes: Estimates of selected parameters from the main specification of the structural model. Same specification as in Table 6, Panel C, not including the months during the Great Recession, from December 2007 until June 2009 (*i.e.*, from December 2007 onwards given my sample). 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 Section 5 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 (no-switching moments) included in this panel correspond to the share of the consumers who did not switch from the car and public transit during the $\tilde{t} = 3$ periods before t, in route r, denoted respectively by $s_{brt|b\tilde{t}}(\vec{p}_t, \vec{\delta}_{rt}; \theta)$ and $s_{crt|c\tilde{t}}(\vec{p}_t, \vec{\delta}_{rt}; \theta)$. These additional moments are computed using the conditional value function and choice probabilities in (15), implemented in the inner subroutine in step 3 as described in Section 4 in the article. See Appendix A for details.

[70] NBER, US Business Cycle Expansions and Contractions, accessed May 19, 2020, https://www.nber.org/cycles.html.

F.5.2 Elasticities

Table A12: Robustness Great Recession. Average Ratios of Elasticities Computed from Different Models Relative to the Long-Run Elasticity Computed from the Dynamic Model.

	Dynamic model short-run temporal	Dynamic model short-run permanent	Static model	Myopic model
Car				
Mean	0.411	0.476	0.323	0.164
Route:				
1	0.465	0.506	0.391	0.193
2	0.383	0.467	0.305	0.142
3	0.426	0.479	0.365	0.169
4	0.401	0.474	0.304	0.162
5	0.379	0.456	0.248	0.153
Public Transit				
Mean	0.353	0.417	0.258	0.052
Route:				
1	0.395	0.452	0.269	0.062
2	0.333	0.405	0.261	0.030
3	0.382	0.439	0.248	0.063
4	0.356	0.407	0.287	0.073
5	0.298	0.380	0.227	0.030
Outside option				
Mean	0.576	0.627	0.375	0.274
Route:				
1	0.507	0.560	0.385	0.197
2	0.549	0.587	0.324	0.256
3	0.725	0.781	0.459	0.348
4	0.552	0.586	0.341	0.331
5	0.546	0.623	0.365	0.240

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 models use the same specifications as in Table 6, but they are all re-estimated not including the months during the Great Recession, from December 2007 until June 2009 (*i.e.*, from December 2007 onwards given my sample). The terms *temporal* and *permanent* refer to the computation of the elasticities for a temporary and a permanent gasoline price increase, respectively, as described in the text. The long-run elasticity from the dynamic model is computed for a permanent gasoline price increase as indicated in the text. 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 costs and additional micro moments. The static model corresponds to the static model with neither switching costs nor random coefficients. The myopic model corresponds to the myopic model with both switching costs and random coefficients. See Section 5 for details about the specifications of the models. See Section 5 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 Section 2.

G Additional Distributional Analysis of a Gasoline Tax Increase

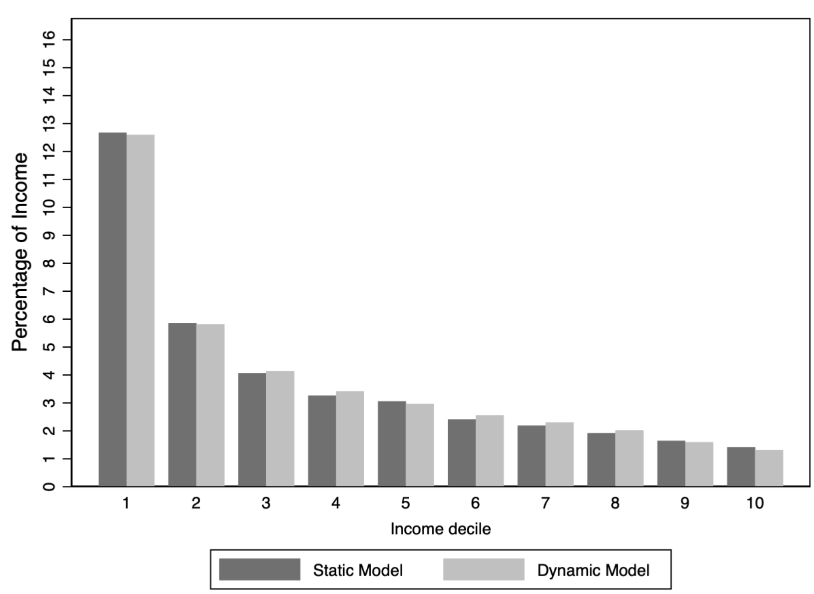
Table A13: Gasoline Expenditures to Income Before/After the Gasoline Tax Increase, by Income Decile.

	Gasoline Expenditure to Income			
	Before	After tax increase		
Income Decile	tax increase (1)	Static Model (2)	Dynamic Model (3)	Ratio (static/dynamic) (4)
1	12.628	15.415	8.316	1.854
2	5.898	7.333	4.703	1.559
3	4.102	5.150	3.615	1.425
4	3.333	4.193	2.982	1.406
5	2.882	3.605	2.499	1.443
6	2.405	3.020	2.126	1.420
7	2.179	2.742	1.964	1.396
8	1.878	2.366	1.731	1.367
9	1.448	1.812	1.285	1.411
10	1.200	1.504	1.079	1.394

Notes: The table displays the gasoline expenditures as a share of income for consumers in different deciles of the pretax income distribution: (1) before the gasoline tax increase from the data, (2) after the gasoline tax increase predicted by the static model, (3) after the gasoline tax increase predicted by the dynamic model, and (4) the ratio of the prediction of the static model relative to the prediction of the dynamic model. See Section 5 for details.

Figure A2: Predictions of Static and Dynamic Models Before Gasoline Tax Increase.

Before Gasoline Tax Increase:
Gasoline expenditure as a share of income computed from the static and dynamic models, by income decile.



Notes: Figure A2 displays the gasoline expenditures as a share of income before the gasoline tax increase predicted by the static and dynamic models for consumers in different deciles of the pretax income distribution. The corresponding values for the static (dynamic) model are: income decile 1, 12.67 (12.59); 2, 5.85 (5.81); 3, 4.06 (4.14); 4, 3.25 (3.41); 5, 3.05 (2.96); 6, 2.40 (2.55); 7, 2.18 (2.30); 8, 1.91 (2.02); 9, 1.64 (1.59); 10, 1.41 (1.31). See Section 5 and Appendix G for details.