VEHICLE SIZE CHOICE AND AUTOMOBILE EXTERNALITIES:

A DYNAMIC ANALYSIS

Clifford Winston — **Brookings Institution**

Jia Yan — **Washington State University**

Abstract. We study the effect of highway congestion on the "arms race" on American roads, which has led to larger and more powerful vehicles that increase the negative safety and fuel consumption externalities from automobile travel. We estimate a dynamic vehicle size choice model and find that congestion delays affect vehicle sizes. We then show that by reducing vehicle sizes, congestion pricing could produce a decline in the vehicle fatality rate that approaches $25 billion in annual benefits, and could result in an improvement in the nation's vehicle fleet fuel efficiency that approaches $10 billion in annual operating cost savings.

April 2017

1. Introduction

The internal combustion engine has been continuously refined since its introduction, enabling motorists to trade off increasingly greater horsepower and fuel efficiency. Since the 1980s, motorists have revealed a strong preference for power. Knittel (2011), for example, reported that the average horsepower of new passenger cars increased by 80 percent from 1980-2004, while fuel economy increased by less than 6.5 percent, and average curb weight increased by 12 percent. That preference has been underscored by the shift from passenger cars to light trucks and SUVs—in the year 2000, 20 percent of new vehicles sold in the United States were SUVs; in 2017, that percentage climbed to 62 percent. And the increasing power and weight of passenger vehicles has occurred despite stricter emissions and fuel economy standards that have created incentives for automobile companies to make lighter vehicles to reduce fuel consumption and to get better fuel economy.

Households' preferences for larger more powerful vehicles have undoubtedly been influenced by demographic changes, such as bigger households, and economic conditions, such as declining real gasoline prices during much of the period. However, another potentially important factor that has been overlooked is the significant growth of traffic congestion and delays on the nation's highways. According to the Texas Transportation Institute's *Urban Mobility Report*, average annual traffic delays in U.S. urban areas have increased from 18 hours in 1982 to 42 hours in 2014. As traffic congestion has increased and trips have become more stressful, it is plausible that travelers have tried to increase their safety, comfort, and privacy by driving larger and more powerful vehicles.[1] White (2004) and Li (2012) report empirical evidence that larger, heavier, and

[1] Motorists place a greater value on the maneuverability of smaller vehicles on local streets and shopping areas than on urban highways.

taller vehicles provide better safety protection to occupants involved in a collision than smaller, lighter, and shorter vehicles do, and that, all else constant, motorists are willing to pay a premium for vehicles with those safety attributes, especially if much of their driving occurs in congested conditions where they have a greater likelihood of being involved in an accident (Yeo, Jang, and Skabardonic (2010)).

A valid causal relationship between highway congestion and vehicle size would indicate that the major automobile externalities are positively related because larger vehicles: (1) consume more fuel than smaller vehicles and generally produce greater emissions, and (2) increase the risk of a fatality to occupants of smaller vehicles in a multi-vehicle crash (White (2004) and Anderson and Auffhammer (2014))—although in a single-vehicle accident, larger vehicles may decrease the risk of a fatality to occupants.[2] Importantly, a policy that reduces highway congestion, such as road pricing, could provide additional social benefits that have not been considered by reducing vehicle sizes, thereby improving automobile safety and fuel economy.

The purpose of this paper is to conduct, to the best of our knowledge, the first disaggregate analysis of the direct effect of highway congestion on vehicle size choice, controlling for other important influences, including vehicle purchase price and operating costs, and to simulate the

[2] It could be argued that light trucks and SUVs increase highway congestion by requiring greater road space. However, highway engineers measure both traffic volume and highway capacity in terms of passenger car equivalents (PCEs) and they have not designated higher PCEs for light trucks and SUVs as they have done for heavy trucks. In addition, significant improvements in engine technology over time have enabled light trucks and SUVs to accelerate and decelerate much faster, which reduces any particular disturbances that those vehicles may have on traffic flows. The risk profiles of the type of drivers who self-select to purchase light trucks and SUVs may also affect traffic flows and congestion if drivers of those vehicles are excessively risky or cautious. But a definitive analysis of this issue requires complete information about drivers, their vehicles, and their trips, which are formidable data requirements. Jacobsen (2013) and Makuch (2015) reach conflicting conclusions about drivers' riskiness by vehicle type; richer data is required to resolve this issue.

effects of congestion pricing on vehicle fatalities and fuel economy.[3] In theory, the gasoline tax, which is currently used to simultaneously address all automobile externalities, has the greatest impact on motorists who are committed to driving the most fuel inefficient vehicles (Langer, Maheshri, and Winston (2017)); thus, it should discourage motorists from purchasing larger and heavier light trucks and SUVs. But this effect has been blunted because: (1) Congress has maintained the federal gasoline tax at its 1993 level of 18.4 cents per gallon, (2) real gasoline prices have declined throughout much of the past few decades and are well below their highest values during the current decade, and (3) technological change has enabled larger vehicles to get better fuel economy.

It is well known that congestion pricing could do a better job than the gasoline tax at reducing peak-period congestion and guiding efficient investment in highway capacity (Lindsey (2012), Winston (2013)), but it has not been documented that it could also do a better job at reducing vehicle sizes and their concomitant externalities. We estimate that an efficient congestion charge would reduce the market share of mid to full-size SUVs from about 31 percent to 23 percent and would reduce average vehicle weight from 3860 pounds to 3730 pounds, which would result in: (1) a 10 percent decline in the vehicle fatality rate that amounts to nearly $25 billion in annual nationwide benefits based on conventional values of life, and (2) a 3 percent improvement in the average fuel efficiency of the nation's vehicle fleet that amounts to nearly $10 billion in annual operating cost savings.

The economic and demographic factors and technological advance that have contributed to households' growing preference for power over fuel economy are unlikely to abate in the near

[3] Brownstone and Golob (2009) analyze the effect of residential density on vehicle use and find that greater density, which is likely to be correlated with greater highway congestion, reduces fleet fuel efficiency through the choice of less fuel-efficient vehicle types.

future. Thus, the potential ability of congestion pricing to efficiently encourage households to reduce their vehicle sizes and to generate social benefits that would be widely shared among the public should improve its political appeal.

2. An Overview of the Sample

Our sample is based on data collected by GfK, a market research firm, on Seattle motorists from 2004 to 2009. We chose Seattle for our geographic unit of analysis because it is a congested city with several bottlenecks created by bridges, and because NAVTEQ (subsequently acquired by Nokia) could provide detailed real-time traffic data for its road network during our period of analysis.

GfK provided a questionnaire to automobile commuters who are members of its survey panel that asked them to indicate the roads that comprise the route that they use to get to work the majority of the time. If another mode, such as a ferry was involved, only the auto portion of the commute was included. Interstate freeways and state routes were identified by their number and direction. Respondents also indicated: (1) the normal departure and arrival time of their commute, (2) any regular stops that they make, (3) whether a High-Occupancy-Vehicle (HOV) lane is used on the freeway portion of the commute, and (4) the vehicle that they normally use for the commute, or the two vehicles that they use if they are consistently alternated. Any changes in the routes and departure and arrival times during the sample period were also noted.

We focus on the congestion delay that commuters experience on the freeway portion of their commute because that delay is likely to influence their vehicle size choice and it could be addressed with public policies like congestion pricing. In addition, congestion delays on local roads and arterials were more difficult to measure accurately than were delays on freeways. To obtain a respondent's commuting information for each year, we used Traffic Message Channel (TMC) codes for the Seattle Metropolitan Urban Area to identify the freeway segments of the commute. We then used data from NAVTEQ to determine the average speeds and commute times on those segments and summed them to obtain the average travel time for the entire freeway portion of the commute.[4] Congestion delays are therefore the difference between those average travel times and commuters' free flow travel times, which were based on the travel time of the freeway portion of their commutes at 2am.

We show in figure 1 the distribution of the percentage of congestion delay in motorists' commute time during our sample period. (The distribution changed very little from year to year.) Some commuters experienced little delay but nearly half experienced non-trivial delays that accounted for roughly 15% or more of the total time of the commute. Our base case specification of congestion delay attempted to capture the effect of those delays on commuters' vehicle size choice behavior, so we interacted a long commute dummy variable, which indicates a commute that takes one hour or more, with an excessive delay dummy variable, which indicates a level of congestion delay on the freeway portion of the commute that accounts for 15% or more of the total time of the commute. We conducted a robustness check by defining the excessive delay dummy variable to indicate congestion delay that accounts for 20% or more of the total time of the commute.

In addition to obtaining information about the respondent's automobile commute, the GfK survey contained information on the respondents' socioeconomic characteristics, including

[4] NAVTEQ reported their travel time data in "waves," where travel times on a given segment were based on a few years of data that were weighted toward the most recent year. For example, travel time on a segment in 2006 was based on travel times from January 2004 to December 2006 that were weighted toward 2006.

household income, household size, age, gender, and education, and their zip code and housing characteristics from which we determined the square footage of the house and the zip code's Zillow Home Value Index, median household income, school quality index, personal crime index, and property crime index. The information on housing and residential location characteristics is important for our identification strategy discussed later.

Final Sample

After eliminating respondents with missing information about their commute and location, we obtained complete information from GfK for 271 respondents. We had to remove 41 of those respondents from the sample because they had lived in the Seattle metropolitan area for only one year while the dynamic choice model that we estimate requires an initial condition for each motorist and at least one subsequent holding and replacement choice. Thus our final sample contained 230 respondents and 866 observations during 2004-2009. All of the respondents lived in the Seattle Metropolitan Area and commuted to work by car as of 2009, although some did not reside in Seattle for the entire 2004-2009 sample period. Roughly one-third of the respondents switched vehicles in our sample, but we could not discern that, in general, those motorists switched to larger (or smaller) vehicles. In addition, none of the respondents switched their residential location.

The Seattle Metropolitan Area is adjacent to the Puget Sound and commuters must often travel over a body of water to get to their workplaces. The bridges that people must cross when they drive into Seattle create many bottlenecks that significantly contribute to congestion in the area. Figures 2a (North of Seattle Tacoma airport) and 2b (South of Seattle Tacoma airport) indicate the residential locations of the commuters in our sample (blue dots) and the major bottlenecks in the road network (black dots), as characterized by the Washington State Department of Transportation.

The figures illustrate that the congestion and delays faced by Seattle commuters are the outcome of their residential location choices (as well as their workplace choices), which expose them or limit their exposure to bottlenecks. They also preview an important identification issue that we address later—specifically, congestion may be correlated with individuals' unobserved characteristics that affect both their vehicle-size and residential location choices.

Vehicle Size Choice Set and Vehicle Attributes

Consistent with the U.S. Department of Transportation, National Highway Transportation and Safety Administration (NHTSA) classifications, we combined automakers' vehicle classes and sizes to define a product in our choice set, and we allowed motorists to select new vehicles and used vehicles. The 13 vehicle size and class combinations in our choice set include:

1. Compact domestic;
2. Compact luxury imported;
3. Compact pickup;
4. Full size SUV;
5. Full size domestic;
6. Full size luxury imported;
7. Midsize SUV;
8. Midsize domestic;
9. Midsize luxury imported;
10. Passenger van;
11. Standard pickup;
12. Sub-compact domestic; and
13. Sub-compact luxury imported.

When motorists decide to replace their current vehicles, we assume they choose among vehicle class and size combination from the most recent 10 model years. Thus, in a given year, a motorist's choice set consists of 130 alternatives.

We used data from Ward's Automotive Yearbook, various editions, to construct vehicle attributes, including purchase price, miles-per-gallon (mpg), body-weight, and horsepower, for each choice alternative by averaging those attributes across vehicles for each combination of vehicle class, size and model year.[5] We measured the operating cost (in dollars per mile) as the ratio of the average gasoline price in Seattle to the average miles per gallon on both highways and local roads. Finally, we obtained vehicle registration data from R.L. Polk, Incorporated to construct the population shares in Seattle of each vehicle class and size combination during the sample period.

We compare the population shares and the sample shares of the vehicle class and size combinations in table 1. Generally, the population and sample shares suggest that the pronounced shift to larger vehicles that occurred in the preceding decades has abated to some extent, perhaps because the period of our analysis includes the Great Recession that began in 2007 and ended in 2009. In any case, midsize luxury vehicles and midsize SUVs comprise the largest vehicle shares in the population and in our sample, although the share of midsize luxury vehicles is somewhat larger in our sample than in the population. Most of the other population and sample shares of the vehicle class and size combinations are aligned, with the exceptions that the population share of compact pickups is larger than the sample share and the sample share of full size SUVs is larger than the population share. The difference in some of the population and sample shares motivate

[5] We accounted for vehicle depreciation when we constructed the purchase prices for used vehicles.

us to perform a robustness check of our findings by re-estimating our base case model by Weighted Exogenous Sample Maximum Likelihood (WESML) to align the sample and population shares of all the vehicle class and size combinations.

We summarize the motorists' socioeconomic and demographic variables in our analysis in table 2. It is interesting that the average household income of commuters in our sample is more than double the median city household income, while nearly two-thirds of the motorists in our sample have a commute that exceeds one hour. Based on U.S. Census data, it could be argued that Seattle's inflated housing market and high cost of living has caused younger people to get started on the American dream by buying a house that is far from their workplace and then commuting long distances to jobs that help them afford their mortgages. We also stress that our sample consists only of people who have a job.

The vehicle attributes that we summarize in table 3 are consistent with the tradeoff that motorists have made since the 1980s of driving vehicles with greater horsepower (and body weight) and little improvement in fuel economy (Knittel (2011)).

3. A Dynamic Model of Vehicle Holdings and Replacement

Our sample enables us to identify motorists' preferences for vehicle attributes, including purchase price, operating cost, weight, horsepower, and vehicle size, when they face different levels of congestion and in two decisionmaking environments: (1) the decision of whether to keep or replace a vehicle, and (2) the decision of which vehicle to purchase when they decide to replace a vehicle.

We model those holding and replacement decisions in a dynamic context because:

● Motorists do not change their vehicles frequently and when they do, they generally sell a vehicle that they currently own in the used-vehicle market.

● As a vehicle that a motorist owns ages, its price depreciates and its maintenance costs increase.

● Vehicle operating costs evolve over time with the fluctuation in gasoline prices.

● Technological advance in the automobile industry causes vehicle attributes to improve over time.

Model Assumptions

Many studies of motorists' preferences for vehicle attributes analyze the one-period utility maximizing decision to purchase a new vehicle from the available makes and models on the market (for example, Train and Winston (2007)). The major challenge to analyze vehicle choice behavior in a dynamic context is the large-dimension of the state space. Schiraldi (2011) employs the Inclusive Value Sufficiency (IVS) assumption, formalized by Gowrisankaran and Rysman (2012) in a study of the demand for camcorders, to reduce the dimensionality in their multinomial logit model of motorists' vehicle choice. They formulate that model in the context of a dynamic optimal stopping problem. Under the IVS assumption, all states that lead to the same inclusive value, which measures a decisonmaker's ex-ante present discounted value of purchasing the preferred vehicle instead of holding on to the current vehicle, are equivalent. Thus, the decisionmaker tracks only the inclusive value instead of the relevant vehicle attributes. Although the assumption simplifies the analysis, it is purely mathematical and has little behavioral justification. This drawback is particularly relevant here because we use disaggregated data and we therefore need information about the state that each motorist faces when making replacement and purchase decisions.

We develop a tractable dynamic choice model by making plausible assumptions, which have clear behavioral interpretations, based on the relevant features of the automobile industry and the available empirical evidence in the literature.

Assumption 1: Consumers do not predict the evolution of the automobile industry's vehicle offerings.

This assumption is plausible because most consumers do not frequently change their vehicles; thus, they pay close attention to the vehicle market only periodically. At the same time, automakers do not significantly change vehicle designs frequently so vehicle attributes tend to be stable for several years. In our model, we construct average attributes based on combining individual vehicles into combinations of class, size, and year, which will be even more stable over time than individual vehicles' attributes. The implication of this assumption is that consumers assume that the same attributes in their current choice set, which we have noted, will be available in their future choice sets.

Assumption 2: Motorists make reasonable predictions of gasoline prices and base their vehicle replacement decisions on those predictions.

Busse, Knittle, and Zettelmeyer (2013) provide support for this assumption with empirical evidence that indicates that consumers predict future gasoline prices based on current gasoline prices and that they make vehicle purchase decisions based on that prediction.

Assumption 3: Consumers track over time the increasing maintenance costs of the vehicles they currently own and the depreciation of their prices.

Given the notable changes in and available information about maintenance costs and resale prices over time, it is plausible that consumers take account of those factors in their vehicle holding and replacement decisions.

Assumption 4: Congestion on urban roads persists and exhibits stable growth in normal macroeconomic environments.

The broad attention given to the Texas Transportation Institute's *Urban Mobility Report*, which every year informs the public of the congestion delays in major U.S. urban areas, supports this assumption.

Model Formulation

We model motorists' vehicle holding and replacement decisions as an optimal stopping problem. We index a sequence of observations for each motorist by year $t = 0,1,...,T$, with a motorist owning a vehicle in the initial period $t = 0$. Given the initial vehicle holding in 2004, the motorist decides whether to hold or to replace the vehicle with another one in each subsequent year starting in 2005. Identification in this dynamic model of the effects of purchase price, operating costs, and congestion on vehicle size choice relies on variation in both vehicle holdings and replacement decisions over time and across motorists, given the initial holdings.

We summarize in figure 3 the information requirements and holdings and replacement decisions in the dynamic model. In the initial year of our analysis, a motorist owns a vehicle and has complete information on the attributes of all the vehicles in the market. The motorist's preferences for vehicles face random shocks in every period that are realized at the beginning of a period. At the same time, the motorist receives information on gasoline prices and formulates expectations of future gasoline prices. Given the information set, the motorist makes a decision of whether to keep the current vehicle or to replace it with another one subject to the preceding four assumptions. The decision generates a utility gain to the motorist in that period and affects the individual's utility in future periods by determining the state—the vehicle that the individual owns—at the beginning of the next period.

Transition of States

Given assumptions 1-4, the information set or the state space of a motorist i, who owns a vehicle j at the beginning of a period t, is $\Omega_{ijt} \equiv (j, a_{jt}, g_t, \varepsilon_{it})$, where a_{jt} is the age of the vehicle the motorist owns; g_t is the price of gasoline; and $\varepsilon_{it} \equiv \{\varepsilon_{ikt}\}_{k \in C_t}$ is the set of random shocks affecting the motorist's preference for vehicles contained in the choice set C_t. The transition of the uncontrolled states $\Lambda_{ijt} \equiv (a_{jt}, g_t, \varepsilon_{it})$ is denoted by $\rho(\Lambda_{ijt+1} | \Lambda_{ijt})$.

We adopt the conditional independence assumption in Rust (1994), thus:

$$\rho(\Lambda_{ijt+1} | \Lambda_{ijt}) = g(\mathbf{S}_{jt+1} | \mathbf{S}_{jt}) \times \pi(\varepsilon_{it+1}), \tag{1}$$

where $\mathbf{S}_{jt} = (a_{jt}, g_t)$ and the vector-valued transition function $g(\mathbf{S}_{jt+1} | \mathbf{S}_{jt})$ is given by:

$$\begin{aligned} a_{jt+1} &= a_{jt} + 1, \text{ which implies } p_{jt+1} = 0.85^{a_{jt+1}} \times p_j^* \\ g_{t+1} &= g_t + \eta_t, \eta_t \sim N(0, 0.28) \end{aligned} \tag{2}$$

The first condition in equation (2) simply says that the age of the vehicle that the motorist owns is increased by one in the next period; accordingly, the manufacturer's price p_j^* of vehicle j depreciates with vehicle age at a rate of 15%, which is consistent with industry standards.[6] Increases in vehicle age also capture the effects of other influences on vehicle size choice, such as maintenance costs, which increase with vehicle age. The second condition indicates that the evolution of the price of gasoline follows a normal random-walk process that is estimated from data on average gasoline prices for U.S. cities from 1981 to 2014.[7] Finally, we assume that the distribution of the random shocks, $\pi(\varepsilon_{it+1})$, is given by a multivariate extreme-value density.

[6] https://www.carfax.com/guides/buying-used/what-to-consider/car-depreciation

[7] We conducted a Dickey-Fuller test and we could not reject the null hypothesis that gasoline prices follow a random-walk process. Empirical evidence in Anderson, Kellogg, and Sallee (2013) also indicates that consumers formulate their expectations of future gasoline prices based on a random-walk process.

One-period Utility and Accounting for Vehicle Price Endogeneity

The one-period indirect utility that motorist i obtains from keeping vehicle j in year t is given by:

$$u_{ijt} = \mathbf{x}_{jt}\mathbf{B}_i + \xi_j + \mathbf{V}_j\mu_{it} + \varepsilon_{ijt} = v_{ijt} + \mathbf{V}_j\boldsymbol{\mu}_{it} + \varepsilon_{ijt}, \tag{3}$$

where $\mathbf{x}_{jt}$ is a vector of vehicle attributes, including vehicle age, size, body-weight, purchase price, and operating cost, which is determined by the price of gasoline and the vehicle's fuel economy (mpg); ξ_j captures omitted attributes of a vehicle in the spirit of Berry, Levinsohn, and Pakes (BLP 1995) so we can avoid the bias caused by endogeneity of the vehicle's purchase price.

Because we have disaggregated data on motorists' vehicle choices over time and because we focus on vehicle-size choice, we can use a fixed-effects specification to control for the omitted product attributes. Specifically, we specify ξ_j with dummy variables for the 13 vehicle class and size combinations in our analysis, and we assume that the omitted attributes in a given class and size combination for different model years are the same. This assumption is plausible because vehicle manufacturers do not significantly change their vehicle designs frequently. Compared with the random-effects BLP specification, which assumes that ξ_j is correlated with the purchase price but uncorrelated with other observed attributes, the fixed-effects specification that we use here has the advantage that it allows omitted vehicle attributes to be correlated with observed vehicle attributes in a flexible way.

Finally, $\mathbf{V}_j$ is a vector of four vehicle classification dummies (large vehicles, luxury vehicles, SUVs, and new vehicles) and $\boldsymbol{\mu}_{it}$ a vector of normal random variables with zero mean. This specification mimics a nested-logit specification in which different vehicle class and size combinations are grouped into four overlapped nests to generate flexible substitution patterns among alternatives, thereby relaxing the restrictive IIA assumption.

If a motorist decides to replace a vehicle with a different one denoted by k, the one-period utility becomes:

$$\begin{aligned} u_{ik(j)t} &= \mathbf{x}_{kt}\mathbf{B}_i - \alpha_i\left(p_{kt} - p_{jt}\right) + \xi_k + \mathbf{V}_k\boldsymbol{\mu}_{it} + \varepsilon_{ikt} \\ &= v_{ik(j)t} + \mathbf{V}_k\boldsymbol{\mu}_{it} + \varepsilon_{ikt} \end{aligned} \tag{4}$$

where p_{jt} is the price received from selling the vehicle *j* that is currently owned and p_{kt} is the price of the replacement vehicle *k*.

Note that a motorist's preferences for vehicle attributes and price are captured by $\mathbf{B}_i$ in equation (3) and by $\mathbf{B}_i$ and α_i in equation (4). Travelers have heterogeneous preferences, which we capture after exploring interactions between price, operating cost, vehicle body weight, and vehicle size dummies with socio-demographic variables denoted by $\mathbf{z}_i$, including age, gender, household income, and household size.

Given our primary interest in the effect of congestion on vehicle size, we interact the delay faced by commuters with the dummy variable for full and medium size SUVs, vehicle body weight, and operating costs. As noted in the introduction, White (2004) and Li (2012) report empirical evidence on the safety benefits of larger vehicles and indicate that drivers are willing to pay for those benefits, especially if much of their driving occurs in congested conditions where they have a greater likelihood of being involved in an accident (Yeo, Jang, and Skabardonic (2010)). We also interact delay with vehicle operating costs because, all else constant, motorists are likely to buy more fuel-efficient vehicles to offset the higher operating costs caused by stop-and-go driving in congested traffic. In sum, our model captures congestion's effects on motorists' vehicle size choice through two channels: a direct channel that is attributable to their desire for self-protection and an indirect channel that arises because vehicle operating costs are affected.

Endogenous congestion

As suggested in figures 2a and 2b, motorists' congestion delays are affected by their residential location choices, which are affected by their preferences for both neighborhood and housing attributes. At the same time, those preferences are likely to be correlated with their vehicle preferences. To take some examples, a motorist who likes large vehicles may also like a spacious house located in an expensive area; a motorist who likes a "trendy" vehicle may also like a cool house in a "trendy" neighborhood; and a motorist who likes a luxury vehicle may also like a stunning house in an exclusive neighborhood. In sum, motorists are likely to sort themselves into different residential locations according to their preferences for housing and location attributes, and because their preferences for vehicle attributes and for housing and location attributes are correlated, motorists living in different locations have different vehicle preferences.

Our modeling approach attempts to identify the effect of congestion on motorists' vehicle-size choices through the variation in their choices under different commuting conditions, as reflected in congestion delays. However, the estimated effect may reflect the heterogeneity in the vehicle preferences of motorists who choose to live in different locations. Attempting to address this problem by modeling residential location choice jointly with vehicle size choice is quite difficult and, to the best of our knowledge, a dynamic disaggregate model of residential location choice does not exist.[8] A fundamental challenge is to accurately characterize the choice set over time, including the non-chosen alternatives and their attributes, and even if that is possible, jointly

[8] Of course, there are many studies of the choice of residential location that use a zip-code as the basic unit of analysis.

estimating both residential location choice and vehicle size choice in a dynamic framework is a daunting task. Because we perform maximum likelihood estimation of a nonlinear model, there is no corresponding approach where we would instrument congestion. Instead, we would have to develop a structural model of congestion and derive the complete data likelihood, which is equivalent to jointly estimating vehicle size and location choices.

Still another strategy to overcome the identification issue caused by endogenous congestion is to treat $\mathbf{B}_i$ as motorist fixed-effects, which are estimated in the dynamic choice model and capture unobserved characteristics of motorists that lead to correlation of their vehicle and location preferences. But the strategy of using motorist dummies faces the incidental parameter problem (Lancaster (2000)), because the time series is short (2004-2009); $\mathbf{B}_i$ cannot be estimated precisely from at most six observations.

Thus our empirical strategy to achieve identification uses selection on observables, where we specify as controls several variables that are well-known in the urban housing and location literature (for example, McFadden (1978)) to affect location and housing choices to capture unobserved preferences, including: (1) the size of the motorists' houses measured by square feet, and (2) several attributes of the location, measured at the zip code where the motorists live, including the Zillow home value index, median household income, school quality index, personal crime index, and property crime index.[9] We pointed out that motorists sort themselves into different residences according to their location and housing preferences, which are correlated with their vehicle preferences. Our empirical strategy implicitly assumes that those observables

[9] Our approach to reduce the potential bias caused by unobservables by controlling for several observables is in the spirit of Altonji, Elder, and Taber (2005). However, we cannot use their approach to test for a critical value where bias is unlikely because we estimate a nonlinear dynamic model instead of a linear regression model.

influence the sorting process so that motorists in locations with similar characteristics have homogeneous vehicle preferences holding congestion constant.

Let $\mathbf{w}_i$ denote those control variables and $\boldsymbol{\beta}_i$ denote the vector of coefficients of the SUVs, vehicle body weight, and operating cost. As before, socioeconomic variables are denoted by $\mathbf{z}_i$, thus:

$$\boldsymbol{\beta}_i = \mathbf{z}_i\boldsymbol{\gamma} + \mathbf{w}_i\boldsymbol{\delta} + delay_i \times \boldsymbol{\theta}, \tag{5}$$

and $\mathbf{w}_i\boldsymbol{\delta}$ in equation (5) captures the preference heterogeneity of motorists living in different locations with different attributes and in different houses with different sizes, and $\boldsymbol{\theta}$ captures the net effect of delay on preferences for vehicle size, vehicle weight, and operating cost.

An important consideration that alleviates concern about the potential endogeneity of congestion delay is that compared with vehicle size choice, location and housing decisions are made over a much longer time horizon and households face greater constraints if they wish to change those decisions. Given our short panel, motorists' location and housing are likely to be pre-determined, thereby enabling the time-invariant control variables in equation (5) to absorb the effects of location choice on vehicle attributes. As noted, none of the motorists in our sample changed their locations during the period of analysis, but one-third of the motorists changed their vehicle holdings.[10] If households could easily change locations to avoid congestion, then we would overestimate the effects of congestion on vehicle size choice.

<u>Motorists' Forward-Looking Decision</u>

Given the information set Ω_{ijt} and our characterization of the transition of states in equation (2), motorists' decide to keep or replace their current vehicle to maximize expected lifetime utility.

[10] The fact that our sample includes the Great Recession limits households' mobility, but we also include a number of years before the Recession.

If a motorist keeps vehicle j instead of replacing it, the present value of the motorist's maximal utility in the planning horizon is given by:

$$\tilde{V}_{it}\left(j,\mathbf{S}_{jt},\varepsilon_{it}\right)=u_{ijt}+\beta EV_{it+1}\left(j,\mathbf{S}_{jt+1},\varepsilon_{it+1}\middle|\, j,\mathbf{S}_{jt}\right), \tag{6}$$

where $EV_{it+1}(\cdot)$ denotes the expected value function and $\beta\in(0,1)$ is the discount factor.

To replace vehicle j with vehicle k, the motorist solves the following problem:

$$\hat{V}_{it}\left(j,\mathbf{S}_{jt},\varepsilon_{it}\right)=\underset{\substack{k\in C_t\\ k\neq j}}{Max}\left\{u_{ik(j)t}+\beta EV_{it+1}\left(k,\mathbf{S}_{kt+1},\varepsilon_{it+1}\middle|\, j,\mathbf{S}_{jt}\right)\right\}. \tag{7}$$

Thus the individual keeps vehicle j in period t if and only if $\hat{V}_{it}\left(j,\mathbf{S}_{jt},\varepsilon_{it}\right)\leq\tilde{V}_{it}\left(j,\mathbf{S}_{jt},\varepsilon_{it}\right)$ such that the valuation-function satisfies the following Bellman equation:

$$V_{it}\left(j,\mathbf{S}_{jt},\varepsilon_{it}\right)=Max\left\{\tilde{V}_{it}\left(j,\mathbf{S}_{jt},\varepsilon_{it}\right),\hat{V}_{it}\left(j,\mathbf{S}_{jt},\varepsilon_{it}\right)\right\}. \tag{8}$$

Because ε_{it} contains a motorist's private information, we integrate the preference shocks out of equation (8) by defining motorist $i's$ expected value of holding any vehicle type j in year t as:

$$\begin{aligned} W_{it}\left(j,\mathbf{S}_{jt}\right)&=\int_{\varepsilon_{it}}V_{it}\left(j,\mathbf{S}_{jt},\varepsilon_{it}\right)\times\pi\left(\varepsilon_{it}\right)d\varepsilon_{it}\\ &=\ln\left[\exp\left(v_{ijt}+\mathbf{V}_j\boldsymbol{\mu}_{it}+\beta E\left(W_{it+1}\left(j,\mathbf{S}_{jt+1}\right)\middle|\, j,\mathbf{S}_{jt}\right)\right)+\exp\left(I_{ijt}\right)\right]\end{aligned} \tag{9}$$

$$I_{ijt}=\ln\left[\sum_{\substack{k\in C_t\\ k\neq j}}\exp\left(v_{ik(j)t}+\mathbf{V}_k\boldsymbol{\mu}_{it}+\beta E\left(W_{it+1}\left(k,\mathbf{S}_{kt+1}\right)\middle|\, j,\mathbf{S}_{jt}\right)\right)\right] \tag{10}$$

The expression for I_{ijt} in equation (10) is the *inclusive value* of choosing $k\in C_t$ ($k\neq j$ if it replaces vehicle j) and it represents the motorist's ex-ante present value of switching to the preferred alternative.

4. Estimation

Estimation of our dynamic vehicle size choice model consists of estimating the probability that motorists replace their vehicle type j in a given year with a different vehicle type and the probability that motorists do not replace their vehicle in a given year. The probability that motorist i replaces vehicle type j with vehicle type k' in year t is:

$$\begin{aligned}
l_{ik'(j)t} &= \int_{\boldsymbol{\mu}_{it}} l_{ik'(j)t}\left(\boldsymbol{\mu}_{it}\right) \times f\left(\boldsymbol{\mu}_{it} \middle| \Xi\right) d\left(\boldsymbol{\mu}_{it}\right) \\
l_{ik'(j)t}\left(\boldsymbol{\mu}_{it}\right) &= \frac{\exp\left(I_{it}\right)}{\exp\left(W_{it}\left(j, \mathbf{S}_{j_t}\right)\right)} \times \frac{\exp\left(v_{ik'(j)t} + \mathbf{V}_{k'}\boldsymbol{\mu}_{it} + \delta E\left(W_{it+1}\left(k', \mathbf{S}_{k't+1}\right) \middle| j, \mathbf{S}_{j_t}\right)\right)}{\exp\left(I_{it}\right)}
\end{aligned} \tag{11}$$

where $f\left(\mu_{it} \middle| \Xi\right)$ is the joint probability density function of $\boldsymbol{\mu}_{it}$, which is assumed to have an i.i.d. normal distribution with zero mean; and Ξ is therefore a 4 by 4 diagonal variance-covariance matrix to be estimated, which captures flexible substitution patterns among the SUV, large, luxury, and new vehicle classifications. The probability that motorist i does not replace vehicle j in year t is:

$$l_{i0t} = \int_{\boldsymbol{\mu}_{it}} \left\{ 1 - \frac{\exp\left(I_{it}\right)}{\exp\left(W_{it}\left(j, \mathbf{S}_{j_t}\right)\right)} \right\} \times f\left(\boldsymbol{\mu}_{it} \middle| \Xi\right) d\left(\boldsymbol{\mu}_{it}\right). \tag{12}$$

Estimation of the parameters in the choice probabilities is achieved by maximizing the simulated likelihood function, where we approximate the integration in the choice probabilities by Monte-Carlo simulation using 250 Halton draws (Train (2003)). Importantly, we need to evaluate the expression in equation (9) of the expected value of holding a vehicle, given parameter values, at each iteration of the optimization procedure. The expected value function has three state variables: (1) vehicle class-size combinations, (2) vehicle age, and (3) the real price of gasoline. As noted, we assume that when motorists decide to replace their current vehicle, they choose one from the most recent 10 model years. Given a motorist could own an older vehicle, the vehicle age space is $[0,15]$; that is, the real vehicle price depreciates 15% annually and stops depreciating after 15 years. Real gasoline prices in dollars per gallon during the period of our sample are in the range of [2.00, 3.80].

We take five equally divided points for vehicle age and the price of gasoline to evaluate the expected value function $E\left(W_{it+1}\left(k, a_{t+1}, g_{t+1} \middle| j, a_t, g_t\right)\right)$, and we interpolate the expected valuation at the 325 (13×5×5) points to a more finely discretized state space (13 vehicles × 16 vehicle ages × 8 gasoline prices). The evaluation is accomplished by backward induction. In the computations, we assume the discount rate is 0.025; evaluate the expectation with respect to g_{t+1} by Monte-Carlo integration; and fix the number of time-periods at 50 years, which is a plausible length given normal life expectancy.

We ensured that we achieved a global optimum of the simulated likelihood function by varying the initial starting values of the parameters. Finally, we discuss the robustness of our parameter estimates later, but we report here that our findings were not affected by our assumptions about the discount rate and the number of time periods.[11]

5. Estimation Results

We estimated a basic specification for both the myopic and dynamic models, which included the important interactions between congestion delays and vehicle characteristics, vehicle attributes, observed and unobserved heterogeneity, and a full set of interactions with demographic variables

[11] We conducted sensitivity analyses of the dynamic baseline model by reducing the discount rate by 50% and increasing the number of time periods to 75 years.

and housing and location characteristics. We then performed a series of robustness checks of our findings on the effects of congestion delays on vehicle size choice.

Baseline Models

The estimated parameters of the myopic and dynamic choice models presented in table 4 are, in general, statistically reliable and have plausible signs. The coefficients of greatest interest capture the effect of interactions between congestion and vehicle characteristics on vehicle size choice. We find that motorists are more likely to acquire a full or mid-size SUV, a heavier vehicle, and a vehicle with lower operating costs as congestion delays increase, indicating that congestion simultaneously affects—in a conflicting manner—motorists' safety and fuel economy considerations.[12] Motorists react more strongly to congestion's interactions with vehicle characteristics in the dynamic model than in the myopic model because the reduction in safety and fuel economy that is caused by congestion persists through time.

Similarly, an examination of the coefficients of the vehicle attributes indicates that motorists place a greater value on reducing operating costs and less value on reducing the vehicle purchase price in the dynamic model compared with their valuations of those attributes in the myopic model. The effects of other vehicle attributes, including horsepower (specified alone and divided by vehicle weight), a dummy variable indicating a new or used vehicle, and a dummy variable indicating whether the age of the vehicle exceeds five years, are similar for the dynamic and myopic models.

We capture observed heterogeneity in the specification by interacting household income per capita with the vehicle purchase price, operating cost, and congestion delay. We also capture

[12] We also estimated a model where we interacted congestion delays with a compact vehicle and we found that motorists were less likely to acquire those vehicles as congestion increased, but the estimated coefficient was statistically insignificant.

preference heterogeneity that indicates that households with higher income per capita have a greater preference for luxury vehicles than other households have and that larger households have a greater preference for passenger vans than other households have. Those preferences are stronger in the dynamic model than in the myopic model because household sizes and incomes often increase over time. Finally, we capture unobserved heterogeneity with random coefficients for the large, luxury, SUV, and new vehicle dummy variables

Robustness Checks

We preformed several robustness checks to shed light on our identification strategy and to test certain assumptions that we have made in the analysis. First, we identified the effects of congestion delays on vehicle size choice by including interactions between housing and residential location characteristics to control for unobserved influences that may affect both location and vehicle size decisions and bias our estimates of congestion's effects. Table 5 shows that removing those interactions reduces the effects of congestion, especially through its interaction with body weight; thus, we avoided a potential downward bias in the parameter estimates of the congestion delay interactions by including the housing and residential location characteristics.

We discussed the assumptions that we used to measure congestion delays and we pointed out that given the distribution of the percentage of delay in motorists' commute time, we could have set a higher threshold for excessive delay and still captured the effect of those delays on motorists' vehicle size choices. Thus, we re-estimated our dynamic choice model and we assumed that the excessive delay dummy took on a value of 1 if congestion delay amounts to 20 percent, instead of 15 percent, or more of total commute time. The estimates in table 5 indicate that the estimated effects of congestion on vehicle size are somewhat lower, but not statistically significantly different from the estimated effects that we obtained under our initial assumptions.

We provided additional perspective on our identification approach by splitting the sample to limit households' mobility and to reduce the endogeneity from commuters' potential changes in residential location. Table 6 presents estimation results for models where we split the sample between commutes that are less than and greater than or equal to ten miles. We find that congestion still affects vehicle size through its interactions with vehicle characteristics for the longer commutes. However, we also find that congestion no longer affects vehicle size through its interactions with SUVs and operating costs for the shorter commutes where, in all likelihood, commuters' considerations of the safety protection and fuel economy provided by vehicle size are less important because the risk of an accident, which is partly determined by exposure in terms of vehicle miles traveled (VMT), and total vehicle operating costs are lower.

In table 7, we present estimation results for models where we split the sample between commuters from households with annual incomes less than and greater than or equal to $50,000. Congestion affects the vehicle size choices of both groups of commuters, but the more affluent commuters respond much more strongly to congestion's interaction with SUVs and body weight, perhaps because they have a higher value of life and place a greater premium on vehicle safety characteristics. At the same time, they also respond more strongly to congestion's interaction with operating costs, in all likelihood to partially offset the higher operating costs they incur by driving heavier and larger vehicles.

Finally, our analysis has proceeded under the assumption that our sample of Seattle commuters is representative of the population of Seattle commuters. However, we pointed out in table 1 that some of the shares of the 13 vehicle classifications in our sample were not closely aligned with the shares in the population. We therefore tested whether this characteristic of our sample had led to any bias by re-estimating our dynamic vehicle size choice model with a

WESMLE estimator that uses weights to align the sample shares of the 13 vehicle classifications with the population shares (Manski and Lerman (1976)). The estimation results presented in table 8 indicate that the WESMLE parameter estimates are generally similar to our baseline parameter estimates, thus providing evidence that our sample has not resulted in biased parameter estimates.

6. Policy Simulations

We have found credible and robust empirical evidence that congestion delays affect vehicle size choice, thus we use our baseline dynamic model to simulate the economic effects of a policy—namely, congestion pricing—to reduce the effects of congestion on traffic safety and fuel consumption caused by larger vehicles. Congestion pricing affects motorists' vehicle size choices by setting a toll on highways that varies by time of day in accordance with traffic volumes, thereby reducing traffic delays during peak travel periods because some motorists decide to avoid the toll by traveling during off-peak times, using less congested routes, and so on. If motorists continue to use the tolled road, their out-of-pocket and thus operating costs increase. At the same time, their vehicle operating costs will decrease because there is less congestion on the road. We capture the effects through the statistically significant coefficient of the interaction of delay and vehicle operating costs. If vehicle operating costs were not affected by delays, then their interaction would not affect vehicle size choice.

The estimation results of our dynamic choice model indicate that a reduction in congestion delays reduces the likelihood that motorists will purchase SUVs and heavy vehicles, but increases the likelihood that they will purchase less fuel-efficient vehicles. Because we measure congestion delays by interacting a dummy variable for a long-distance commute with a dummy variable for excessive delay, we operationalize the effect of congestion pricing by assuming the existence of a toll that would eliminate excessive delay, which would change the excessive delay dummy variable from one to zero.[13] Using the estimation results of the baseline dynamic model, we simulate that instituting such a congestion toll would cause the market share of mid to full-size SUVs to drop from roughly 31% to 23% and would cause average vehicle weight to drop from 3860 pounds to 3730 pounds.

Anderson and Auffhammer (2014) find that reductions in both the share of mid to large-size SUVs and vehicle weight significantly reduces the fatality rate in vehicle collisions—specifically, their estimates imply that the 8 percent decrease in the share of SUVs and the 130 decrease in vehicle weight from reducing congestion would lead to nearly a 10 percent drop in the automobile fatality rate. Given some 315 automobile fatalities in the Seattle-Tacoma Metropolitan Area in 2013, that is a reduction of 32 deaths, which at a value of life of $9.2 million, results in a gain of nearly $300 million in that year.[14] Because Seattle is representative of U.S. cities in terms of the risk it poses to drivers,[15] it is reasonable for us to apply our calculation to the roughly 25,000 auto fatalities in all U.S. metropolitan areas, which implies that 2,500 lives could be saved for a total value to the nation of $23 billion in 2013.[16] Note that we are underestimating the benefits from

---

[13] For example, based on the results in Small, Winston, and Yan (2006), a $0.50 per mile charge could reduce travel time 25 percent. Calfee and Winston (1998) find that such a charge could reduce travel time even more.

[14] Our figure for the value of life is from the U.S. Department of Transportation Guidelines for valuing the reduction of fatalities:
https://www.transportation.gov/sites/dot.gov/files/docs/VSL_Guidance_2014.pdf

[15] By representative we mean Seattle is not an excessively risky or safe place to drive. Based on data from *Vital Statistics* and the U.S. Department of Transportation, Seattle's annual fatalities per billion vehicle-miles-traveled (VMT) is 34.6, which is lower than the most risky U.S. urban areas to drive (for example, New York City, Washington, D.C., and Philadelphia) that have fatality rates per billion VMT that exceed 100 and it is above less risky U.S. urban areas (for example, Portland, Spokane, and Wichita) that have fatality rates per billion VMT around 20.

[16] Auto fatalities in all U.S. metropolitan areas, defined as having populations greater than 100,000, are reported in *Vital Statistics*.

improved vehicle safety because we are not including the potential benefits of reduced non-fatal injuries. Green, Heywood, and Navarro (2016) suggest that our findings may be more than a hypothetical policy simulation because they find that the London congestion charge contributed to a decline in fatal traffic accidents by causing commuters to change modes, which may have also reduced vehicle sizes.

Reductions in both the share of mid to large-size SUVs and vehicle weight would also reduce fuel consumption. Because VMT in our analysis accrues through commuting and because we do not assume that motorists adjust their commute distances, we assume VMT is fixed in this policy simulation. Based on the data on fuel economy in our analysis, the change in the shares of SUVs and vehicle weight would increase average fuel efficiency in the vehicle fleet from 21.32 miles per gallon (the average value in the sample) to 21.89 miles per gallon, or a 3 percent increase. Given this improvement in fuel economy, motorists' annual expenditures on fuel based on their VMT in the nation's urban areas would be reduced by some $10 billion. Note that any environmental benefits from less fuel consumption would be in addition to the reduction in annual operating costs.

7. Final Comments

Technological change has spurred an "arms race" on American roads, leading to larger and more powerful vehicles that increase the negative safety and fuel consumption externalities from automobile travel. And that arms race is likely to continue for the foreseeable future. For example, SUVs are likely to improve their fuel economy because of new technology that allows for the engine to shut down and restart when the vehicle is idling.

We have developed a dynamic model of vehicle size choice and found that the congestion externality is related to the safety and fuel consumption externalities through its effect on vehicle size, which indicates that congestion pricing could: (1) simultaneously decrease all of the automobile externalities by reducing vehicle sizes and congestion, (2) generate additional social benefits that have not previously been estimated, and (3) improve its political attractiveness because, in contrast to conventional criticisms that congestion pricing primarily benefits affluent motorists with a high value of travel time, its benefits would be widely shared among all motorists and the public.

Importantly, our findings add support to VMT taxes that are currently being tested and under serious consideration by several states on the west and east coasts because it would not be difficult to allow those taxes to vary on different roads in accordance with traffic volumes (Langer, Maheshri, and Winston (2017)). Looking to the future, implementing such taxes has taken on additional importance because societies are preparing for the adoption of autonomous vehicles, which are expected to dramatically cut auto fatalities. But their effect on congestion and the environment is less clear because of uncertainties over whether they will spur additional driving; thus, an externality-based VMT tax would be socially desirable while contributing to more socially optimal vehicle sizes.

References

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy*, volume 113, February, pp. 151-184.

Anderson, Michael L. and Max Auffhammer. 2014. "Pounds that Kill: The External Costs of Vehicle Weight," *Review of Economic Studies*, volume 81, April, pp. 535-571.

Anderson, Soren T., Ryan Kellogg, and James M. Sallee. 2013. "What Do Consumers Believe About Future Gasoline Prices?," *Journal of Environmental Economics and Management*, volume 66, November, pp. 383-403.

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. "Automobile Prices in Market Equilibrium," *Econometrica*, volume 63, July, pp. 841-890.

Brownstone, David and Thomas F. Golob. 2009. "The Impact of Residential Density on Vehicle Usage and Energy Consumption," *Journal of Urban Economics*, volume 65, January, pp. 91-98.

Busse, Meghan R., Christopher R. Knittle, and Florian Zettelmeyer. 2013. "Are Consumers Myopic? Evidence from New and Used Car Purchases," *American Economic Review*, volume 103, February, pp. 220-256.

Calfee, John and Clifford Winston. 1998. "The Value of Automobile Travel Time: Implications for Congestion Policy," *Journal of Public Economics*, volume 69, July, pp. 83-102.

Gowrisankran, Gautam and Marc Rysman. 2012. "Dynamics of Consumer Demand for New Durable Goods," *Journal of Political Economy*, volume 120, December, pp. 1173-1219.

Green, Colin P., John S. Heywood, and Maria Navarro. 2016. "Traffic Accidents and the London Congestion Charge," *Journal of Public Economics*, volume 133, January, pp. 11-22.

Jacobsen, Mark R. 2013. "Fuel Economy and Safety: The Influences of Vehicle Class and Driver Behavior," *American Economic Journal*: *Applied Economics*, volume 5, July, pp. 1-26.

Knittel, Christopher R. 2011. "Automobiles on Steroids: Product Attribute Trade-Offs and Technological Progress in the Automobile Sector," *American Economic Review*, volume 101, December, pp. 3368-3399.

Lancaster, Tony. 2000. "The Incidental Parameter Problem Since 1948," *Journal of Econometrics*, volume 95, April, pp. 391-413.

Langer, Ashley, Vikram Maheshri, and Clifford Winston. 2017. "From Gallons to Miles: A Disaggregate Analysis of Automobile Travel and Externality Taxes," unpublished paper.

Li, Shanjun. 2012. "Traffic Safety and Vehicle Choice: Quantifying the Effects of the 'Arms Race'on American Roads," *Journal of Applied Econometrics,* volume 27, January/February, pp. 34-62.

Lindsey, Robin. 2012. "Road Pricing and Investment," *Economics of Transportation,* volume 1, December, pp. 49-63.

Makuch, Kim. 2015. "The External Congestion Costs of Light Trucks," unpublished paper.

Manski, Charles F. and Steven R. Lerman. 1977. "The Estimation of Choice Probabilities from Choice Based Samples," *Econometrica,* volume 45, November, pp. 1977-1988.

McFadden, Daniel. 1978. "Modelling the Choice of Residential Location," in A. Karlqvist et al., editors, *Spatial Interaction Theory and Residential Location*, North Holland, Amsterdam, pp. 75-96.

Rust, John. 1994. "Structural Estimation of Markov Decision Processes," in James J. Heckman and Edward E. Leamer, editors, *Handbook of Econometrics*, Volume 4, North Holland, Elsevier Press, Amsterdam, pp. 3081-3143.

Schiraldi, Pasquale. 2011. "Automobile Replacement: A Dynamic Structural Approach," *RAND Journal of Economics,* volume 42, Summer, pp. 266-291.

Small, Kenneth A., Clifford Winston, and Jia Yan. 2006. "Differentiated Road Pricing, Express Lanes, and Carpools: Exploiting Heterogeneous Preferences in Policy Design," *Brookings-Wharton Papers on Urban Affairs*, pp. 53-96.

Train, Kenneth E. 2003. *Discrete Choice Models with Simulation,* Cambridge University Press, Cambridge, United Kingdom.

Train, Kenneth and Clifford Winston. 2007. "Vehicle Choice Behavior and the Declining Market Share of U.S. Automakers," *International Economic Review,* volume 48, November, pp. 1469-1496.

White, Michelle J. 2004. "The 'Arms Race' on American Roads: The Effect of Sport Utility Vehicles and Pickup Trucks on Traffic Safety," *Journal of Law and Economics,* volume 47, October, pp. 333-355.

Winston, Clifford. 2013. "On the Performance of the U.S. Transportation System: Caution Ahead," *Journal of Economic Literature*, volume 51, September, pp. 773-824.

Yeo, Hwasoo, Kitae Jang, and Alexander Skabardonis. 2010. "Impact of Traffic States on Freeway Collision Frequency," http://escholarship.org/uc/item/1v9383k9

Table 1: Summary statistics on market shares of vehicle Class and Size Combinations

Vehicle type	Population market share from vehicle registration data	Sample market share
Compact domestic		
2004	0.039	0.068
2005	0.041	0.065
2006	0.039	0.071
2007	0.039	0.067
2008	0.040	0.061
2009	0.040	0.052
Compact luxury		
2004	0.074	0.067
2005	0.081	0.058
2006	0.086	0.065
2007	0.092	0.077
2008	0.098	0.087
2009	0.098	0.089
Compact pickup		
2004	0.098	0.030
2005	0.089	0.026
2006	0.080	0.024
2007	0.074	0.026
2008	0.069	0.030
2009	0.069	0.045
Full size domestic		
2004	0.061	0.030
2005	0.056	0.032
2006	0.050	0.029
2007	0.047	0.041
2008	0.044	0.035
2009	0.044	0.037
Full size luxury		
2004	0.065	0.023
2005	0.067	0.033
2006	0.069	0.041
2007	0.072	0.046
2008	0.073	0.039
2009	0.072	0.041
Full suv		
2004	0.052	0.120
2005	0.051	0.110
2006	0.051	0.124
2007	0.053	0.103
2008	0.050	0.096
2009	0.049	0.093
Midsize domestic		
2004	0.046	0.015
2005	0.044	0.026
2006	0.041	0.024
2007	0.039	0.031
2008	0.038	0.030
2009	0.037	0.033

Table 1 continued: Summary statistics on market shares of vehicle Class and Size Combinations

Midsize luxury		
2004	0.138	0.241
2005	0.144	0.240
2006	0.155	0.229
2007	0.160	0.217
2008	0.165	0.222
2009	0.165	0.219
Midsize suv		
2004	0.160	0.196
2005	0.163	0.195
2006	0.168	0.188
2007	0.167	0.201
2008	0.170	0.200
2009	0.170	0.175
Passenger van		
2004	0.089	0.068
2005	0.086	0.065
2006	0.082	0.065
2007	0.078	0.077
2008	0.076	0.074
2009	0.077	0.074
Standard pickup		
2004	0.088	0.061
2005	0.091	0.052
2006	0.095	0.047
2007	0.097	0.041
2008	0.095	0.048
2009	0.095	0.048
Subcompact domestic		
2004	0.022	0.023
2005	0.020	0.019
2006	0.019	0.023
2007	0.020	0.015
2008	0.021	0.013
2009	0.021	0.019
Subcompact luxury		
2004	0.034	0.060
2005	0.034	0.078
2006	0.035	0.071
2007	0.035	0.057
2008	0.037	0.065
2009	0.036	0.074

Table 2 Summary Statistics of Motorists' Socio-demographic Variables

Variables	Dynamic choice
Household income	108,108 (48,779)
Household size	3.03 (1.38)
House square footage (000's)	1.89 (0.78)
Zillow home price index at the zip code level	2.89 (1.05)
Median city household income	50,515 (11,554)
People crime index at the zip code level	3.02 (1.49)
Property crime index at the zip code level	3.62 (1.40)
School index at the zip code level	0.64 (0.22)
Percentage of motorists with long commute time (greater than 1 hour)	65%
Number of motorists	230
Number of observations	866

Note: Standard deviations in parentheses.

Table 3 Summary Statistics of Vehicle Attributes

Variables	Dynamic choice
Manufacturer's Price	35,552 (18,157)
Miles per gallon	21.32 (3.87)
Horse power	229 (54)
Body weight in pounds	3,860 (849)

Note: Standard deviations in parentheses.

Table 4 Baseline Results [1]

Variables	Myopic choice	Dynamic choice
Price/Household Income per Capita	-0.9472	-0.7220
	(0.1031)	(0.1121)
Operating Cost/Household Income per Capita	-0.0383	-0.0601
	(0.0105)	(0.0215)
Horse Power	0.9802	1.0630
	(0.4873)	(0.7916)
Horse power/Body Weight	-1.1826	-1.1484
	(0.2349)	(0.4429)
New Vehicle	0.6115	0.4925
	(0.1222)	(0.1111)
Vehicle more than 5 years old	-0.6163	-0.6209
	(0.1130)	(0.1735)
Household Size × Passenger Van	0.7846	1.1545
	(0.1103)	(0.3523)
Household Income per Capita × Luxury vehicle	0.6543	0.7844
	(0.3405)	(0.3661)
Delay × Full or mid-size SUV	0.6875	0.8182
	(0.1988)	(0.3012)
Delay × Body Weight	3.5638	4.9771
	(1.7290)	(2.6111)
Delay × (Operating cost / Household income per capita)	-0.1915	-0.3633
	(0.0471)	(0.0691)
Std. of large vehicle dummy [2]	0.1884	0.1141
	(0.1787)	(0.1553)
Std. of luxury vehicle dummy [3]	0.3618	0.3166
	(0.2448)	(0.3288)
Std. of SUV dummy	0.2640	0.3865
	(0.1312)	(0.1774)
Std. of new vehicle dummy	0.5444	0.7231
	(0.2630)	(0.3340)
Alternative constants included	YES	YES
Interactions between Full or mid-size SUV/body weight / operating cost per capita and demographic variables included [4]	YES	YES
Interactions between Full or mid-size SUV/body weight/ operating cost per capita and housing/residential location characteristics included [5]	YES	YES
Number of commuters	230	230
Number of observations	866	866

Notes:

1. Vehicle holding in 2004 is treated as the initial condition in estimation.
2. Large vehicles include medium or full-size sedan, medium or full-size SUV, standard pickup and passenger van.
3. Luxury vehicles include imported luxury vehicles of all sizes.
4. Demographic variables include gender, young (age <= 35), household size and household income-per-capita.
5. Housing and residential location characteristics include house square footage, Zillow Home Value Index of the zip code, median house income of the zip code, school index of the zip code, personal crime index of the zip code and property crime index of the zip code.

Table 5. Robustness Checks

Variables	Baseline dynamic model	Removing the interactions between Full or mid-size SUV, body weight, operating cost per capita and housing and residential location characteristics	Excessive delay dummy takes 1 if delay accounts for 20% or more in total commute time (15% in baseline specification)
Delay × Full or mid-size SUV	0.8182	0.6132	0.6022
	(0.3012)	(0.1855)	(0.2167)
Delay × Body Weight	4.9771	0.4766	3.6980
	(2.6111)	(1.9603)	(2.3219)
Delay × (Operating cost / Household income per capita)	-0.3633	-0.1604	-0.2970
	(0.0691)	(0.0895)	(0.1039)

Table 6. Split sample by trip distance to work [1]

Variables	Greater than or equal to 10 miles	Less than 10 miles
Price/Household Income per Capita	-0.8760	-0.3800
	(0.1334)	(0.1827)
Operating Cost/Household Income per Capita	-0.0956	-0.0527
	(0.0329)	(0.0413)
Horse power	-0.4557	4.2533
	(1.0978)	(1.4013)
Horse Power/Body Weight	-0.2720	-3.2926
	(0.4683)	(0.5996)
New Vehicle	0.5567	0.4192
	(0.0907)	(0.1509)
Vehicle more than 5 years old	-0.5599	-0.6912
	(0.1541)	(0.1690)
Household Size × Passenger Van	0.8959	1.2892
	(0.2357)	(0.2968)
Household Income per Capita × Luxury vehicle	0.1349	1.9444
	(0.4449)	(0.6676)
Delay × Full or mid-size SUV	1.2895	0.1046
	(0.2826)	(0.3855)
Delay × Body Weight	4.7978	5.3983
	(2.9324)	(2.8829)
Delay × (Operating cost / Household income per capita)	-0.6150	-0.0730
	(0.1869)	(0.1844)
Std. of large vehicle dummy [2]	0.1988	0.0088
	(0.1756)	(0.2103)
Std. of luxury vehicle dummy [3]	0.2163	0.4221
	(0.2777)	(0.3230)
Std. of SUV dummy	0.3180	0.5104
	(0.1725)	(0.2530)
Std. of new vehicle dummy	0.8520	0.4301
	(0.3888)	(0.3130)
Interactions between Full or mid-size SUV/body weight/ operating cost per capita and demographic variables included [4]	YES	YES
Interactions between Full or mid-size SUV/body weight/ operating cost per capita and housing/residential location characteristics included [5]	YES	YES
Number of commuters	127	103
Number of observations	469	397

Notes:

1. Vehicle holding in 2004 is treated as the initial condition in estimation.
2. Large vehicles include medium or full-size sedan, medium or full-size SUV, standard pickup and passenger van.
3. Luxury vehicles include imported luxury vehicles of all sizes.
4. Demographic variables include gender, young (age <= 35), household size and household income-per-capita.
5. Housing and residential location characteristics include house square footage, Zillow Home Value Index of the zip code, median house income of the zip code, school index of the zip code, personal crime index of the zip code and property crime index of the zip code.

Table 7. Split sample by median household income in community [1]

Variables	Greater than or equal to $50,000	Less than $50,000
Price/Household Income per Capita	-0.8142	-0.6795
	(0.3059)	(0.2241)
Operating Cost/Household Income per Capita	-0.1243	-0.0412
	(0.0370)	(0.0333)
Horse power	0.5120	2.9795
	(1.4637)	(1.1957)
Horse Power/Body Weight	-1.0247	-2.2602
	(0.6152)	(0.5039)
New Vehicle	0.5878	0.4186
	(0.1588)	(0.1524)
Vehicle more than 5 years old	-0.5444	-0.6614
	(0.1952)	(0.2037)
Household Size × Passenger Van	1.1211	1.2516
	(0.3774)	(0.3408)
Household Income per Capita × Luxury vehicle	-0.8777	1.5414
	(0.5279)	(0.5863)
Delay × Full or mid-size SUV	1.6315	0.8506
	(0.4522)	(0.2581)
Delay × Body Weight	8.7413	1.8168
	(4.0149)	(2.8203)
Delay × (Operating cost / Household income per capita)	-0.4980	-0.2745
	(0.0912)	(0.0988)
Std. of large vehicle dummy [2]	0.1016	0.1604
	(0.2120)	(0.1633)
Std. of luxury vehicle dummy [3]	0.4219	0.1314
	(0.2832)	(0.1543)
Std. of SUV dummy	0.4995	0.3423
	(0.2339)	(0.1990)
Std. of new vehicle dummy	0.6205	0.7881
	(0.3002)	(0.2885)
Interactions between Full or mid-size SUV/body weight/ operating cost per capita and demographic variables included [4]	YES	YES
Interactions between Full or mid-size SUV/body weight/operating cost per capita and housing/residential location characteristics included [5]	YES	YES
Number of commuters	74	156
Number of observations	299	567

Notes:

1. Vehicle holding in 2004 is treated as the initial condition in estimation.
2. Large vehicles include medium or full-size sedan, medium or full-size SUV, standard pickup and passenger van.
3. Luxury vehicles include imported luxury vehicles of all sizes.
4. Demographic variables include gender, young (age <= 35), household size and household income-per-capita.
5. Housing and residential location characteristics include house square footage, Zillow Home Value Index of the zip code, median house income of the zip code, school index of the zip code, personal crime index of the zip code and property crime index of the zip code.

Table 8 WESMLE Results[1]

Variables	Baseline dynamic choice	WESMLE
Price/Household Income per Capita	-0.7220	-0.6502
	(0.1121)	(0.2613)
Operating Cost/Household Income per Capita	-0.0601	-0.0417
	(0.0215)	(0.0214)
Horse Power	1.0630	1.8066
	(0.7916)	(1.1316)
Horse power/Body Weight	-1.1484	-1.8366
	(0.4429)	(0.7299)
New Vehicle	0.4925	0.6978
	(0.1111)	(0.1848)
Vehicle more than 5 years old	-0.6209	-0.7232
	(0.1735)	(0.2309)
Household Size × Passenger Van	1.1545	0.9737
	(0.3523)	(0.4517)
Household Income per Capita × Luxury vehicle	0.7844	0.8290
	(0.3661)	(0.3988)
Delay × Full or mid-size SUV	0.8182	0.8538
	(0.3012)	(0.3319)
Delay × Body Weight	4.9771	4.0723
	(2.6111)	(2.3533)
Delay × (Operating cost / Household income per capita)	-0.3633	-0.3627
	(0.0691)	(0.1027)
Std. of large vehicle dummy [2]	0.1141	0.1721
	(0.1553)	(0.1830)
Std. of luxury vehicle dummy [3]	0.3166	0.2056
	(0.3288)	(0.3777)
Std. of SUV dummy	0.3865	0.4405
	(0.1774)	(0.2238)
Std. of new vehicle dummy	0.7231	0.6889
	(0.3340)	(0.3245)
Alternative constants included	YES	YES
Interactions between Full or mid-size SUV/body weight / operating cost per capita and demographic variables included [4]	YES	YES
Interactions between Full or mid-size SUV/body weight/ operating cost per capita and housing/residential location characteristics included [5]	YES	YES
Number of commuters	230	230
Number of observations	866	866

Notes:

1. Vehicle holding in 2004 is treated as the initial condition in estimation.
2. Large vehicles include medium or full-size sedan, medium or full-size SUV, standard pickup and passenger van.
3. Luxury vehicles include imported luxury vehicles of all sizes.
4. Demographic variables include gender, young (age <= 35), household size and household income-per-capita.
5. Housing and residential location characteristics include house square footage, Zillow Home Value Index of the zip code, median house income of the zip code, school index of the zip code, personal crime index of the zip code and property crime index of the zip code.

Figure 1. Distribution of the Percentage of Congestion Delay in Motorists' Commute Time

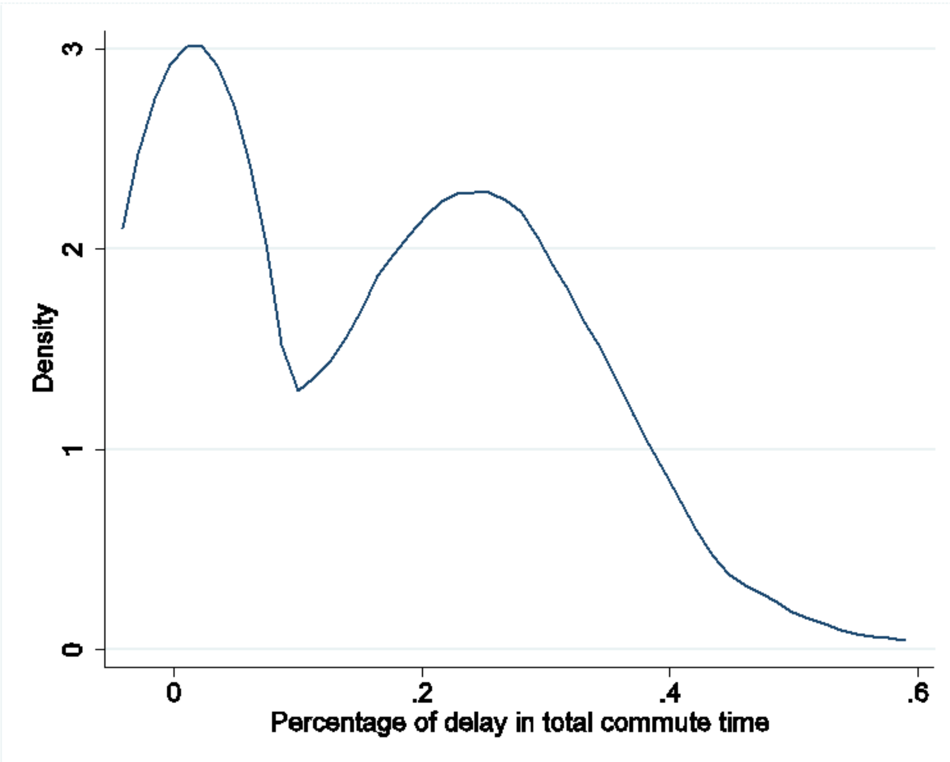

Figure 2a. Sample Household Residences and Bottlenecks
North of Seattle Tacoma Airport (Sea-Tac)

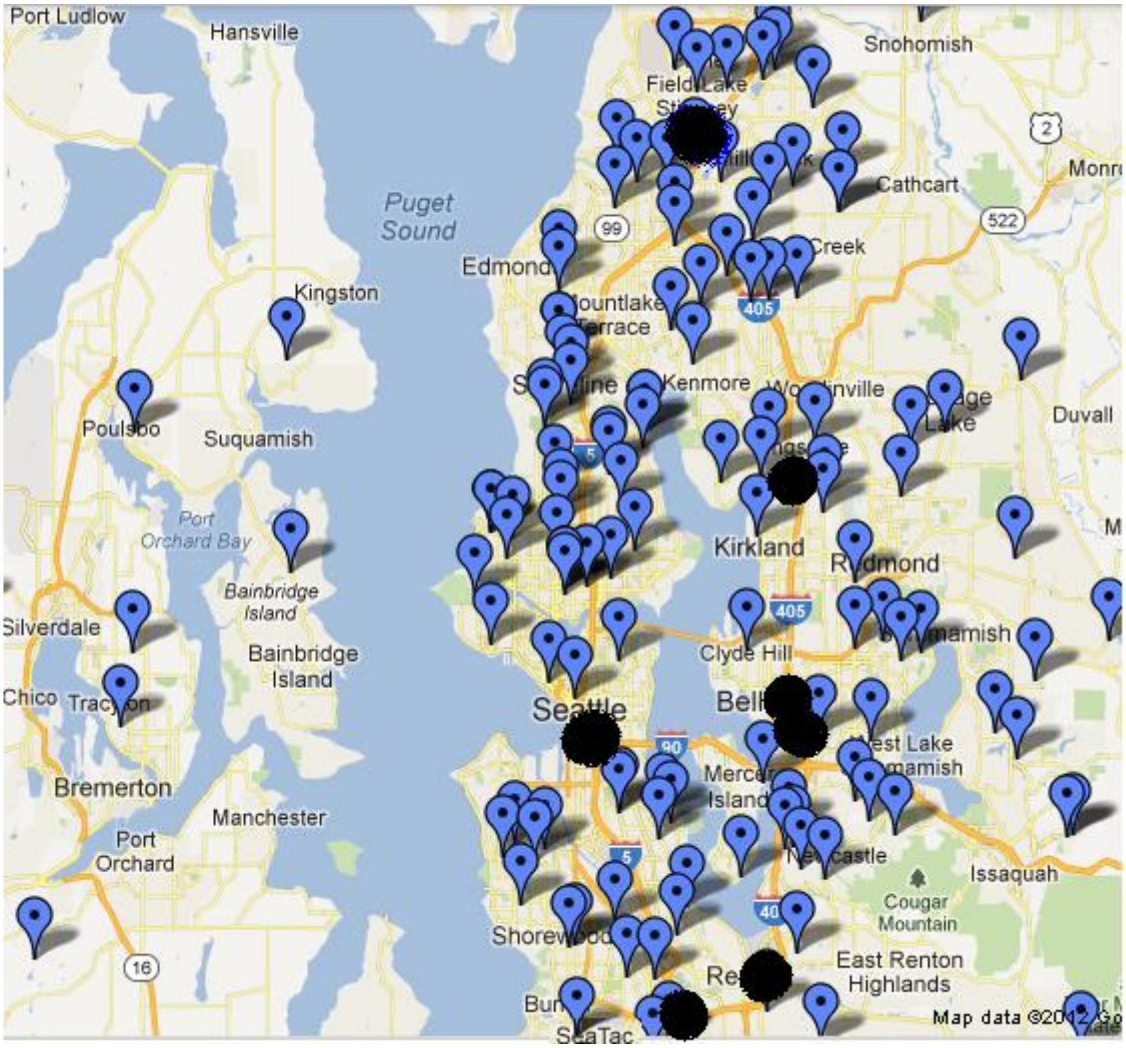

Figure 2b. Sample Household Residences and Bottlenecks South of Seattle Tacoma Airport (Sea-Tac)

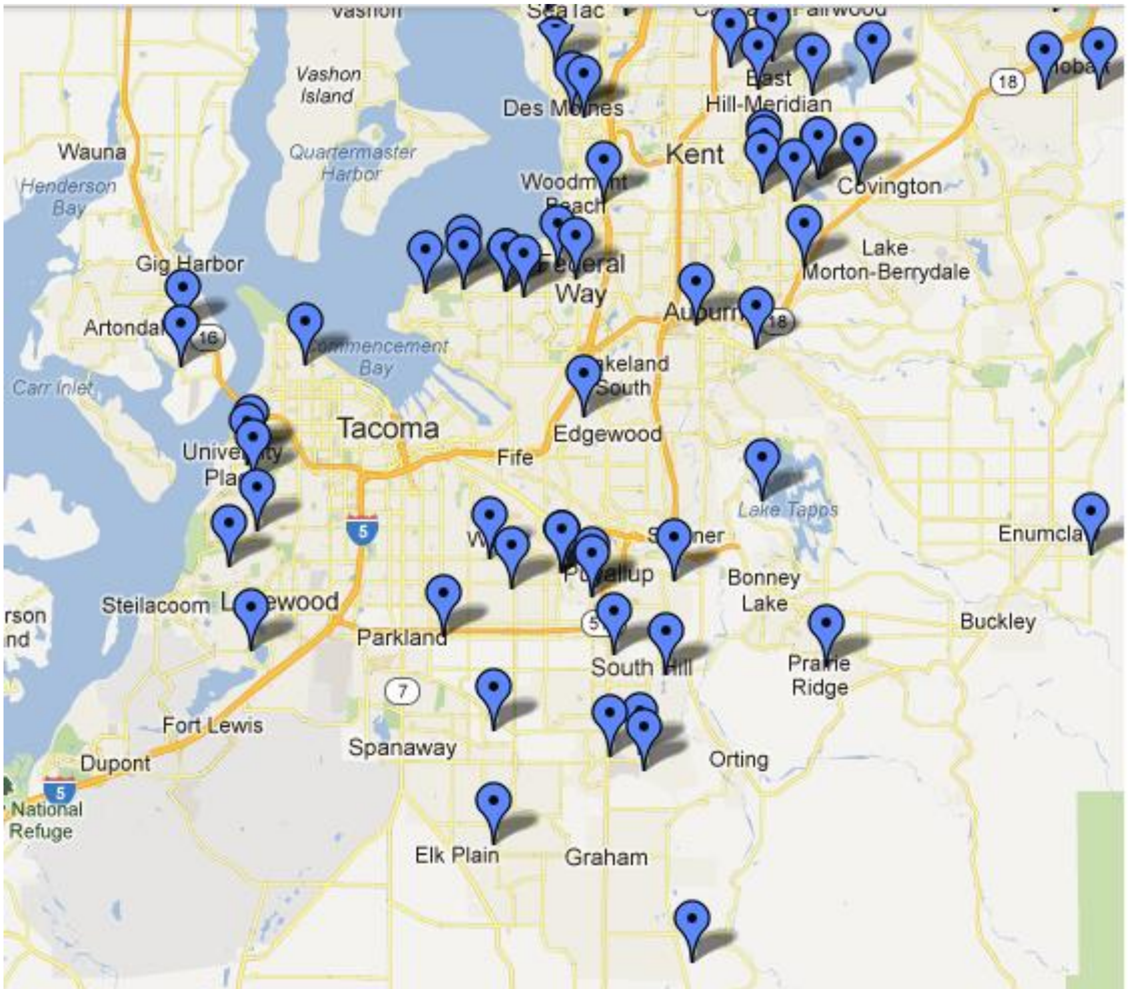

Figure 3. A Dynamic Model of a Decisionmaker's Vehicle Holding and Replacement Decisions