Green Sharing and Spatial Distribution of Automobile Demand: The Role of Income, Public Transit, and Portfolio Considerations

Yoshifumi Konishi*
Department of Policy and Planning Sciences
Faculty of Engineering, Information and Systems
University of Tsukuba

and

Shunsuke Managi
Department of Urban and Environmental Engineering
School of Engineering, Kyushu University

VERY PRELIMINARY

*Corresponding author. Mailing address: 1-1-1 Tennodai, Tsukuba, Ibaragi, 305-8573, Japan. E-mail: ykonishi@sk.tsukuba.ac.jp. Phone: +81-29-853-5072. Fax: +81-29-853-5070.

Abstract: Demand for vehicle transport varies over geographic space, so does the economic impact of a policy to control vehicle emissions, due to spatial heterogeneity in income and public transit density. We explore this proposition empirically in a discrete-continuous choice framework that accounts for: (i) the value of vehicle transport that depends on public transit density, (ii) correlation due to portfolio considerations between multiple-vehicle holdings, and (iii) correlation due to unobservables between vehicle ownership and utilization. Our findings are two-hold. First, the effect of income, public transit, and portfolio considerations are critical in resolving two empirical puzzles in Japan: Overall price elasticity of demand for vehicle ownership increases with vehicle size; invariance of demand for hybrid vehicles with respect to public transit density. Second, we demonstrate that a policy to promote sharing of eco-friendly vehicles has a far better carbon-reduction potential than a tax incentive policy, particularly in low density areas.

JEL Codes: H23, H31, L62, Q54
Key Words: Automobiles, carbon emissions, control function, discrete-continuous choice, energy efficiency, feebates, hybrid vehicle, portfolio effect, public transportation, vehicle demand

1. Introduction

Greenhouse gas (GHS) emissions pose a significant risk of climate change. Road transportation is the major and rapidly growing contributor to the risk today. In theory, an efficient tax on gasoline or carbon emissions can fully restore economic efficiency since it forces firms and consumers to internalize the social cost of climate change in all economic margins: from residential/employment choice to transport mode choice, and to vehicle ownership and utilization. Economists have been increasingly aware, however, of the difficulties associated with reducing carbon emissions from road transportation [see Anderson *et al.* (2011) and Knittel (2012) for a review of the issues]. In the U.S., for example, vehicle miles travelled nearly doubled from 1970 to 2009 while the fleet average fuel economy (of only new cars) improved at a moderate rate over the same period (Knittel, 2012). Associated with this increase in vehicle miles is the rapid urban sprawl: i.e., urban development in geographically sparse, low density areas (Glasser and Kahn, 2010). The form of cities, roads, and public transit networks we observe today is the result of this development, so is the demand for vehicle transport [e.g., Beaudoin and Lawell (2018); Bento *et al.* (2005); Boarnet and Crane (2001); Duranton and Turner (2011)].

The thesis of the manuscript is an important conjecture that arises naturally from this: the economic impact of a (price-based) carbon-reduction policy can be highly spatially heterogeneous due to geographic variation in the price elasticity of demand for vehicle ownership and utilization. The key here is the geographic distribution of incomes and public transit networks. On one hand, demand for vehicle ownership and utilization is low in urban areas with high public transit density. Hence, *ceteris paribus*, the demand is elastic with respect to both the cost of ownership (e.g., car price) and the cost of utilization (e.g., gasoline price). On the other hand, income tends to be high in urban areas where abundant employment opportunities exist (even after adjusting for the cost of living).[1] This tends to make the demand less price elastic. The situation reverses in rural areas. Thus, the intricate interaction between the two effects tends to generate substantial heterogeneity in demand elasticity across geographic areas. As a result, the overall effectiveness of a carbon-reduction policy may depend critically on the joint distribution of household income and public transit density over geographic space.

We quantitatively examine this conjecture in Japan, which present two empirical puzzles of economic significance in the current context. First, the estimates of the product-level demand in Konishi and Zhao (2017) suggest that in Japan, the price elasticities of new car demand are larger for small-sized vehicles than for large-sized vehicles. This is puzzling in

[1] There is a large literature showing that real wages are not equalized across regions.

that low-income households tend to buy smaller cars, and hence, the demand for these cars must be price-elastic, in principle. Second, in our data (to be described below), we observe the share of hybrid car owners is virtually flat with respect to the density of public transit. This is also puzzling because the demand for vehicle transport increases quite sharply, so does the average fuel economy of owned cars, as public transit becomes less dense. The key to these two puzzles indeed lies in the spatial distribution of vehicle demand. Households in low density areas tend to buy keicars[2], which are cheap and fuel efficient, instead of hybrid cars, which are more fuel efficient yet are more expensive. These households also have high demand for vehicle ownership and utilization, which tends to make their demand for keicars price inelastic.

With this observation, we attempt to achive three goals in this manuscript. Our first goal is to estimate the structural parameters of the demand for vehicle ownership and utilization in Japan, explicitly accounting for its dependence on public transit density. The focus here is on the parameters that can fully explain the two empirical puzzles mentioned above. Our second goal is to evaluate how the economic performance of a price-based incentive policy for vehicle carbon emissions may depend on the geographic distribution of public transit network. Here, the key is the distribution of price elasticities over geographic space, which are the results of the intricate interaction between household-level demand elasticities and the distribution of public transit and socioeconomic variables. Lastly, we seek to evaluate the supply of a car-sharing platform as an alternative to the price-based incentive policy. Over the last decade, car-pooling has emerged as an innovative transit solution to the conflicting needs: reducing the cost of public transportation infrastructure while reducing carbon emissions from vehicle transport (e.g., Ostrovskyand Schwarz, 2018). Our empirical strategy is suited to evaluate the essential trade-off consumers make in choosing a car-sharing option. Car-sharing allows consumers to reduce the cost of ownership and utilization, but its downside is all the hussle and inflexibility of scheduling their driving needs with others. Hence, 'buying' the n-person sharing option allows one to reduce the rental cost of ownership by $1/n$, but forces her to spend $1/n$ less time behind wheel. The question of who buys this option essentially boils down to identifying the parameters of marginal utility terms with respect to income and vehicle utilization.

To achieve these goals, we develop and estimate a model of vehicle ownership and utilization using household-level microdata in Japan. The model is based on the discrete-continuous choice model of Dubin and McFadden (1984) and Train (1993). As in Train (1993), our model accounts for a sequence of three choices: the number of vehicles owned, the class/type of

[2] Kei-car is an extremely small car segment with displacement levels of 660 cc or less. Kei-cars account for roughly 20% of the total vehicle sales in Japan.

each vehicle owned, and the vehicle kilometers traveled (VKT) for each vehicle owned. What is new here is that we incorporate three new aspects into the model. First, we introduce portfolio considerations in a manner analogous to Gentzkow (2007) and Wakamori (2015). That is, we explicitly model the correlation between choice of the first car and that of the second car. This is accomplished by addting the terms that capture utility from having a particular combination of vehicles. Adding the portfolio effect allows us to model intricate behavioral responses that are likely important in our empirical context. For example, when the price of a hybrid car decreases, a consumer who prefer to own a certain combination of vehicles, say, due to a large household size, may end up owning that combination of vehicles with at least one hybrid. Likewise, when the price of minivans increases, a consumer who prefers to own a combination of a minivan and a regular car may still continue to own the same combination by buying a cheaper regular car. This tends to yeild some 'stickiness' of vehicle demand. Second, following the spirit of random-coefficient logit, we allow the parameters on (indirect) utility to depend explicitly on a measure of public transit density.[3] With this formulation, a consumer who has a high valuation of fuel economy because she lives in an area with limited public transit would substitute to a less expensive alternative, such as a keicar, that has less but similar fuel economy when a hybrid car is too expensive. Third, the resulting model produces the high demensional sample correction terms that enter the vehicle utilization equation. To address it, we employ Dahl's control function approach (2003). By this, we account for two types of correlation in the VKT equation due to unobservables: one between the uses of multiple cars and the other between car ownership and utilization. We estimate this model, using a large, nation–wide internet survey in 2016. The survey contains a usable sample of approximately 100,000 households, and hence, we have a sufficiently large subsample within each decile of public transit density. The covariates' variations that come with it allows us to estimate the demand parameters that vary by that density.[4]

As with all analogous studies using household survey data [e.g., Bento *et al.* (2005), Bento *et al.* (2009), Goldberg (1998), Train (1993), West (2002)], the identification of the model parameters is challenging. It requires exogenous variation in access to public transit, the rental price of vehicle ownership, and the operating cost of vehicle utilization as well as

[3] We explain how we construct this measure in **Section 3**. A fractional polinomial regression of this meausure against population density is fit quite well.

[4] Japan is known for its highly efficient public transit network. Yet, public transit is only sparsely distributed in most non-urban areas. In Tokyo, for instance, only 31.7% of the working population drives to work (MLIT, 1998). This number is surprisingly small even compared to 65.7% in New York (American Community Survey, 2009). In Toyoma, a moderately populated prefecture in Japan, the share of the working population who drives to work comes at 83.8% (MLIT, 1999). This number is close to what we observe in most U.S. counties.

an exclusion restriction. To ensure exogenous sources of variation in public transit density, we use road density, public bus access, and railway access as of 1980. To control for the endogeneity of car prices, we exploit the fact that the Japanese government implemented a variety of incentives for eco-friendly vehicles since 2001. We use information on purchase year/month to adjust the rental prices of *all* vehicles in each household's choice set. Dahl's control function approach requires instruments that vary at the household level and affect vehicle purchase decisions, yet do not affect vehicle utilization decisions directly. In our case, this is trivially satisfied because the vehicle characteristics of *all* vehicles in the choice set enter the ownership choice while only those of the chosen car enter in the utilization equation. The exclusion restriction is further strengthened by the fact that we adjust the vehicle attributes of each consumer's choice set based on her purchase year/month. Lastly, in estimating the car utilization equation, we restrict the sample to those who purchsed the cars only after 2012. This allows us to remove the superious correlation that arise between our key variables and the unobservables due to the endogenous duration of car holding.

We have four sets of results. First, the estimated price elasticity of vehicle ownership is indeed smaller for keicars than for hybrid cars, overall as well as by public transit density, yet the elasticity declines monotonically with public transity density. This finding is consistent with our conjecture above: Households living in low density areas tend to have lower real incomes, and thus, their demand tends to be more price-elastic, yet their demand for keicars are relatively inelastic precisely because keicars offer economical options — inexpensive and fuel-efficient. Thus, the result offers the household-level evidence in support for the product-level demand estimated with the market-level data in Konishi and Zhao (2017). Second, the effects of income, public transit, and portfolio considerations, all taken together, can also explain why the share of hybrid cars is roughly constant over geographic space. The effects of income and public transit alone can, in part, explain why the demand for hybrid cars is relatively low in low density areas and high in high density areas. Yet, these two effects alone tend to overstate the demand for hybrid cars in low density areas while understaing it in high density areas. The model incorporating the portfolio effect corrects these prediction errors by making consumer's demand 'sticky' — consumers who prefer a certain combination tend to stick to that combination in face of relative price changes. Third, with this estimated model, a feebates policy is shown to be highly ineffective — the social cost of carbon (set at \$60 per ton) is too small to induce a meaningful change in CO_2 emissions from driving. This is despite the fact that the policy is estimated to influence important economic margins in an intricate way. Feebates indeed tend to *increase* car holdings and car utilization rates, due to income and rebound effects. In contrast, the car-sharing policy is estimated to *decrease* the number of cars owned, the monthly VKT, and driving-related carbon emissions. The effects are most pronounced in low density areas. This contrast between the two policies is indeed due to an intricate interaction between the effects of income, public transit, and portfolio considerations (see our discussion in **Section 7**).

Our work complements three strands of literature: (a) empirical studies that investigate the relationship between urban structures and demand for vehicle transport, either using city-level observations [e.g., Levinson and Kumar (1997); Glasear and Kahn (2010)] or household-level observations (e.g., Beaudoin and Lawell (2018); Bento *et al.* (2005); Boarnet and Crane (2001); Train (1986)); (b) the economic incentives for efficiently controlling emissions from mobile sources [see Knittel (2012) and Anderson *et al.* (2011) for a comprehensive review on the topic]; and empirical studies that estimate the discrete-continuous decision model on car ownership and utilization, with applications to the effect of gasoline tax [Bento *et al.* (2009), Train (1986), and West (2004)], to the effect of CAFE standards [Goldberg (1998)], and to the effect of feebates [D'Haultfoeuille *et al.* (2014)]. To the best of our knowledge, however, none has explicitly examined the role of income, public transit, and portfolio considerations in understanding the spatial heterogeneity in demand for vehicle transport and its implications for the design of a carbon-reduction policy. Moreover, ours is probably the first to quantify the environmental outcomes of car-sharing in a real empirical context.

2. A Statistical Overview of Vehicle Demand in Japan

We start by presenting a statistical overview of the relationship between access to public transit and vehicle-related household choice in Japan. For ease of visualization, we report household choice against a single composite index of public transit accessibility. We construct this index by an (unweighted) average of two district-level measures of railway transit network (incl. cable cars, surface rails, and subways). The first measure is the kilometers of railways per square kilometer and is intended to measure ease of access to destinations via railways. The second is the percentage of the habitable area within a district that has at least one train station within 15-min walking distance and is intended to measure ease of access to railways. We do not include the bus network in this index because in Japan, bus network is highly developed and even rural residents have access to a bus station within a walking distance. To confirm, **Figure 1** displays two scatter plots: the composite index above (top panel) and a similar index using bus network (bottom panel), both against the population density using district-level observations. The figure demonstrates that the composite index using railways has an increasing, but non-linearly relationship to population density while the index using buses has very little geographic variation and its inclusion would misleadingly overstate public transit accessibility. Hence, we use this index as the measure of public transit density.[5]

Figure 2 plots (a) the rate of vehicle ownership, (b) monthly vehicle kilometers traveled (VKT), (c) fuel economy ratings, and (d) the share of hybrid cars, all against public transit density.[6] In line with the theory and the findings around the world, vehicle ownership rate, VKT, and fuel-economy ratings decline sharply with public transit density. There is also a sign of self-selection: Only households with high enough demand for driving own cars, and thus, an average VKT tends to be high in the most dense area despite its public transit availability. Interestingly, however, demand for hybrid vehicles does not seem quite responsive to the density of public transit. Why don't households with limited access to public transit not own hybrid vehicles, despite their demand for driving seems twice as high as those in high density areas?

A key to resolving this puzzle lies with the fact that there is a large gain in fuel economy by downsizing cars (Knittel, 2011). Households in low density areas tend to buy small-sized vehicles, that are both cheaper and more fuel efficient, for daily transportation. **Figure 3** plots (a) average car size, (b) the share of *kei*-cars, (c) average car price, and (d) average household income, against public transit density.[7] The figure demonstrates that both average car size and average car price decrease quite sharply with a decline in public transit density — with the share of keicars also rising sharply. Combined with the fact that average household income is lower in low density areas, the figure leads us to an observation that low-income households in low density areas are buying low-cost cars that offer high fuel-economy performance. This observation may still come as a surprise. Low density areas tend to have much wider roads and higher rates of traffic accidents. Hence, there is also a consumer gain from up-sizing cars. The key to resolving the puzzle, therefore, boils down to identifying the essential trade-offs consumers make over fuel economy, size/safety, acceleration, and price of a vehicle, and how these trade-offs differ by income and public transit availability. Our empirical strategy attempts to get at this by accounting for heterogeneous demand elasticities with respect of these vehicle attributes.

[5] We think that the frequency of service would give us a more reliable measure of public transit accessibility in case of bus service. Unfortunately, we do not have access to such data. However, we have a sense that the frequency of bus service would generate essentially the same geographic variation as our composite measure.

[6] We have detailed information on up to two most frequently used vehicles each household owns. Only about 5% of households own three or more vehicles. Hence, we use the sum over two cars for VKT and the (unweighted) average for fuel economy ratings and the hybrid vehicle shares. In figures are the averages of these over households.

[7] Keicars are extremely small passenger vehicles with displacement of 660 cc or less. They are highly popular in Japan. Keicars account for roughly 30% of sales in the domestic sales in Japan. The Japanese government offer a variety of tax incentives for these vehicles. As in **Figure 2**, we use the (unweighted) average over the two most frequently used cars for car size, keicar share, and car price.

Our data also point to another important economic margin, which may become critical in identifying the demand parameters. Car owners rarely own two vehicles of the same type. Hence, households seem to make portfolio considerations in buying a vehicle rather than considering each vehicle in isolation. In our data, of those who own two (or more) cars, only 36% own the same type of vehicles. The remaining 64% own a combination of either regular-keicar (24.5%), regular-minivan (16.3%) or keicar-minivan (23.2%). What seems critical for our empirical analysis is that the portfolio of vehicle holdings also varies sharply with the density of public transit. **Figure 4** demonstrates that the share of households who own two vehicles of the same type (i.e., regular-regular, keicar-keicar, or minivan-minivan) rises with public transit density whereas the share of those who hold a keicar and another vehicle type (i.e., either regular or minivan) declines. Whether this phenomenon can be explained purely by geographic or household-level variation or by explicit modeling of the portfolio effect is another important empirical question. Hence, we incorporate this aspect into our empirical model.

3. The Empirical Model

There is a large literature in transportation research that empirically examines consumer's vehicle ownership and utilization in the spirit of the continuous-discrete choice framework à la Dubin and McFadden (1984). In that literature, consumer's choice is modeled as a two-stage decision process. In the first stage, the consumer chooses whether to own a car or not, and if she does, which type of car to own. In the second stage, the consumer chooses how much to drive over a given period of time. Our empirical framework follows this tradition, but extends it in several important ways.

We start by describing the second stage. Following the convention [e.g., Bento *et al.* (2005), Bento *et al.* (2009), Goldberg (1998), West (2002)], we assume that monthly driving distance m (in log) of consumer i who lives in area s and who owns vehicle j is:

$$\ln m_{ijs} = \alpha_s \ln(y_i - r_{ij}) + \beta_s \ln p_{ij} + \mathbf{X}_i' \boldsymbol{\delta}_{js} + \eta_{ijs}, \tag{1}$$

where y_i is i's household income, r_{ij} and p_{ij} are, respectively, the annual rental price of vehicle ownership and the operating cost of utilization per unit of driving distance for car j for household i, and $\mathbf{X}_i$ is a vector of other household characteristics (e.g., age, education, household size, work status). The primary parameters of interest are the income and price elasticities, α_s and β_s, of vehicle utilization. We allow these parameters to depend on the geographic characteristics $\mathbf{S}_s$:

$$\alpha_s = \alpha + \mathbf{S}_s'\gamma_\alpha; \quad \beta_s = \beta + \mathbf{S}_s'\gamma_\beta.$$

We are particularly interested in learning how these elasticities depend on access to public transit.

In the first stage, the consumer chooses whether or not/which vehicle to own given her expectation about her vehicle utilization. Following McFadden's conditional logit framework, consumer i's indirect utility from choosing car model j is given by

$$u_{ijs} = \rho \ln(y_i - r_{ij}) + E[v(m_{ijs})|\mathbf{X}_i, \mathbf{Z}_j, \mathbf{S}_s] + \epsilon_{ijs}, \tag{2}$$

where y_i and r_{ij} are defined in eq. (1), ϵ_{ijs} is i.i.d. errors, and $E[v(m_{ijs})|\mathbf{X}_i, \mathbf{Z}_j, \mathbf{S}_s]$ is consumer i's expected utility from driving, which we assume to depend explicitly on household characteristics $\mathbf{X}_i$, vehicle characteristics $\mathbf{Z}_j$, and geographic characteristics $\mathbf{S}_s$. Eq. (2) embodies the idea that consumers make a trade-off between money spent on buying a car j versus the utility of driving that car. Note that although the parameter on net income ρ is not allowed to vary by household, the income and price elasticities of demand still differ across households since $\partial u/\partial y = \rho/(y-r)$ and $\partial u/\partial p = -\rho/(y-r)$. Furthermore, because the consumer demand is a non-linear function of observables, these demand elasticities also depend on $\mathbf{X}_i, \mathbf{Z}_j$, and $\mathbf{S}_s$. In contrast, eq. (1) is the (reduced-form) equation, and thus, without explicitly allowing for interaction terms, the income/price elasticity will be constant across areas.

We posit that $E[v(m_{ijs})|\mathbf{X}_i, \mathbf{Z}_j, \mathbf{S}_s]$ is linear in parameters and is given by $E[v(m_{ijs})|\mathbf{X}_i, \mathbf{Z}_j, \mathbf{S}_s] = \mathbf{Z}_j'\boldsymbol{\lambda}_{is}(\mathbf{X}_i, \mathbf{S}_s)$. We emphasize here that because we observe driving distance only for those who own cars, it is not the *observed* or *realized* driving distance, but rather the *expected* driving distance, that enters this indirect utility. Provided that we correctly specify $\mathbf{Z}_j'\boldsymbol{\lambda}_{is}(\mathbf{X}_i, \mathbf{S}_s)$, inclusion of these observable characteristics allows us to capture the ideas that the utility of vehicle ownership is higher for those with higher demand for vehicle utilization than for those with lower demand. Following the spirit of random-coefficient logit, we specify $\boldsymbol{\lambda}_{is}$ as

$$\boldsymbol{\lambda}_{is}(\mathbf{X}_i, \mathbf{S}_s) = \boldsymbol{\lambda}_0 + \mathbf{X}_i'\boldsymbol{\lambda}_1 + \mathbf{S}_s'\boldsymbol{\lambda}_2.$$

Thus, the marginal utility of a vehicle attribute varies by household as well as geographic characteristics. This formulation is analogous to, but is substantially more flexible than, nested logit. Note that this formulation is quite analogous to mixed logit, except that we do not include the interaction terms with unobservables. We are fairly confident that we have a sufficiently rich set of covariates with a large enough sample size, which give us sufficient variation in observables to control for the unobservable terms. **Indeed, when we estimate a mixed-logit version of the model, none of the standard deviation parameters become statistically significant [I might have done this incorrectly. I need to further investigate on this point].**

It is known that OLS regresstion of eq. (1) would generally give us biased estimates of parameters due to sample selection because we observe each consumer's driving behavior only for the car model chosen, but not for car models that had not been chosen. To see this in our empirical setup, note that we can write $v(m_{ijs}) = E[v(m_{ijs})|\mathbf{X}_i, \mathbf{Z}_j, \mathbf{S}_s] + e_{ijs}$, and hence, the error term ϵ_{ijs} in eq. (2) is confounded with another error term e_{ijs}, forming the joint error term $\mu_{ijs} \equiv e_{ijs} + \epsilon_{ijs}$. Consequently,

$$
\begin{aligned}
&E[\eta_{ijs}|\mathbf{X}_i, \mathbf{Z}_j, \mathbf{S}_s, j \text{ is chosen}] \\
&\quad = E[\eta_{ijs}|V_{ijs} + \mu_{ijs} \geq V_{iks} + \mu_{iks} \text{ for all } k] \neq 0
\end{aligned}
$$

where V_{ijs} is the observable part of the indirect utility, and the last inequality follows because η_{ijs} contains some of the information in e_{ijs}, the unobserved part of utility from driving car model j. Simply put, consumers would enjoy driving cars of their favarites and not so much for others.

To address this selection problem, previous studies either assumed a joint distribution of errors (μ_{ijs}, η_{ijs}) or used a selectivity correction à la Dubin and McFadden (1984). The former is known to place severe restrictions on the selection process, while the latter is known to become imprecise or infeasible when there are many alternatives in the first stage decision. We instead use Dahl (2002)'s control function approach to correct for this selection bias in the case of many alternatives. Specifically, Dahl showed that, in case of high-dimensional alternatives, eq. (2) can be consistently estimated using estimates of individual purchase probabilities:

$$
\ln m_{ijs} = \alpha_s \ln(y_i - r_{ij}) + \beta_s \ln p_{ij} + \mathbf{X}_i' \boldsymbol{\delta}_{js} + \sum_{j=1}^{J} M_{ij} \times T_{ij}(P_{i0}, P_{i1}, \ldots, P_{iJ}) + \upsilon_{ij}, \qquad (3)
$$

where $T_{ij}(\cdot)$ is some unknown function of purchase probabilities $P_{i1,} \ldots, P_{iJ}$. Dahl suggests that, in practice, we may include only a few probabilities such as the probabilities of the first-best choice, the second-best choice, and the outside option. We follow this advice in our estimation.

Portfolio Effect: We augment this model by allowing for the decision to own multiple-car holdings and the dependence between the choices on these multiple holdings. In the augmented model, choices are restricted to two vehicles per household since we have detailed information only on two, most frequently used cars.[8] Following Wakamori (2015), consumer i's utility from owning a pair of cars j and k is given by

$$\begin{aligned} u_{i(j,k)s} &= \rho \ln(y_i - r_{ij} - r_{ik}) + E[v(m_{ijs})|\mathbf{X}_i, \mathbf{Z}_j, \mathbf{S}_s] \\ &\quad + E[v(m_{iks})|\mathbf{X}_i, \mathbf{Z}_k, \mathbf{S}_s] + \Gamma(j,k;\mathbf{X}_i) + \epsilon_{i(j,k)s}, \end{aligned} \tag{4}$$

where $\Gamma(j,k;\mathbf{X}_i)$ is the portfolio-effect, which captures the idea that households derive utility from owning a particular combination of vehicle types. It is quite plausible, for example, households with children may prefer owning a minivan for recreational use, yet may prefer owning a sedan or kei-car for daily commuting use. As discussed in **Section 2**, we indeed observe such portfolio considerations in our data — a majority of households own two vehicles of different types, and its prevalence varies across regions.

As in Wakamori (2015), we consider three mutually exclusive sets of vehicle types: i.e., kei-cars $\mathcal{K}$, sedan/regular cars $\mathcal{R}$, and minivans $\mathcal{M}$. Then the portfolio effect is given by

$$\Gamma(j,k;\mathbf{X}_i) = \sum_l \kappa_{(j,k)} x_{il},$$

where x_{il} denotes l-th characteristic of household i and $\kappa_{(j,k)}$ is the combination-specific parameter for $(j,k) \in (\mathcal{K},\mathcal{K}),(\mathcal{K},\mathcal{R}),(\mathcal{K},\mathcal{M}),(\mathcal{R},\mathcal{R}),(\mathcal{R},\mathcal{M})$, or $(\mathcal{M},\mathcal{M})$. There are two advantages of modeling the portfolio considerations this way. First, as discussed in Wakamori (2015), unlike other previous studies, the approach does not assume products are either complements, substitutes, or independent, and instead, allow the estimates of parameters to flexibly capture complementarity patterns observed in the data. Second, we can estimate this model using conventional conditional logit routines available in most statistical softwares. Consumer would choose to own a pair (j,k) if and only if her utility of owning the pair is higher than owning any other pairs, owning a vehicle j, or not owning any.

We can incorporate the portfolio effect into the sample correction terms in the vehicle milage equation (3) by slightly modifying our notation. To see this, let J_1 and J_2 be the sets of products for her first and second cars, respectively. Let us augment J_2 by including 'zero', an option to own no second car. Adjoining these two sets and an outside option to own no car, we create the joint choice set J, which contains $1 + \#J_1 \times \#J_2$ alternatives. The consumer chooses an alternative j from this adjoined set. That is, one may choose to own no car $(0,0)$, choose to own one car $(j_1, 0)$, or choose to own two cars (j_1, j_2). with a slight abuse of the notation, (j,k) in place of j, the model described by (2) and (3) is essentially

[8]As discussed in **Section 3**, only 6% of the households in our data hold three or more cars.

intact.

There is one subtle, yet important, issue in estimating the VKT regression (3) — we observe VKT for *each* of the vehicles owned, and hence, (3) must be estimated separately for each vehicle, accounting for that vehicle's attributes. Here, the difficulty is that the households who own multiple cars are likely to decide on how often to use one vehicle jointly with other vehicles. Consequently, the utilization levels are likely to be correlated across car holdings. The literature to date seems silent as to how to address this issue. In her seminal work, Goldberg (1998) estimates the VKT regression using observations on newly purchased cars only, ignoring this correlation in vehicle utilization. Bento *et al.* (2005) instead use the VKT per vehicle, averaged over vehicles owned, as an dependent variable, excluding vehicle-specific regressors from the list of independent variables.

In the context of the present paper, the policy impact on the second car's VKT is quite important. Hence, we address this issue as follows. We estimate the VKT regression, pooling all VKT observations on the two most frequently used cars, with a dummy indicating a second car. This ensures that the same sample correction terms enter the VKT regression for the two cars owned by the same household, yet accounting for the fact that one of the observation is on the second car. This allows multiple-car owners' vehicle utilization decisions to be correlated across their vehicle holdings, either through observable household-level characteristics or through (unobservable) selectivity terms. Because our model of car ownership accounts for portfolio effects, the selectivity correction terms in (3) control for the unobservable correlations that are specific to the same household who decide to own a particular combination of cars.[9]

4. Data

Our study relies on a large cross-sectional sample of households from a nationwide internet survey conducted in November 2016 in Japan. In designing the survey, we aimed for two goals. The first is to obtain a sufficiently large sample, with variations in household-level characteristics, for each population density decile. This is essential for our study because we need comparable households to separetely identify the demand parameters that vary by

[9] One could, instead, estimate the seemingly unrelated regression or the second-car's VKT equation independently. Both approaches resulted in parameter estimates that are hard to interpret. For example, the estimated elasticity on net income was negative. We would think that this occurs precisely because of the substitution in vehicle utilization between the two cars. The households with high incomes primarily drive the first car for daily use, keeping the second car only for luxurious use. The households with low incomes, on the other hand, are likely to own the second car for primary use, and hence, they drive the second car more. The estimates may be simply capturing this correlation.

income and public transit: i.e., those with different levels of income, yet with the same level of access to public transit as well as those with the same level of income, yet with different levels of access to public transit. The second is to collect sufficiently detailed information on each household's vehicle ownership and utilization that is comparable to the U.S. Consumer Expenditure Survey (CES). In particular, we aim to collect information such as the number of vehicles, the vehicle type (fuel economy, engine/fuel type, horsepower, make, size, weight, vintage), the year/month of purchase, and the vehicle kilometers traveled since the purchase. Such detailed information on vehicle ownership and utilization is not available in national consumer surveys in Japan.

The survey was administered under the contract with Nikkei Research Inc. to the pool of registered internet monitors. The survey resulted in a sample of 105,000 usable respondents with complete responses. As with other internet-based surveys, we did not have direct control over the sampling process. Our usable sample, however, covers a sufficiently large number of households in every prefecture, with sufficient variation in key socioeconomic variables such as age and income. In the appendix, we compare our sample distribution against the population distribution by prefecture. The geographic distribution of our survey respondents by prefecture is sufficiently close to the population distribution, though populated prefectures (e.g., Tokyo and Kanagawa) are over-represented while less populated prefectures (e.g., prefectures in Kyushu region) are under-represented. As expected, average household incomes in our sample are slightly higher than in the population for most prefectures, although we do not see significant differences in average household sizes. Our results may be somewhat biased toward households with relatively higher incomes.[10]

We supplement the survey with the data from various sources. First, we use the GIS datasets on city boundaries, bus stops, train stations, train networks, hospitals, road length, and public parks from the National Land Numerical Information Download Service, made available online by the Ministry of Land, Infrastructure, and Transportation (MLIT). We use the coodinates of train stations and the line data on train networks to construct the composite measure of public transit density at the 'city-district' level (see **Section 2**). Our definition of 'city-district' follows that of the Ministry of Internal Affairs and Communications (MIAC). As of 2018, there are 1,724 city districts in Japan. Second, we also use the car catalog data from the carsensor.com, one of the largest online car retailers in Japan. The survey respondents are asked to provide detailed information on each of the cars they own (up to their second car): i.e., model year/month, purchase year/month, make, model name,

[10] There is a large literature in environmental economics, examining the extent of bias in demand estimation that may arise due to the internet-based survey. The results are mixed. Comparing the internet survey versus other modes of survey, some (Lindhjem *et al.*, 2011 and Nielsen, 2011) report no or small bias while others (Boyle *et al.*, 2016) report a non-negligible bias.

displacement level, curb weight, and mileage. We use these to match their cars with those listed in the carsensor catalog to obtain other vehicle characteristics such as fuel economy ratings, horsepower, size, and transmission. Third, we use the district-level population estimates (both current and forecast) provided by the National Institute of Population and Social Security Research (IPSS). Lastly, we also obtain a measure of prefecture-level road congestion from MLIT, regional consumer price and gasoline price series from the Ministry of Economy, Trade and Industry (METI), and historical discount rates from the Bank of Japan. Detailed descriptions on the variables used in the manuscript are available in the **Appendix**.

Table 1 reports the means and standard deviations of key variables by population density. The table confirms substantial variations both within and across population density quintiles, which we exploit in our estimation. First, all measures of public transit sharply decrease as population density declines. Not only that, we have substantial variation in these measures within each density quintile, and interestingly, more so in low-density quintiles: coefficients of variation are 0.58, 0.50, 0.52, 0.36, and 0.27 for the lowest, 4th, 3rd, 2nd and the highest population densities. This is in sharp contrast to household characteristics. Average household incomes decline as population density declines, yet the coefficients of variation stay roughly the same across all quintiles. The same is true with household size. This 'within' variation in public transit measures helps us identify the effects of public transit on vehicle ownership/utilization. Second, as we have seen, the rate of car ownership rises quickly as population density declines, possibly in response to declines in public transit availability. Interestingly, however, the coefficient of variation for car ownership declines as population density declines. Instead, the coefficient of variation for the number of cars owned rises, from 0.33 in the highest density quintile to 0.52 in the lowest density quintile. This point is also closely related to our next observation. Third, we observe a smaller variation in vehicle utilization than vehicle ownership: after taking logs, the coefficients of variation for monthly VKT range from 0.17 to 0.19 for the first car (= most frequently used car), and from 0.18 to 0.24 for the second most used car. Combined, these two observations are suggestive of the tendency that households absorb the impact of public transit (un)availability by adjusting the number of cars owned rather than by adjusting the vehicle utilization. We take this as suggesting that it is indeed important to account for correlation between ownership decisions and utilization decisions. Lastly, the sample characteristics of the first car seem to differ substantially from those of the second car. The second cars are cheaper, more fuel-efficient, and smaller on average in virtually all density quintiles (while coefficients of variation are similar between the first cars and the second cars). Interestingly, for their first cars, households in low density areas are more likely to own hybrid cars than in high density areas. Yet, the opposite is true with their second cars. These points seem to re-confirm the existence of the portfolio effect discussed in **Section 2**.

5. Estimation and Identification Strategy

The discrete-continuous choice model we develop in **Section 3** is estimated in two steps. In the first step, we estimate the discrete choice model, assuming the form of indirect utility as in eq. (4). For this step, estimation is done by Stata's alternative-specific conditional logit routine. In the second step, we estimate the VKT regression in eq. (3), pooling all VKT observations for all cars owned by households in the sample. In this step, we use polynomials of predicted probabilities from the first step as selection control terms as in Dahl (2002). We experiment with a polynomial of up to third degree, using the probabilities of the chosen, the no-car, the highest-likelihood, and the lowest-likelihood options. Based on the fitness and the sign/significance of key variables (i.e., net income and cost per kilometer of driving), we end up using the third-degree polynomials of the highest and the lowest probabilites. Furthermore, with this approach, the conventional covariance estimator is biased (Dahl, 2002). Hence, we use bootstrapped standard errors, with 500 draws, for inference.

Because we use one-shot household survey for both steps, the identification of the parameters relies on cross-sectional variation at both the household and the district levels in economic/geographic variables. Though this poses a challenge in identification, this is typical of studies that estimate the discrete-continuous choice model of car-holding decisions using survey data (see Goldberg, 1998; Bento *et al.*, 2005, and Bento *et al.*, 2009). In the literature, four identification challenges are discussed: (1) endogeneity of measures of public transit and (2) endogeneity of rental price of car ownership in the first-stage choice of car ownership; (3) endogeneity of operating cost of car utilization (due to sample selection) in the second-stage choice of car utilization. In addition to these, there is an issue with the endogenous duration of car ownership. Below, we discuss each of these and our strategies to address them in order.

Public transit: For this issue, we impose a conventional assumption that public transit measures are at least predetermined prior to their car-holding decisions. Indeed, these measures are known to work better than other forms of urban structures (see Bento *et al.*, 2005). Even with this assumption, however, it is quite hard to ensure that households' intrinsic preferences for car holdings are uncorrelated with measures of public transit — households may make residential choice in conjunction with choice of car holding. To address this concern, we also exploit the idea that the past publit transit is a good instrument (e.g., Duranton and

Turner, 2011) and use the two-stage residual inclusion (2SRI) method. Specifically, we construct the composite index of public transit density as of 1980 as an instrument in the stage 'zero' and then the residuals are added to the first-stage vehicle ownership equation.[11] The procedure is known to produce consistent estimates in nonlinear models with endogenous regressors (Terza *et al.*, 2008; Wooldridge, 2015).

Rental price of car ownership: Studies on the automobile demand estimation are often concerned with the endogeneity of car prices. There may be product attributes consumers see but researchers do not, such as brand images, styles, and non-price incentives. Since they are demand-shifters, they may as well be correlated with car prices. Because ours is based on household data while car prices are mostly determined at the market level, this concern is less serious. In addition, we include maker, car-type, fuel-type, and used-car dummies to control for unobservable product characteristics. This identification strategy is analogous to Goldberg (1998) and Bento *et al.* (2009). However, individual households also negotiate prices at the dealer level. Hence, there may be some measurement error in our price variable that may be correlated with vehicle attributes at the local level. To take care of the concern, we use *time-varying* car-related tax incentives as exogenous price shifters. Specificatlly, we use information on purchase year/month to adjust the rental prices of *all* vehicles in each household's choice set. The Japanese government implemented a variety of incentives for eco-friendly vehicles since 2001 [see Konishi and Zhao (2017) for more detail]. This not only gives exogenous price variation, but also implicitly restrict each household's choice set.

Sample selection: For the second-stage choice of car utilization, we essentially have two identification issues. Both issues are closely related to the endogeneity of p, i.e., the operating cost of vehicle utilization. The first issue concerns the sample selection we discussed in **Section 3**. Because the fuel cost per kilometer is a function of the fuel economy rating of the car chosen, this variable is clearly correlated with unobservable demand for driving distance. To control for this, we use Dahl's control function approach. This approach requires the exclusion restriction: i.e., we need instruments that vary at the household level and affect vehicle purchase decisions, yet do not affect vehicle utilization decisions directly. In our case, this is trivially satisfied. Vehicle utilization in eq. (3) depends only on the attributes of the car that is actually owned, but not those of alternatives, while vehicle ownership choice in eq. (4) depends not only on the attributes of the chosen car but also on those of the other alternatives. Previous studies essentially use the same argument in implementing Dubin-McFadden type correction. The identifying condition is further strengthened by the fact that we use the prices of alternative vehicles actually observed at the time of purchase

[11] We make sure that all of the households included in the estimation do not own cars sold prior to 1980.

instead of using contemporaneous values observed today, as discussed above.

Duration of car ownership: The second identification issue for the vehicle utilization regression concerns the duration of car ownership. The problem here is that those who own cars longer tend to be those with low incomes and own cars with low fuel economy ratings (both because cars sold in the past tend to be fuel-inefficient and because fuel-efficient cars tend to be expensive), while at the same time, these households continue to hold cars despite their low incomes precisely because they have high demand for driving. Hence, households who hold the same car for a long time tend to be those with lower incomes and higher costs of vehicle utilization. This results in spurious correlations that bias the parameter estimates in the opposing direction — i.e., negative correlation between net income and VKT and positive correlation between the cost of driving and VKT. Our first-stage model of car ownership allows us to account for the economic margins that affect "whether or not", "how many", and "what type of cars" to own, but not "how long" to own. Naturally, the selection correction terms cannot control for the endogeneity that arises from the duration of ownership. To address this concern, we restrict the observations to those on cars purchased after 2012.

6. Results

6.1. Vehicle Ownership

We first report on the first-stage discrete choice model of car ownership in **Table 2**. Two sets of results are reported in the table: The model with and without portfolio effects. For each model, the first row exhibits the estimates of the mean parameters for our key variables while the second and the third columns present their interaction effects with transit availability and household size. These interactions allow us to account for cross-region as well as within-region heterogeneity in demand for car ownership. The estimated model also includes maker dummies, a used-car dummy, fuel-type dummies (hybrid, diesel), maker dummies (Toyota, Honda) and vehicle-type dummies (keicar, minivan) as well as their interactions with metropolitan dummies to tease out the effects of unobserables.

Virtually all parameters are statistically highly significant, and their signs are consistent with economic theory as well as previous studies that estimated similar models. First, the parameter on the logged net income is significantly positive, which implies that consumers with higher incomes are more likely to own cars and that consumers prefer cheeper cars, holding all else constant. Second, the mean parameter on the fuel cost per kilometer (YPK) is signigicantly negative, which implies that consumers on average value fuel economy. However, its interaction terms suggest that consumers with access to transit density or with large familiy tend to care less about fuel economy (even after controlling for vehicle size and vehicle types). Third, consumers on average prefer high acceleration. Interestingly, consumers with large family size tend to value acceleration much less. Lastly, though not reported, hybrid and keicar dummies are significantly positive, whereas diesel car dummy is significantly negative. The Japanese consumers thus prefer hybrid or keicar over regular gasoline cars, yet prefer gasoline cars over diesel cars. This occurs because in Japan, diesel cars are not popular as they are often perceived as unsafe (due to its low ignition temperature) or unclean (due to its high sulfur content before desulfurization process).

There is one anomaly that may seem at odds with previous studies in the U.S. — the mean parameter on car size is significantly negative, implying that in Japan, consumers on average value smaller cars. Note, however, that we obtain these estimates *after* controlling for the vehicle types (i.e., keicars, minivans etc) and portfolio effects, which also vary by household size. Hence, the negative sign on car size should be capturing the preferences for the compactness of vehicles within, but not across, vehicle class. Our interpretation therefore is that because the roads and parking spaces are narrow virtually everywhere in Japan (even in rural areas compared to roads in U.S.), consumers on average prefer smaller cars, given their preferred vehicle class. Furthermore, its interaction terms suggest that consumers living in areas with high public transit availability value smaller cars even more, whereas households with large family tend to value them less. This is consistent with the findings in Konishi and Zhao (2017).

Table 2 also signifies the importance of accounting for portfolio considerations — both mean and interaction parameters on many of the portfolio terms are statistically significant. The mean parameter on the keicar-keicar combination is significantly positive, suggesting that consumers, on average, value the keicar-keicar combination more than others, relative to the regular-regular combination, which we take as the base combination. Somewhat surprisingly, the mean parameters on all other portfolio combinations are negative, which implies that consumers, on average, do not prefer having two vehicles of different types (over the base combination). However, looking at their interaction terms with household size, we see that consumers with large family tend to value the keicar-keicar combination significantly less while valuing the combinations of different vehicle types more. This is consistent with the idea that consumers adjust their vehicle portfolios according to the households' needs. Furthermore, the interaction terms with public transit measure suggest that in general, the value of having a keicar in the portfolio increases with public transit availability. This make sense because in areas with public transit, consumers tend to drive short distance mainly for daily shopping and pickups, for which small-sized cars are more convenient.

To gauge the importance of accounting for the portfolio effects, we also compare the predicted shares of alternative vehicle choices against the observed shares by transit density density. **Figure 5** reports the results of this exercise for two vehicle types: hybrid and keicars. Unlike in **Figures 2 and 3**, this figure report the 'unconditional' shares: i.e., not conditioned on owning a car. Hence, the figure evaluates the predictive performance on two margins: whether or not to own a car and which type of car to own. The figure demonstrates that the model with portfolio effects does a consistently better job than the model without portfolio effects. The model without portfolio effects predicts these shares to respond more sharply to public transit density, failing to explain the flat hybrid demand discussed in **Section 2**. The accuracy of prediction on the ownership shares of these two vehicle types is quite important in simulating the counterfactuals as these offer the most viable options for reducing gasoline use. Given these, we use the model with portfolio effects in the subsequent analyses.

6.2. Vehicle Utilization

Next, we turn to the vehicle utilization regression. Our focus is on demand elasticities with respect to (net) income and operating cost, and on the influence of public transit density on these elasticities. The first four columns of **Table 4** report the results without the congestion variable interacted with the key variables: income and operating cost (in log). The next four columns include these interaction terms. We estimate each of these regressions with varying levels of controls: metropolitan dummies and selection corection terms. Though the parameter estimates are not reported, all regressions control for other vehicle characteristics (fuel type and car type dummies), demographic characteristics (age, household size, marital status, number of cars owned, work status, distance to work, years of education) and urban structures (district-level population density, access to hospital, and access to public parks). As discussed in **Section 5**, we only report the results with the third-degree polynomials of the highest and lowest estimated choice probabilities for selectivity correction.

The estimate of the mean income elasticity is always positive, but become larger and more statistically significant when we include the congestion interaction term. This makes intuitive sense. When a household's income increases, the household would increase her time to allocate for leisure, but how much she would increase time to spend on driving depends on how congested roads are. She would drive more if roads are less congested. Furthermore, the income elasticity is smaller for consumers living in high density areas. This makes sense since recreational value of driving would be larger for consumers who have limited access to in high public transit. Note that we can focus on leisure-related arguments as we already control for distance to work.

The estimate of the mean price elasitity is negative and statistically significant across all specifications. This is consistent with economy theory, but suggests the success of our control strategy — in studies that use cross-sectional household-level data, this elasticity is often estimated with bias toward zero or even positive [e.g., Goldberg (1998)]. The interaction with public transit density is positive and marginally significant. This suggests that the demand for driving is less price-elastic in high density areas. This may seem somewhat counter-intuitive at first. When the price of gasoline increases, for example, consumers who have access to public transit can use public transit instead of driving, and therefore, we would expect the demand for driving to be more price elastic in areas with high public transit density. This logic ignores the income effect. Consumers with access to public transit use cars primarily for a recreational purpose, and the fuel cost accounts for a relatively small portion of the recreational expenditures. On the other hand, consumers with limited acess to public transit use cars for daily use, and the fuel cost accounts for a larger share of consumer expenditures. Hence, this makes the demand for driving more price-elastic for consumers in areas with low public transit density. The inclusion of the congestion interaction term makes the interaction term with transit density become less statistically significant and smaller in magnitude. Our reasoning is that congestion and transit density tend to co-move, and thus, are picking up similar effects.

With all specifications, the selectivity correction terms are jointly highly significant. This is despite the fact that we already control for a number of observables. Moreover, inclusion of these correction terms generally improves the statistical significance of our key parameters. This implies that there is selection on unobservables, and thus, omitting selection correction terms is likely to bias our estimates. When translating these parameter estimates into elasticity estimates by public transit quintile, this gain in consistency does seem to matter. Hence, we are in general in favor of models with selection correction. On one hand, the estimate of the mean income elasticity parameter is positive and gets larger in magnitude after selection correction. This means that unobservable demand factors for vehicle utilization, which the selectivity terms are meant to control for, are negatively correlated with household (net) incomes. On the other hand, the estimate of the mean price elasticity parameter does not seem to change much, both in magnitude and in statistical significance.

6.3. Elasticity Estimates

Table 5 reports the estimates of various elasticities of car ownership and utilization by transit density quintile. We estimate the price elasticity ϕ_{dt} of car ownership with respect to rental price r for fuel type t (t = hybrid, diesel, keicar) for transit density quintile s as follows:

$$\phi_{dt} = \frac{\partial s_{dt}}{\partial r_{dt}} \cdot \frac{r_t}{s_{dt}},$$

where s_{dt} and $\frac{\partial s_{dt}}{\partial r_t}$ are calculated as

$$s_{dt} = \frac{1}{N_d} \sum_{i \in I_d} \hat{P}_{it} \quad \text{and} \quad \frac{\partial s_{dt}}{\partial r_{dt}} = -\frac{1}{N_d} \sum_{i \in I_d} \frac{\hat{\rho}}{y_i - r_{it}} \hat{P}_{it}[1 - \hat{P}_{it}],$$

where $\hat{P}_{it}$ denotes the estimate of household i's probability of holding car models of fuel type t, $\hat{\rho}$ is the estimate of the parameter on net income, $y_i - r_i$ is household i's observed net income, N_d is the number of sample households in d, and I_d is the set of households in d. For the price elasticity of VKT, we simply evaluate the parameter estimates $\hat{\alpha}_d$ and $\hat{\beta}_d$ for $\ln(y - p)$ and $\ln(YPK)$ at the means of congestion and transit density measures for each transit density quintile d, using the estimates from our preferred specification **(6)** in **Table 3**.

The price elasticities of car ownership range from -0.224 to -1.304. The numbers are in line with, but slightly lower than, those reported in analogous studies (Bento *et al.*, 2009; Konishi and Zhao, 2017). As expected, the car ownership elasticities get smaller (more inelastic) with transit density — consumers in low density areas are more price elastic than those in high density areas. Moreover, hybrid cars face the most elastic demand. In contrast, Keicars face the most price inelastic demand. We believe this is the key to resolving the opening puzzle. Consumers in low density areas have lower incomes, and hence, are more price elastic. Yet, even these consumers have substantially more price elastic demand for hybrid cars than Keicars. Consequently, consumers in low density areas tend to buy keicars for cheaper prices despite their high level of vehicle utilization. This also offers support for the earlier findings of Konishi and Zhao (2017) who also find that Keicars face more price inelastic demand, despite that Keicars typically serve low-income groups whose demand tend to be more elastic *ceteris paribus*.

Vehicle utilization is not quite elastic with respect to (net) income. Interestingly, the income elasticities decline with public transit density, and turn practically zero (i.e., statistically highly insignificant) in the third to first transit density quintiles. We believe this reflects the competing effects of income. As consumer's income rises, the consumer increases the time to spend on leisure, which tends to increase vehicle utilization, but at the same time, it also increases the opportunity cost of time. With the high congestion level in urban areas where public transit availability is also high, the consumer faced with the high opportunity cost of time tends to opt for public transit rather than drive their cars. In Japan, roads in even moderately populated areas are still congested, compared to the U.S. Hence, it is quite reasonable to observe that the income elasticities in such areas tend toward zero (or even negative).

The price elasticities of VKT are negative and statistically significant in all transit density quintiles, but slightly lower in magnitude than those reported in Bento *et al.* (2009). Because our car utilization regression focuses on a subsample of households who have bought cars since 2012, much of the variation in YPK comes mainly from cross-sectional variation in fuel economy ratings rather than that of gasoline price. Hence, our price elasticity estimate is essentially capturing the rebound effect: i.e., buying a more fuel efficient car makes a consumer drive more. Our results are consistent with the argument that the rebound effect is often overstated (Gillingham *et al.*, 2013). Our estimates suggest that the rebound effect is even smaller in Japan than in the U.S. However, the use of cross-section data is known to bias the price elasticity estimate toward zero due to the selection bias [for example, Goldberg (1998) obtain a positive and statistically insignificant price elasticity estimate]. Hence, there is a possibility that our elasticity estimates may still be biased toward zero even after controlling for the selectivity terms.

7. Counterfactual Analysis

7.1. Counterfactual Scenarios

We now use the estimated demand system to simulate the economic outcomes of several counterfactual policy scenarios. The goal of this counterfactual analysis is three-fold: (i) to quantify how the economic impacts of a carbon-reduction policy may vary over geographic space, (ii) to understand how the income, the public transit, and the portfolio effects may interact to generate this spatial heterogeneity, and (iii) to quantitatively evaluate the carbon-reduction potential of car-sharing relative to a more conventional tax-incentive policy, particularly in non-urban areas. With these goals in mind, we simulate the effects of (a) tax incentive policy and (b) car-sharing policy on the following economic outcomes: car ownership, car utilization, and carbon emissions from driving.

There are a number of reasons why these are non-trivial economic questions in the present context. First, we take the 'holding' approach as opposed to the 'transactions' approach in our demand estimation [see Goldberg (1998) or Bento *et al.* (2009) on the difference between the two]. With this approach, we implicitly get at the long-run demand elasticities, which tend to be smaller than the short-run elasticities. For example, a majority of households who buy an eco-friendly car when tax incentives are in place may be those who are thinking to replace their cars anyway, and hence, represent a small share of overall car holders. Second, we take into account portfolio considerations in multiple car holdings. There is a reason to believe that eco-car tax incentives may not be effective in view of the portfolio considerations. A consumer may buy an eco-friendly car for one of her vehicle holdings, yet with her reduced net income, may end up buying a cheaper, less fuel efficient car for another use. Third, we also take into account demand heterogeneity due to differences in income as well as access to public transit. As shown above, consumers have more price-elastic demand for hybrid cars than keicars and in low density areas than in high density areas. Given this, it is not clear how the demand response varies over geographic space, for example, when the price of a hybrid car falls relative to the price of a keicar.

Before discussing the details on each policy, a few caveats are in order. Our ability to simulate economic outcomes is limited to those that make use of demand-side parameters only. We have neither data nor policy relevant variations to estimate the supply-side parameters on the firm-level cost/behavior with respect to public transit density or introduction of a car-sharing platform. Therefore, in the analysis below, we assume perfectly elastic supply for both policies. Consequently, we do not attempt to simulate car prices, gasoline prices, and any other economic outcomes that would require supply-side parameters. These are simply treated as 'fixed' in our simulation analysis. We also assume no revenue recycling (or revenue accounting) because it would only obscure the essence of the main analysis.

Counterfactual 1. Eco-car Incentives

We first consider tax incentives on car holdings. Tax and subisidy incentives are a common policy apparatus for inducing purchase of eco-friendly cars around the world. Such incentive schemes are more generally termed as 'feebates' since they impose fees (or taxes) on less fuel-efficient vehicles and rebates (or subsidies) on more fuel-efficient vehicles (See Anderson *et al.*, 2011). Institutional details on these incentive schemes vary by country. A number of empirical studies have credibly quantified the economic impacts of such incentive policies in a variety of contexts: the U.S. (Beresteanu and Li, 2011), France, Germany, and Sweden (Klier and Linn, 2015), and Japan (Konishi and Zhao, 2017). Our purpose here is not to repeat these studies; rather, to understand how consumer's response may vary over geographic space due to the intricate interaction of income, public transit access, and portfolio considerations. Hence, we do not take any particular country's incentive scheme, and instead, consider a general feebates schme: tax and subsidy in proportion to its carbon footprint. Specifically, the tax t_j (or the subsidy) is charged to the holding of car j according to the formula:

$$t_j = \tau \times E[v] \times (EPK_j - x),$$

where τ is the social cost of carbon emissions, $E[v]$ is the expected annual VKT, EPK_j is the carbon emissions per kilometer of driving distance, which is estimated as the emissions per liter of gasoline divided by the fuel economy (km/L), and x is the base emissions rate beyond which consumers pay a tax and below which consumers receive a subsidy. Greenstone *et al.* (2015) estimate the social cost of carbon emissions can be as large as \$60 per ton. We use this as an upper bound estimate. Note that the Japanese government offered a variety of tax and subsidy incentives since 2009. Our demand estimation accounts for these, and hence, this incentive scheme is implemented on top of these existing incentives.

Counterfactual 2. Eco-car Sharing

We implement this scenario as the introduction of a new ownership option by the government. Specifically, the government supplies, perfectly elastically, a platform for n-person sharing of ownership of an eco-friendly car. That is, the consumer pays for the rental price on the shared vehicle itself, but the fixed cost of establishing and maintaining the car-sharing platform is paid by the government. Hence, from the consumer's eyes, the supply of this option is perfectly elastic. Then, the consumer faces the essential trade-off between the shared cost of vehicle ownership/utilization versus the reduced rate of utilization. The question is, who 'buys' this option, where and how much?

Our structural approach can get at this question. n-person sharing enables consumers to co-own cars, reducing the rental cost of vehicle ownership by $1/n$. However, n-person sharing causes some inconvenience to the sharing users, reducing the rate of utilization as much as $1/n$.[12] Given our empirical model in **Section 3**, we can simulate the consumer's indirect utility of an n-person car-sharing option as

$$u_{ijs} = \rho \ln(y_i - r_{ij}/n) + (1/n)E[v(m_{ijs})|\mathbf{X}_i, \mathbf{Z}_j, \mathbf{S}_s] + \epsilon_{ijs}. \tag{5}$$

In implementation, we simply add this new option to the set of alternatives while varying the size of n. From (5), it is clear that a consumer who has a low valuation for the benefit of vehicle utilization relative to the value of income tends to buy this option. Hence, con-

[12] There is some uncertainty as to the utility loss from car sharing. On one hand, the utilization loss may not be as much as $1/n$. A majority of consumers use vehicles for limited times of a day and a week. By scheduling the times of use, n users may be able to satisfy all driving needs. This may be particularly true if the information and autonomous vehicle technologies allow consumers to costlessly schedule their use in the future. On the other hand, there may be some mental cost associate with scheduling friction, which may be added to the loss of utility. This utility loss can be large when the size of platform users is small, particularly in non-urban areas, but may become negligible if the size of users get large (the economy of agglomeration). Hence, to be conservative, we assume the maximum utility loss of $1/n$ for this simulation scenario.

sumer behavior under this counterfactual depends on the estimates of the marginal utility parameters with respect to these two terms.

There is a remaining question as to what type of vehicles should be promoted as the sharing platform. We focus on a hybrid vehicle. Hybrid cars offer substantially better mileage per liter of gasoline than the keicars, holding car displacement, size, weight and other on-vehicle amenities. Because consumers in low density areas rely on cars for daily commuting and transportation, allowing those consumers to use hybrid cars instead of keicars is likely to reduce pollution and increase welfare. Nonetheless, we have seen that the consumers tend to own keicars instead of hybrid vehicles and that the estimated demand elasticities can explain this puzzle. Then an interesting question is, Is the sharing of a hybrid vehicle effective in reducing vehicle-related CO_2 emissions?

7.2. Results of Counterfactuals

In **Table 5**, we report the results of the two counterfactual simulations: Tax incentives (or feebates) **in Panel I** and Eco-car sharing policy in **Panel II**. For comparison, we also report, in the top panel, the predicted outcomes under the observed conditions. For ease of interpretation, we report the counterfactual outcomes in average per household for each subsample sorted by population density quintile as of the 2015 level. Note that all statistics reported in this table are *unconditional* averages; this is in contrast to **Table 1** where the descriptive statistics are reported as *conditional* averages (i.e., of those who own cars only). Furthermore, our demand estimation is restricted to the holding of at most two cars. Hence, the counterfactual results are directly comparable to the top panel of **Table 5**, but not to **Table 1**.

Panel I reports the results of the feebates policy in which the base emissions rate x is set at the median of all observations. We have two important observations. First, unfortunately, this incentive policy is estimated to increase both the average car ownership rate and the average VKT, and hence, increase overall CO_2 emissions from driving. This result may be surprising, but is indeed consistent with economic theory. Feebates shift the consumer demand from less fuel-efficient vehicles to more fuel-efficient ones. But fuel-efficient vehicles are relatively cheaper, both because they tend to be smaller (e.g., keicars) and because they are subsidized. This tends to increase car holdings in two ways: it induces some consumers to start owning a car; it frees up some of the consumer's budget and when this income effect is large enough, some consumers may end up owning a second car. This is known as the income effect of the incentive policy. Furthermore, the operating cost of driving is lower for fuel-efficient cars, which induces consumers to drive more. This is known as the rebound effect of the incentive policy.

Second, as expected, we observe substantial heterogeneity in these two effects. The feebates policy is estimated to increase the car ownership rate, but decrease the car utilization rate in the lowest population quintile. The income effect seems largest in the highest population quintile whereas the rebound effect seems largest in the middle quintile. This heterogeneity essentially concerns the question of "who is induced into which type of vehicles". In the lowest quintile, households generally have lower incomes and are car owners to start with due to low public transit availability, and the feebates tend to induce these households into holding keicars rather than hybrid cars. Keicars offer less comfort in driving, and hence, drivers of keicars tend to drive less. In contrast, in the highest quintile, households generally have higher incomes and are not car owners. The feebates tend to induce them into holding a car. The average car utilization rises because there are more car owners. These results highlight the difficulty associated with controlling vehicle-related carbon emissions via an incentive policy as well as the importance of the interplay between the effects of income, public transit, and portfolio considerations.

We now turn to the results of the eco-car sharing policy in **Panel II**. The table reports the results with $n = 10$: the policy-induced supply of 10-person sharing of a hybrid vehicle, with the attributes of the shared hybrid held at the means of the hybrid cars we observe today. For ease of quantification, we count the car-sharing user as a non-car owner. The results are more encouraging. First, the car-sharing platform is estimated to decrease both the average car ownership rate and the average utilization rate, reducing the overall CO_2 emissions from driving, relative to the observed state. Interestingly, the demand for the car-sharing option is higher in areas with lower population density: the share of households who choose the car-sharing option is 0.1% in the highest quintile and rises monotonically to 3.4% in the lowest quintile. By assumption, the users of n-person car sharing drives the car by $1/n$ of time. Consequently, the average VKT also declines. However, in the highest quintile, the car-sharing platform offers a low-cost ownership option. Hence, it induces some non-car-owners to start owning and driving the shared vehicle, and hence, increase the average VKT. We emphasize here, however, that the counterfactual simulation assumes no friction in scheduling, over time as well as over geographic space. Ttherefore, the results indicate that in such a frictionless world, the households in the *lowest* quintile benefits the most from the car-sharing option. This is somewhat counter-intuitive, and has an important policy implication. In areas with low public transit density, the demand for vehicle transport is high, and hence, one may be tempted to infer that consumers in such areas are less likely to benefit from car-sharing. Our results indicate this inference is incorrect. In fact, there are many marginal driviers in these areas whose incomes are relatively low and are forced into driving for necessity. These marginal drivers prefer driving less if such an option is available at a lower cost. This tells us where the government should allocate more resources in establishing such a system and removing the scheduling friction.

Lastly, we examine the impacts of perturbing x (from x = 10th percentile to x = 100th percentile; 100th percentile corresponds to a carbon tax on car holdings) and n (from $n = 1$ to $n = 20$). **Figure 5** displays, for each policy, the scatter plots describing two relationships: (A) car ownership and utilization and (B) user share and CO_2 emissions from driving. Except for the user shares, the numbers are in percentange changes relative to the base scenario, which we take as x = 100th percentile for the feebates policy and $n = 1$ for the car-sharing policy. The figure demonstrates an interesting contrast between the two policies with respect to these relationships. First, as we reduce the base emissions rate x, more vehicles are offered rebates, so that the rate of car ownership increases. At the same time, the rebates induce consumers to own cheaper, more fuel-efficient cars, so that the rate of car utilization tends to increase. These effects together tend to increase CO_2 emissions from driving. Interestingly, the share of hybrid cars are inresponsive to the changes in x (*and* the tax incentives in general).

On the other hand, as the number of users n increases, the share of households who choose the car-sharing option tends to increase. This occurs because the utility gain from a decline in ownership/utilization cost due to an increase in n tends to exceed the utility loss from the decline in vehicle utilization. As expected, there is a decreasing marginal return to n — the share of car-sharing users becomes non-responsive to changes in n at some point. Interestingly, the effects are highly heterogeneous across regions. The share of car-sharing users responds much more sharply to an increase in n in areas with lower population density. Indeed, in the highest density quintile, the relationship is reversed: the share of car-sharing users becomes smaller as n increases. This result is somewhat counter-intuitive — it means that the utility gain from the decline in ownership/utilization cost is larg relative to the utility loss from the decline in vehicle utilization in non-urban areas, but is small in urban areas. Furthermore, as discussed above, this increase in the share of car-sharing users translates into a (small) decrease in CO_2 emissions from driving, due to decreases in both the number of car holdings and driving distance. Interestingly, the average driving distance also declines in the highest density area, despite the decrease in the share of car-sharing users, because an increase in n tends to induce more households with high demand for driving into the car-sharing option.

There is another take-away message from this exercise. It is, in general, difficult to generate meaningful carbon reduction via tax incentives — the social cost of carbon emissions is too small to influence the demand for vehicle transport. The car-sharing policy has a far better potential; it seems to affect all economic margins in much larger magnitudes than tax incentives. We emphasize here that this general take-away is unlikely to depend on the unbiasedness of our price elasticity estimates — what is driving the difference between the two policies is the wedge between the change in price versus the change in utility from vehicle utilization.

8. Conclusion

Public transportation infrastructure is not economically viable in low density areas. How best to design incentive policies for ownership and sharing of eco-friendly vehicles is the key to efficiently control vehicle-related carbon emissions in low density areas. We empirically investigate this issue, exploiting a rich cross-sectional survey data collected in Japan in 2016. The survey covers a large enough sample of households in each density level, which allows us to identify the demand parameters by income and access to public transportation. In estimation, we augment the discrete-continuous choice model of vehicle ownership and utilization to incorporate portfolio considerations in multiple-vehicle holding, and use a control function approach to correct for multi-dimensional sample selection bias in vehicle utilization. The focus of our empirical research is on an apparent economic puzzle: the vehicle utilization rate rises quickly as access to public transportation declines, whereas the ownership rate of eco-friendly vehicles does not. We show evidence that the key to resolving the puzzle is the differential income and price elasticities of demand between low and high density areas. With the estimated demand, we simulate the economic outcome of a policy to promote sharing of eco-friendly vehicles. Our findings indicate that such a policy is more effective in reducing carbon emissions than the conventional tax incentive policy, and therefore, can be is a viable substitute for public transportation infrastructure in low density areas.

References

[1] Anderson, Soren T., Ian W. H. Parry, James M. Sallee, and Carolyn Fischer. 2011. Automobile Fuel Economy Standards: Impacts, Efficiency, and Alternatives. *Review of Environmental Economics and Policy* 5 (1): 89-108.

[2] Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. Automobile Prices in Market Equilibrium. *Econometrica* 63 (4): 841-890.

[3] Bento, Antonio, Lawrence Goulder, Mark Jacobsen, and Roger von Haefen. 2009. Distributional and Efficiency Impacts of Increased US Gasoline Taxes. *American Economic Review* 99 (3): 667-699.

[4] Bento, Antonio, Maureen Cropper, Ahmed Mobarak, and Katja Vinha. 2005. The Effects of Urban Spatial Structure on Travel Demand in the United States. *The Review of Economics and Statistics* 87(3): 466–478.

[5] Beaudoin, Justin and Cynthia Lin Lawell. 2018. The Effects of Public Transit Supply on the Demand for Automobile Travel. *Journal of Environmental Economics and Management* 88: 447-467.

[6] Boarnet, Marlon and Randall Crane. 2001. The Influence of Land Use on Travel Behavior: Specification and Estimation Strategies. *Transportation Research Part A: Policy and Practice* 35 (9): 823-845.

[7] Dahl, Gordon. 2002. Mobility and the Return to Education: Testing A Roy Model with Multiple Markets. *Econometrica* 70 (6): 2367-2420.

[8] D'Haultfœuille, Xavier, Pauline Givord, and Xavier Boutin. 2014. The Environmental Effect of Green Taxation: The Case of the French Bonus/Malus. *The Economic Journal* 124 (578): 444–480.

[9] Dubin, Jeffrey and Daniel McFadden. 1984. An Econometric Analysis of Residential Electric Appliance Holdings and Consumption. *Econometrica* 52 (2): 345-362.

[10] Duranton, Gilles and Matthew Turner. 2011. The Fundamental Law of Road Congestion: Evidence from US Cities. *The American Economic Review* 101(6): 2616–2652.

[11] Gentzkow, Matthew. 2007. Valuing New Goods in a Model with Complementarity: Online Newspapers. The American Economic Review 97 (3): 713-744.

[12] Glaeser, Edward and Matthew Kahn. 2010. The Greenness of Cities: Carbon Dioxide Emissions and Urban Development. *Journal of Urban Economics* 67 (3): 404-418.

[13] Goldberg, Pinelopi Koujianou. 1998. The Effects of the Corporate Average Fuel Efficiency Standards in the US. *Journal of Industrial Economics* 46 (1): 1-33.

[14] Knittel, C. R. 2012. Reducing Petroleum Consumption from Transportation. *Journal of Economic Perspectives* 26 (1): 93-118.

[15] Konishi, Yoshifumi and Meng Zhao. 2017. Can Green Car Taxes Restore Efficiency? Evidence from the Japanese New Car Market. *Journal of the Association of Environmental and Resource Economists* 4(1): 51-87.

[16] Terza, Joseph, Anirban Basu, and Paul Rathouz. 2008. Two-stage Residual Inclusion Estimation: Addressing Endogeneity in Health Econometric Modeling. *Journal of Health Economics* 27 (3): 531-543.

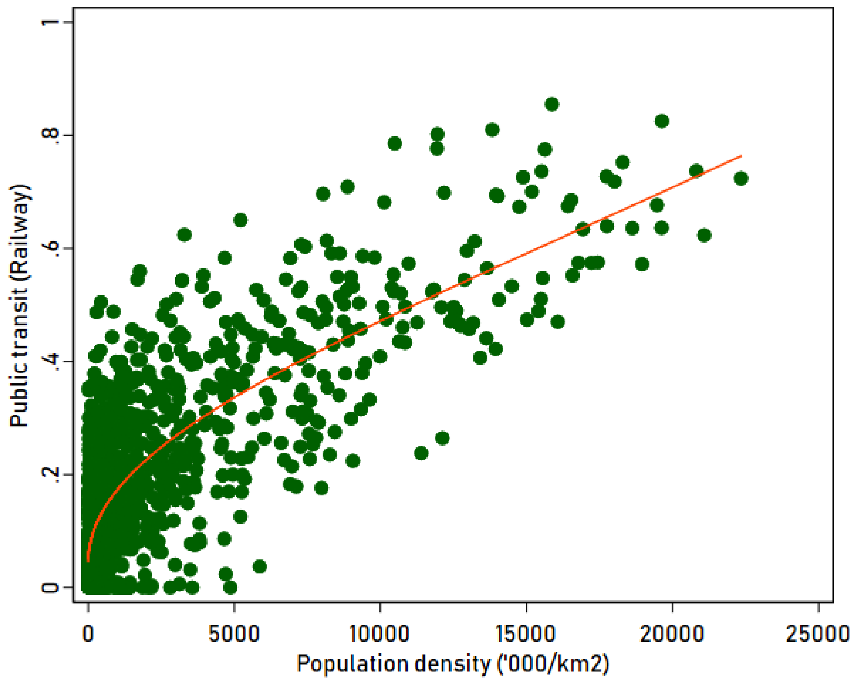
[17] Train, Kenneth. 1986. *Qualitative Choice Analysis: Theory, Econometrics, and an Application to Automobile Demand*. Cambridge: MIT Press.

[18] Wakamori, Naoki. 2015. Portfolio Considerations in Differentiated Product Purchases: An Application to the Japanese Automobile Market. *SFB/TR Discussion Paper 499.*

[19] West, Sarah E. 2004. Distributional Effects of Alternative Vehicle Pollution Control Policies. *Journal of Public Economics* 88: 735– 757.

[20] Wooldridge, Jeffrey. 2015. Control Function Methods in Applied Econometrics. *Journal of Human Resources* 50 (2): 420-445.

Figure 1. Measures of Access to Public Transit versus Population Density



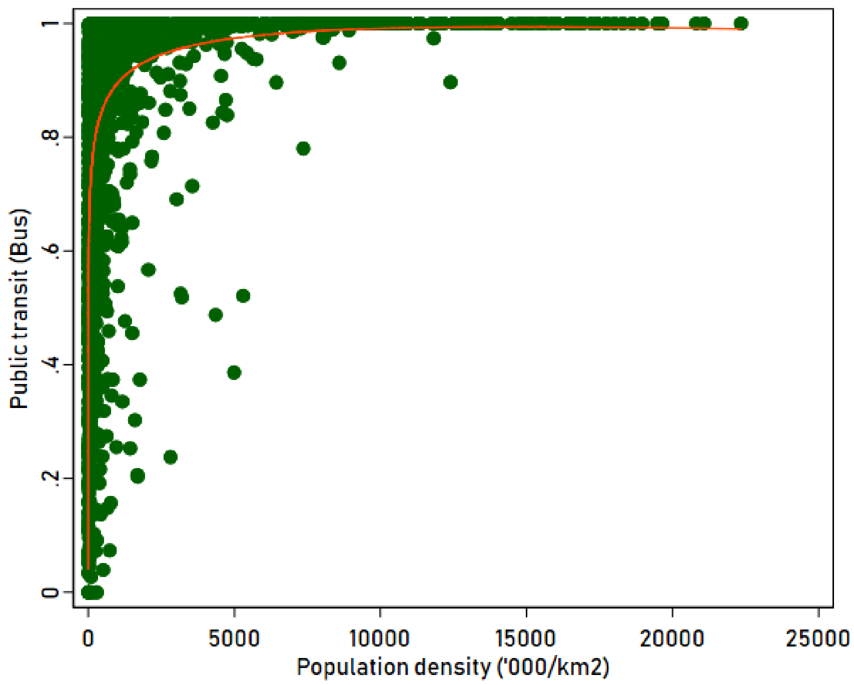
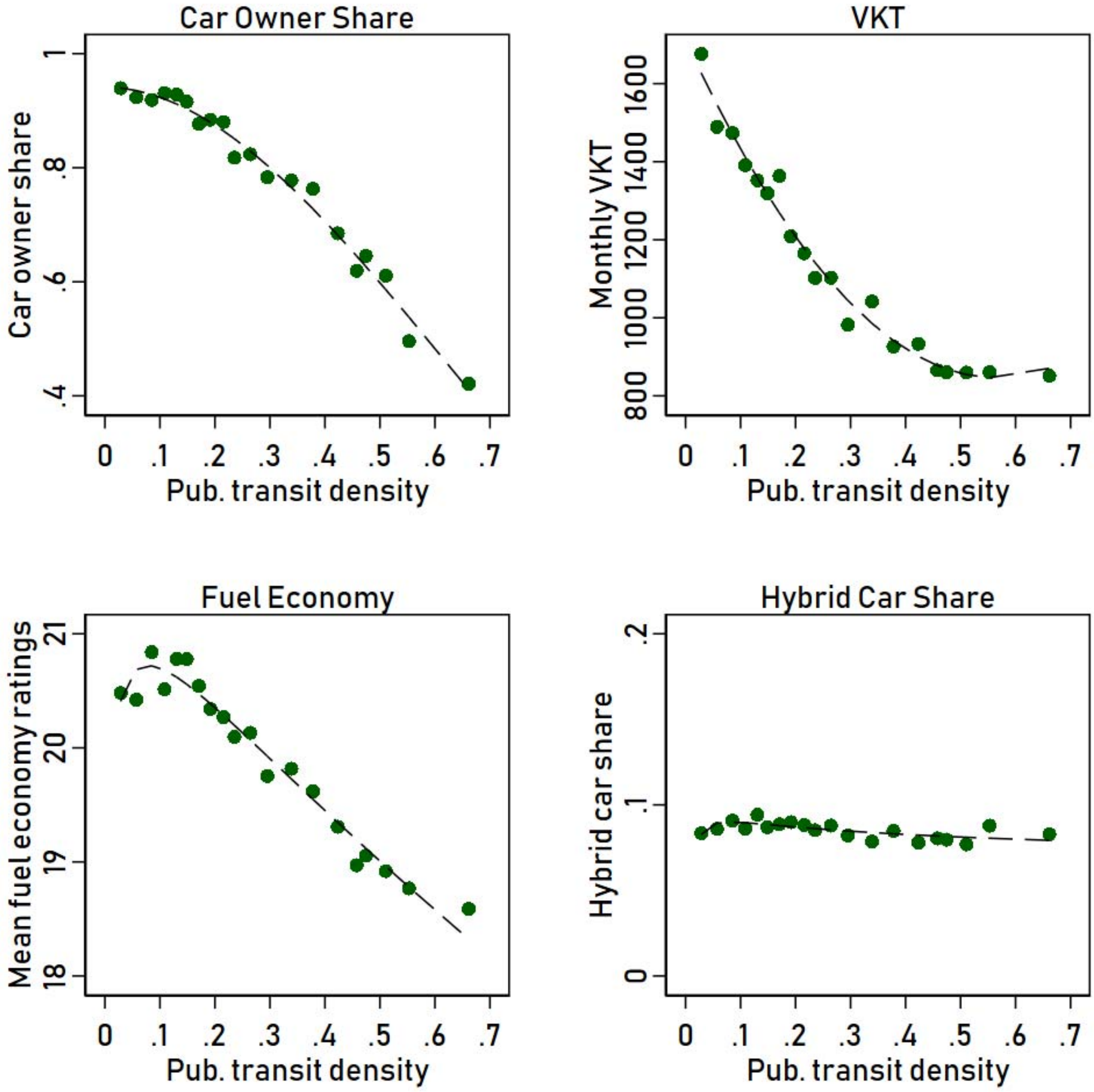
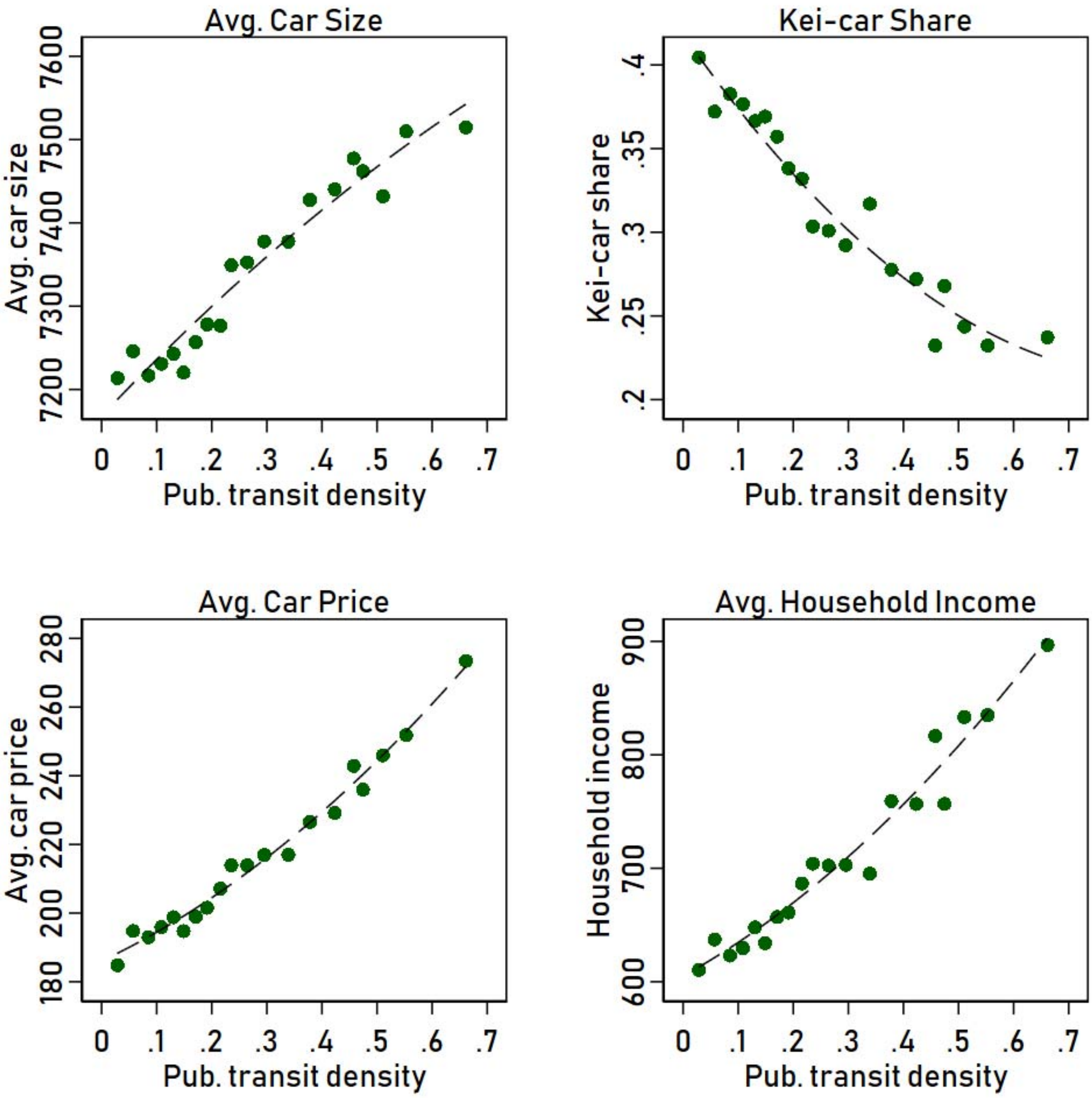


Figure 2. Vehicle-related Consumer Choice by Public Transit Density



Note: Car owner share is the percent of households who own at least one car; Vehicle kilometer traveled (VKT) is the sum over two most frequently used cars owned by each household; Fuel economy ratings and hybrid vehicle shares are the (unweighted) averages over the two cars. All figures report the averages of these over households in each public transit density bin.

Figure 3. Vehicle/Household Characteristics by Public Transit Density



Note: Car size is calculated as the sum of length, height, and width in millimeter. Car price and household income are in millions of Japanese yen. We use the (unweighted) average over the two most frequently used cars for car size, the share of keicars, and car price. All figures report the averages of these over households in each public transit density bin.

Figure 4. Portfolio Effect:
Ownership Rates of Alternative Vehicle Combinations

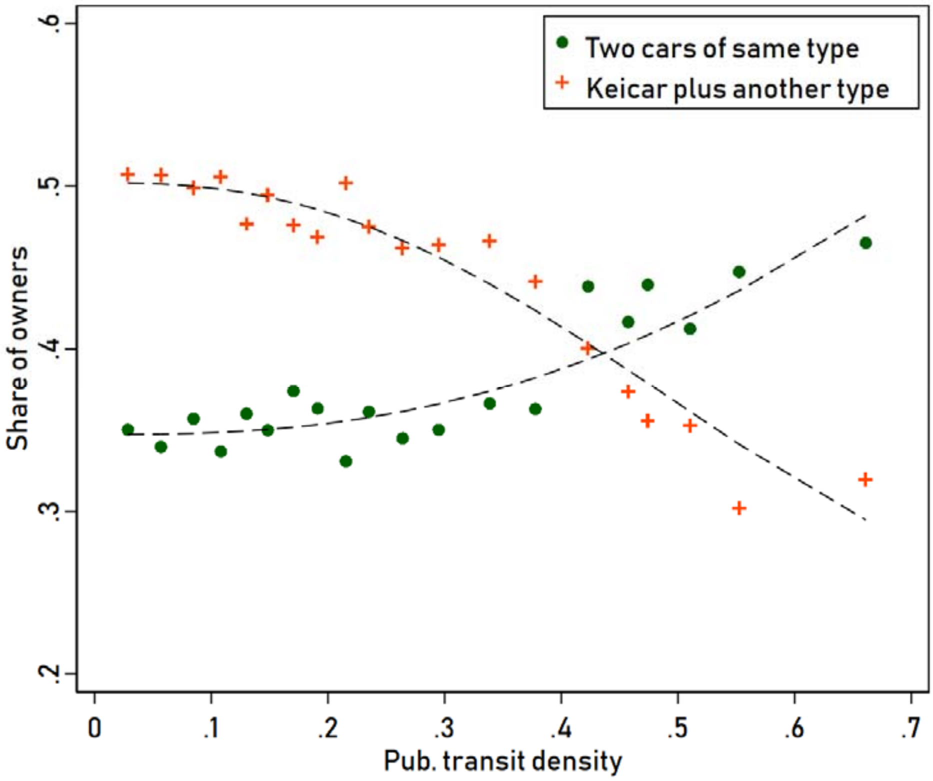
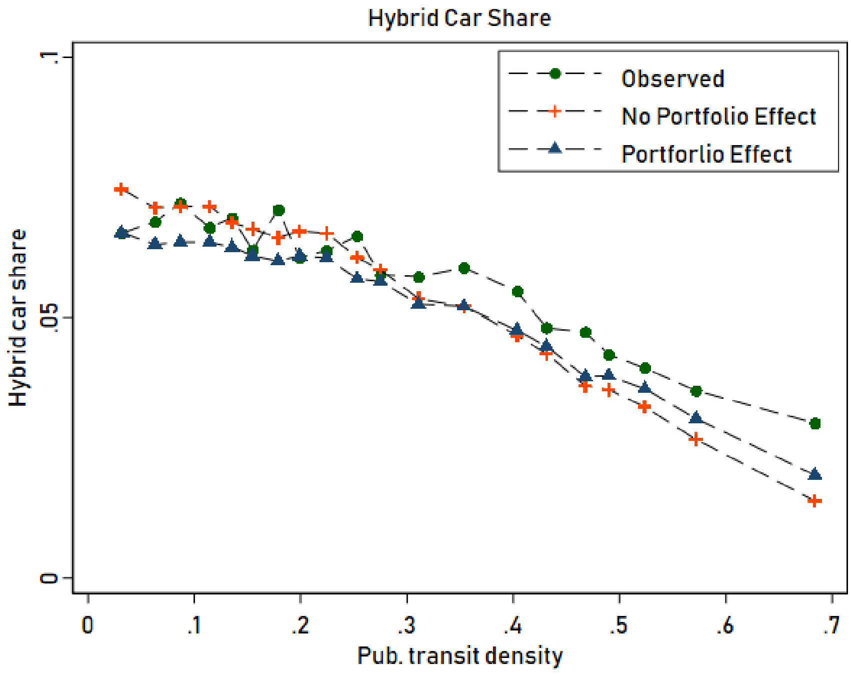
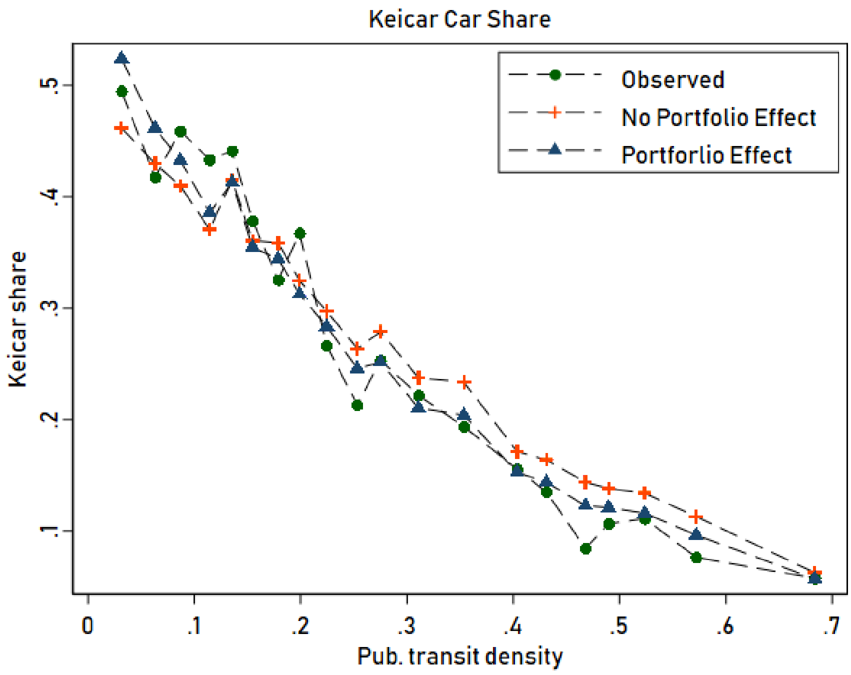


Figure 5. Prediction Performance:
Portfolio Effect versus No Portfolio Effect

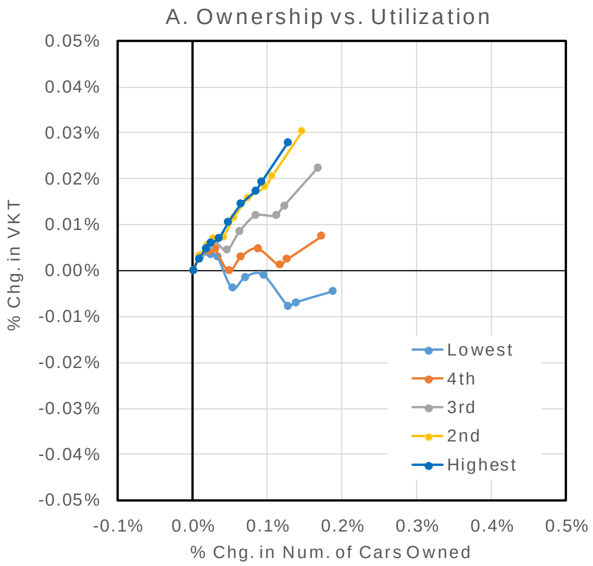


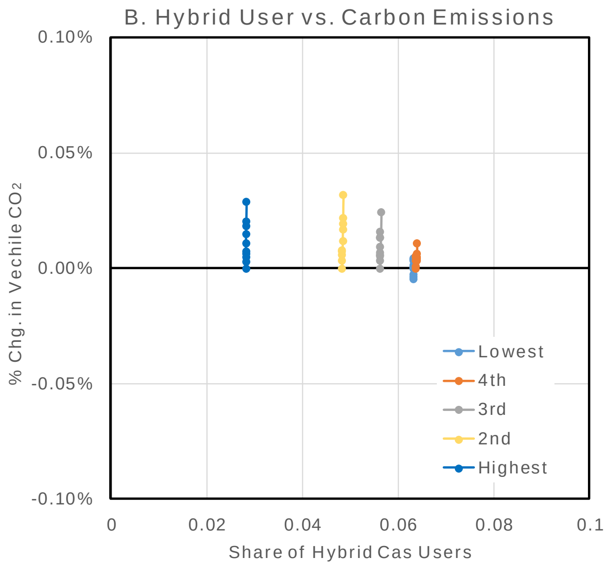


Note: Predicted shares in this figure are calculated as *unconditional* shares: i.e., not conditional on car holding unlike in Figure 2 and 3.

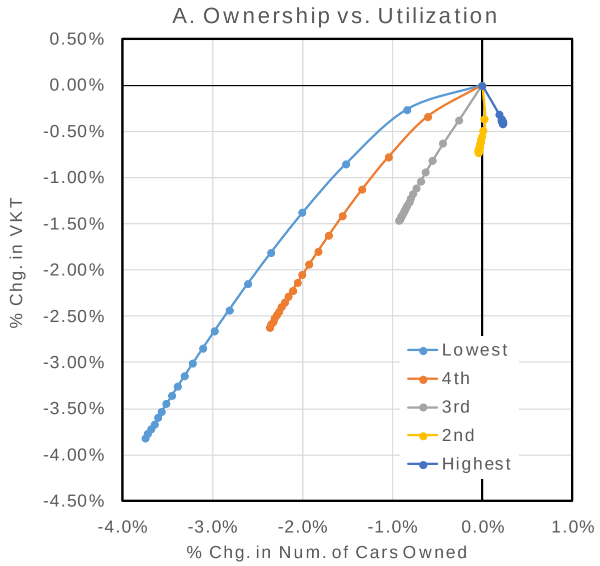
Figure 6. Impacts of Perturbations to Counterfactuals

(1). Eco-car Incentive Policy





(2). Eco-car Sharing Policy



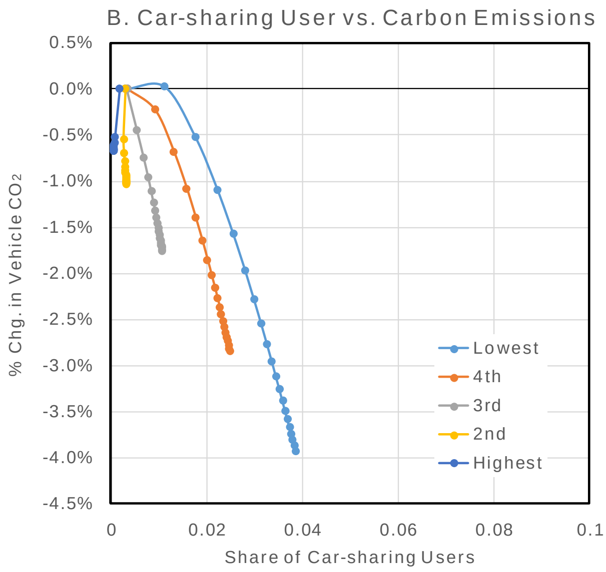


Table 1. Descriptive Statistics by Population Density Quintiles

	Population Density Quintiles									
	Lowest		4th		3rd		2nd		Highest	
Number of obs. (households)	20,963		20,851		21,041		21,078		21,033	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Public transit density	0.112	0.064	0.178	0.080	0.277	0.125	0.392	0.110	0.559	0.119
Congestion measure	0.595	0.108	0.689	0.105	0.753	0.136	0.813	0.103	0.899	0.062
Household income (10,000 yen)	601.6	410.1	634.6	427.1	656.1	442.5	684.1	460.7	746.1	530.1
Household size [Num. of people in household]	2.94	1.39	2.88	1.32	2.81	1.27	2.70	1.24	2.52	1.25
Own car? [1 = yes]	0.938	0.241	0.908	0.289	0.829	0.376	0.730	0.444	0.497	0.500
Drive to work? [1 = yes]	0.548	0.498	0.434	0.496	0.266	0.442	0.154	0.361	0.070	0.256
Num. of cars owned [of those who own car]	1.73	0.90	1.56	0.78	1.32	0.61	1.17	0.44	1.09	0.36
Most frequently used car										
Monthly VKT	1,162.7	2,378.0	1,000.6	1,916.8	892.3	1,884.1	785.6	1,837.6	730.3	1,785.4
Price (10,000 yen)	199.6	129.3	206.5	135.4	218.9	140.0	229.7	150.7	261.7	182.1
Fuel economy (km/L)	20.4	7.6	20.5	7.7	19.8	7.6	19.3	7.7	18.5	7.7
Vehicle size (mm) [length + width + height]	7,276.2	819.0	7,297.1	801.1	7,394.9	783.4	7,457.4	793.8	7,537.6	844.1
Hybrid [1 = yes]	0.135	0.342	0.146	0.353	0.146	0.353	0.150	0.357	0.150	0.357
Keicar [1 = yes]	0.349	0.477	0.327	0.469	0.289	0.454	0.254	0.436	0.221	0.415
Second most used car										
Monthly VKT	964.4	1,920.3	940.7	2,231.6	937.0	2,318.8	941.9	2,856.1	1,065.5	3,399.3
Price (10,000 yen)	170.8	116.4	180.8	128.6	194.7	151.9	215.9	184.1	283.5	268.5
Fuel economy (km/L)	21.1	7.0	21.0	7.1	21.0	7.3	20.2	7.6	19.0	7.3
Vehicle size (mm) [length + width + height]	7,037.7	760.9	7,058.4	777.2	7,068.4	838.0	7,192.5	843.2	7,234.9	1,050.5
Hybrid [1 = yes]	0.039	0.195	0.035	0.183	0.023	0.150	0.015	0.122	0.009	0.096
Keicar [1 = yes]	0.531	0.499	0.501	0.500	0.480	0.500	0.420	0.494	0.329	0.470

Table 2. Demand for Vehicle Ownership

	Without Portfolio Effects			With Portfolio Effects		
		Interaction Terms			Interaction Terms	
	Mean Parameter	Transit Density	Household Size	Mean Parameter	Transit Density	Household Size
ln(y - r)	1.7986 ***			1.9212 ***		
	(0.0390)			(0.0433)		
YPK	-0.1092 ***	0.0853 ***	0.0426 ***	-0.1194 ***	0.1566 ***	0.0492 ***
	(0.0063)	(0.0116)	(0.0014)	(0.0072)	(0.0132)	(0.0017)
HP/W	13.1612 ***	25.5427 ***	-8.3499 ***	13.9062 ***	4.0971 *	-8.3720 ***
	(1.0733)	(1.9223)	(0.2436)	(1.2087)	(2.1639)	(0.3037)
Size	-0.3259 ***	-0.6994 ***	0.0686 ***	-0.2544 ***	-0.0835 ***	0.0286 ***
	(0.0096)	(0.0166)	(0.0020)	(0.0112)	(0.0206)	(0.0027)
Portfolio Effects						
Kei-Kei				0.8942 ***	9.6611 ***	-0.5368 ***
				(0.0680)	(0.1602)	(0.0181)
Kei-Regular				-0.8361 ***	0.8479 ***	0.0259
				(0.0811)	(0.2151)	(0.0189)
Kei-Minivan				-0.1767 ***	0.0766	0.0348 **
				(0.0644)	(0.1586)	(0.0158)
Regular-Minivan				-0.0818	0.3305	0.1311 ***
				(0.0962)	(0.2142)	(0.0242)
Minivan-Minivan				-0.0648	-0.0361	0.0720 ***
				(0.0788)	(0.1787)	(0.0181)
Used-car dummy x Metropolitan dummies	Yes			Yes		
Maker dummies x Metropolitan dummies	Yes			Yes		
Fuel-type dummies x Metropolitan dummies	Yes			Yes		
Vehicle-type dummies x Metropolitan dummies	Yes			Yes		
# of Obs.	15,684,191			15,684,191		
# of Cases	95,604			95,604		

Table 3. Vehicle Utilization Equation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln(y - r)	0.0171	0.0166	0.0229	0.0465 **	0.0595 *	0.0824 **	0.0553 *	0.0892 ***
	(0.0143)	(0.0185)	(0.0156)	(0.0195)	(0.0320)	(0.0345)	(0.0326)	(0.0348)
× Transit Density	-0.0786 **	-0.0900 **	-0.0826 **	-0.1056 ***	-0.0739 *	-0.0813 *	-0.0777 *	-0.0931 **
	(0.0346)	(0.0369)	(0.0376)	(0.0405)	(0.0409)	(0.0438)	(0.0415)	(0.0443)
× Congestion					-0.0557	-0.0764 †	-0.0457	-0.0621 †
					(0.0456)	(0.0472)	(0.0459)	(0.0454)
ln(YPK)	-0.3217 ***	-0.3243 ***	-0.3224 ***	-0.3192 ***	-0.3170 ***	-0.3463 ***	-0.3201 ***	-0.3496 ***
	(0.0653)	(0.0699)	(0.0655)	(0.0658)	(0.1118)	(0.1063)	(0.1114)	(0.1044)
× Transit Density	0.2033 *	0.1957 *	0.2024 *	0.2276 **	0.1891 †	0.1717 †	0.1875 †	0.1866 †
	(0.1099)	(0.1091)	(0.1134)	(0.1164)	(0.1301)	(0.1291)	(0.1287)	(0.1337)
× Congestion					-0.0005	0.0502	0.0032	0.0591
					(0.1430)	(0.1468)	(0.1446)	(0.1378)
HP/W	1.9199 *	1.9607 *	1.9416 †	1.7906 †	1.9778 †	1.9050 *	1.9653 †	1.7933 †
	(1.1630)	(1.1847)	(1.2973)	(1.3205)	(1.2563)	(1.1586)	(1.2375)	(1.2440)
Size	0.1496 ***	0.1524 ***	0.1498 ***	0.1536 ***	0.1496 ***	0.1533 ***	0.1498 ***	0.1538 ***
	(0.0340)	(0.0366)	(0.0358)	(0.0380)	(0.0360)	(0.0344)	(0.0356)	(0.0362)
Second Car	-0.2443 ***	-0.2372 ***	-0.2442 ***	-0.2367 ***	-0.2441 ***	-0.2365 ***	-0.2443 ***	-0.2366 ***
	(0.0284)	(0.0286)	(0.0286)	(0.0291)	(0.0290)	(0.0302)	(0.0294)	(0.0314)
Demographic controls	✓	✓	✓	✓	✓	✓	✓	✓
Urban structure controls	✓	✓	✓	✓	✓	✓	✓	✓
Metropolitan dummies			✓	✓			✓	✓
Selection controls		✓		✓		✓		✓
Num. of obs.	25,953	25,953	25,953	25,953	25,953	25,953	25,953	25,953
R^2	0.075	0.076	0.077	0.078	0.076	0.077	0.077	0.078
F-stat. on selection terms	--	18.93 **	--	30.35 ***	--	20.08 **	--	23.85 ***

Table 4. Elasticity Estimates

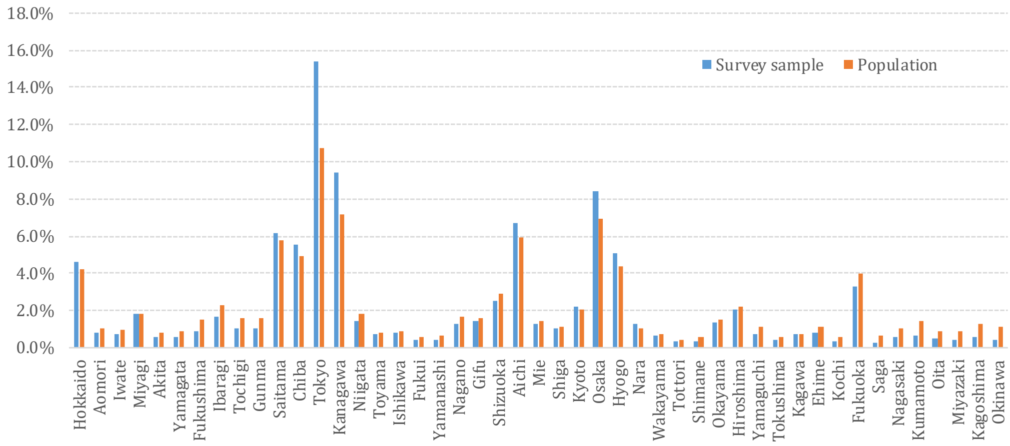
	Public transit quintiles				
	5th (Lowest)	4th	3rd	2nd	1st (Highest)
Car ownership elasticity w.r.t. rental price					
Hybrid cars	-1.281	-1.063	-0.805	-0.564	-0.346
Diesel cars	-0.953	-0.871	-0.700	-0.572	-0.379
Kei-cars	-1.021	-0.874	-0.608	-0.386	-0.223
VKT elasticity w.r.t. net income	0.0429 **	0.0314 **	0.0165	0.0016	-0.0196
	(0.0174)	(0.0161)	(0.0165)	(0.0165)	(0.0199)
VKT elasticity w.r.t. price of driving	-0.2978 ***	-0.2781 ***	-0.2544 ***	-0.2255 ***	-0.1882 **
	(0.0639)	(0.0636)	(0.0596)	(0.0651)	(0.0740)

Table 5. Impacts of Counterfactual Policies

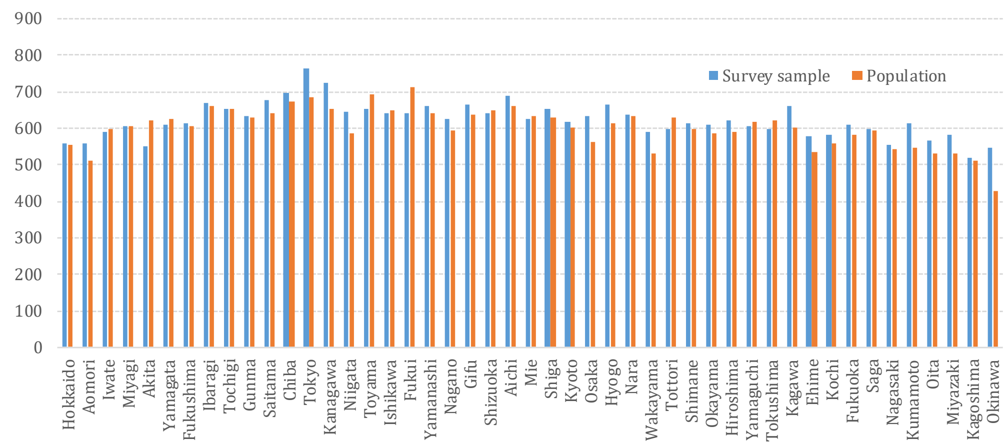
	Population density quintiles (as of 2015)					
	5th (Lowest)	4th	3rd	2nd	1st (Highest)	All
Observed						
Pop. density (1000/km2)	0.273	1.082	3.409	7.414	14.412	5.318
Pub. transit density	0.112	0.178	0.277	0.392	0.559	0.303
Num. of cars owned	1.250	1.130	0.945	0.796	0.491	0.922
Car/ownership type (%)						
Hybrid	0.063	0.064	0.056	0.048	0.028	0.052
Diesel	0.010	0.011	0.011	0.010	0.007	0.010
Keicar	0.454	0.346	0.242	0.164	0.084	0.258
Car-sharing	0.000	0.000	0.000	0.000	0.000	0.000
Monthly VKT (km)	561.0	485.8	404.7	332.7	186.1	394.069
CO_2 from driving (kg/mo.)	76.3	66.9	56.7	47.7	27.5	55.006
Counterfactual I: Feebates (x = 50th percentile)						
Num. of cars owned	1.252	1.133	0.954	0.796	0.496	0.926
Car/ownership type (%)						
Hybrid	0.063	0.064	0.056	0.048	0.028	0.052
Diesel	0.010	0.011	0.011	0.011	0.007	0.010
Keicar	0.458	0.350	0.247	0.163	0.085	0.261
Car-sharing	0.000	0.000	0.000	0.000	0.000	0.000
Monthly VKT (km)	556.9	486.7	408.7	334.3	187.7	394.879
CO_2 from driving (kg/mo.)	75.7	67.0	57.2	47.9	27.7	55.084
Counterfactual II: Car-sharing of hybrid cars (n = 10)						
Num. of cars owned	1.209	1.106	0.943	0.793	0.496	0.909
Car/ownership type (%)						
Hybrid	0.061	0.062	0.056	0.048	0.028	0.051
Diesel	0.008	0.009	0.009	0.009	0.006	0.008
Keicar	0.442	0.342	0.245	0.163	0.085	0.255
Car-sharing	0.034	0.022	0.010	0.003	0.001	0.014
Monthly VKT (km)	541.4	477.4	404.4	332.6	187.1	388.584
CO_2 from driving (kg/mo.)	73.7	65.7	56.6	47.6	27.5	54.219

Appendix. Sample versus Population Distribution by Prefecture

(a) Number of respondents/population



(b) Household income



(c) Household size