The Role of Consumer Beliefs in Consumer Undervaluation of Automotive Fuel Cost Savings

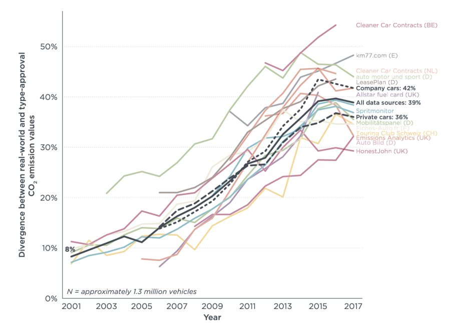
Rin Futara, Takahiro Hoshino

January 30, 2026

University of Maryland at College Park, Keio University

Consumers make choices based on their beliefs

- Researchers rely on official ratings, but
- Official ratings can differ from consumer beliefs
    - Knowledge/experience: Official rating (based on test cycle) shows more divergence from the real world (Tietge et al., 2017)



    - Driving behaviors: Assumptions in test-cycle regarding driver behaviors (Karabasoglu and Michalek, 2013; Greene et al., 2017)
- **Our paper**: Show the data pattern of the future fuel cost belief

Introduction: Implications

Economics and the discrete choice model:

Consumer Myopia: Consumers' decisions on the tradeoff between willingness to pay (WTP) for vehicles and future fuel cost-saving:

- Do consumers underestimate or correctly estimate future fuel costs?
- Consumer's belief affects its valuation parameter
- **Our paper**: Reveal the econometrics mechanisms of ignoring beliefs

Policy:

- Fuel economy standard: The presumption of undervaluation is rational for tightening standards (National Research Council, 2015; National Highway Traffic Safety Administration, 2018)
- Test-cycle: Authority (e.g., EPA) justifies an official fuel consumption as information for consumers' choice and a policy basis for standards
- **Our paper**: Estimate valuation parameter & Discuss the policy designs

Concept: Consumer Myopia

Suppose the trade-off between WTP for a vehicle and future fuel cost saving when choosing a car:

$$\Delta p = \gamma \cdot \Delta G$$

where p: WTP, G: future fuel cost, γ: **valuation parameter**.

1. $\gamma = 1$: Consumers correctly evaluate future fuel cost
2. $\gamma < 1$: Consumers undervaluate future fuel cost
    - i.e. Consumers are myopic

Literature

Valuation parameter (γ) estimates: Undervaluate future costs

- Vehicle (Allcott and Wozny, 2014; Grigolon et al., 2018; Leard et al., 2023; Gillingham et al., 2021)
- Appliance and housing (Myers, 2019; Houde and Myers, 2021, 2025)

Assumptions in the literature

1. No belief heterogeneity in future fuel cost savings (uniform within vehicles/times)
   - Exceptions: Allcott (2013) and Allcott and Knittel (2019)
2. No systematic relations between choice and consumer type (e.g., more and less intensive drivers)

Data

Single source of vehicle purchase and usage microdata ⇒ Capture relations among vehicle choices and consumer belief before purchase

	Micro Purchase Data	Surveyed Data
Period	Jan, 2016 $\sim$ Mar, 2020	—
Sample	48,549	Approximate 10,000
Provider	Intage, Inc.	Intage, Inc.
Main Variables	Fuel consumption belief (post), Vehicle transaction price, Average monthly VKT, Last vehicle's info, Contract, Demographics, etc.	Fuel consumption belief (pre), Expected gasoline price, Expected ownership duration, Maintenance cost, Insurance cost, etc.
Timing of Survey	One-month after contract	One-shot (Nov, 2022)

- Merge the micro-purchase data (Nielsen type data) and the survey data by member ID, enabling tracking of the same consumer

▸ Questions: Fuel Consumption ▸ Questions: Fuel Price ▸ Survey Timing: Car Delivery

Specification: Future Fuel Costs

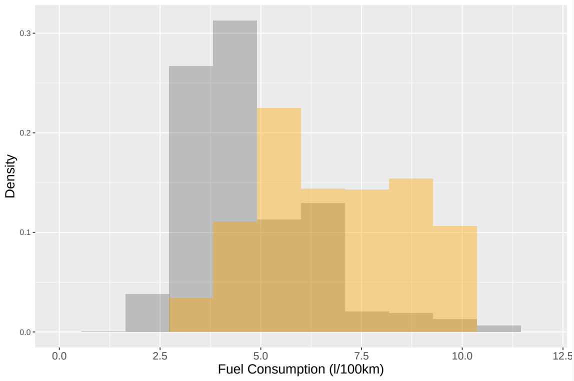
Future fuel costs = Sum of discounted present fuel costs comprised of:

1. Fuel consumption belief of car (l/km; the inverse of fuel efficiency)
2. Fuel price belief that consumer faces
3. Expected vehicle ownership duration
4. Individual discount rate
5. Expected annual vehicle kilometers traveled (VKT)

Data I: Fuel Consumption Belief

- Average fuel consumption gap (Official vs. Belief) across fuel consumption measures: 28.7% (30.5% in Tanaka (2020))
- Almost no difference in belief before and after purchasing vehicles ▸ Appendix

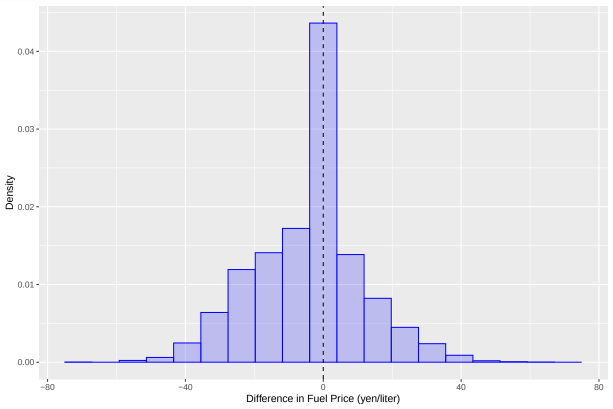
Figure 1: Density of Fuel Consumption Measures



Data II: Fuel Price Belief

- Difference between fuel price belief and fuel price when a consumer purchased a vehicle is almost normally distributed (consistent with Anderson et al. (2013))
  ▸ Over-time
- Less concerns with a survey timing effect ▸ Appendix

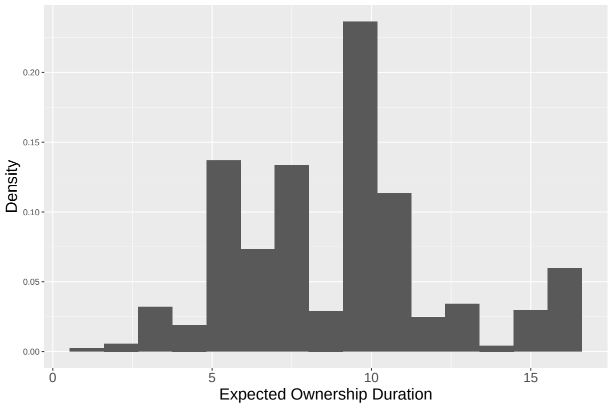
Figure 2: Difference between belief and fuel price at purchase



Data III: Ownership Duration Belief and Resale Values

- Average expected vehicle ownership duration: 9 years
- We account for resale values ▸ Resale Values

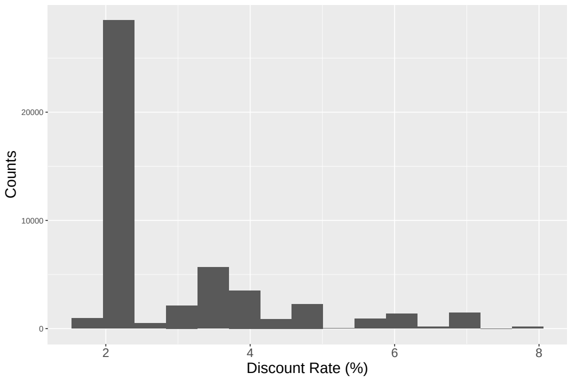
Figure 3: Histogram of Ownership Duration Belief



Data IV: Discount Rates

- Discount rates are defined by opportunity costs (Allcott and Wozny, 2014)
    - Pay by cash = discount rate in an average opportunity cost 2% (in Japan)
    - Dealer finance and Other finance (e.g., Bank/Credit union loan): Varying across firms, regions, time, and vehicles ▸ Data

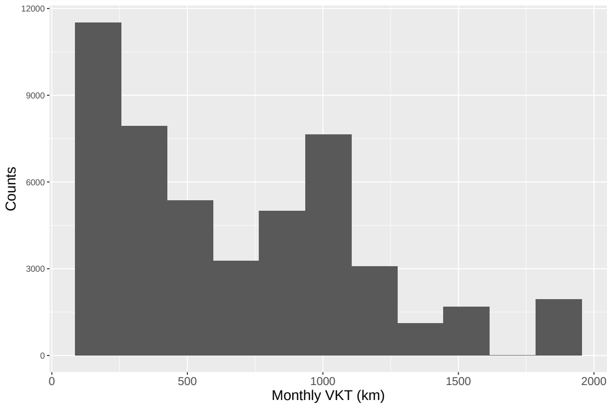
Figure 4: Histogram of Discount Rates



Data V: VKT

- Positive correlation with income ▸ Figure

Figure 5: Histogram of VKT

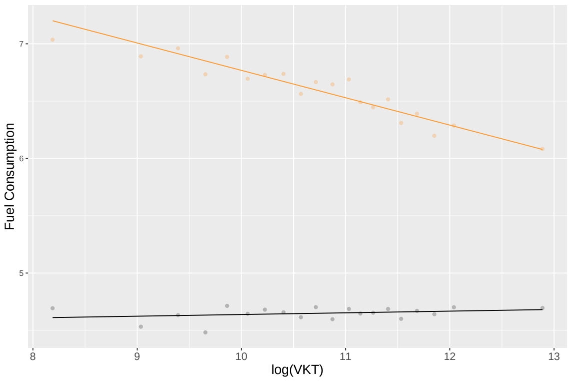


Data: Consumer Type

Correlation (A):

- Negative correlation between fuel consumption belief and VKT
    - No correlation between official fuel consumption and VKT

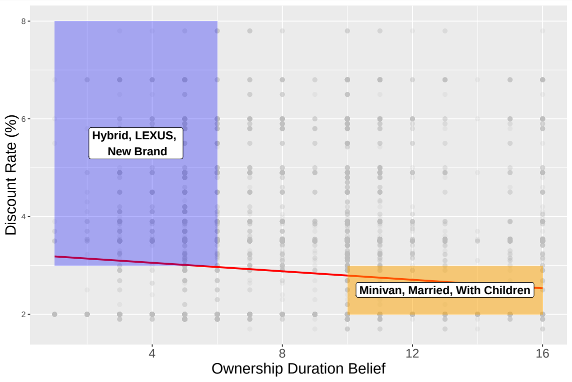
Figure 6: Fuel Consumption and log(VKT)



Correlation (B):

- Correlation between discount rates and ownership duration belief (negative)
  - Vehicle selection: Car Enthusiast vs. Parenting generation

Figure 7: Vehicle Selection



Theoretical Insights

Two types of distortions:

1. Heterogeneities in cost components
2. Consumer-type characterized by consumer-types

$\Rightarrow$ Both affect variations in future fuel costs

Intuition:

- Within-individual: Large variation in fuel cost heterogeneities within-individual leads to a small valuation parameter (i.e., attenuation)
    - Consumer-type correlation: Valuation parameter depends on whether the correlation exacerbates the heterogeneity or not
- Cross-sectional: Not only heterogeneity but also the correlation across fuel costs, choice probabilities, and vehicle prices affect valuation parameter estimates

Details

Estimation: Method

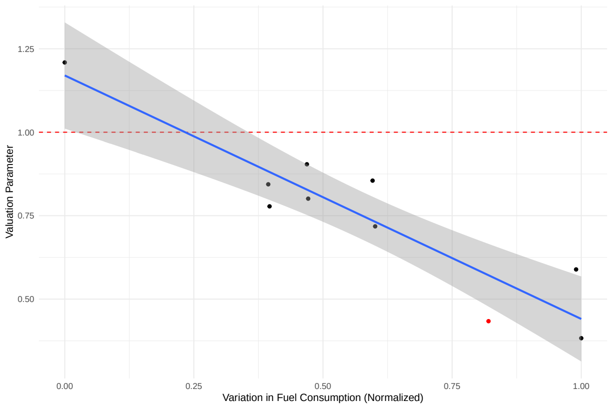
- Apply a discrete model ▸ Technical Details
  1. Within-individual variation: Limited implication
  2. Cross-sectional variation [On-going]
- Issues: Did not observe fuel consumption beliefs for all non-purchased vehicles
  - Prediction: Using machine learning techniques ▸ Appendix

$\Rightarrow$ Potential systematic bias in fuel consumption between purchased/non-purchased vehicles

  - Consumers may misunderstand fuel consumption of non-purchased vehicles
  - **Solution**: Additional survey on January 15-19 ▸ Details

Results: Within-individual Variation

Figure 8: Fuel Consumption Variations and Valuation Parameters



- Large heterogeneity in (residual) variations of fuel consumption $\Rightarrow$ Small valuation parameter

▸ Table

Counterfactual Plans

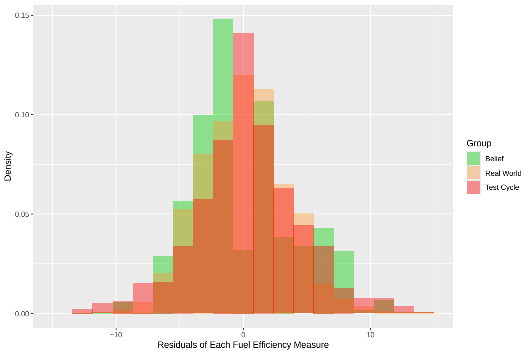
Two plans:

1. Biases in welfare from ignoring consumers' beliefs
2. Value of test-cycle variations ▸ Test-cycle

Value of test-cycle variations:

- Focus only on belief heterogeneity in fuel consumption (i.e., abstracting from consumer-type distortions)
- Compare welfare differences in the discrete choice model by using estimated parameters and replacing the fuel consumption variations
  - (a) official figures (no within-trim variation)
  - (b) real-world variation
  - (c) test-cycle variation
- Given the distribution of real-world fuel consumption (as a ground truth), we can discuss the welfare implications of variations.

Counterfactual Plans: Test-Cycle Variation



Note: The histogram plots the residuals after controlling for constant and brand-fixed effects.

- Pooled data: 7 Vehicle Brands (from `FASTSim` and `D3`)
- Residual variations: Belief (16.64), Test-Cycle (28.39), Real World (11.94)
- Sample size: Belief (1, 458), Test-Cycle (1, 372), Real World (69, 591)

Summary and Plan

Findings:

- New data about consumer beliefs in future fuel costs
- Theoretical insights of econometric bias without beliefs
  - Belief heterogeneity in fuel cost components
  - Consumer type: Systematic correlations across belief components in vehicle choice
- Empirical support for the importance of belief heterogeneity variation
  - Within-individual: Large heterogeneity in variations of fuel consumption leads to a small valuation parameter

Plans:

- Receive the results of an additional survey (February) and check the data
- Estimation using a cross-sectional variation
- Theoretical expectation using a cross-sectional variation
- Counterfactual simulations

Appendix

Appendix - Data Construction: Average On-road Fuel Efficiency

Collects on-road fuel consumption data through a mobile phone application (Tietge et al., 2017; Tanaka, 2020).

- Website (e-nenpi): `https://e-nenpi.com/`
- Membership: 650,000, Number of posted on-road fuel efficiency: 13 million

- Two ways to report on-road fuel efficiency for car drivers:

1. Report odometer values and the amount of refueled gasoline
2. Upload the receipt for refueling at gas stations

$\Rightarrow$ Cannot report directly fuel efficiency: distinct feature from `fueleconomy.gov` (Greene et al., 2017)

- The median of reported numbers of on-road fuel efficiency for each vehicle is 658.

Return

Appendix - Data Description

Table 1: Data Description

	Micro			Macro (JP Population)		
Variable	Mean	St. Dev.	Median	Mean	St. Dev.	Median
Inside Share (%)	0.127	0.126	0.085	0.049	0.071	0.015
Store Price (hundred thousand yen)	310.731	125.458	295.015	328.231	285.299	228.492
Official Fuel Efficiency (km/liter)	23.234	6.958	23.727	20.321	7.371	20.400
On-road Fuel Efficiency (km/liter)	14.986	3.887	14.983	—	—	—
Fuel Efficiency Belief (km/liter)	15.481	4.916	15	—	—	—
Annual VKT (km)	7,338.609	5,407.091	5,400	6,398.126	326.041	6,361
Life-Time of Vehicle (years)	8.635	4.272	7	7.1	—	—
Discount Factor	3.436	1.852	2	—	—	—
Horsepower (ps)	82.685	37.465	73	98.416	65.168	80
Vehicle Weight (kg)	1,314.241	338.906	1,350	1,323.777	415.525	1,280
HEV Dummy	0.447	0.496	0	0.331	0.469	0

Notes: Macro data is obtained by merging share data and automobile attribute data except ownership duration. The macro data on the ownership duration of vehicles is from the survey of the passenger car market by JAMA (Japan Automobile Manufacturers Association). The number of consumers does not weight inside share in micro data but a pooled average. On-road fuel efficiency is the average on-road fuel efficiency of each vehicle. Fuel efficiency belief comes from microdata.

Appendix - Data Description

Table 2: Data Description (Demographic)

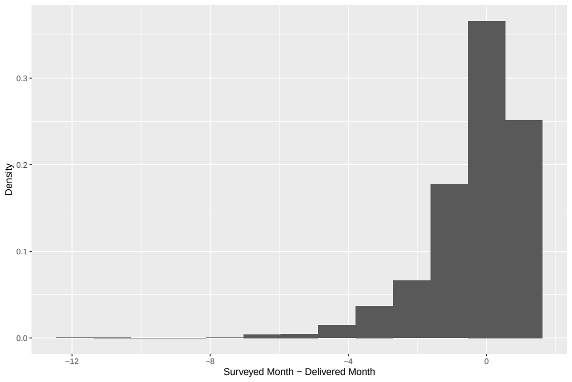
Variable	Micro Mean	Micro St. Dev.	Micro Median	Macro (JP Population) Mean
Gender (Male, %)	69.9%	0.459	1	52%
Age	49.311	11.663	49	51.6
Martial Status (Married %)	78.7%	0.410	1	86%
With Children (%)	54.1%	0.498	1	—
Family Size	3.100	1.333	3	—
Household Income Q1	0.053	0.225	0	—
Household Income Q2	0.174	0.379	0	—
Household Income Q3	0.324	0.468	0	—
Household Income Q4	0.449	0.497	0	—

Notes: Household income class: Q1 (under 299 million yen), Q2 (300 million yen to 499 million yen), Q3 (500 million yen to 799 million yen), Q4 (over 800 million yen). Macro data is retrieved from "Market Trend Survey: Passenger Vehicles" by JAMA (Japan Automobile Manufacturer Association). Notice that macro data is based on not purchase data but main driver data.

Appendix - Timings of Survey and Delivery

- While the survey took place one month after the contract, the timing of the car delivery is not the same across consumers
- 73% consumer took a survey before delivered or in the same month

Figure 9: Histogram of Difference between Surveyed and Delivered Months



Return

Appendix - Survey: Fuel Efficiency Belief

- How did we ask?

- Fuel Efficiency Belief (before purchase):
  - Please choose the real-world fuel efficiency that you had expected at the purchase. (* Please choose the closest alternatives.)
  - Note: First, we check whether consumers see official fuel efficiencies
- Fuel Efficiency Belief (after purchase):
  - Please select your evaluation of the real-world fuel efficiency (per 1 litre) of your vehicle after you purchased the vehicle.

- Alternatives for both before/after purchase:

- Below 10km/l, 10-12 km/l, ..., 32-34 km/l, over 34 km/l, Do not know

For consistency, we asked several questions about fuel efficiency beliefs:

- Whether respondents consider official and real-world fuel efficiency to be the same
- The percentage of differences between official and evaluation of fuel efficiency: $-70\% \sim 30\%$

Return

Appendix - Survey: Fuel Efficiency Belief

We asked fuel consumption beliefs of 10 non-purchased vehicles to all individuals:

- Please choose the real-world fuel efficiency that you expected (* Please choose the closest alternatives.)

Alternatives are the same as purchased vehicles:

- Below 10km/l, 10-12 km/l, ..., 32-34 km/l, over 34 km/l, Do not know

Return

Appendix - Fuel Consumption Difference between Before and After

Figure 10: Density of Belief Before/After Purchasing

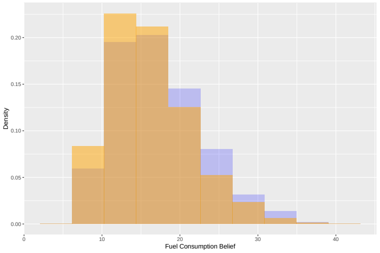
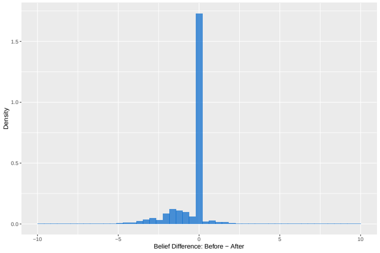


Figure 11: Histogram of Difference between Belief Before—After



- No large difference between belief before and after purchasing vehicles
  - Observable covariates (vehicle, demographics, city) do not explain the difference

Appendix

Appendix - Fuel Consumption Difference between Before and After

The difference in beliefs between before and after purchase:

Figure 12: Difference: Rate

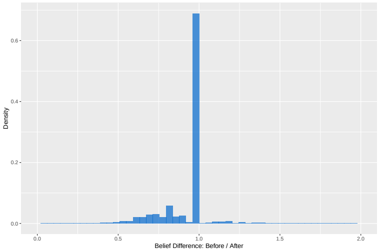
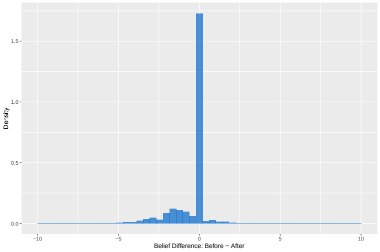


Figure 13: Difference: Value



Return

Appendix - Fuel Consumption Difference between Before and After

- Observable variables are not correlated with the difference in fuel consumption between before and after purchase

Table 3: Correlations with difference in fuel consumption before and after purchase

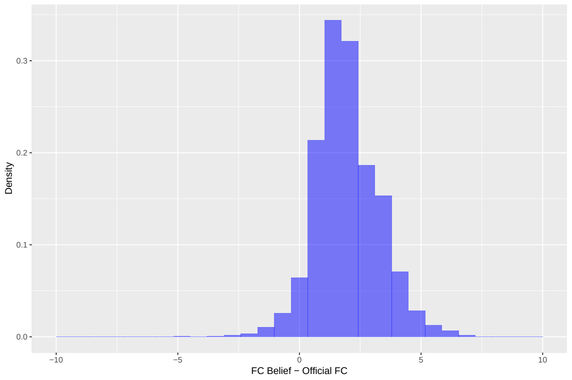
Regression Type	R-Squared: Rate	R-Squared: Value
All Variables	0.039	0.040
Vehicle Attribute (with Brand FE)	0.019	0.019
Vehicle Attribute (without Brand FE)	0.013	0.012
Demographics	0.003	0.002
City (with Prefecture FE)	0.007	0.008
City (without Prefecture FE)	0.002	0.004

Return

Appendix - Fuel Consumption Gap (Histogram)

- Normally distributed

Figure 14: Individual-Level Difference between Belief and Official



Appendix - Fuel Consumption Gap (Scatter-plot)

Figure 15: On-Road FC / Official FC

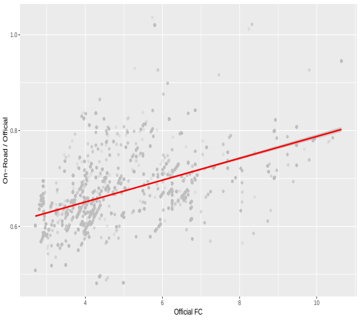
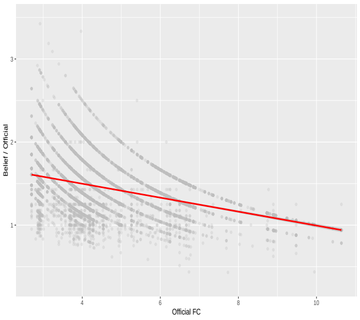


Figure 16: FC Belief / Official FC



▸ Return

Appendix - Fuel Consumption Gap: Summary

Difference between belief and official ratings:

- Automaker: Small gap (Toyota), Large gap (Mazda and Subaru)
- Body-type: Large gap (SUV, Sedan, and Wagon)
- Fuel-type: HEVs are believed to have a smaller gap than ICEs
- Demographics: Elderly and those who have children believe more fuel-consuming
- City information: The livings in the areas with small areas, low employment, and educated believe more fuel-consuming

$\Rightarrow$ However, these differences are very small

Appendix - Fuel Consumption Gap (Test-Drive)

Table 4: Difference: Test Drive

	Fuel Consumption Gap (Rate)			
	(1)	(2)	(3)	(4)
Test Drive	0.004		0.005	0.002
	(0.006)		(0.006)	(0.006)
Check Catalog Fuel Efficiency		−0.009	−0.010	−0.006
		(0.007)	(0.007)	(0.007)
Vehicle Brand FE	Y	Y	Y	Y
Month FE	Y	Y	Y	Y
Prefecture FE	N	N	N	Y
Num. obs.	9523	9523	9523	9523

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Appendix - Survey: Fuel Price Belief

- How did we ask?

- What was the average fuel price you had expected when you purchased the vehicle?

For consistency, we asked several questions about a fuel price:

- Whether the average fuel price was as/more/less than expected
- Actual fuel price before the Covid
- Expected annual fuel costs

Note:

- We do not assume that the consumer has calculated future fuel costs based on very frequent predictions (e.g., per-month fuel price prediction). That is, consumers have only one specific fuel price belief.

Return

Appendix - Data: Fuel Price Belief (Histogram)

Figure 17: Histogram of Fuel Price Measures

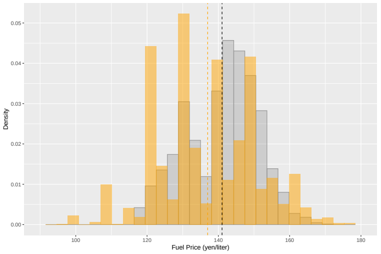
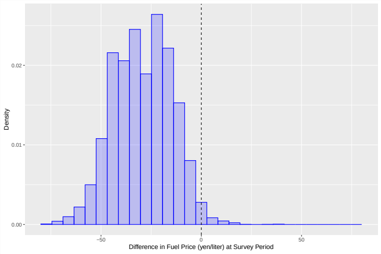


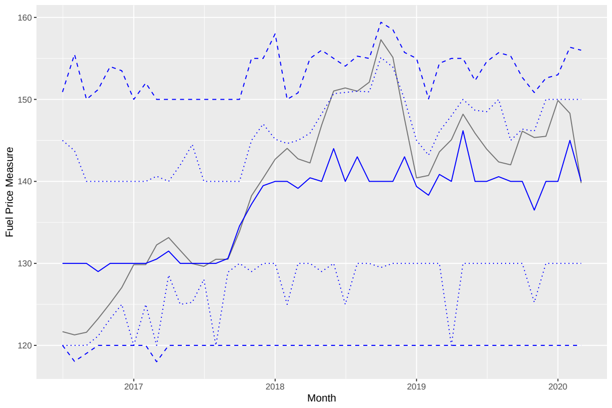
Figure 18: Difference between Belief and Price at Surveyed Period



Return

Appendix - Data: Fuel Price Belief (Over-time)

Figure 19: Fuel Price Belief over Time



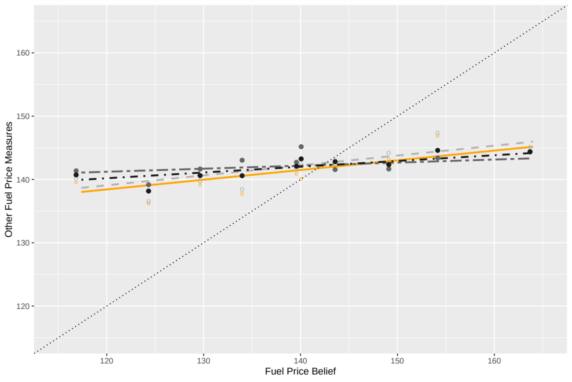
Note: Samples in this figure omit consumers who purchase a vehicle using high-octane gasoline. The black solid line is the median price of actual gasoline. The blue lines represent the fuel price belief (solid: median, dotted: 25% and 75% percentiles, dashed: 10% and 90% percentiles).

▸ Return

Appendix - Data: Fuel Price Belief (Correlation across Measures)

- Fuel price when consumers purchased and other fuel price measures (moving-average, seasonally-adjusted prices)

Figure 20: Binscatterplot: Correlation across Fuel Price Measures



Return

Appendix - Data: Fuel Price Belief (Summary)

The difference between belief and price when consumers purchased the vehicles:

- Automaker: Drivers with Lexus have larger fuel price belief gaps
- Other heterogeneities: Not significant

$\Rightarrow$ Fuel price belief is random to the observed characteristics

Appendix - Data: Fuel Price Belief (Test Drive)

Table 5: Test Drive

	Fuel Price Gap			
	Rate	Rate	Value	Value
Test Drive	-0.003^{*}	-0.003^{*}	-0.381^{*}	-0.372^{*}
	(0.001)	(0.001)	(0.210)	(0.211)
Vehicle Brand FE	Y	Y	Y	Y
Month FE	Y	Y	Y	Y
Prefecture FE	N	Y	N	Y
Num. obs.	9523	9523	9523	9523

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Appendix - Data: Fuel Price Belief (Correlation with Fuel Consumption)

- Compare two measures: Fuel price belief and Fuel price when a consumer purchased

Figure 21: Correlation with Fuel Consumption Belief

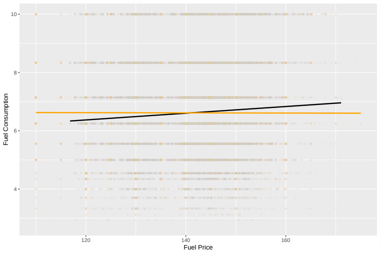
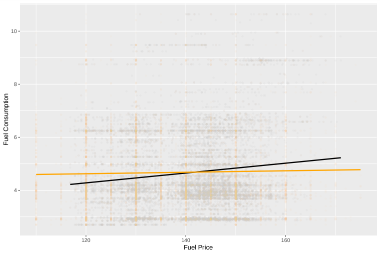


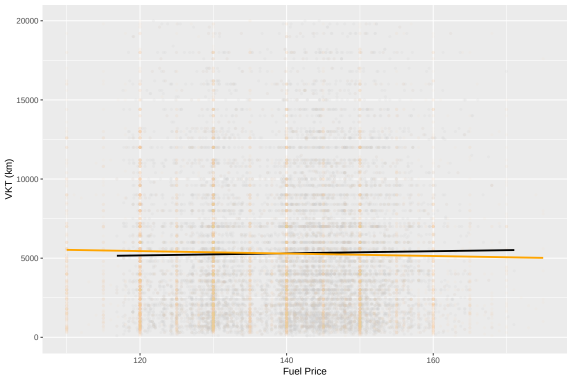
Figure 22: Correlation with Official Fuel Consumption



Appendix - Data: Fuel Price Belief (Correlation with VKT)

- Compare two measures: Fuel price belief and Fuel price when a consumer purchased

Figure 23: Correlation with VKT



Appendix - Data: Ownership Duration (Summary)

- Automaker: Longer ownership (Mitsubishi and Suzuki), Short ownership (Toyota, Nissan, and Lexus)
- Body-type: Longer ownership (Minivan), Short ownership (Light and SUV)
- Demographics: Longer ownership (Married), Short ownership (High income, male, and elderly)

Appendix - Data: Ownership Duration Belief (Correlation with Income)

Figure 24: Correlation with Household Income

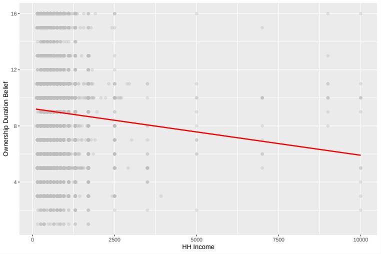
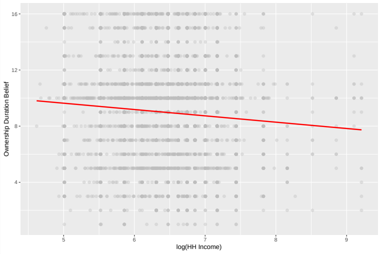


Figure 25: Correlation with Household Income (log)



Return

Appendix - Data: Ownership Duration Belief (Correlation with Fuel Consumption and Price)

Figure 26: Correlation with Fuel Consumption

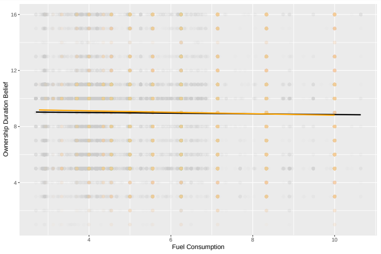
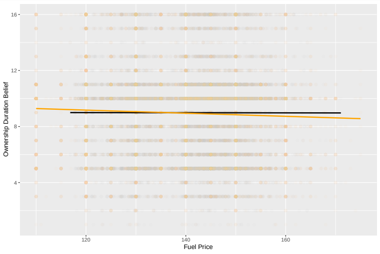
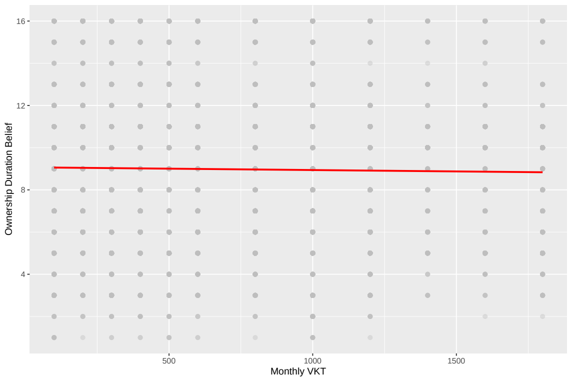


Figure 27: Correlation with Fuel Price



Appendix - Data: Ownership Duration Belief (Correlation with VKT)

Figure 28: Correlation with VKT



Appendix - Data: Discount Rates

- Similar idea with Allcott and Wozny (2014):

  - Pay by cash = discount rate in an average opportunity cost, 2.07% $\approx$ 2%
  - Dealer finance: Varying across auto-dealers, regions, time, and specific vehicles
  - Other finance (e.g., Bank/Credit union loan): Varying across regions

Table 6: Heterogeneity in Loan Rate

Payment Method	Share of Vehicles	Discount Rate (S.D.)
Lump-sum Cash	58.11%	2.00% (—)
Dealer Finance	33.66%	4.41% (1.33)
Other Finance	8.22%	3.25% (0.51)

- Consistent with national survey by Japan Automobile Manufactures Association (57% for those who purchased vehicles between 2018-2019 choose lump-sum cash payment)
- Bank/Credit union Loan is preferred by SUV and expensive vehicle buyers

▸ Return

Appendix - Data: Discount Rates (Summary)

- Automaker: High (Toyota, Nissan, and Mitsubishi), Low (Mazda and Subaru)
- Body-type: High (Light), Low (Minivan, Sedan, SUV, and Wagon)
- Demographics: High (Middle income, male, with children, and many family members), Low (eldery and married)
- City information: Low (Many employment)

Appendix - Data: Discount Rates (Correlation with Fuel Consumption and Price)

- Correlations with beliefs in fuel consumption

Figure 29: Correlation with Fuel Consumption

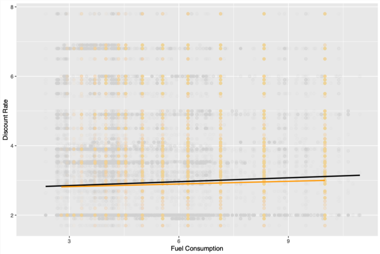
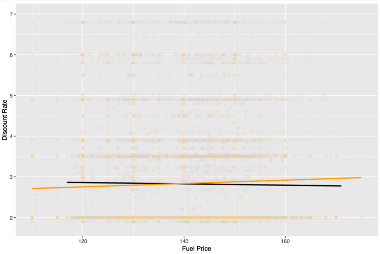
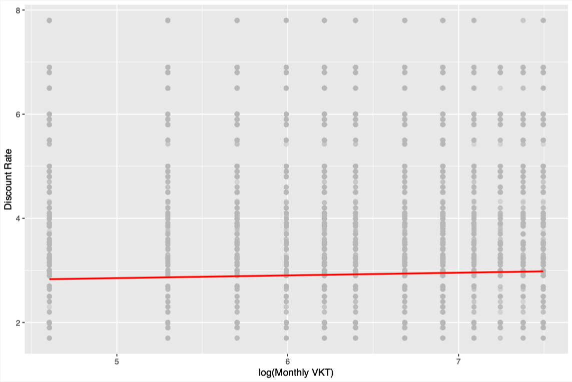


Figure 30: Correlation with Fuel Price



Appendix - Data: Discount Rates (Correlation with VKT)

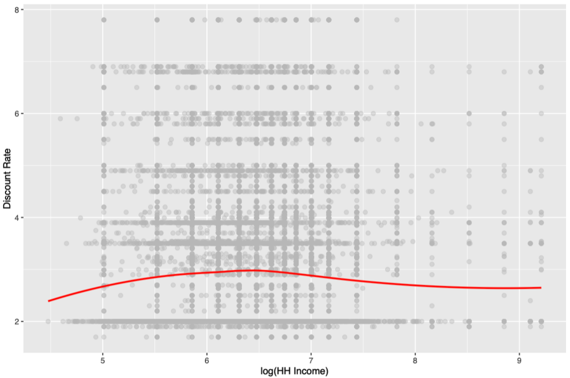
Figure 31: Correlation with log(VKT)



Appendix - Data: Discount Rates (Income)

- Lower/higher-income HHs are older than middle-income ▸ Appendix
  - Some lower-income HHs may be retired and have large assets

Figure 32: Correlation with Income (log)

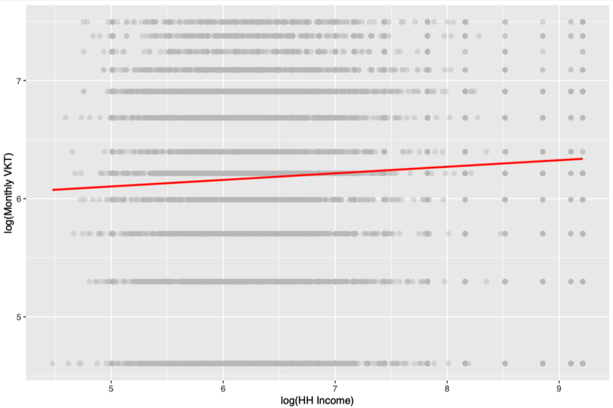


Appendix - Data: VKT (Summary)

- Automaker: Long (Suzuki), Short (Toyota, Nissan, and Lexus)
- Body-type: Long (Wagon), Short (Light and Minivan)
- Demographics: Long (High income, male and with many family members), Short (elderly, married, and with children)
- City information: Long (many employments), Short (more population and educated)

Appendix - Data: VKT (Correlation with Income)

Figure 33: Correlation: VKT (log) and Income (log)



▸ Return

Appendix - Data: Income (Summary)

Higher-income car buyers are more likely to:

- Married, having children, and more family members
- Prefer automakers of TOYOTA and LEXUS
- Prefer Sedan, Minivan, and SUV
- Buy expensive and more fuel-consuming, hybrid, more weighted, large horsepower, and gas emission vehicles

Lower-income car buyers are more likely to:

- Prefer SUZUKI and DAIHATSU vehicles
- Prefer Compact or Light vehicles

Both lower/higher-income buyers are more likely to:

- Older than middle income

Return to DR

Appendix - Data: Income and Demographics

Table 7: Income and Demographics

Income Group	Age	Martial Status	With Children	Family Members
Q1	51.33	0.54	0.32	2.54
Q2	50.69	0.63	0.35	2.62
Q3	48.39	0.68	0.43	2.75
Q4	47.50	0.74	0.51	2.99
Q5	49.62	0.71	0.48	3.16
Q6	46.86	0.78	0.57	3.12
Q7	47.71	0.81	0.59	3.20
Q8	48.17	0.83	0.61	3.24
Q9	49.17	0.84	0.64	3.41
Q10	50.87	0.84	0.64	3.54

Appendix - Data: Income and Vehicle Selection

Table 8: Income and Automakers

Income Group	TOYOTA	HONDA	NISSAN	MAZDA	SUBARU	MITSUBISHI	SUZUKI	DAIHATSU	LEXUS
Q1	0.33	0.23	0.10	0.04	0.03	0.02	0.13	0.12	0.01
Q2	0.34	0.24	0.10	0.04	0.03	0.02	0.11	0.10	0.01
Q3	0.35	0.24	0.11	0.05	0.03	0.01	0.10	0.09	0.01
Q4	0.36	0.25	0.11	0.05	0.04	0.02	0.09	0.08	0.01
Q5	0.38	0.21	0.08	0.05	0.05	0.02	0.08	0.10	0.02
Q6	0.37	0.25	0.11	0.05	0.04	0.01	0.09	0.07	0.01
Q7	0.38	0.23	0.12	0.05	0.04	0.02	0.08	0.07	0.02
Q8	0.39	0.21	0.11	0.06	0.06	0.02	0.08	0.06	0.01
Q9	0.38	0.23	0.11	0.06	0.05	0.02	0.06	0.06	0.03
Q10	0.41	0.20	0.10	0.06	0.05	0.01	0.06	0.04	0.06

Appendix - Data: Income and Vehicle Selection

Table 9: Income and Body-Types

Income Group	Sedan	Wagon	Minivan	SUV	Open-Coupe	Compact	Light
Q1	0.07	0.02	0.23	0.10	0.01	0.17	0.40
Q2	0.08	0.03	0.26	0.09	0.00	0.18	0.37
Q3	0.07	0.03	0.31	0.10	0.01	0.16	0.32
Q4	0.07	0.03	0.35	0.11	0.01	0.16	0.28
Q5	0.08	0.03	0.32	0.13	0.01	0.12	0.31
Q6	0.07	0.03	0.37	0.13	0.01	0.14	0.26
Q7	0.07	0.03	0.38	0.13	0.00	0.14	0.25
Q8	0.09	0.03	0.37	0.14	0.01	0.14	0.23
Q9	0.10	0.03	0.35	0.15	0.01	0.15	0.20
Q10	0.15	0.03	0.32	0.18	0.01	0.14	0.16

Appendix - Data: Income and Vehicle Selection

Table 10: Income and Vehicle Price+Attributes

Income Group	Transaction Price	Official Fuel Consumption	Hybrid	Diesel	Weight	Horse Power	Gas Emission
Q1	2155760	4.29	0.34	0.02	1151	67.48	1217
Q2	2235202	4.36	0.35	0.02	1181	69.77	1266
Q3	2288231	4.51	0.36	0.02	1220	73.64	1333
Q4	2436746	4.63	0.36	0.03	1266	77.77	1405
Q5	2500699	4.55	0.35	0.03	1252	77.37	1377
Q6	2491604	4.75	0.36	0.03	1297	81.27	1455
Q7	2519394	4.75	0.38	0.03	1311	81.42	1476
Q8	2563134	4.81	0.37	0.04	1327	84.00	1515
Q9	2675759	4.80	0.41	0.04	1347	86.41	1560
Q10	2937483	4.87	0.46	0.04	1395	93.48	1684

Appendix - Data: Resale Values

- Data record resale prices of the previous vehicles for $26,723$ samples
- LEXUS vehicle has a large resale value
- The number of holding years is a strong predictor of resale values
- Prediction is sufficiently strong ($R^2 = 0.668$) and consistent with the report by a market research company

Figure 34: Histogram of log(Resale Values)

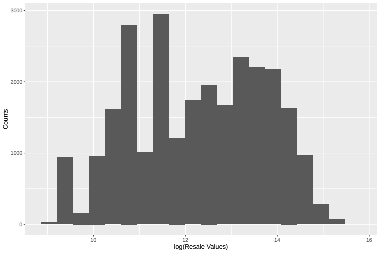
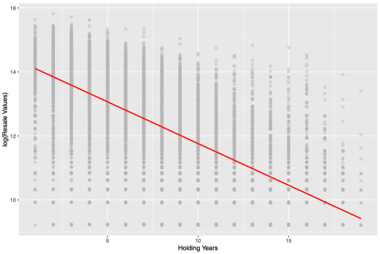


Figure 35: Correlation: Holding Years and log(Resale Values)



Back

Appendix - Data: Resale Values (Prediction)

- Holding years and vehicle attributes explain most variations in resale values
- Predicted resale values: 571,365 yen on average (481,383 yen in median), consistent with the report by a market research company

Table 11: Which Variations Explain Most?

Regression Type	R-Squared: Value
All Variables	0.668
All Variables (Except for Holding Years)	0.435
Vehicle Attribute (with Brand FE)	0.373
Vehicle Attribute (without Brand FE)	0.278
Demographics (with Preference)	0.029
Demographics (without Preference)	0.013
City (with Prefecture FE)	0.040
City (without Prefecture FE)	0.029

Appendix - Data: Construction of Future Fuel Cost

Following Allcott and Wozny (2014) and Grigolon et al. (2018),

$$G_{ij} = \mathbb{E}\left[\sum_{s=0}^{S_i}(1+r_i)^{-s}\left(m_i e_{ij} g_{is} + I_i\right)\right]. \tag{1}$$

where

- r_i: discount factor
- m_i: consumer i's expected annual VKT
- e_{ij}: fuel consumption (or its belief) of car j (l/km; the inverse of fuel economy)
- g_{is}: fuel cost that consumer i faces at period s
- S_i: consumer's expected vehicle ownership duration
- I_i: Insurance and maintenance costs

- Two Assumptions:

1. Inelasticity of VKT to fuel price (Small and Van Dender, 2007; Hughes et al., 2008; Gillingham, 2014); we confirm it Appendix
2. Random-walk of fuel price (Anderson et al., 2013)

Appendix - Data: Overview of Price/Future Values

- Insurance and maintenance costs (including equipment costs, mandatory inspection fees, etc) are surveyed

Table 12: Examples of Prices and Costs (Yen)

Car Brand	Transaction Price	MSRP	Future Fuel Cost	Insurance	Maintenance	Resale Value
TOYOTA: Corolla	2409525	2480446	502707	452542	485612	500679
	(620868)	(228305)	(398775)	(325368)	(512735)	(309498)
HONDA: Odyssey	3810829	4016663	688084	631924	668941	833439
	(897477)	(238007)	(513020)	(616859)	(854444)	(293955)
NISSAN: Rogue	3033528	3493203	741003	652741	617249	755248
	(739281)	(249188)	(529788)	(852648)	(597675)	(296029)
MAZDA: CX	2974698	3277312	664395	571946	529254	817892
	(692744)	(415707)	(566986)	(611599)	(526797)	(291225)
SUBARU: Forester	3036340	3329543	774649	621940	493078	828210
	(791771)	(127886)	(606213)	(587008)	(550687)	(321388)
LEXUS: NX	4779612	5647367	470787	471037	493285	2754952
	(1525446)	(336139)	(443567)	(357586)	(676277)	(242531)
MITSUBISHI: eK	1572795	1730245	500035	534654	555839	130998
	(420207)	(215131)	(457082)	(643658)	(627066)	(168819)
SUZUKI: Wagon R	1491009	1539774	394045	423202	507628	228308
	(286576)	(149405)	(314480)	(338591)	(519381)	(200269)

Appendix - Data: Overview of Price/Future Values

- Operation costs include insurance and maintenance costs
- Ownership duration belief and resale values are incorporated at the same time

Table 13: Future Fuel Costs

Belief Components	Value	SD
No Belief	553388	(199518)
Belief: Fuel Cons (FC)	781220	(241624)
Belief: Fuel Price (FP)	536496	(194357)
Belief: Lifetime (LT) + Resale	947717	(511315)
Belief: Discount Rate (DR)	670588	(244141)
Belief: VKT	541032	(440078)
Belief: No FC	973153	(615600)
Belief: No FP	1139336	(673121)
Belief: No LT+Resale	865350	(645284)
Belief: No DR	1054162	(627288)
Belief: No VKT	1157518	(501845)
Belief: All	1120400	(660830)
Belief: All + Operating Costs	2203209	(1156094)

Appendix - Data: Histogram of Fuel Efficiency Ratios

Figure 36: On-Road FE / Official FE

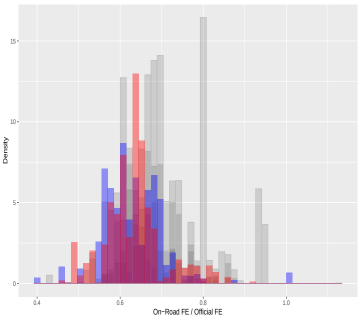
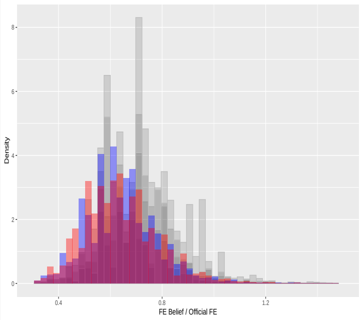


Figure 37: FE Belief / Official FE



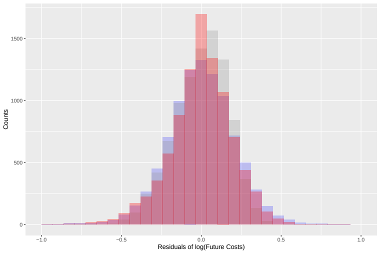
Note: Blue histogram shows compact vehicle, red histogram shows minicars ("kei"-cars).

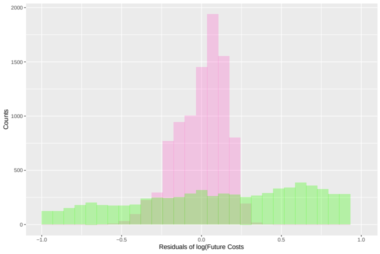
Appendix - Data: Future Fuel Costs

- Compute residuals after controlling for vehicle attributes and fixed effects
- Show five histograms with only one belief component: (a) FC Belief, (b) FP Belief, (c) OD Belief + Resale, (d) DR Belief, (e) VKT Belief

⇒ VKT creates the large variations

Figure 38: Histogram of Residuals in Additional One Belief





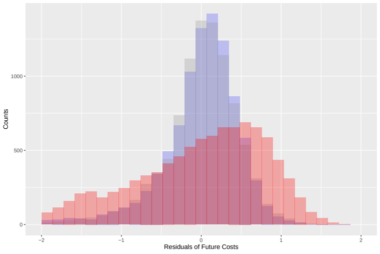
Appendix - Data: Future Fuel Costs

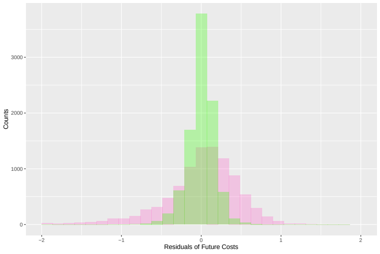
- Show five histograms with four belief components: (a) No FC, (b) No FP, (c) No OD+Resale, (d) No DR, (e) No VKT

$\Rightarrow$ Omission of FC, FP, and VKT lost variations

$\Rightarrow$ Beliefs of fuel consumption, fuel price, and VKT are important

Figure 39: Histogram of Residuals in Additional One Belief





- Note: This decomposition exercise ignores the covariance across belief components

Appendix - Data: Summary

- Large variations in belief heterogeneity:
  - Fuel consumption, fuel price, and VKT
- Notable correlations:
  - Fuel consumption and VKT: Driving behavior
  - Discount rate and vehicle ownership duration: Vehicle selection

Table 14: Correlation Matrix across Belief Components

	Belief Components					Income
Variable	Fuel Cons	Fuel Price	Lifetime	Discount	VKT	
Fuel Consumption					$-$	$+$
Fuel Price						
Lifetime				$-$		$-$
Discount Rate			$-$			U-inverted
VKT	$-$					$+$

Appendix - Theoretical Insights

How does the omission of the belief component affect the valuation parameter estimates?

1. Heterogeneity in belief
   - Within-individual: Large heterogeneity ⇒ Small
   - Cross-sectional: Depends on the correlation across fuel costs, market shares, and prices
2. Consumer type: Within-individual
   A. Correlation between VKT and fuel consumption (driving style) ⇒ Large
   B. Correlation between discount rate and ownership duration (vehicle selection) ⇒ Small
3. Consumer type: Cross-sectional
   - [Working]

- Analogous to nudge theory:

The choice distortion affects the nudge welfare effect (Allcott et al., 2025):

1. Heterogeneity in belief → Heterogeneity in nudge effects
2. Consumer type → Correlations between nudge effects and consumer bias

▸ Proof 1 ▸ Proof 2-A ▸ Proof 2-B

▸ Back

Appendix - Theory: (i) Alternative Fuel Consumption Measure

- Case of within-individual variations

  - Variation only arises from fuel consumption belief

According to the concept, suppose two vehicles A and B (same attributes except for fuel consumption), the valuation parameter of official FC (γ^o) is:

$$
\begin{aligned}
(p_A - p_B) &= \gamma^o \cdot (G_B^o - G_A^o) \\
\Leftrightarrow \quad \gamma^o &= \frac{(p_A - p_B)}{(G_B^o - G_A^o)}
\end{aligned}
$$

Suppose fuel consumption belief measure, $G_i^b = G_i^o \cdot \frac{FC_i^b}{FC_i^o}$:

$$
\frac{\gamma^b}{\gamma^o} = \frac{G_B^o - G_A^o}{G_B^b - G_A^b} = \frac{FC_B^o - FC_A^o}{FC_B^b - FC_A^b}
$$

Variations of "'Belief/Official" matter

- If belief heterogeneity is large, the valuation parameter using belief becomes smaller

Return

Appendix - Theory: (i) Alternative Fuel Consumption Measure

- Case of cross-sectional variations

Abstracting away from random coefficients and endogeneity issues, the simplified estimated equation is:

$$s = \alpha p + \alpha\gamma G + \varepsilon$$

Denote

- γ^o and G_o: valuation parameter and future fuel cost without belief heterogeneity
- γ^b and G_b: valuation parameter and future fuel cost with belief heterogeneity

$$\frac{\gamma^b}{\gamma^o} = \frac{(\sum p^2)(\sum G_b s) - (\sum pG_b)(\sum ps)}{(\sum p^2)(\sum G_o s) - (\sum pG_o)(\sum ps)} \cdot \frac{(\sum G_o^2)(\sum ps) - (\sum pG_o)(\sum G_o s)}{(\sum G_b^2)(\sum ps) - (\sum pG_b)(\sum G_b s)}$$

Which valuation parameters (with/without belief heterogeneities) are large depends on correlations across future fuel costs, market shares, and prices.

Appendix - Theory: (ii) Consumer Type

- Correlation between VKT and FE (driving style):

  - Individuals expecting to drive intensively have higher expected fuel cost savings for a particular efficiency improvement
  - They will prefer vehicles with low fuel consumption

- Correlation between discount and ownership duration (vehicle selection):

  - Car enthusiast: Short periods and loan with high rates
  - Parenting generation: Long periods and cash payment

Applying the same logic as "(i) alternative fuel efficiency measure",

$$\gamma^{corr} > \gamma^{no-corr}$$

▸ Data

Appendix - Theory: (ii) Consumer Type - Correlation Between FC and VKT

Suppose that $G_i = H \cdot \frac{m_i}{FE_i}$ (H is another component of future fuel costs). The valuation parameter can be expressed as:

$$\begin{aligned} \gamma &= \frac{(p_A - p_B)}{(G_B - G_A)} \\ &= \frac{(p_A - p_B)}{H \cdot (\frac{m_B}{FE_B} - \frac{m_A}{FE_A})} \end{aligned}$$

With correlations between VKT (m_i) and FE_i, $m_A^{non-corr} < m_A^{corr}$ and $m_B^{non-corr} > m_B^{corr}$. That implies:

$$\begin{aligned} \frac{\gamma^{corr}}{\gamma^{non-corr}} &= \frac{(\frac{m_B^{non-corr}}{FE_B} - \frac{m_A^{non-corr}}{FE_A})}{(\frac{m_B^{corr}}{FE_B} - \frac{m_A^{corr}}{FE_A})} \\ &> 1 \end{aligned}$$

$\because \frac{m_B^{non-corr}}{FE_B} > \frac{m_B^{corr}}{FE_B}$ and $-\frac{m_A^{non-corr}}{FE_A} > -\frac{m_A^{corr}}{FE_A}$.

Therefore,

$$\gamma^{corr} > \gamma^{non-corr}$$

Return

Appendix - Theory: (ii) Consumer Type - Correlation Between DR and LT

Suppose that $G_{ij} = C_{ij} \cdot \sum_{s=0}^{S_i}(1+r_i)^{-s}$ (C is a fuel cost per year). The valuation parameter can be expressed as:

$$\begin{aligned}\gamma &= \frac{(p_A - p_B)}{(G_B - G_A)} \\ &= \frac{(p_A - p_B)}{(C_A - C_B) \cdot \sum_{s=0}^{S^{nc}}(1+r_i^{nc})^{-s}}\end{aligned}$$

Then,

$$\frac{\gamma^{corr}}{\gamma^{nc}} = \frac{\sum_{s=0}^{S^{nc}}(1+r_i^{nc})^{-s}}{\sum_{s=0}^{S^{corr}}(1+r_i^{corr})^{-s}}$$

Suppose $S^{corr} > S^{nc}$ and $r^{corr} < r^{nc}$ (given negative correlation), $\sum_{s=0}^{S^{corr}}(1+r_i^{corr})^{-s} > \sum_{s=0}^{S^{nc}}(1+r_i^{nc})^{-s}$.

Therefore,

$$\gamma^{corr} < \gamma^{non-corr}$$

Return

Appendix - Estimation: Method and Issues

- MWTP for future fuel cost is biased without random coefficients (Bento et al., 2012) Linear logit

1. Simulated Maximum Likelihood (random coefficient mixlogit estimation)
    - Similar to Ito and Sallee (2018) and Huse and Koptyug (2022)
    - Applicability is limited (Correlation effect cannot be tested)
2. Random coefficient model with random variables [Plan]
    - Consider the one-step likelihood estimation using 2nd choice information
    - Similar strategy with Goolsbee and Petrin (2004); Vatter (2025)

Common Issues:

- Cannot observe fuel consumption beliefs for all non-purchased vehicles $\Rightarrow$ Prediction
    - Systematic bias in fuel consumption between purchased/non-purchased vehicles may appear

Back

Appendix - Estimation: Fuel Consumption Belief Prediction

Predicting belief keeping variations:

- Mean prediction by random forests + Matching through PMM (Predictive Mean Matching) or LRD (Local Residual Draws) ▸ Method
  - Random forest can achieve precise predictions, but loses variations. PMM/LRD can keep the same level of variation as beliefs ▸ Variation

Selection issues:

- Fuel consumption in the 2nd (or more later) choice vehicles may be overvalued (i.e., more fuel consuming)
- Additional survey: We have conducted an additional survey on Jan 15-19, 2026, asking about fuel consumption beliefs for non-purchased vehicles Note: Add survey structure in appendix and put the link

▸ Back

Appendix - Empirical Method I: Simulated Maximum Likelihood

Estimated equation:

$$U_{ijt} = \alpha_i p_{jt} + \alpha_i \gamma_i G_{ijt} + \nu_{it} + \phi_j + \varepsilon_{ijt}$$

where p_{jt}: MSRP, G_{ijt}: fuel cost belief, and ε_{ijt} follows type-I extreme value distribution.

Sample and estimation of SML:

1. Consideration set:
   - Data has information on vehicle consideration set (up to seven)
   - Consumers have 2.7 vehicles in a consideration set, on average
   - On-going: Try the case where I use all products in the market as a choice set
2. Fuel consumption beliefs of vehicles in consideration sets
   - We only ask about the fuel consumption beliefs on vehicles purchased
   - Predict fuel consumption belief Prediction

Appendix - Results: Simulated Maximum Likelihood

Table 15: Simulated Maximum Likelihood

	Official	Belief				
		RF	PMM (1)	PMM (2)	LRD (1)	LRD (2)
Price	-0.216	-0.264	-0.244	-0.236	-0.244	-0.241
	(0.068)	(0.072)	(0.070)	(0.069)	(0.070)	(0.069)
Fuel Cost	-0.112	-0.320	-0.190	-0.090	-0.206	-0.142
	(0.048)	(0.087)	(0.059)	(0.044)	(0.059)	(0.045)
Future Valuation	**0.434**	**1.209**	**0.778**	**0.383**	**0.844**	**0.589**
	(0.272)	(0.477)	(0.332)	(0.219)	(0.345)	(0.253)
Price (S.D.)	0.005	0.008	0.009	0.006	0.007	0.009
	(0.010)	(0.013)	(0.011)	(0.011)	(0.011)	(0.010)
Fuel Cost (S.D.)	0.010	3.483	2.074	1.326	2.093	1.495
	(0.035)	(0.351)	(0.250)	(0.233)	(0.249)	(0.224)
Num. obs.	100,141	100,141	100,141	100,141	100,141	100,141
Num. consumers	32,117	32,117	32,117	32,117	32,117	32,117

- Concerns: Endogeneity (can be addressed by Petrin and Train (2010)), Size of choice sets

Back Role of random coefficients

Appendix - Empirical Method II: Random Coefficient with Random Variables

Utility function:

$$U_{ijt} = \alpha_i p_{jt} + \alpha_i \gamma_i G_{ijt} + \beta X_{jt} + \phi_j + \xi_t + \varepsilon_{ijt}$$

Composite probabilities of vehicle purchase choice and 2nd choice (c_{1i}: first choice, c_{2i}: second choice):

$$\Pr\left[(c_{1i} = j) \cap (c_{2i} = l) \mid \theta\right] = \frac{\exp(V_{ij}(\theta_i))}{\sum_{k \in J} \exp(V_{ik}(\theta_i))} \cdot \frac{\exp(V_{il}(\theta_i))}{\sum_{k' \in J \setminus \{j\}} \exp(V_{ik'}(\theta_i))} \quad (2)$$

Maximize log-likelihood with the following two constraints ▸ Likelihood:

- Share equalization (Berry et al., 1995): Data=Prediction in the share variables for product j (same spirit as Goolsbee and Petrin (2004); Vatter (2025))

$$s_j = \frac{1}{N} \sum_{i \in N} \hat{\Pr}\left[(c_{1i} = j) \mid \eta\right]$$

- Selection on non-purchased vehicles for product j:

$$\text{Avg. of FC Belief for purchased}_j - \text{Avg. of FC Belief for non-purchased}_j = \text{Selection}_j$$

Appendix - Utility and Consumer Myopia

Utility function specification in a discrete choice model is consistent with the consumer myopia concept:

E.g., Suppose two vehicles (A, B) with the same vehicle attributes except fuel efficiency and same unobserved fixed effect δ_j. What happens, when consumer (i) compare two vehicles when they decide purchase vehicle A?

$$\begin{aligned}
& V^p_{iAt} > V^p_{iBt} \\
\Leftrightarrow \quad & -\alpha(p_{iAt} + \gamma G^p_{iAt}) > -\alpha(p_{iBt} + \gamma G^p_{iBt}) \\
\Leftrightarrow \quad & p_{iAt} - p_{iBt} < \gamma(G^p_{iBt} - G^p_{iAt}) \\
\Leftrightarrow \quad & \Delta p_{ijt} < \gamma \Delta G^p_{ijt}
\end{aligned}$$

γ: Valuation parameter

Appendix - Issues of Estimation in BLP

Can our data fit the estimation of BLP?

- Our data tries to estimate the random coefficients for G_{ijt} (individual i) to incorporate idiosyncratic beliefs to future fuel costs.

Decompose individual-level data into:

- $p_{ijt} = \bar{p}_{.,jt} + \Delta p_{ijt}$
- $G_{ijt} = \bar{G}_{.,jt} + \Delta G_{ijt}$

$$\begin{aligned} u^p_{ijt} &= -\alpha_i(\bar{p}_{.,jt} + \Delta p_{ijt}) - \alpha_i \tilde{\gamma}_i(\bar{G}_{.,jt} + \Delta G_{ijt}) + \boldsymbol{\beta}^X x_{jt} + \delta_j + \lambda_t + \varepsilon_{ijt} \\ &= \delta_{jt}(x_{jt}, \bar{p}_{.,jt}, \bar{G}_{.,jt}, \xi_j, \Delta_{jt}; \theta_1) + \mu_{ijt}(p_{ijt}, G_{ijt}, v_i; \theta_2) + \varepsilon_{ijt} \end{aligned}$$

where

$$\delta_{jt} = -\alpha \bar{p}_{.,jt} - \alpha \tilde{\gamma} \bar{G}_{.,jt} + \beta^X x_{jt} + \xi_j + \Delta \xi_{jt}$$

$$\begin{aligned} \mu_{ijt} &= -\alpha \Delta p_{ijt} + \sigma^p v_i(\bar{p}_{.,jt} + \Delta p_{ijt}) - \alpha \tilde{\gamma} \Delta G_{ijt} \\ &+ \left\{\sigma^p v_i(\alpha + \sigma^G v_i) + (\sigma^G v_i)^2\right\}(\bar{G}_{.,jt} + \Delta G_{ijt}) \end{aligned}$$

Appendix - Issues of Estimation in BLP

BLP Estimation Procedure:

1. Guess a set of parameters of (σ^p, σ^G)
2. Contraction mapping solves for δ_{jt} sets predicted shares equal to actual shares and gets mean utilities δ_{jt}
3. Recover linear parameters $(\hat{\alpha}, \hat{\gamma}, \hat{\beta}^X)$ by IV-GMM given (σ^p, σ^G)
4. Compute the moment conditions using the residuals from the above IV-GMM

Issues:

- Estimation: μ_{ijt} includes α and $\tilde{\gamma}$ (not separable)
- Cannot compute μ_{ijt} at Step 2 given (σ^p, σ^G) (use a given μ_{ijt} when calculating individual/market shares in the inner loop at Step 2)
- To compute μ_{ijt} at Step 2, we need to guess $(\hat{\alpha}, \hat{\gamma})$, but it may disable the linear IV-GMM when estimating linear parameters.

▸ Return

Appendix - Estimation: Fuel Efficiency Belief Prediction

Predictive Mean Matching (PMM) (Rubin, 1986; Little, 1988):

1. Compute predictions μ_j for observed samples (purchased vehicles)
2. Find the k nearest donors by $|\mu_j - \mu_i|$ (i.e., local nearest matching)
3. Set $\tilde{Y}_i = Y_j$: Y_J is observed data sampled among the k nearest donors

Local Residual Draws (Morris et al., 2014):

1. Compute residuals $e_j = Y_j - \mu_j$ for observed samples (instead of predictions, μ_j)
2. Find the k nearest donors from the set of μ (i.e., local nearest matching)
3. Sample residuals among the k nearest donors and construct $\tilde{Y}_i = \mu_i + e_J$

Application examples in applied economics

- Similar imputation as PMM (Kuhn, 2020; White et al., 2018; Davis and Heller, 2020)

Back

Appendix - Estimation: Fuel Efficiency Belief Prediction

Table 16: Prediction Method

Prediction Method	R-Squared	Raw S.D.	Residual S.D.	Prediction Error Mean	Prediction Error Median	Prediction Error S.D.
Regression	0.574	3.42	1.55	3.28%	1.43%	18.96
Gradient Boosting	0.650	3.40	1.80	3.00%	1.30%	17.31
Random Forest (RF)	0.848	3.37	1.82	2.02%	0.96%	11.26
PMM with RF (1)	0.749	3.74	3.41	1.89%	0.29%	23.29
PMM with RF (2)	0.588	4.20	5.83	1.89%	0.29%	23.29
LRD with RF (1)	0.749	3.75	3.40	1.94%	0.27%	23.36
LRD with RF (2)	0.585	4.21	5.79	1.89%	0.29%	23.29

A clear tradeoff between prediction precision and variations:

- Random forest has more precise predictions, but small variations
- PMM/LRD increases variations, but the prediction falls

⇒ Try estimation with three types of predictions: (1) Random forest (RF), (2) PMM with RF, and (3) LRD with RF

  - PMM and LRD have two cases: (i) take an average over 25 predictions, (ii) take an average over 2 predictions

Appendix - Estimation: Fuel Efficiency Belief Prediction

Table 17: Residual Variations across Prediction Method

Fuel Cons. Measures		Sample + Variables		
		Purchased + Car FEs	Purchased + All Demo. Variables	All + Ind./Car FEs
Data	Official	9.33	7.39	5.11
	Belief	13.21	10.87	—
Prediction	Regression (mean)	4.04	2.39	1.55
	GB (mean)	4.43	2.92	1.80
	RF (mean)	5.42	4.07	1.82
	PMM with RF (1)	7.92	6.39	3.41
	PMM with RF (2)	11.48	9.74	5.83
	LRD with RF (1)	7.95	6.39	3.40
	LRD with RF (2)	11.77	10.04	5.79

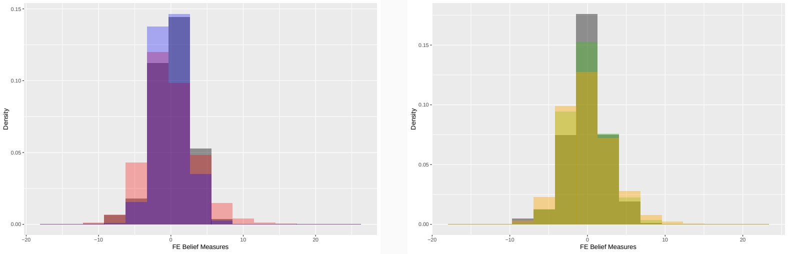
- The variations in belief are slightly larger than PMM/LRD with RF (2)

⇒ Estimated valuation parameters with PMM/LRD with RF (2) are upper bounds in using belief

Back

Appendix - Estimation: Fuel Efficiency Belief Prediction

Figure 40: Distribution of Residuals in FE (Belief) Measures



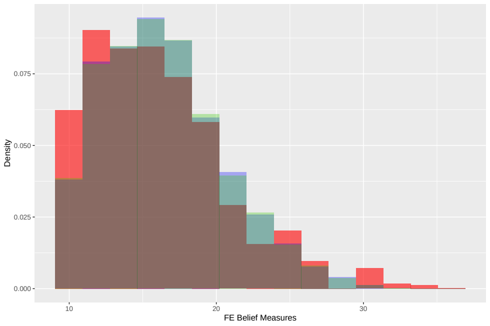
Note: The sample is limited to the purchased vehicles. Colors: official, belief (data), random forest (mean prediction), RF with PMM (1), and RF with PMM (2).

- Belief has the largest variations after controlling all relevant variables
  - S.D. Official: 7.39
  - S.D. Belief: 10.87
  - S.D. RF (mean): 4.07
  - S.D. RF with PMM (1): 6.39
  - S.D. RF with PMM (2): 9.74

▸ Appendix: In detail

Appendix - Estimation: Fuel Efficiency Belief Prediction

Figure 41: Distribution of FE Belief and FE Belief Prediction with PMM



Note: The red, blue, and green show belief (data), and predicted belief (PMM, and LRD) in fuel efficiency.

- Applying PMM (Predictive Mean Matching) introduces the randomness

Appendix - Estimation: Fuel Efficiency Belief Prediction

- Vehicle attribute can explain most variations in the fuel efficiency gap
- Rate = "fuel efficiency belief (after purchase)" / "catalog fuel efficiency"

Table 18: What factors drive variations in fuel efficiency gap (surveyed sample)

Regression Type	R-Squared: Rate	R-Squared: Value
All Variables	0.243	0.538
Vehicle Attribute (with Brand FE)	0.196	0.511
Vehicle Attribute (without Brand FE)	0.155	0.498
Demographics	0.013	0.040
City (with Prefecture FE)	0.034	0.040
City (without Prefecture FE)	0.025	0.024

Return

Appendix - Estimation: Simulated Maximum Likelihood

- Data Type:

  - # of Samples: $100,141$
  - # of Consumers: $32,117$
  - Utilize information of consideration sets: microdata asks which vehicles consumers consider as alternatives to purchased vehicles (up to 7)

  $\Rightarrow$ Average # of Consideration Sets: 3.11 (Min: 2 and Max 7)

- How to create the variation within a consumer:

  - Predict belief by vehicle attributes in the data
  - Most variations of belief are explained by vehicle attributes (not individual-level variables)
  - Note: Cannot use cross-sectional variation $\Rightarrow$ Cannot test bias induced by systematic correlation

Appendix - Results: Simulated Maximum Likelihood

Table 19: Simulated Maximum Likelihood

	Official	Belief				
		RF	PMM (1)	PMM (2)	LRD (1)	LRD (2)
Price	-0.216	-0.225	-0.226	-0.227	-0.226	-0.226
	(0.068)	(0.067)	(0.067)	(0.067)	(0.067)	(0.067)
Fuel Cost	-0.094	-0.352	-0.207	-0.101	-0.221	-0.153
	(0.048)	(0.074)	(0.054)	(0.042)	(0.054)	(0.042)
Future Valuation	**0.435**	**1.570**	**0.915**	**0.444**	**0.976**	**0.675**
	(0.272)	(0.578)	(0.363)	(0.227)	(0.378)	(0.274)
Price (S.D.)	0.005	0.005	0.005	0.005	0.005	0.005
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Num. obs.	100,141	100,141	100,141	100,141	100,141	100,141
Num. consumers	32,117	32,117	32,117	32,117	32,117	32,117

- Drop the random coefficients for future fuel costs
- As the theory predicts, the variation drives future valuation parameter estimate magnitudes

$\Rightarrow$ Estimates with PMM/LRD (2) are upper bounds

- But, why do the omissions of random coefficients increase the parameters in fuel costs? Opposite observations from Bento et al. (2012).

Appendix - Results: Simulated Maximum Likelihood

Why do the opposite results from Bento et al. (2012) come out?

- Bento et al.'s framework:

$$\begin{aligned} u_{ij} &= \beta_i \cdot fc_j + \varepsilon_{ij} \\ &= (\bar{\beta} + \tilde{\beta}_i) \cdot fc_j + \varepsilon_{ij} \\ &= \bar{\beta} \cdot fc_j + e_{ij} \end{aligned}$$

where $e_{ij} = \tilde{\beta}_i \cdot fc_j + \varepsilon_{ij}$

By rescaling, the issue of heteroskedasticity turns into the omitted variable bias:

$$\frac{u_{ij}\sigma_\varepsilon}{\sqrt{\sigma_\beta^2 fc_j^2 + \sigma_\varepsilon^2}} = \frac{\bar{\beta} fc_j \cdot \sigma_\varepsilon}{\sqrt{\sigma_\beta^2 fc_j^2 + \sigma_\varepsilon^2}} + \frac{e_{ij}\sigma_\varepsilon}{\sqrt{\sigma_\beta^2 fc_j^2 + \sigma_\varepsilon^2}}$$

$$\Leftrightarrow \qquad \tilde{u}_{ij} = \bar{\beta} fc_j + \bar{\beta} fc_j \left[\frac{\sigma_\varepsilon}{\sqrt{\sigma_\beta^2 fc_j^2 + \sigma_\varepsilon^2}} - 1 \right] + \tilde{e}_{ij}$$

$$\Leftrightarrow \qquad \tilde{u}_{ij} = \bar{\beta} fc_j + \bar{\beta} fc_j (A_j - 1) + \tilde{e}_{ij}$$

and $Cov((A_j - 1), fc_j) > 0$. The omission of z_j induces the upward bias to β (closer to zero).

Appendix - Results: Simulated Maximum Likelihood

- Add individual-specific fc_{ij}:

$$
\begin{aligned}
u_{ij} &= \beta_i \cdot fc_{ij} + \varepsilon_{ij} \\
&= (\bar{\beta} + \tilde{\beta}_i) \cdot (1 + \theta_i) fc_j + \varepsilon_{ij} \\
&= \bar{\beta} \cdot fc_{ij} + e_{ij}
\end{aligned}
$$

By rescaling, the issue of heteroskedasticity turns into the omitted variable bias:

$$
\frac{u_{ij}\sigma_\varepsilon}{\sqrt{\sigma_\beta^2 fc_{ij}^2 + \sigma_\varepsilon^2}} = \frac{\bar{\beta} fc_{ij} \cdot \sigma_\varepsilon}{\sqrt{\sigma_\beta^2 fc_{ij}^2 + \sigma_\varepsilon^2}} + \frac{e_{ij}\sigma_\varepsilon}{\sqrt{\sigma_\beta^2 fc_{ij}^2 + \sigma_\varepsilon^2}}
$$

$$
\Leftrightarrow \qquad \tilde{u}_{ij} = \bar{\beta} fc_{ij} + \bar{\beta} fc_j \left[\frac{(1 + \theta_i) \cdot \sigma_\varepsilon}{\sqrt{\sigma_\beta^2 fc_{ij}^2 + \sigma_\varepsilon^2}} - 1 \right] + \tilde{e}_{ij}
$$

$$
\Leftrightarrow \qquad \tilde{u}_{ij} = \bar{\beta} fc_{ij} + \bar{\beta} fc_j \left[\tilde{A}_{ij} - 1 \right] + \tilde{e}_{ij}
$$

$\tilde{A}_{ij}$ can affect the results if $Cov(\theta_i, fc_{ij}) > 0$ and it dominates the denominator effect size (Note: endogeneity still may affect too. So, implement the control function approach).

- Data show: $Cov(\theta_i, fc_{ij}) > 0$

$\Rightarrow$ That leads to our counter-results from Bento et al. (2012)

Appendix - Results: Simulated Maximum Likelihood

Table 20: $Cov(\theta_i, fc_{ij})$

	Belief Deviations in Fuel Cost			
	PMM: Full	PMM: Subset	LRD: Full	LRD: Subset
log(Fuel Costs)	0.721***	0.403***	0.726***	0.412***
	(0.003)	(0.004)	(0.003)	(0.004)
Car Price	Y	Y	Y	Y
Vehicle Brand FE	Y	Y	Y	Y
Num. obs.	100141	32117	100141	32117

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Back

Appendix - Belief Prediction Extrapolation

Given fuel cost specification in (**??**), the cross-product variable is only e_{ij}.

- We can predict fuel consumption belief, but the prediction quality will matter (predictions have less variation)
- Possible selection: fuel consumption in the 2nd choice vehicles may be undervalued

How about adding a selection parameter to be estimated and random noise for the fuel consumption belief?

$$e_{ij} = \begin{cases} e_{ij} & j \text{ chosen vehicle} \\ \hat{e}_{ij'} + \kappa_{j'} & j' \text{ non-chosen vehicle} \end{cases}$$

where $\kappa_{2,j'}$ can be identified from the correlation between first choice (or automaker, vehicle types, ...) and measurement errors.

- Plan: We plan to run an additional survey, which elicits both fuel consumption beliefs of first- and second-choices, and use that information as micro-moments

Appendix - Empirical Method II: Random Coefficient with Random Variables

Sample log-likelihood:

$$L(\eta) = \sum_{i=1}^{N} w_i \cdot \log \left[\frac{\exp(V_{ij}(\theta_i))}{\sum_{k \in J} \exp(V_{ik}(\theta_i))} \cdot \frac{\exp(V_{il}(\theta_i))}{\sum_{k \in J \setminus \{j\}} \exp(V_{ik}(\theta_i))} \right] f(\theta \mid \eta)$$

where w_i is a weight for each individual i.

Estimation: Draw R taste vectors $\{\theta_i^{(r)}\}_{i=1}^{R}$ from distribution $f(\theta \mid \eta)$ and take an average

$$\hat{\Pr}\left[(c_{1i} = j) \cap (c_{2i} = l) \mid \eta\right] = \frac{1}{R} \sum_{i=1}^{R} \left[\frac{\exp(V_{ij}(\theta_i^{(r)}))}{\sum_{k \in J} \exp(V_{ik}(\theta_i^{(r)}))} \cdot \frac{\exp(V_{il}(\theta_i^{(r)}))}{\sum_{k \in J \setminus \{j\}} \exp(V_{ik}(\theta_i^{(r)})} \right]$$

and

$$\hat{L}(\eta) = \sum_{i=1}^{N} w_i \cdot \log\{\hat{\Pr}\left[(c_{1i} = j) \cap (c_{2i} = l) \mid \eta\right]\}$$

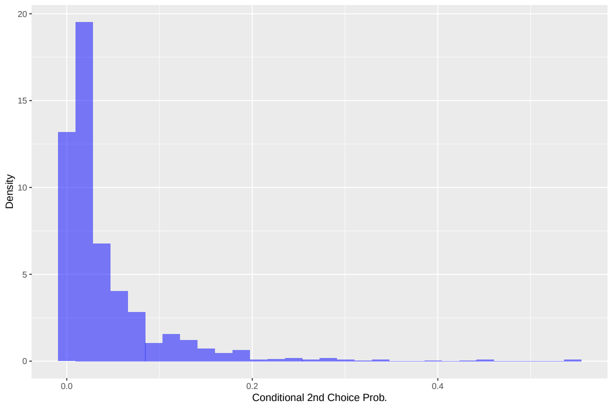
- Maximize this log-likelihood function.

▸ Back

Appendix - 2nd Choice Data

- Check variations in 2nd choice

Figure 42: Histogram of Cond. 2nd Choice Prob. (Pooled)



Appendix - 2nd Choice Data

Table 21: Cond. Choice Prob. for Top 10 vehicles in Pooled Data

1st Choice	2nd Choice (Best Match)	Cond. Choice Prob.	2nd Choice (2nd-Best Match)	Cond. Choice Prob.
HONDA: N-BOX	DAIHATSU: Tanto	0.24	SUZUKI: Spacia	0.18
HONDA: Freed	TOYOTA: Sienta	0.54	HONDA: STEP WGN	0.09
TOYOTA: Sienta	HONDA: Freed	0.28	TOYOTA: Aqua	0.08
TOYOTA: Aqua	TOYOTA: Prius	0.19	TOYOTA: Vitz	0.14
TOYOTA: Voxy	TOYOTA: Noah	0.19	NISSAN: Serena	0.18
NISSAN: Serena	TOYOTA: Voxy	0.24	TOYOTA: Noah	0.14
HONDA: Fit	TOYOTA: Aqua	0.17	NISSAN: Note	0.14
NISSAN: Note	TOYOTA: Aqua	0.19	HONDA: Fit	0.12
NISSAN: Dayz	HONDA: N-BOX	0.15	NISSAN: Note	0.12
HONDA: Vezel	TOYOTA: C-HR	0.24	MAZDA: CX	0.08

Appendix - 2nd Choice Data

- 2nd choice is relatively consistent over time (but may not have large variations across time)

Table 22: Cond. Choice Prob. for Top 10 vehicles (2016 Q3)

1st Choice	2nd Choice (Best Match)	Cond. Choice Prob.	2nd Choice (2nd-Best Match)	Cond. Choice Prob.
TOYOTA: Aqua	HONDA: Fit	0.22	TOYOTA: Prius	0.22
HONDA: N-BOX	DAIHATSU: Tanto	0.24	NISSAN: Dayz	0.12
TOYOTA: Sienta	HONDA: Freed	0.30	TOYOTA: Aqua	0.13
HONDA: Fit	TOYOTA: Aqua	0.26	HONDA: Freed	0.13
HONDA: Freed	TOYOTA: Sienta	0.48	HONDA: Odyssey	0.09
NISSAN: Serena	TOYOTA: Voxy	0.27	NISSA: Elgrand	0.14
TOYOTA: Voxy	TOYOTA: Vellfire	0.25	TOYOTA: Alphard	0.10
HONDA: STEP WGN	NISSAN: Serena	0.26	TOYOTA: Voxy	0.26
NISSAN: Note	TOYOTA: Aqua	0.19	NISSAN: March	0.19
NISSAN: Dayz	SUZUKI: Alto	0.13	TOYOTA: Sienta	0.13

Table 23: Cond. Choice Prob. for Top 10 vehicles (2020 Q1)

1st Choice	2nd Choice (Best Match)	Cond. Choice Prob.	2nd Choice (2nd-Best Match)	Cond. Choice Prob.
HONDA: N-BOX	DAIHATSU: Tanto	0.30	SUZUKI: Spacia	0.17
TOYOTA: Sienta	HONDA: Freed	0.34	TOYOTA: Aqua	0.07
HONDA: Freed	TOYOTA: Sienta	0.56	HONDA: Fit	0.15
HONDA: Fit	TOYOTA: YARIS	0.21	TOYOTA: Aqua	0.14
NISSAN: Dayz	HONDA: N-BOX	0.19	HONDA: N-WGN	0.15
TOYOTA: Voxy	TOYOTA: Noah	0.35	NISSAN: Serena	0.15
NISSAN: Serena	TOYOTA: Voxy	0.30	TOYOTA: Sienta	0.15
TOYOTA: Aqua	TOYOTA: Prius	0.28	TOYOTA: Corolla	0.17
TOYOTA: Noah	TOYOTA: Alphard	0.19	TOYOTA: Voxy	0.19
TOYOTA: RAV4	TOYOTA: C-HR	0.29	TOYOTA: Corolla	0.14

Appendix - Inelasticity of VKT to fuel price

Table 24: Elasticity between VKT and fuel price

	VMT		
	OLS	OLS	IV
log(Gasoline Price)	-0.002	-0.070	-0.018
	(0.114)	(0.069)	(0.069)
Constant	6.291***	7.370***	6.356***
	(0.561)	(1.283)	(0.434)
Olea Montiel-Pflueger F-Statistics			10896.540
Vehicle Characteristics		○	○
Demographic Variables		○	○
City FE		○	○
City Level Economic Conditions		○	○
Month FE		○	○
Num. obs.	37,405	37,405	37,405

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Instrumental variable is the six-months-ahead gasoline price.

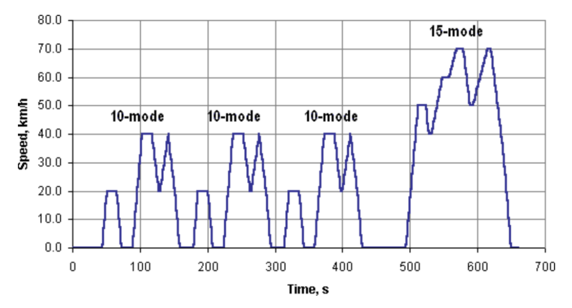
Appendix - Test-Cycle

Figure 43: Example of Fuel Consumption Test

Note: From Japan Automobile Research Institute

Appendix - Test-Cycle

Figure 44: Example of Test Cycle

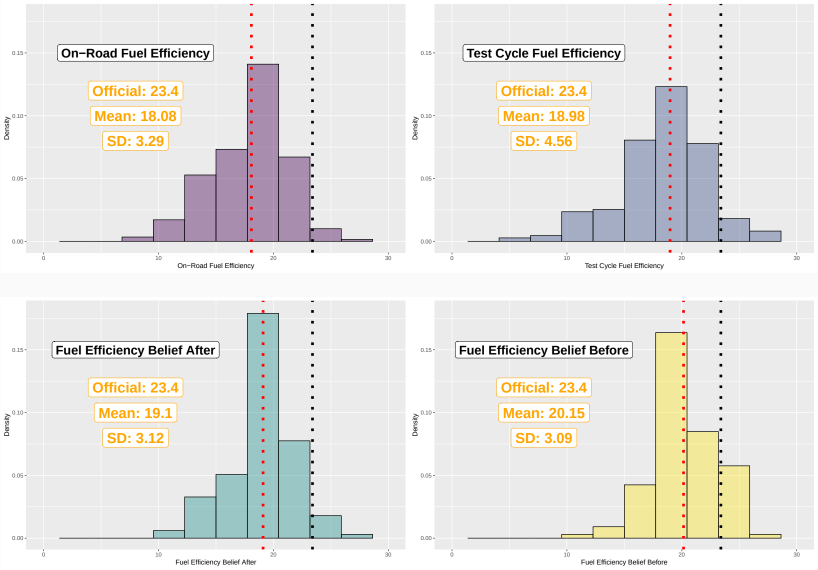


Note: From DieselNet

Back

Appendix - Test-cycle uncertainty

Figure 45: Example of Each Fuel Efficiency Measure (Case of TOYOTA Camry)



Appendix - Use Only Macro Data

Consistent with Grigolon et al. (2018).

Table 25: Macro BLP

	Logit (IV)			Logit (GMM)		GRV
	IV (1)	IV (2)	IV (3)	GMM (1)	GMM (2)	
MSRP	−9.8158	−7.6779	−28.9956	−7.6779	−12.4367	−4.52
	(0.4564)	(0.3761)	(0.7799)	(0.3762)	(4.0079)	(0.19)
Fuel Cost	−8.8437	−7.1389	−10.5411	−7.1389	−7.5392	
	(0.4089)	(0.2706)	(1.2453)	(0.2706)	(1.2204)	
Weight	16.4969	14.3183	36.9011	14.3183	109.0206	
	(0.6144)	(0.6212)	(6.5174)	(0.6214)	(5.0286)	
HP	2.2688	1.2827	−1.1766	1.2827	8.6622	2.28
	(0.4053)	(0.1899)	(1.0555)	(0.1899)	(2.2056)	(0.14)
Future Valuation	0.9009	0.9298	0.3635	0.9298	0.6062	0.89
	(0.0688)	(0.0624)	(0.0457)	(0.0623)	(0.1540)	(0.06)
BLP IV	∘	∘	∘	∘	∘	∘
Cost-Shifter IV	∘	∘	∘	∘	∘	∘
ABR IV	∘	∘	∘	∘	∘	
Cragg-Donald F-Stat.	955.744	592.546	493.895	592.215	171.022	
Time FE	∘	∘	∘	∘	∘	∘
Engine Type FE	∘					
Manufacturer FE		∘		∘		
Model FE			∘		∘	∘
Num. obs.	16,146	16,146	16,146	16,146	7,380	82,151

Notes: MRSP (Manufacturer Suggested Retail Price). ABR-type IV means attributed-based regulation IV.