

Driving Behavior and the Price of Gasoline: Evidence from Fueling-Level Micro Data

Christopher R. Knittel*

Shinsuke Tanaka†

Supplementary Materials

*Massachusetts Institute of Technology and NBER. Email: knittel@mit.edu.

†Tufts University. Email: Shinsuke.Tanaka@tufts.edu.

Contents

A Additional Information on Data	2
B Additional Information on Price Elasticities of Demand for Gasoline	9
C Additional Information on Learning Effect	13
D Additional Information on Heterogenous Price Elasticities	18

A. Additional Information on Data

The primary data we use for the analysis were obtained and can be purchased from the IID, Inc. Group, the private company that operates the mobile phone application, called *e-nenpi* (*nenpi* means fuel economy). The application is free to download and use for users. Figure A.1 illustrates the sample screenshots of the application. Using this application, users report the amount of gasoline purchased, odometer values, and gasoline prices paid at every time they refuel. The information can be uploaded by simply taking the photographs of the receipt and the odometer to minimize the typing errors. Using the date and time of the gasoline purchase, we computed the daily (e.g., 24 hours) gasoline consumed and distance traveled between the two consecutive refuels. Further, the actual fuel-economy figures were computed by dividing the distance traveled by the gasoline consumption.

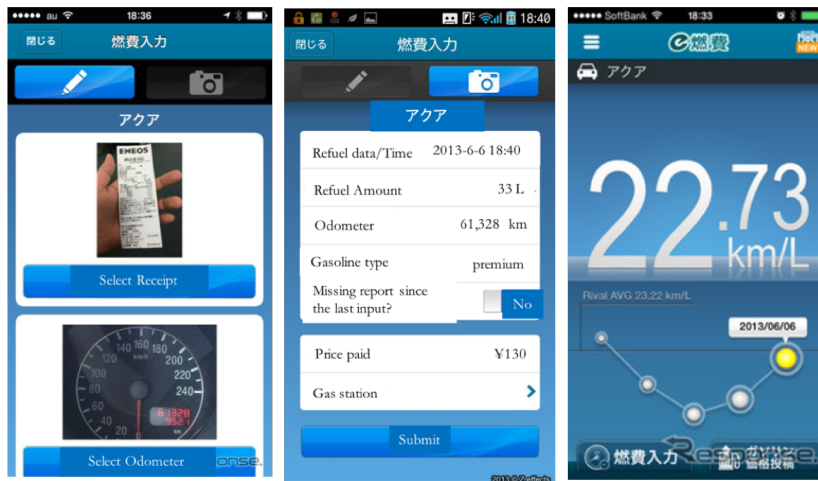
The original data we obtained included 5,884,179 observations at the refueling level. The data used for the analysis are limited to gasoline-powered passenger vehicles and minicars by domestic automakers. Minicars (called “*kei-cars*”) constitute one of the primary classifications of vehicles in Japan. They are tiny vehicles whose displacement is 660 cubic centimeter (cc) or lower and are popular because roads are typically narrow, and automobile-related taxes are substantially lower than passenger vehicles. Passenger cars include light duty vehicles whose displacement is under 2000 cc for gasoline-powered vehicles and regular vehicles but excludes motorcycles, buses, and trucks. According to the report by the Japan Automobile Dealers Association, passenger cars and minicars together account for close to 84% of all new vehicles sold (slightly more than 5 million vehicles) within the domestic economy, 30 percentage points of which are minicars. This process effectively removes trucks (0.42% of the original sample), hybrid cars (3.13%), foreign automakers (8.12%), and other fuel types (1.84% of diesel) as shown in Table A.1.

As a guard against extreme values and potential typing errors, we removed outliers in terms of the bottom and top one percentile of vehicle distance traveled and the actual fuel-economy. Because daily vehicle distance traveled was highly skewed to the right tail, the main analysis is limited to observations where travel is less than 100 kilometers (km) per day to focus on a range of daily lives. As a robustness check, we expand observations with up to 250 km a day (adding about 7% more observations). Lastly, because our model includes the driver-vehicle fixed effects, we removed single observations at this level. Ultimately, the sample for the analysis came down to 4,088,789 observations.

The national daily average gasoline prices are also collected by and obtained from the IID, Inc. Group. Their data start from 2010, and thus for years before 2009, we use the weekly gasoline retail price at the prefecture level reported by the Institute of Energy Economics,

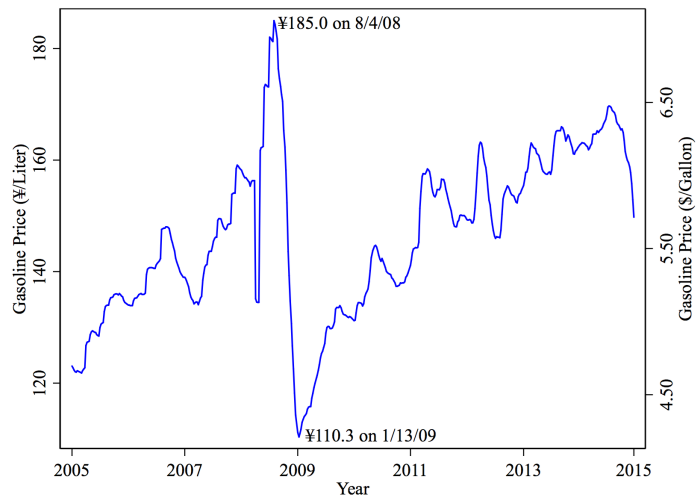
Japan. We computed the daily fuel prices for non-reported days by taking the arithmetic averages of the two most recently reported values. All fuel prices are converted into the January 2010 yen using the consumer price index.

Figure A.1: Sample Screenshots of the Application



Notes: These pictures were obtained from different sources only to illustrate how the application can be operated and do not necessarily reflect actual reports. The author translated the language originally in Japanese to English.

Figure A.2: Gasoline Price Trend



Notes: This figure plots the weekly average price of the regular and high octane gasoline between 2005 and 2014. The units are Japanese yen per liter in January 2010 value. Average exchange rate during this period is US\$ = ¥99.2.

Table A.1: Share of the original sample

Category	Percentage
Truck	0.42
Hybrid	3.13
Foreign maker	8.12
Fuel type	
Regular	66.84
High octane	31.31
Electricity	0.00
LP Gas	0.00
Diesel	1.84

Notes: The original sample includes 5,884,179 observations at the driver-vehicle level.

Table A.2: Summary Statistics

	Mean	Std.	N	US standard
<i>Panel A: Individual report level</i>				
Gasoline consumption per day (L/day)	3.535	2.034	4,088,789	0.93 gallon/day
Vehicle distance traveled per day (km/day)	36.70	20.66	4,088,789	22.81 mile/day
Actual fuel economy (km/L)	10.85	3.538	4,088,789	25.51 MPG
Gasoline price paid (¥/L)	136.55	15.24	4,088,789	\$5.21Gallon
# of days b/w refueling	14.12	25.54	4,088,789	
Odometer (km)	66,359.8	48,788.9	4,088,789	41,234.2 mile
Prefecture (%)				
Saitama	5.30		4,088,789	
Kanagawa	5.23		4,088,789	
Aichi	4.81		4,088,789	
Tokyo	4.35		4,088,789	
Chiba	4.04		4,088,789	
Osaka	3.15		4,088,789	
Report year (%)				
2005	8.69		4,088,789	
2006	10.13		4,088,789	
2007	10.27		4,088,789	
2008	11.12		4,088,789	
2009	11.00		4,088,789	
2010	11.87		4,088,789	
2011	10.51		4,088,789	
2012	9.59		4,088,789	
2013	8.27		4,088,789	
2014	8.54		4,088,789	
<i>Panel B: Driver-vehicle level</i>				
# of reports	48.59	49.43	90,411	
Initial year	2008.7	2.906	90,411	
<i>Panel C: Driver level</i>				
Total number of drivers			71,263	
Male	0.889	0.314	33,804	
Age	35.514	8.277	33,428	

Summary Statistics cont.

	Mean	Std.	N	US standard
<i>Panel D: Vehicle level</i>				
Manufacturing year	1,999.7	7.130	3,932	
Vehicle price (¥10,000)	191.0	101.5	3,663	\$19,249
Dummy for regular gasoline (vs. highoctane)	0.760	0.427	3,932	
Dummy for passenger vehicle (vs. minicars)	0.695	0.461	3,932	
Seating capacity	4.838	1.130	3,830	
Dummy for automatic transmission	0.710	0.454	3,932	
Vehicle weight (kg)	1192.9	336.9	3,783	2,630 lb
Displacement (cc)	1631.4	851.2	3,931	
Official fuel economy (km/L)	14.71	4.966	3,707	34.6 MPG
Automaker (%)				
Toyota	21.44		3,932	
Nissan	16.91		3,932	
Suzuki	11.72		3,932	
Honda	11.06		3,932	
Mitsubishi	10.40		3,932	
Subaru	9.16		3,932	
Daihatsu	9.21		3,932	
Mazda	9.00		3,932	
Isuzu	0.56		3,932	
Lexus	0.53		3,932	

Notes: Prefectures are shown only for the six largest shares, and prefecture is unknown for about 30.9% of the sample.

B. Additional Information on Price Elasticities of Demand for Gasoline

Table B.1: The Price Elasticity of Vehicle Distance Traveled

Dep. Var.	ln(VKT)				
	(1)	(2)	(3)	(4)	(5)
ln(Price)	0.179*** (0.00643)	-0.0408*** (0.00839)	-0.0551*** (0.0136)	-0.302*** (0.0367)	-0.199*** (0.0242)
Model	OLS	IV	IV	IV	Reduced
Driver-vehicle FE	Y	Y	Y	Y	Y
Time FE	Year + month	Year + month	Year × quarter	Year × month	Year × month

Notes: The outcome variables are the log of vehicle-kilometer traveled (VKT) per day. All models except Column (1) and (5) are estimated by the instrumental variable approach, whose first stage results are presented below. All specifications include the driver-vehicle fixed effects and variant time fixed effects specified in each column. The number of observations is 4,088,789. Standard errors clustered at the driver-vehicle level are reported in the parentheses.

*** $p < 0.01$

Table B.2: The Price Elasticity of Actual Fuel Economy

Dep. Var.	ln(KPL)				
	(1)	(2)	(3)	(4)	(5)
ln(Price)	0.0399*** (0.00164)	0.0336*** (0.00209)	0.158*** (0.00295)	0.0695*** (0.00735)	0.0459*** (0.00485)
Model	OLS	IV	IV	IV	Reduced
Driver-vehicle FE	Y	Y	Y	Y	Y
Time FE	Year + month	Year + month	Year × quarter	Year × month	Year × month

Notes: The outcome variables are the log of real-world fuel economy (in km/liter) (KPL) obtained by dividing gasoline consumption by vehicle distance traveled. All models except Column (1) and (5) are estimated by the instrumental variable approach, whose first stage results are presented below. All specifications include the driver-vehicle fixed effects and variant time fixed effects specified in each column. The number of observations is 4,088,789. Standard errors clustered at the driver-vehicle level are reported in the parentheses.

*** $p < 0.01$

Table B.3: The First Stage Results of the IV Estimates

Dep. Var.	ln(Price paid)		
	(1)	(2)	(3)
ln(Price)	0.965*** (0.00109)	0.888*** (0.00149)	0.660*** (0.00237)
Driver-vehicle FE	Y	Y	Y
Time FE	Year + month	Year × quarter	Year × month
F -stat	779,684	353,091	77,504

Notes: This table presents the first stage results from the IV estimates with variant time fixed effects as specified in each column. The F -statistics of the excluded instrument are also reported. The number of observations is 4,088,789.

*** $p < 0.01$

Table B.4: Robustness: Including Long Distance traveled

Dep. Var.	ln(Gasoline)					ln(VDT)	ln(KPL)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ln(Price)	0.264*** (0.00676)	-0.0482*** (0.00846)	-0.123*** (0.0135)	-0.225*** (0.0358)	-0.148*** (0.0235)	-0.142*** (0.0380)	0.0832*** (0.00738)
Model	OLS	IV	IV	IV	Reduced	IV	IV
Driver-vehicle FE	Y	Y	Y	Y	Y	Y	Y
Time FE	Year + month	Year + month	Year × quarter	Year × month	Year × month	Year × month	Year × month

Notes: This table presents the analogous results to Table 2 in the main text except that the sample includes observations whose vehicle distance traveled is up to 250 km per day. The number of observations is 4,304,452.

*** $p < 0.01$

Table B.5: Robustness: Alternative Specifications

Dep. Var.	Gasoline					VKT	KPL
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<i>Panel A: Level-level</i>							
Price	0.00158*** (0.000131)	-0.00289*** (0.000177)	-0.00569*** (0.000266)	-0.0160*** (0.00144)	-0.00698*** (0.000630)	-0.106*** (0.0149)	0.00769*** (0.00126)
<i>Panel B: Semi-log</i>							
Price	0.000770*** (0.0000426)	-0.000560*** (0.0000515)	-0.00132*** (0.0000791)	-0.00399*** (0.000430)	-0.00174*** (0.000188)	-0.00334*** (0.000477)	0.000650*** (0.000129)
Model	OLS	IV	IV	IV	Reduced	IV	IV
Driver-vehicle FE	Y	Y	Y	Y	Y	Y	Y
Time FE	Year + month	Year + month	Year × quarter	Year × month	Year × month	Year × month	Year × month

Notes: This table presents the results of the robustness check in Table 2 to alternative specifications: the level-level specification in Panel A and the semi-log specification in Panel B. The dependent variables are in levels in Panel A and in logs in Panel B. The number of observations is 4,088,789.

*** $p < 0.01$

Table B.6: Robustness: Using Average Daily Prices between Refuels

Dep. Var.	ln(Gasoline)			ln(VKT)	ln(KPL)
	(1)	(2)	(3)	(4)	(5)
ln(Price)	-0.117*** (0.00704)	-0.288*** (0.0108)	-0.684*** (0.0317)	-0.642*** (0.0337)	0.0420*** (0.00717)
Model	Reduced	Reduced	Reduced	Reduced	Reduced
Driver-vehicle FE	Y	Y	Y	Y	Y
Time FE	Year + month	Year × quarter	Year × month	Year × month	Year × month

Notes: The outcome variables are the logs of gasoline consumption (in liter) per day (i.e. 24 hours) in Columns (1)–(3), vehicle-kilometer traveled (VKT) per day in Column (4), and the real-world fuel economy (in km/liter) (KPL) obtained by dividing gasoline consumption by vehicle-kilometer traveled in Column (5). The main independent variable reflects the average fuel price of all days since the last refuel until this time. All specifications include the driver-vehicle fixed effects and variant time fixed effects as specified in each column. The number of observations is 4,088,789. Standard errors clustered at the driver-vehicle level are reported in the parentheses.

*** $p < 0.01$

Table B.7: The Price Elasticity of Vehicle Distance Traveled Using Configuration FE

Dep. Var.	ln(VKT)				
	(1)	(2)	(3)	(4)	(5)
ln(Price)	0.146*** (0.0135)	0.00252 (0.0137)	-0.0187 (0.0237)	-0.246*** (0.0644)	-0.133*** (0.0355)
Model	OLS	IV	IV	IV	Reduced
Configuration FE	Y	Y	Y	Y	Y
Time FE	Year + month	Year + month	Year × quarter	Year × month	Year × month

Notes: This table presents the analogous results to Table B.1 with including the vehicle configuration fixed effects in place of driver-vehicle fixed effects.

*** $p < 0.01$

C. Additional Information on Learning Effect

Table C.1: Correlation Coefficients

	$Price_t$	$Price_{t-1}$	$Price_{t-2}$	$Price_{t-3}$	$Price_{t-4}$	$Price_{t-5}$
$Price_t$	1.000					
$Price_{t-1}$	0.959	1.000				
$Price_{t-2}$	0.899	0.960	1.000			
$Price_{t-3}$	0.832	0.901	0.961	1.000		
$Price_{t-4}$	0.765	0.836	0.903	0.961	1.000	
$Price_{t-5}$	0.699	0.769	0.837	0.902	0.959	1.000

Notes: This table presents the correlation coefficients of the six most recent prices, of which the most recent one at time t is the price paid for the current trip.

Table C.2: Learning Effect of Price on Driving Behavior

Dep. var.	ln(KPL) (1)	ln(VKT) (2)	ln(GPD) (3)
<i>Panel A: Up to 50th obs.</i>			
β	0.0575*** (0.00310)	-0.204*** (0.0148)	-0.262*** (0.0142)
λ	0.747*** (0.0187)	0.814*** (0.0156)	0.810*** (0.0115)
Test: $\lambda = 1$	$p=0.000$	$p=0.000$	$p=0.000$
<i>Panel B: Up to 100th obs.</i>			
β	0.0621*** (0.00293)	-0.311*** (0.0136)	-0.373*** (0.0129)
λ	0.778*** (0.0135)	0.833*** (0.00846)	0.830*** (0.00659)
Test: $\lambda = 1$	$p=0.000$	$p=0.000$	$p=0.000$

Notes: The table reports the estimated β and λ based on Equation (6) for the dependent variable specified at the column head using up to 50th refuel from the initial one in Panel A and up to 100th one in Panel B. The test statistics for the null hypothesis: $\lambda = 1$ are also reported. The numbers of observations are 2,705,006 (66% of total observations) and 3,565,208 (87%), respectively.

*** $p < 0.01$

Table C.3: Learning Effect of Price on Driving Behavior

	(1) 50th	(2) 100th	(3) 200th
β	0.0575*** (0.00310)	0.0621*** (0.00293)	0.0679*** (0.00308)
λ	0.747*** (0.0187)	0.778*** (0.0135)	0.793*** (0.0108)
Test: $\lambda = 1$	$p=0.000$	$p=0.000$	$p=0.000$
N	2,705,006	3,565,208	4,024,897
Share of N	0.66	0.87	0.98

Notes: The table reports the estimated β and λ for the log of actual fuel economy based on Equation (6) in the main text for three subsamples: up to the first 50th observations for each driver in column (1), 100th in column (2), and 200th in column (3). The test statistics for the null hypothesis: $\lambda = 1$ are also reported. Each share of the observations to the total observations is reported at the bottom.

*** $p < 0.01$

Table C.4: Robustness: Distributed Lag Model

Variable	(1)	(2)	(3)	(4)	(5)	(6)
$\ln(\text{Price}_t)$	0.0459*** (0.00485)	0.0271*** (0.00508)	0.0217*** (0.00518)	0.0215*** (0.00524)	0.0208*** (0.00528)	0.0201*** (0.00532)
$\ln(\text{Price}_{t-1})$		0.0198*** (0.00356)	0.0132*** (0.00417)	0.00784* (0.00427)	0.00759* (0.00431)	0.00651 (0.00435)
$\ln(\text{Price}_{t-2})$			0.0175*** (0.00331)	0.0220*** (0.00409)	0.0174*** (0.00419)	0.0178*** (0.00424)
$\ln(\text{Price}_{t-3})$				0.00141 (0.00318)	0.0129*** (0.00403)	0.00881** (0.00413)
$\ln(\text{Price}_{t-4})$					-0.00791** (0.00308)	-0.000571 (0.00402)
$\ln(\text{Price}_{t-5})$						-0.00415 (0.00304)
N	4,088,789	3,998,378	3,907,967	3,818,660	3,730,273	3,642,848

Notes: This table reports the coefficients of the six most recent prices paid based on the distributed lag model as specified by Equation (3) in the main text. The regressions are based on the reduced-form, where the fuel price on the day of purchase is used as the instrument. The dependent variable is the log of actual fuel economy. The regressions include the year-by-month fixed effects and the driver-vehicle fixed effects. The standard errors clustered at the driver-vehicle level are reported in the parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.5: Exploring the Potential Mechanisms

Dep. var	ln(GPD) (1)	ln(VKT) (2)	ln(KPL) (3)
$\ln(Price_t)$	-0.165*** (0.0285)	-0.128*** (0.0290)	0.0373*** (0.00389)
$\ln(Price_{t+1})$	-0.184*** (0.0323)	-0.164*** (0.0344)	0.0197*** (0.00744)

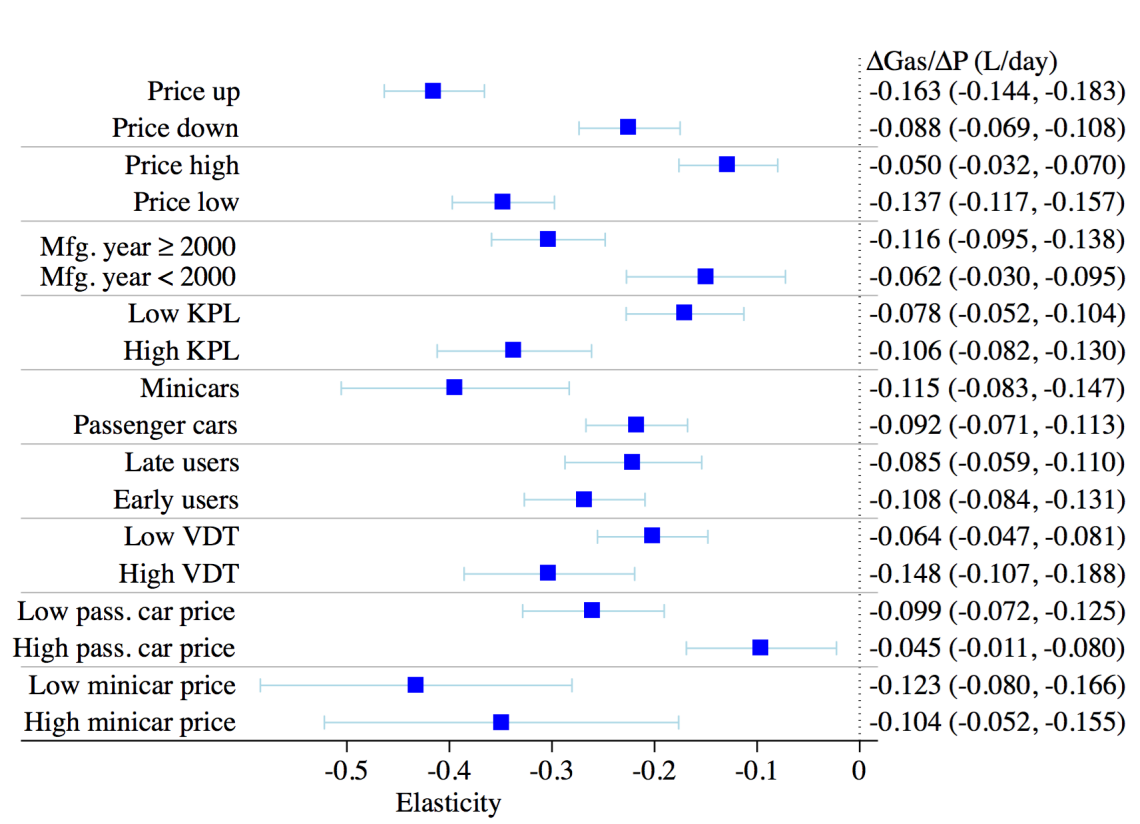
Notes: The dependent variables are the logs of gasoline consumption per day (L/day) in Column (1), vehicle-kilometers traveled per day (km/day) in (2), and the actual fuel economy (km/L) in (3). The independent variables are the logs of prices paid for the current trip ($Price_t$) and for the next trip ($Price_{t+1}$). All specifications include the driver-vehicle fixed effects, and the standard errors clustered at the customer-vehicle level are reported in the parentheses.

*** $p < 0.01$

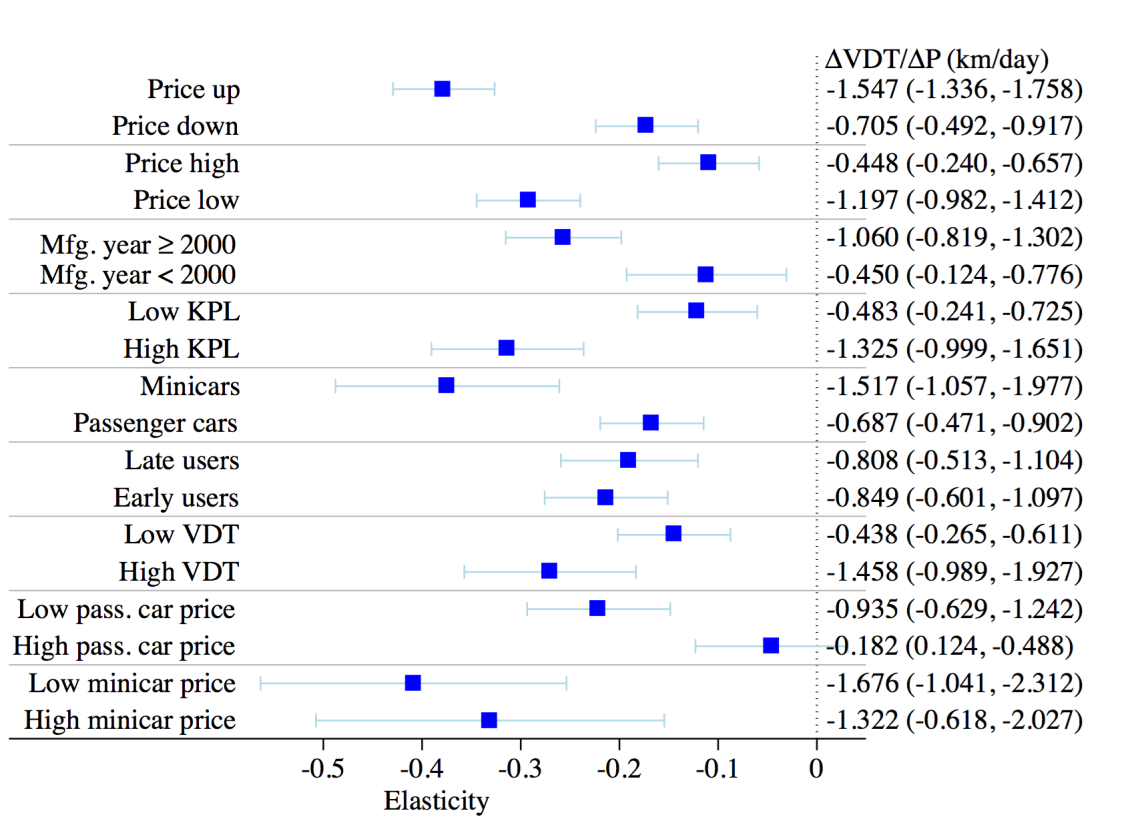
D. Additional Information on Heterogenous Price Elasticities

Figure D.1: Asymmetric Price Elasticities of Demand for Gasoline

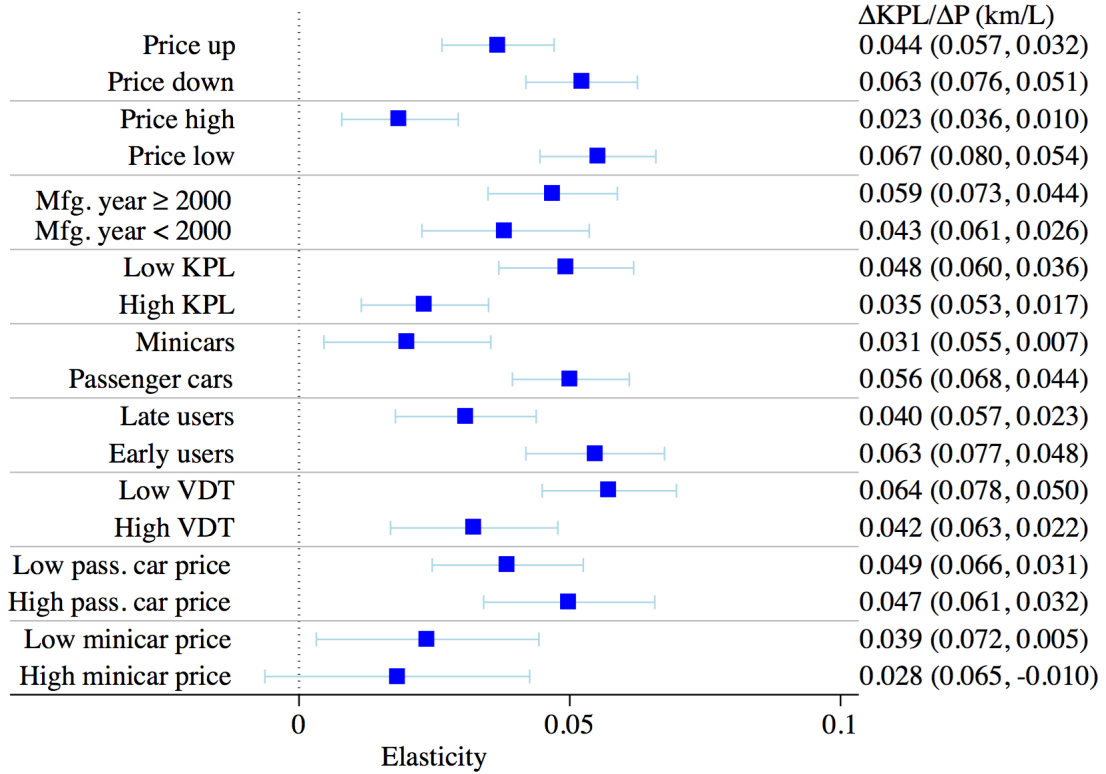
Panel A: Gasoline Consumption



Panel B: Vehicle Kilometers Traveled



Panel C: Actual Fuel Economy



Notes: Each panel presents the estimated price elasticities for the respective dependent variable. The square dots represent the coefficients, and the lines indicate the associated 95% confidence interval. The asymmetric effects for fuel prices going up/down and fuel prices high/low are estimated from the single regressions, whereas other specifications are separately estimated for the relevant subgroup. The effect size indicated on the right side of the figure represents the changes and the associated 95% confidence interval of gasoline consumption per day in liter in response to a one standard deviation increase in gasoline price (¥15.24).

Table D.1: Heterogeneities in Price Elasticities

Dep. var. Subsample	ln(Gasoline)		ln(VKT)		ln(KPL)	
	Reduced	IV	Reduced	IV	Reduced	IV
Price going up	-0.415*** (0.0249)	-0.593*** (0.0363)	-0.378*** (0.0262)	-0.538*** (0.0382)	0.0367*** (0.00528)	0.0547*** (0.00770)
Price going down	-0.224*** (0.0251)	-0.351*** (0.0361)	-0.172*** (0.0264)	-0.278*** (0.0380)	0.0522*** (0.00525)	0.0731*** (0.00756)
Price high	-0.128*** (0.0246)	-0.00162 (0.0323)	-0.109*** (0.0259)	-0.00444 (0.0342)	0.0186*** (0.00547)	-0.00281 (0.00779)
Price low	-0.347*** (0.0254)	-0.851*** (0.0560)	-0.292*** (0.0268)	-0.714*** (0.0590)	0.0552*** (0.00546)	0.136*** (0.0124)
Mfg. year \geq 2000	-0.304*** (0.0283)	-0.454*** (0.0422)	-0.257*** (0.0299)	-0.384*** (0.0447)	0.0469*** (0.00610)	0.0702*** (0.00914)
Mfg. year $<$ 2000	-0.150*** (0.0395)	-0.227*** (0.0599)	-0.112*** (0.0413)	-0.170*** (0.0627)	0.0381*** (0.00790)	0.0579*** (0.0120)
Low KPL	-0.170*** (0.0293)	-0.272*** (0.0468)	-0.121*** (0.0309)	-0.193*** (0.0494)	0.0493*** (0.00634)	0.0789*** (0.0102)
High KPL	-0.337*** (0.0253)	-0.401*** (0.0458)	-0.313*** (0.0267)	-0.373*** (0.0469)	0.0232*** (0.00551)	0.0277*** (0.00714)
Minicars	-0.394*** (0.0567)	-0.456*** (0.0656)	-0.374*** (0.0579)	-0.433*** (0.0669)	0.0200** (0.00785)	0.0231** (0.00911)
Passenger vehicles	-0.217*** (0.0253)	-0.340*** (0.0395)	-0.167*** (0.0267)	-0.261*** (0.0417)	0.0502*** (0.00551)	0.0785*** (0.00863)
Late users	-0.221*** (0.0340)	-0.301*** (0.0464)	-0.190*** (0.0354)	-0.259*** (0.0483)	0.0308*** (0.00665)	0.0420*** (0.00909)
Early users	-0.268*** (0.0300)	-0.429*** (0.0480)	-0.213*** (0.0318)	-0.342*** (0.0508)	0.0547*** (0.00651)	0.0875*** (0.0104)
Low VKT	-0.202*** (0.0275)	-0.308*** (0.0418)	-0.145*** (0.0291)	-0.220*** (0.0443)	0.0573*** (0.00633)	0.0873*** (0.00966)
High VKT	-0.303*** (0.0424)	-0.468*** (0.0654)	-0.270*** (0.0444)	-0.418*** (0.0684)	0.0323*** (0.00788)	0.0500*** (0.0122)
Low passenger price	-0.260*** (0.0352)	-0.366*** (0.0496)	-0.221*** (0.0370)	-0.311*** (0.0521)	0.0386*** (0.00712)	0.0543*** (0.0100)
High passenger price	-0.0958** (0.0373)	-0.154** (0.0599)	-0.0459 (0.0394)	-0.0737 (0.0632)	0.0499*** (0.00807)	0.0801*** (0.0130)
Low minicar price	-0.432*** (0.0775)	-0.494*** (0.0885)	-0.409*** (0.0790)	-0.467*** (0.0902)	0.0238** (0.0105)	0.0271** (0.0120)
High minicar price	-0.349*** (0.0881)	-0.411*** (0.103)	-0.331*** (0.0900)	-0.389*** (0.106)	0.0182 (0.0125)	0.0214 (0.0147)

Notes: This table presents the estimated coefficients and standard errors of price elasticities for gasoline consumption (L/day), vehicle-kilometers traveled (km/day), and actual fuel economy (km/L). For each dependent variable, both the reduced form and IV estimates are presented. In particular, they are elasticities when a price goes up or down, a price is above or below the mean price paid, among vehicles whose manufacturing years are before or after 2000, vehicles with greater or lower than the mean official fuel economy level, minicars vs. passenger vehicles, users who started using the application before or after the average, and vehicle prices above or below the mean prices separately for passenger cars and minicars. The asymmetric effects for fuel prices going up/down and fuel prices high/low are estimated from the single regressions, whereas other specifications are estimated from separate regressions for the respective subsample.

** $p < 0.05$, *** $p < 0.01$