Online appendix for

THE RESPONSE OF CONSUMER SPENDING TO CHANGES IN GASOLINE PRICES*

Michael Gelman[a], Yuriy Gorodnichenko[b,c], Shachar Kariv[b], Dmitri Koustas[b],

Matthew D. Shapiro[c,d], Dan Silverman[c,e], and Steven Tadelis[b,c]

Appendix A: Construction of spending in the app data

This appendix discusses the data and provides details on how we prepare the data for analyses.

We received anonymized data directly from the personal financial management service provider (app). The process by which the company acquires the data can differ across users, account providers (e.g., Bank of America, Wells Fargo) and time. For some account providers, the data are scraped from the website of an account provider, and in other cases a direct feed is received from the account provider. All account numbers and other personal identifying information is removed by the app company before we receive the data. Otherwise, we receive the data exactly as it is received by the app. The table below summarizes the key variables in the data that are used in our analysis:

User_id	-	Anonymous identifier constructed by the personal financial management service
Posted_date	-	Date a transaction was recorded
Account_Provider_Id	-	An identifier for a specific account provider (e.g. Bank of America)
Account_Type	-	An indicator for whether an account is a checking account, savings account, credit card, or other account.
Transaction_Amount	-	The amount of the transaction
Is_Credit	-	Whether the transaction was a credit or a debit
Transaction_Description	-	A string variable describing the transaction.

Our cleaning processes proceeds in steps outlined below:

I. Remove likely duplicates +/- 3 days

Because the data may include pending transactions, a given spending may show up multiple times in different transactions. For instance, if a transaction was pending on one day, and posted the next day, we could see a duplicate recording of the same transaction in the data, which would not reflect actual spending.

Some account providers indicate whether a transaction is pending or posted, and we first remove all transactions that are flagged as pending, or contain the word "pending" in the transaction string. Since many account providers do not indicate whether a transaction is pending, and since this information also varies across time, we deal with this problem by removing transactions that are duplicates on the dimensions of {User_Id, Account_Provider_Id, Account_Type, Is_Credit, Transaction_Amount, Transaction_Description} over a 3 day window. This removes approximately 5% of transactions. Some of these transactions could be nonduplicates (for instance, if someone buys the exact same item every day), and so these transactions will also be removed by this procedure. Using the data with likely duplicates removed, we next proceed to calculate total spending and total income, which we aggregate to the weekly level.

II. Construct variables used in analysis and aggregate to weekly level

The transactions contain every single inflow and outflow from a household's account, some of which are not "consumption." Two problematic types of transactions are transfers across accounts and credit card payments. In most cases, transfers across accounts can be identified from the transaction strings, since they are commonly flagged as "transfer," "xfer," "tfr," "xfr" or "trnsfr." We remove all transactions with these words appearing in their description.

Credit card payments reflect lagged spending that we have already included in our measure of total spending, since we can see the individual purchases that make up the credit card payment on the credit card. Therefore, we wish to identify and remove these payments. We identify credit card payments as debits appearing on a non-credit-card account that also appears as a credit to a credit card, and remove these.

We also remove the largest transaction greater than $1,000 in a weekly window, since these transactions appear to be predominantly credit card payments and transfers missed by our procedure. As a caveat, this likely also removes mortgage payments (committed spending), extremely large durables purchases (such as a down payment on a car, although we would still see car payments), and payments on tax liabilities. To summarize, our measure of "total spending" used in this paper is defined as:

{Total spending} = {Total Account Debits} – {Flagged Duplicates} – {Transfers} – {Credit Card Payments} – {Largest Transaction > $1,000 (if any)}.

Finally, we address the issue of accounts that become unlinked from the app or are not properly synchronising. If an account goes out of sync for a period of up to two weeks, the app will generally be able to backfill these transactions. Longer periods will result in missing spending. Unfortunately, there is no indicator in the data when an account is not syncing. We identify non-syncing credit cards as cards that carry an account balance for longer than a month, but have no interest charges or payments. To ensure the quality of our spending data, we drop users in the weeks where credit cards that ever amounted to 10 percent or more of their overall weekly spending are flagged as nonsyncing based on the above criteria.

Appendix B: Machine Learning Classification of Transactions

As discussed in Appendix 1, the data we receive contain raw transaction strings. These transaction strings differ across account providers in their context. We wish to identify spending that comes from gasoline. Identifying to which of a set of categories an observation belongs, based on information in the transaction descriptions, is a classic "classification" problem in machine learning.

We seek a simple machine learning (ML) model to identify gasoline spending in the data. For this to work, we require a "training" set of data containing observations whose category membership is known. Fortunately, two account providers in our data categorize the transactions into merchant category codes (MCCs) directly in the transaction strings. These two cards represent about 3% of all transactions. As discussed in the text, it is virtually impossible to separate out our main MCC of interest, 5541, "Automated Fuel Dispensers" from MCC code 5542, "Service Stations," which in practice covers gasoline stations with convenience stores.[1] Because distinguishing gasoline purchases classified as 5542 or 5541 is nearly impossible with the information in transaction descriptions,[2] we group transactions with these two codes together.

Before proceeding with the details of the machine learning model, it is useful to discuss an alternative approach that identifies all gasoline stations in the data through string matching techniques. To see why this is infeasible, consider that the 100 most popular gasoline station strings cover approximately 50% of the total market share in the transactions where we know the MCC codes. Scaling up is costly: to get 90 percent of the market share, we would need to search for over 30,000 strings (Appendix Figure 1). Moreover, since other spending can often have similar transaction descriptions, it is hard to know what strings minimize noise while maximizing predictive power. The machine learning algorithm thus helps discipline the approach of what transaction strings contain the most useful information. The machine learning procedure proceeds in 3 steps: training, testing, and application.

Machine learning requires both a "training" data set—data actually used to fit a classification model-—and a "testing" data set to evaluate the out of sample performance of the model. In the training step, we build a prediction model using data with the MCC codes (i.e. data

[1] To be clear, "Service Stations" do not include services such as auto repairs, motor oil change, etc.
[2] E.g., a transaction string with word "Chevron" or "Exxon" could be classified as either MCC 5541 or MCC 5542.

where classification is known). We use the larger of the two account providers as the training data set, and test the performance of the model on the smaller account. We explicitly set aside the second card as the training data set because transaction strings, which we will feed into the model to classify the data, can differ across account providers. Therefore, if we train on data from the two accounts, we may fit our two cards extremely well, but we may have a poor "out of sample" fit of our model.

The classification algorithm we use is known as a random forest classifier, which fits a number of separate decision trees. A decision tree is a series of classification rules that ultimately lead to a classification of a purchase as gasoline or not. The rules, determined by the algorithm, minimizes the decrease in accuracy when a particular model "feature" is removed. The features we use are the transaction values and individual words in the transaction strings (this approach is known as "bag of words"), after some basic string cleaning. We limit the number of features to 20,000 words, and transaction amounts rounded to the nearest 50 cents. An example decision tree is shown in Appendix Figure 2.

In this example, the most important single word is "oil." If a transaction string contains the word oil, the classification rule is to move to the right, otherwise the rule is to move to the left. If the string does not contain the word oil, the next most important single word is "exxonmobil." The tree keeps going until all the data are classified.

Whether a transaction is classified as gasoline spending or not is simply the majority vote over a number of decision trees. This is known as a "white box algorithm" because the model determines optimal decision rules that we can see. We use prebuilt packages from the python machine learning toolkit.[3]

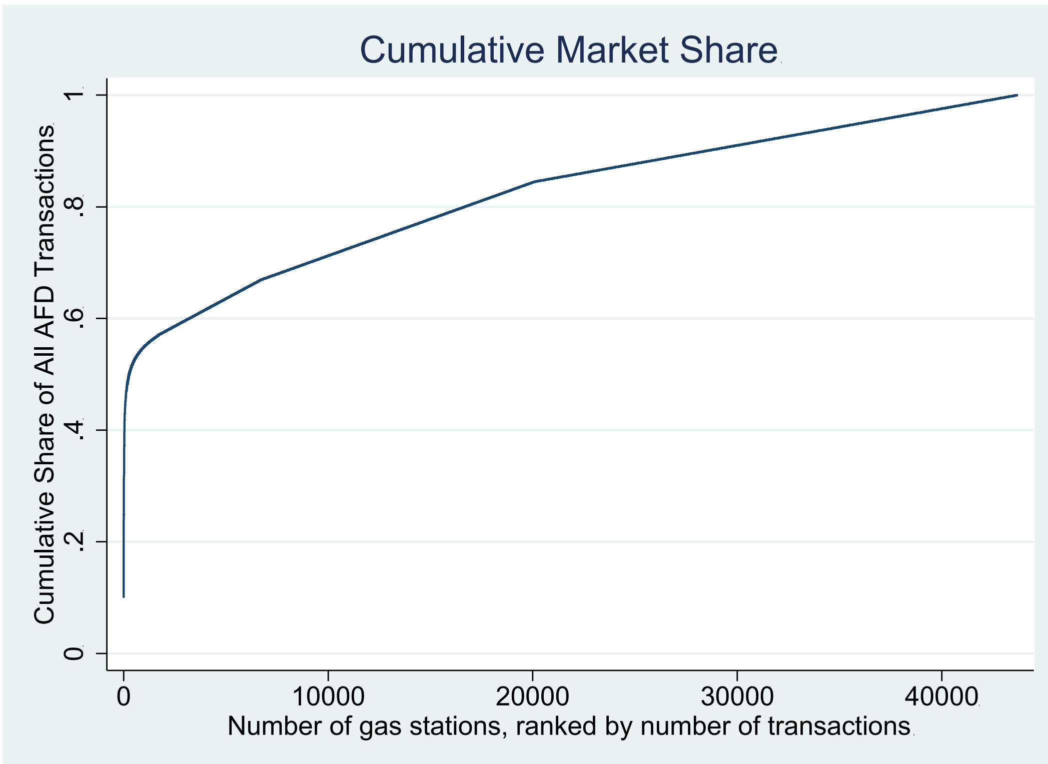
The results of the model are shown in Appendix Table B.1. The model predicts 292,997/(292,997+26,553) = 92% of automated fuel dispenser and service station transactions. The ratio of misclassfications to correct classifications is (30,080+26,553)/292,997=19%.

In summary, the ML approach is able to correctly classify over 90% of gasoline spending in the test data. If a human were to do this, she would need to identify over 30,000 strings. In

[3] Scikit-learn: Machine Learning in Python, Pedregosa *et al.*, JMLR 12, pp. 2825-2830, 2011.

addition, the model correctly classifies over 99.5% of the gasoline stations that would have been captured in an alternative approach of identifying the 100 largest gasoline stations by market share.

Appendix Figure B.1



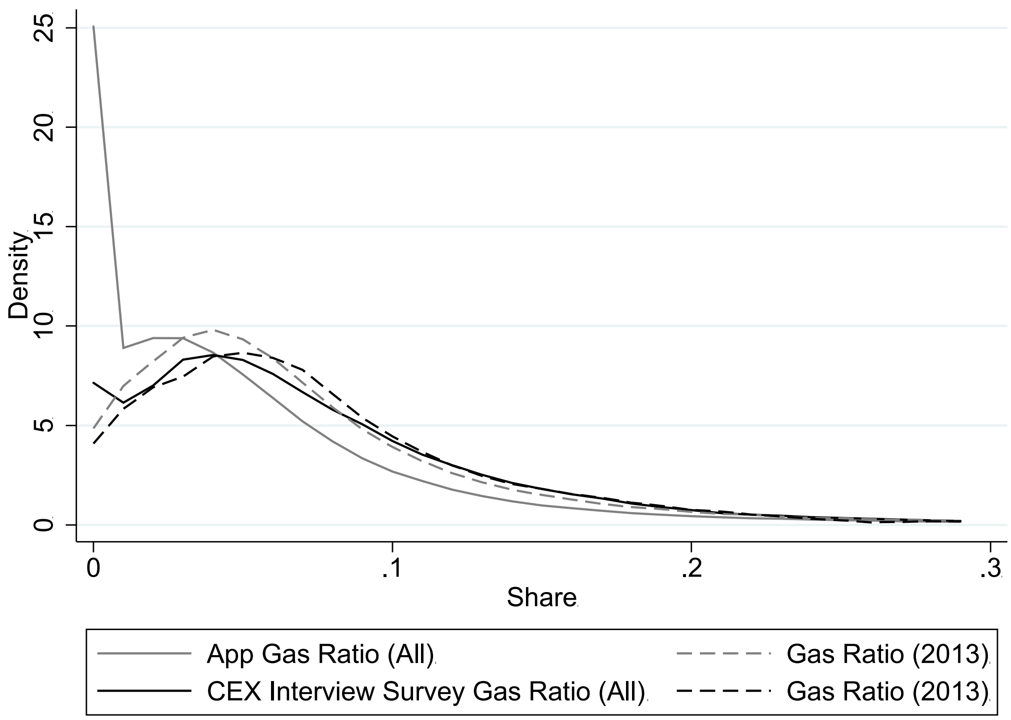
Appendix Table B.1. Confusion Matrix

		Actual gasoline spending	
		No	Yes
Predicted gasoline spending	No	2,741,524	26,553
	Yes	30,080	292,997

Notes: Table shows the four possible outcomes for our testing data set which is not used in any way to train the model, as described in the text. The rows "Predicted gasoline spending" refer to the binary prediction from the model as not gasoline, "no," or gasoline, "yes". Actual gasoline refers to the "truth," which is known for the case of our testing dataset.

Appendix C. Additional Tables and Figures

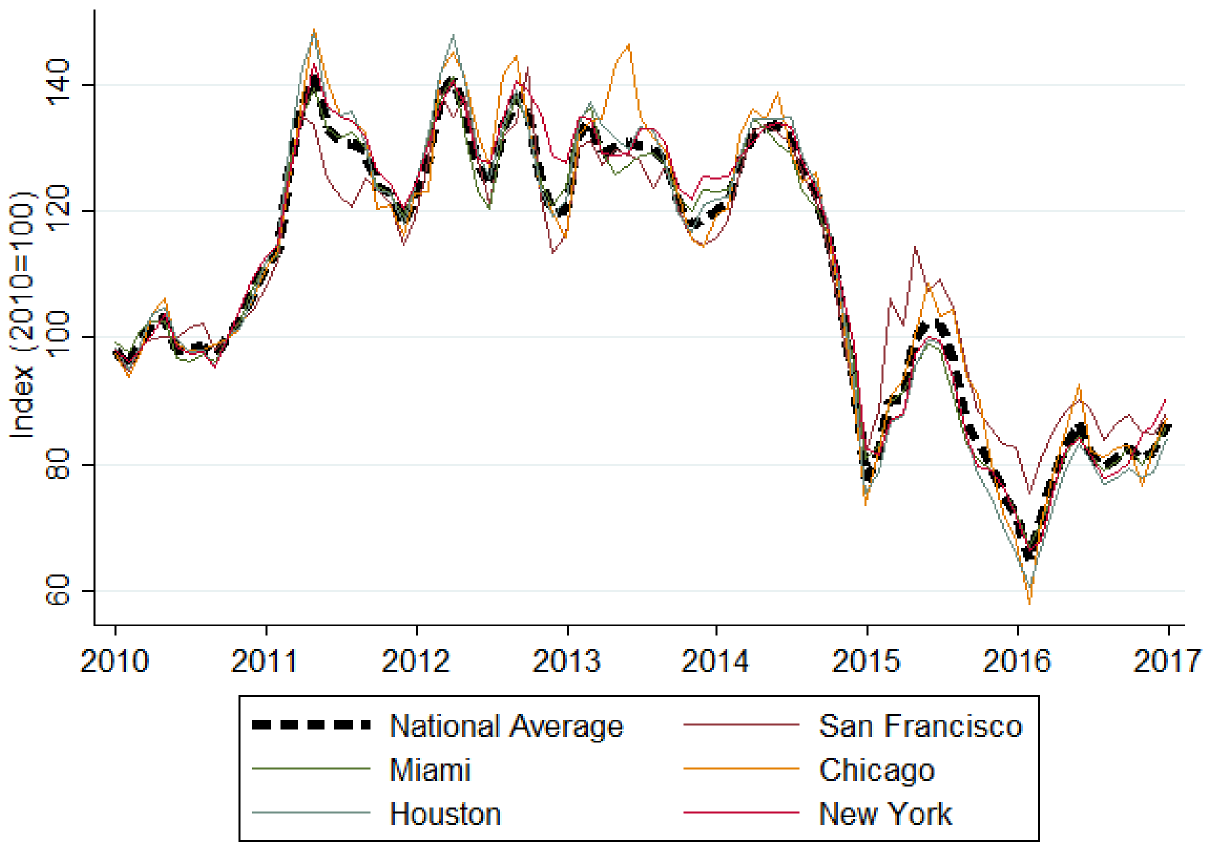
Appendix Figure C.1. Distribution of Ratio of Gasoline to Non-Gasoline Spending, 2013Q1-2014Q4



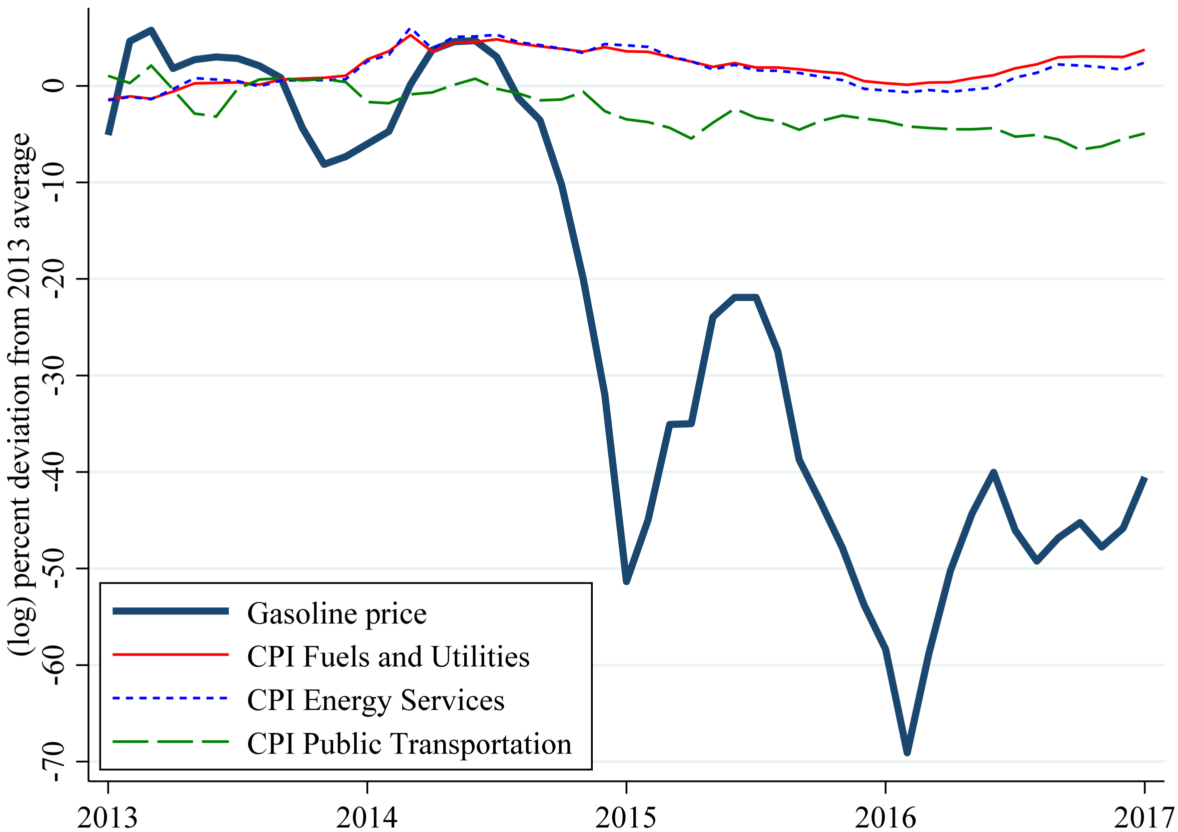
Note: the figure shows the quarterly gasoline to non-gasoline spending distribution in the app data and the CEX interview survey (solid lines), and the same ratio calculated over all of 2013 (dashed lines).

Appendix Figure C.2. Dynamics of gasoline prices across metropolitan areas

Notes: The figure plots time series of gasoline prices (all types) to major metropolitan areas. All series are from the FRED© database (mnemonics: CUUR0000SETB01, CUURA422SETB01, CUURA320SETB01, CUURA207SETB01, CUURA318SETB01, CUURA101SETB01) and normalized to be equal to 100 in 2010.

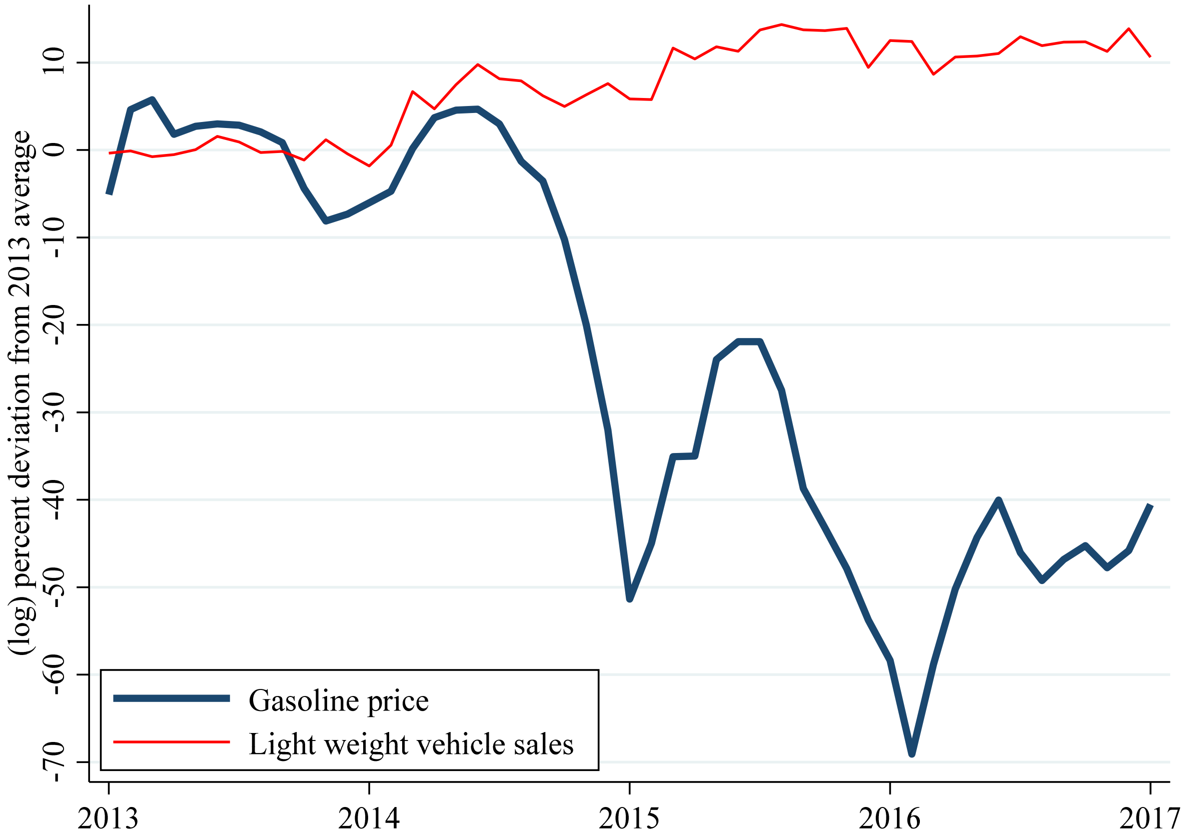


Appendix Figure C.3. Dynamics of prices for gasoline, fuel and utilities, and energy services.



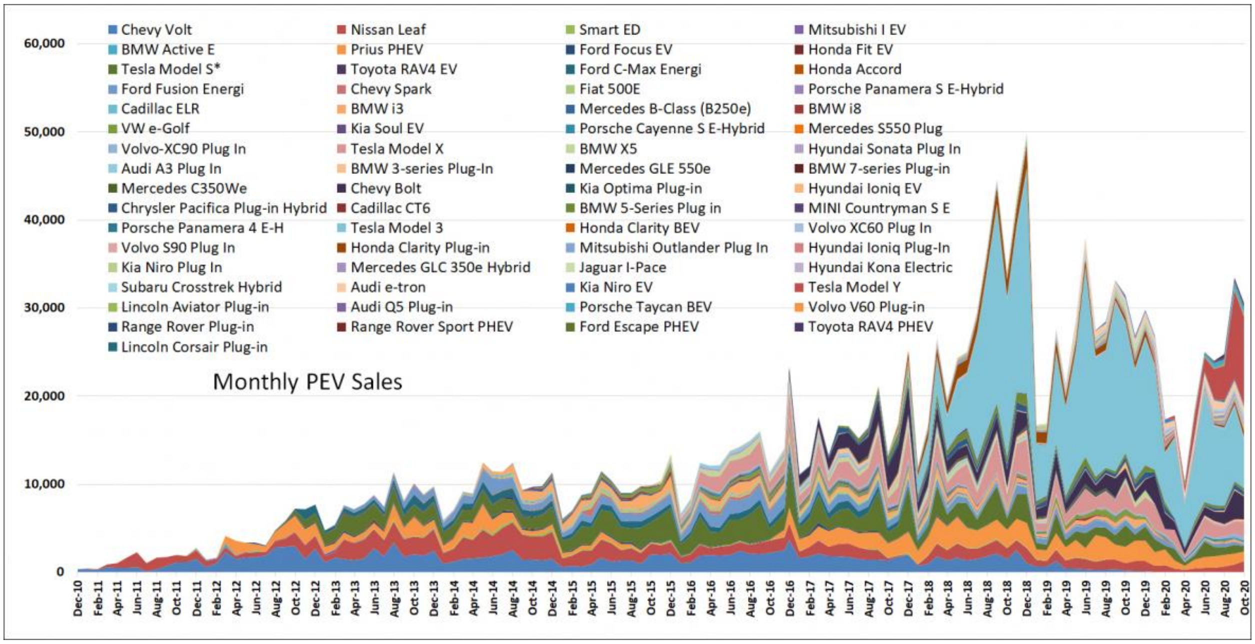
Notes: the figure show the log percent deviation of prices (or price indices) from the average value of the corresponding price (price index) in 2013. The gasoline price is taken from the FRED database (APU000074714). CPI Fuel and Utilities is a subindex of the Consumer Price Index that covers fuel oil, propane, kerosine, firewood, electricity, and piped gas service (FRED name: CUSR0000SAH2). CPI Energy Services is a subindex of the Consumer Price Index that covers electricity and piped gas service (FRED name: CUSR0000SEHF). CPI Public Transportation is a subindex of the Consumer Price Index that covers the cost of public transportation (FRED name: CUSR0000SETG).

Appendix Figure C.4. Dynamics of the gasoline price and light weight vehicle sales.



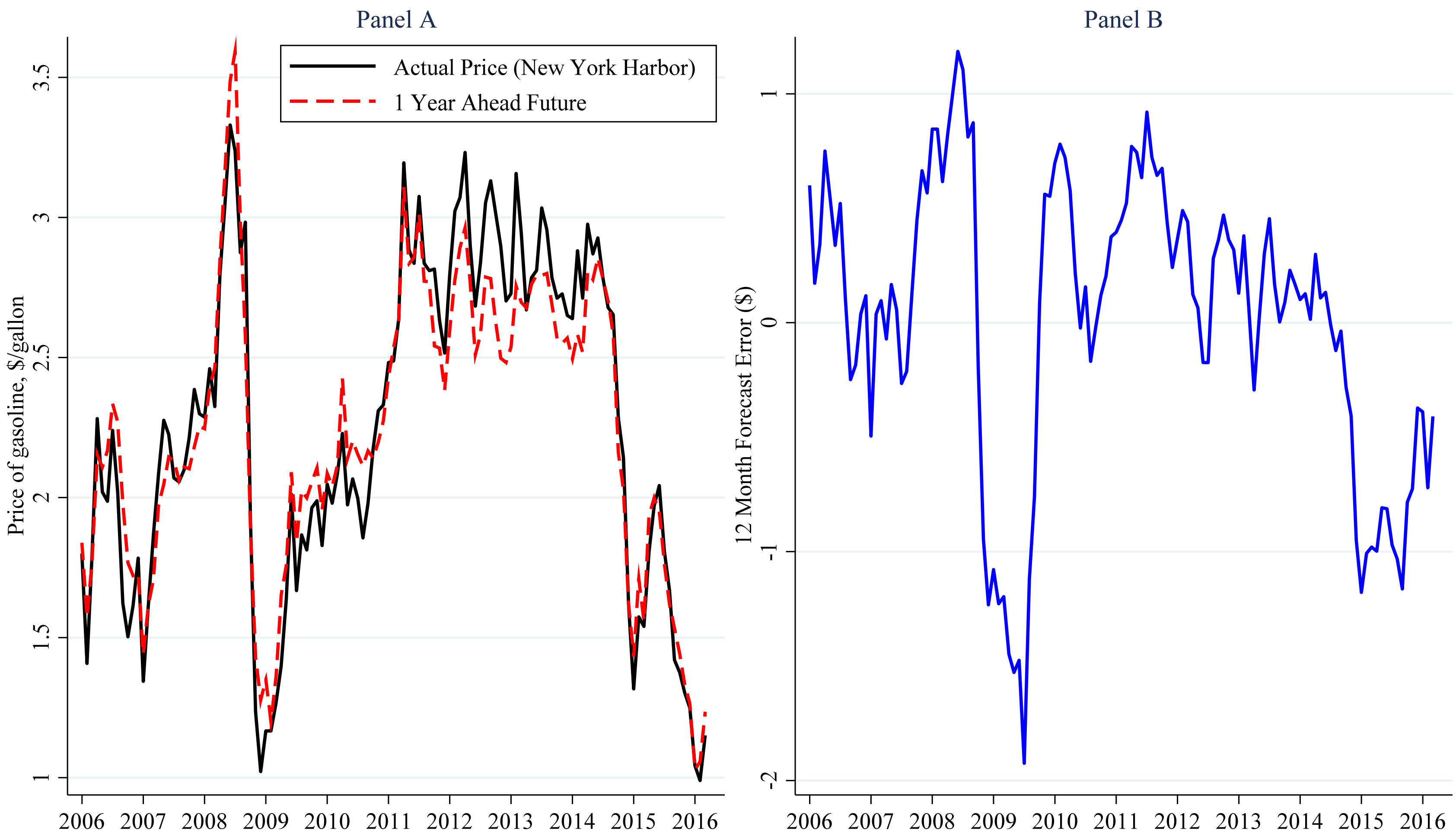
Notes: the figure show the log percent deviation of gasoline and light weight vehicle (cars, light trucks) sales from the average value of the corresponding series in 2013. The gasoline price is taken from the FRED database (APU000074714). Light weight vehicle sales are seasonally adjusted (FRED name: ALTSALES).

Appendix Figure C.5. Purchases of electric cars.



Source: Argonne National Laboratory; https://www.anl.gov/es/light-duty-electric-drive-vehicles-monthly-sales-updates.

Appendix Figure C.6. Gasoline prices and futures.



Notes: Panel A shows the New York Harbor spot price and the 1-year-ahead futures price. Panel B shows the 1 year ahead forecast error, defined as the difference between the realization of the spot price and the forecast 1 year earlier.

Appendix Table C.1. Gelman et al. (2014)

Table 1. Check versus ACS demographics (percent). The sample size for Check is 59,072, 35,417, 28,057, and 63,745 for gender, age, education, and region, respectively. The sample size for ACS is 2,441,532 for gender, age, and region and 2,158,014 for education.

	Check	ACS
Sex		
Male	59.93	48.59
Female	40.07	51.41
Age		
18–20	0.59	5.72
21–24	5.26	7.36
25–34	37.85	17.48
35–44	30.06	17.03
45–54	15.00	18.39
55–64	7.76	16.06
65+	3.48	17.95
Highest degree		
Less than college	69.95	62.86
College	24.07	26.22
Graduate school	5.98	10.92
Census Bureau region		
Northeast	20.61	17.77
Midwest	14.62	21.45
South	36.66	37.36
West	28.11	23.43

Source: Gelman et al. (2014)

Appendix D: Marginal Propensity to Consumer in Partial Equilibrium versus General Equilibrium Effects

In this appendix, we link the marginal propensity to consume (MPC) to structural parameters and explore how partial equilibrium MPC is potentially related to general equilibrium effect of gasoline price changes. In this exercise, we vary the price of gasoline holding everything else constant.

Partial equilibrium

Consumer a household that solves the following problem

$$\max m(c_1) + u(c_2) - v(L)$$
$$s.t.\ \ c_1 + p_2 c_2 = wL$$

where c_1 is the numeraire good (or "non-gasoline spending"; we normalize $p_1 = 1$), c_2 is "gasoline", L is labor, w is wages, and functions m, u, v describe how the household values goods and leisure. In a popular case, $m(c_1) = c_1$ so that the utility is quasi-linear. Because this is a partial equilibrium model, we take wages as given.

The first-order conditions yield:

$$m'(c_1) = \lambda$$
$$u'(c_2) = p_2\lambda$$
$$v'(L) = w\lambda$$

After log-linearization of these FOCs and the budget constraint, we have (assume w is fixed)

$$\check{c}_1 = \epsilon_1\check{\lambda}$$
$$\check{c}_2 = \epsilon_2(\check{\lambda} + \check{p}_2)$$
$$\frac{v''(L)L}{v'(L)}\check{L} = \eta^{-1}\check{L} = \check{\lambda}$$
$$\check{c}_1 = \frac{wL}{c_1}\check{L} - \frac{p_2c_2}{c_1}(\check{p}_2 + \check{c}_2) = (1+s)\check{L} - s(\check{p}_2 + \check{c}_2)$$

where $s \equiv \frac{p_2c_2}{c_1}$ and $\frac{wL}{c_1} = \frac{c_1+p_2c_2}{c_1} = 1 + s$ and $\eta = \left(\frac{v''(L)L}{v'(L)}\right)^{-1}$ is the Frisch labor supply elasticity, $\epsilon_2 \equiv \frac{u'(c_2)}{u''(c_2)c_2} < 0$ and $\epsilon_1 \equiv \frac{m'(c_1)}{m''(c_1)c_1} < 0$. Note that the first two conditions imply that $\check{c}_2 = \frac{\epsilon_2}{\epsilon_1}\check{c}_1 + \epsilon_2\check{p}_2$.

It follows that

$$\check{c}_1 = (1+s)\check{L} - s(\check{p}_2 + \check{c}_2) = (1+s)\eta\check{\lambda} - s\left(\frac{\epsilon_2}{\epsilon_1}\check{c}_1 + (1+\epsilon_2)\check{p}_2\right)$$
$$= \frac{(1+s)\eta}{\epsilon_1}\check{c}_1 - s\left(\frac{\epsilon_2}{\epsilon_1}\check{c}_1 + (1+\epsilon_2)\check{p}_2\right) = \left\{\frac{(1+s)\eta}{\epsilon_1} - s\frac{\epsilon_2}{\epsilon_1}\right\}\check{c}_1 - s(1+\epsilon_2)\check{p}_2 \Rightarrow$$
$$\check{c}_1 = -\frac{s(1+\epsilon_2)}{1 - \frac{(1+s)\eta}{\epsilon_1} + s\frac{\epsilon_2}{\epsilon_1}}\check{p}_2 = -\frac{s(1+\epsilon_2)}{\left\{1 - \frac{(1+s)\eta}{\epsilon_1} + s\frac{\epsilon_2}{\epsilon_1}\right\} \times s \times (1+\varepsilon)} \times s \times (1+\varepsilon) \times \check{p}_2 \Rightarrow$$
$$dlog\ C = -\frac{s(1+\epsilon_2)}{\left\{1 - \frac{(1+s)\eta}{\epsilon_1} + s\frac{\epsilon_2}{\epsilon_1}\right\} \times s \times (1+\varepsilon)} \times s \times (1+\varepsilon) \times dlog\ P_{gas}$$

where $(1+\varepsilon)$ is estimated from the regressing $dlog\,(p_2 c_2)$ on $dlog\, p_2$. Note that this equation provides structural interpretation of our specification (2) in the paper.

We know from the derivation above that

$$dlog\,(p_2 c_2) = \check{c}_2 + \check{p}_2 = \frac{\epsilon_2}{\epsilon_1}\check{c}_1 + (1+\epsilon_2)\check{p}_2$$
$$= -\frac{\epsilon_2}{\epsilon_1}\frac{s(1+\epsilon_2)}{1-\frac{(1+s)\eta}{\epsilon_1}+s\frac{\epsilon_2}{\epsilon_1}}\check{p}_2 + (1+\epsilon_2)\check{p}_2$$
$$= \frac{(1+\epsilon_2)\left(1-\frac{(1+s)\eta}{\epsilon_1}\right)}{1-\frac{(1+s)\eta}{\epsilon_1}+s\frac{\epsilon_2}{\epsilon_1}}\check{p}_2$$

So that $(1+\varepsilon) = \frac{(1+\epsilon_2)\left(1-\frac{(1+s)\eta}{\epsilon_1}\right)}{1-\frac{(1+s)\eta}{\epsilon_1}+s\frac{\epsilon_2}{\epsilon_1}} > 0$. Also note that $(1+\varepsilon) < (1+\epsilon_2)$.

Hence, marginal propensity to consume (MPC) is equal to

$$MPC = \frac{s(1+\epsilon_2)}{\left\{1-\frac{(1+s)\eta}{\epsilon_1}+s\frac{\epsilon_2}{\epsilon_1}\right\}\times s\times(1+\varepsilon)}$$
$$= \frac{s(1+\epsilon_2)}{\left\{1-\frac{(1+s)\eta}{\epsilon_1}+s\frac{\epsilon_2}{\epsilon_1}\right\}\times s}\times\frac{1-\frac{(1+s)\eta}{\epsilon_1}+s\frac{\epsilon_2}{\epsilon_1}}{(1+\epsilon_2)\left(1-\frac{(1+s)\eta}{\epsilon_1}\right)}$$
$$= \frac{1}{1-\frac{(1+s)\eta}{\epsilon_1}} > 0$$

In the case of quasi-linear utility, $\epsilon_1 = -\infty$ and $MPC = 1$.

General equilibrium

In the general equilibrium version of the model, we consider two sectors. The first sector (sector A) produces consumer goods. The second sector (sector B) produces gasoline. Gasoline is consumed by households working in both sectors. Gasoline is also a production input in the first sector. We assume that households cannot move across sectors, which is likely a reasonable approximation in the short-to-medium run. The mass of households in sector A is q. The mass of households in sector B is $1-q$. The economy is closed.

Households in sector A solve the following maximization problem:

$$\max m(c_1^A) + u(c_2^A) - v(L^A)$$
$$s.t.\ \ c_1^A + p_2 c_2^A = w^A L^A$$

where c_1 is the numeraire good (or “non-gasoline spending”; we normalize $p_1 = 1$), c_2 is “gasoline”, L is labor, w is wages. After log-linearization of first-order conditions and the budget constraint, we have

$$\check{c}_1^A = \epsilon_1^A \check{\lambda}^A$$
$$\check{c}_2^A = \epsilon_2^A(\check{\lambda}^A + \check{p}_2)$$
$$\frac{v''(L)L}{v'(L)}\check{L}^A = \eta^{A^{-1}}\check{L}^A = \check{\lambda}^A + \check{w}^A$$

$$\check{c}_1^A = \frac{w^A L^A}{c_1^A}\left(\check{L}^A + \breve{w}^A\right) - \frac{p_2 c_2^A}{c_1^A}(\check{p}_2 + \check{c}_2^A) = (1+s^A)\left(\check{L}^A + \breve{w}^A\right) - s^A(\check{p}_2 + \check{c}_2)$$

where $s^A \equiv \frac{p_2^A c_2^A}{c_1^A}$ and $\frac{w^A L^A}{c_1^A} = \frac{c_1^A + p_2^A c_2^A}{c_1^A} = 1 + s^A$ and $\eta^A = \left(\frac{v''(L)L}{v'(L)}\right)^{-1}$ is the Frisch labor supply elasticity, $\epsilon_2^A \equiv \frac{u'(c_2)}{u''(c_2)c_2} < 0$ and $\epsilon_1^A \equiv \frac{m'(c_1)}{m''(c_1)c_1} < 0$. Note that the first two conditions imply that $\check{c}_2^A = \frac{\epsilon_2^A}{\epsilon_1^A}\check{c}_1^A + \epsilon_2^A \check{p}_2$.

Households in sector B solve the following maximization problem:

$$\max m(c_1^B) + u(c_2^B) - v(L^B)$$
$$s.t.\quad c_1^B + p_2 c_2^B = w^B L^B$$

After log-linearization of first-order conditions and the budget constraint, we have

$$\check{c}_1^B = \epsilon_1^B \check{\lambda}^B$$
$$\check{c}_2^B = \epsilon_2^B(\check{\lambda}^B + \check{p}_2)$$
$$\frac{v''(L)L}{v'(L)}\check{L}^B = {\eta^B}^{-1}\check{L}^B = \check{\lambda}^B + \breve{w}^B$$
$$\check{c}_1^B = \frac{w^B L^B}{c_1^B}\left(\check{L}^B + \breve{w}^B\right) - \frac{p_2 c_2^B}{c_1^B}(\check{p}_2 + \check{c}_2^B) = (1+s^B)\left(\check{L}^B + \breve{w}^B\right) - s^B(\check{p}_2 + \check{c}_2)$$

where $s^B \equiv \frac{p_2 c_2^B}{c_1^B}$ and $\frac{w^B L^B}{c_1^B} = \frac{c_1^B + p_2 c_2^B}{c_1^B} = 1 + s^B$ and $\eta^B = \left(\frac{v''(L)L}{v'(L)}\right)^{-1}$ is the Frisch labor supply elasticity, $\epsilon_2^B \equiv \frac{u'(c_2)}{u''(c_2)c_2} < 0$ and $\epsilon_1^B \equiv \frac{m'(c_1)}{m''(c_1)c_1} < 0$. Note that the first two conditions imply that $\check{c}_2^B = \frac{\epsilon_2^B}{\epsilon_1^B}\check{c}_1^B + \epsilon_2^B \check{p}_2$.

Production in sector A is characterized by constant return to scale and perfect competition:

$$Y_1 = \left({L^A}^{\alpha} O^{1-\alpha}\right) \Rightarrow \check{Y}_1 = \alpha \check{L}^A + (1-\alpha)\check{O}$$

where O is gasoline used in production of good 1. Perfect competition means $p_1 = MC_1$. Given normalization $p_1 = 1$, we have $MC_1 = 1$. Given the Cobb-Douglass production function, we find that $\widetilde{MC} = \alpha\breve{w}^A + (1-\alpha)\check{p}_2$. This means that

$$\breve{w}^A = -\left(\frac{1-\alpha}{\alpha}\right)\check{p}_2.$$

Production in sector B is also characterized by constant return to scale (production function is linear in labor) and perfect competition:

$$Y_2 = L^B \Rightarrow \check{Y}_2 = \check{L}^B$$
$$w^B = MPL = MC = p_2 \Rightarrow \breve{w}^B = \check{p}_2$$

Market clearing in sector A:

$$Y_1 = qc_1^A + (1-q)c_1^B \Rightarrow \check{Y}_1 = \left(\frac{qc_1^A}{Y_1}\right)\check{c}_1^A + \left(\frac{(1-q)c_1^B}{Y_1}\right)\check{c}_1^B$$

Market clearing in sector B:

$$Y_2 = qc_2^A + (1-q)c_2^B + qO \Rightarrow \check{Y}_2 = \left(\frac{qc_2^A}{Y_2}\right)\check{c}_2^A + \left(\frac{(1-q)c_2^B}{Y_2}\right)\check{c}_2^B + \left(\frac{qO}{Y_2}\right)\check{O}$$

where qO is the total amount of oil consumed in production of good 1 (each firm in this sector consumes O and q is the mass of firms in the sector).

Now we derive MPC for each group of households:

$$\check{c}_1^A = -\frac{(1+s^A)(1+\eta_A)\left(\frac{1-\alpha}{\alpha}\right) + s^A(1+\epsilon_2^A)}{1 - \frac{(1+s^A)\eta^A}{\epsilon_1^A} + \frac{s^A\epsilon_2^A}{\epsilon_1^A}}\check{p}_2$$

Clearly, $\frac{\partial \check{c}_1^A}{\partial \check{p}_2} < 0$.

$$\check{L}_A = \eta\left(\check{\lambda}^A + \check{w}^A\right) = -\eta \times \frac{\left(\frac{1-\alpha}{\alpha}\right)(1+\epsilon_1^A) - \frac{s^A}{\alpha}(1+\epsilon_2^A)}{\epsilon_1^A - (1+s^A)\eta^A + s^A\epsilon_2^A}\check{p}_2$$

We can generate $\frac{\partial \check{L}_A}{\partial \check{p}_2} < 0$ if demand for good "1" is sufficiently elastic (i.e., $\epsilon_1^A < -1$), which seems a reasonable assumption. With utility quasi-linear c_1, we have $\check{L}_A = -\eta\left(\frac{1-\alpha}{\alpha}\right)\check{p}_2$.

The sensitivity of group A's total spending on gasoline to the price of gasoline is

$$\check{p}_2 + \check{c}_2^A = \frac{(1+\epsilon_2^A)\epsilon_1^A - \epsilon_2^A(1+s)\left(\frac{1-\alpha}{\alpha}\right) - (1+s)\eta^A\left(1+\frac{\epsilon_2^A}{\alpha}\right)}{\epsilon_1^A - (1+s^A)\eta^A + s^A\epsilon_2^A}\check{p}_2$$

Hence, $(1+\varepsilon^A) = \frac{(1+\epsilon_2^A)\epsilon_1^A - \epsilon_2^A(1+s)\left(\frac{1-\alpha}{\alpha}\right) - (1+s)\eta^A\left(1+\frac{\epsilon_2^A}{\alpha}\right)}{\epsilon_1^A - (1+s^A)\eta^A + s^A\epsilon_2^A}$.

The MPC we define in the paper is

$$\check{c}_1^A = -\frac{(1+s^A)(1+\eta_A)\left(\frac{1-\alpha}{\alpha}\right) + s^A(1+\epsilon_2^A)}{1 - \frac{(1+s^A)\eta^A}{\epsilon_1^A} + \frac{s^A\epsilon_2^A}{\epsilon_1^A}}\check{p}_2$$

$$= -\frac{(1+s^A)(1+\eta_A)\left(\frac{1-\alpha}{\alpha}\right) + s^A(1+\epsilon_2^A)}{1 - \frac{(1+s^A)\eta^A}{\epsilon_1^A} + \frac{s^A\epsilon_2^A}{\epsilon_1^A}} \times \frac{1}{(1+\varepsilon^A)} \times \frac{1}{s^A} \times s^A \times (1+\varepsilon^A) \times \check{p}_2$$

That is,

$$MPC^A = \frac{(1+s^A)(1+\eta_A)\left(\frac{1-\alpha}{\alpha}\right) + s^A(1+\epsilon_2^A)}{1 - \frac{(1+s^A)\eta^A}{\epsilon_1^A} + \frac{s^A\epsilon_2^A}{\epsilon_1^A}} \times \frac{1}{(1+\varepsilon^A)} \times \frac{1}{s^A}$$

$$= \frac{1 + \frac{(1+s^A)}{s^A}\frac{(1+\eta_A)}{1+\epsilon_2^A}\left(\frac{1-\alpha}{\alpha}\right)}{1 - \frac{\epsilon_2^A}{\epsilon_1^A} \times \frac{(1+s^A)}{(1+\epsilon_2^A)} \times \frac{1-\alpha+\eta}{\alpha} - \frac{(1+s^A)}{(1+\epsilon_2^A)}\frac{\eta}{\epsilon_1^A}}$$

$$= \frac{1 + \frac{(1+s^A)}{s^A}\frac{(1+\eta_A)}{1+\epsilon_2^A}\left(\frac{1-\alpha}{\alpha}\right)}{1 - \frac{\eta^A(1+s^A)}{\epsilon_1^A} - \frac{\epsilon_2^A}{\epsilon_1^A} \times \frac{(1+s^A)}{(1+\epsilon_2^A)} \times \frac{1-\alpha}{\alpha}}$$

For comparison, in the partial equilibrium model (effectively $\alpha = 1$) we had

$$MPC = \frac{1}{1 - \frac{(1+s)\eta}{\epsilon_1}} < 1$$

Note that the general equilibrium MPC for this group is greater than the partial equilibrium MPC because $\frac{(1+s^A)}{s^A}\frac{(1+\eta_A)}{1+\epsilon_2^A}\left(\frac{1-\alpha}{\alpha}\right) > 0$ and $\frac{\epsilon_2^A}{\epsilon_1^A} \times \frac{(1+s^A)}{(1+\epsilon_2^A)} \times \frac{1-\alpha}{\alpha} > 0$ (provided $\epsilon_2^A > -1$).

Doing a similar derivation for group B, we find that

$$\check{c}_1^B = \frac{(1+s^B)(1+\eta_B) - s^B(1+\epsilon_2^B)}{\epsilon_1^B - (1+s^B)\eta^B + s^B\epsilon_2^B}\epsilon_1^B\check{p}_2$$

Note that $\check{c}_1^B$ <u>increases</u> in $\check{p}_2$.

Employment for these agents increases in response to a shock in p_2 if their demand for good “1” is sufficiently elastic:

$$\check{L}^B = \eta\frac{1+\epsilon_1^B}{\epsilon_1^B - (1+s^B)\eta^B + s^B\epsilon_2^B}\check{p}_2$$

With utility quasi-linear in c_1 (i.e., $\epsilon_1^B = -\infty$) , we have $\check{L}^B = \eta^B\check{p}_2$.

Now the sensitivity of group B’s total spending on gasoline to the price of gasoline is

$$\check{p}_2 + \check{c}_2^B = \check{p}_2 + \frac{\epsilon_2^B}{\epsilon_1^B}\check{c}_1^B + \epsilon_2^B\check{p}_2$$

$$= (1+\epsilon_2^B)\check{p}_2 - \frac{\epsilon_2^B}{\epsilon_1^B}\frac{(1+s^B)(1+\eta^B) - s^B(1+\epsilon_2^B)}{\epsilon_1^B - (1+s^B)\eta^B + s^B\epsilon_2^B}\epsilon_1^B\check{p}_2$$

$$= (1+\epsilon_2^B)\left\{\frac{\epsilon_1^B - (1+s^B)\eta^B + \frac{\epsilon_2^B(1+s^B)(1+\eta^B)}{(1+\epsilon_2^B)}}{\epsilon_1^B - (1+s^B)\eta^B + s^B\epsilon_2^B}\right\}\check{p}_2$$

Hence,

$$(1+\varepsilon^B) = (1+\epsilon_2^B)\left\{\frac{\epsilon_1^B - (1+s^B)\eta^B + \frac{\epsilon_2^B(1+s^B)(1+\eta^B)}{(1+\epsilon_2^B)}}{\epsilon_1^B - (1+s^B)\eta^B + s^B\epsilon_2^B}\right\} > (1+\epsilon_2^B)$$

It follows that MPC for type B is

$$MPC^B = -\frac{\left(\frac{1+s^B}{s^B}\right)\left(\frac{1+\eta^B}{1+\epsilon_2^B}\right) - 1}{1 - (1+s^B)\frac{\eta^B}{\epsilon_1^B} + \left(\frac{\epsilon_2^B}{\epsilon_1^B}\right)\frac{(1+s^B)(1+\eta^B)}{(1+\epsilon_2^B)}}$$

The denominator is positive ($\epsilon_2^B > -1$). The numerator is positive too ($\epsilon_2^B > -1$). Thus, this MPC is negative and can be greater than one in absolute magnitude. For example, with infinitely elastic demand for good "1" (i.e., quasi-linear utility in c_1), we have

$$MPC^B = -\left\{\left(\frac{1+s^B}{s^B}\right)\left(\frac{1+\eta^B}{1+\epsilon_2^B}\right) - 1\right\} < 0$$

which can be less than -1 provided that the share good "2" in the consumption basket of type B agent (s^B) is sufficiently small.

The aggregate employment depends on the relative strength of employment responses across sectors:

$$\check{L} = q\check{L}^A + (1-q)\check{L}^B$$

For the case with quasi-linear utility, we have $\check{L}_A = -\eta^A\left(\frac{1-\alpha}{\alpha}\right)\check{p}_2$ and $\check{L}^B = \eta^B\check{p}_2$. Output is likely to decrease in response to a hike in oil prices. First, one can proxy inelastic supply of gasoline with inelastic supply of labor, that is $\eta^B \approx 0$ and hence $\check{L}_A$ (which has a clear sign) drives aggregate employment. Second, $1-q$ is small and, thus, it would take very large employment effects in sector 2 to drive aggregate employment.

The aggregate $\overline{MPC} = qMPC^A + (1-q)MPC^B$. Note that MPC^A and MPC^B have different signs. Depending on parameter values, partial equilibrium MPC can be greater or smaller than the aggregate $\overline{MPC}$.

Appendix E: Additional Heterogeneity Results

Liquidity constraints

Macroeconomic theory predicts that the responses of consumers to changes in income (or prices) could be heterogeneous with important implications for macroeconomic dynamics and policy. For example, Kaplan and Violante (2014) present a theoretical framework where "hand-to-mouth" (HtM) consumers with liquidity constraints should exhibit a larger MPC to *transitory*, *anticipated* income shocks than non-HtM consumers for whom these constraints are not binding. Kaplan and Violante (2014) document empirical evidence consistent with these predictions and quantify the contribution of consumer heterogeneity in terms of liquidity holdings for the 2001 Bush tax rebate. In a similar spirit, Mian and Sufi (2014), McKay, Nakamura and Steinsson (2016), and many others document that consumers' liquidity and balance sheets can play a key role for aggregate outcomes.

The conventional focus in this literature is the consumption response to transitory, anticipated income shocks because the behavior of HtM and non-HtM consumers should be particularly different in this case. First, HtM consumers spend an income shock when it is realized rather than when it is announced, while non-HtM consumers respond to the announcement and exhibit no change in spending at the time the shock is realized. Second, the MPC of non-HtM consumers should be small (this group smooths consumption by saving a big fraction of the income shock), while the MPC of HtM consumers should be large (the income shock relaxes a spending constraint for these consumers). This sharp difference in the responses hinges on the temporary, anticipated nature of the shock. For other shocks, the responses may be alike across HtM and non-HtM consumers. For example, when the shock is permanent and unanticipated, HtM and non-HtM consumers should behave in the same way (Mankiw and Shapiro 1985): both groups should have $MPC = 1$ at the time of the shock. Intuitively, non-HtM consumers have $MPC = 1$ because their lifetime resources change permanently and, accordingly, these consumers adjust their consumption by the size of the shock when the shock happens. HtM consumers have $MPC = 1$ because they are in a "corner solution" and would like to spend away every dollar they receive in additional income the moment they receive it. Thus, macroeconomic theory predicts that, in this case, the MPC should be similar across HtM and non-HtM consumers and that the MPC should be close to one. We focus this section on testing these two predictions.

For these tests one needs to identify HtM and non-HtM consumers. This seemingly straightforward exercise has proved to be a challenge in applied work due to a number of data limitations, which have made researchers use proxies for liquidity constraints. As a result, estimated MPCs should be interpreted with caution and important caveats. For example, Kaplan and Violante (2014) argue that identification of HtM consumers requires information on consumers' liquidity holdings *just before* they receive pay checks.[4] Because the Survey of Consumer Finances (SCF), the dataset used in Kaplan and Violante (2014), reports *average*

[4] Intuitively, hand-to-mouth consumers do not carry liquid assets from period to period. Hence, just before receiving a pay check (an injection of liquidity), a hand-to-mouth consumer should have zero liquid wealth.

balances for a household as well as *average* monthly income, Kaplan and Violante are forced to make assumptions about payroll frequency (also not reported in the SCF) and behavior of account balances (e.g., constant flow of spending). Given heterogeneity in payment cycles (i.e., weekly, biweekly, monthly) and spending patterns across consumers, this procedure can mix HtM and non-HtM consumers and, thus, yield an attenuated estimate of MPC.

In contrast, the app data allow us to take Kaplan and Violante (2014)'s definition literally. We identify the exact day of a consumer's payroll income (if any), and examine bank account and credit card balances of the consumer the day before this payment arrives. If a consumer has several pay checks per month, we treat these as separate events. A consumer is classified as HtM in a given month if, for any pay check events in the previous month, the consumer has virtually no liquid assets (less than \$100 in the consumer's checking or savings accounts net of credit card debt), or the consumer is in debt (the sum of the consumers' liquid assets and available balance on credit cards is negative) and is within \$100 of the consumer's credit card limits. Denote the dummy variable identifying hand-to-mouth consumers at this frequency with D_{it}^{*}. We find that, in the app data, roughly 20% of consumers are HtM, which is similar to the estimate reported in Kaplan and Violante (2014) for a nationally representative sample of U.S. households in the Survey of Consumer Finances.[5]

To allow for heterogeneity in the MPC by liquidity, we add interaction terms to the baseline specifications (4) and (5):

$$\begin{aligned}\Delta_{\mathrm{k}} \log C_{it} = {} & \beta_1 \times s_i \times \Delta_{\mathrm{k}} \log P_t + \beta_2 \times s_i^{gas} \times \Delta_{\mathrm{k}} \log P_t \times D_{it} \\ & + \mu_0 \times D_{it} + \mu_1 \times s_i \times D_{it} + \psi_t + \omega_t \times D_{it} + \varepsilon_{it} \qquad (6)\end{aligned}$$

$$\Delta_{\mathrm{k}} \log PQ_{it} = \delta_1 \times \Delta_{\mathrm{k}} \log P_t + \delta_2 \times \Delta_{\mathrm{k}} \log P_t \times D_{it} + \xi \times D_{it} + u_{it} \qquad (7)$$

where D_{it} is a variable measuring the presence/intensity of liquidity constraints identifying HtM consumers, and $\omega_t \times D_{it}$ is the time fixed effect specific to HtM consumers.

We have several options for D_{it}. One could use a dummy variable equal to one if a consumer is liquidity constrained in period $t-k-1$ (recall that Δ_{k} operator calculates the growth rate between periods $t-k$ and t). We denote this "lagged" measure of HtM with $\widetilde{D}_{it} \equiv D_{i,t-k-1}^{*}$ where D_{it}^{*} is a dummy variable equal to one if consumer i at time t satisfies the Kaplan-Violante HtM criteria and zero otherwise. Alternatively, because liquidity constraints may be short-lived, one may want to use measures that are calculated over a longer horizon to identify "serial" HtM consumers. To this end, we construct three measures on the 2013 sample which are not used in the estimation of MPC and ϵ. Specifically, for each month of data available for consumer i in 2013, we use three metrics to classify consumers as HtM or not. We consider the average value of D_{it}^{*} (this continuous variable provides a sense of frequency of liquidity constraints; we denote this measure with $\overline{D}_{i,2013}$), the modal value of $D_{i,t}^{*}$ (most frequent value;[6] we denote this measure with $\breve{D}_{i,2013}$), or the minimum

[5] While the app data are close to ideal for identification of hand-to-month (i.e., low liquidity) consumers, the app data are not suitable for further disaggregation of consumers into wealthy hand-to-mouth and poor hand-to-mouth because the app does not collect information on consumer durables (e.g., vehicles), housing and other illiquid assets which are not backed by corresponding loans and mortgages.

[6] We classify a household as HtM if there is a tie.

value of $D^*_{i,2013}$ during the 2013 part of the sample. The latter measure, which we denote with $\widehat{D}_{i,2013}$, is equal to one only if a consumer is identified as HtM in *every* month in 2013.

Irrespective of which measure we use, we find in results reported in Appendix Table E1 that estimated MPCs are very similar for HtM and non-HtM consumers. Although the point estimates for HtM consumers tend to be larger at short horizons (e.g., 5 weeks), we generally cannot reject the null of equal MPCs across the groups or the null that estimated MPCs are equal to one, which is consistent with the PIH predictions.

Nonlinearity

We test for nonlinearities in the MPC and the elasticity of demand. We do this by examining responses by deciles of the 2013 gasoline share, s_i. The means of s_i within each decile are given in Appendix Table E2.

To examine heterogeneity in the elasticity of demand, we interact $\Delta \log P_t$ in specification (5) with deciles of s_i. Because there is no time fixed effect in this specification, we can separately identify each of these interactions. However, since we examine gasoline spending in logs, no elasticity can be estimated for those with \$0 in gasoline spending. Since few people in the first quintile have any gasoline spending, we combine the small number of individuals in decile 1 who have gasoline spending over the 2014-2016 period with the second decile. We estimate a main effect, which will be the average elasticity for these two "lowest gas share" quintiles, and estimate 8 additional interactions for quintiles 3-10: $\sum_{q=3}^{10} \delta^q \mathbb{1}\{q(i) = q\} \times \Delta \log p_t$. The interpretation of the δ^q coefficients is the average difference in $(1 + \epsilon)$ for decile q, relative to deciles 1-2.

To examine heterogeneity in the MPC, we replace s_i in (4), with indicators for deciles of s_i, $\sum_{q=2}^{10} \beta^q \mathbb{1}\{q(i) = q\} \times \Delta \log p_t$. The interpretation of the coefficient then becomes the average difference for decile q, relative to decile 1.

Appendix Figure E3 plots the results of this exercise run on our baseline estimation sample at horizons 5, 15, and 25 weeks. There is little evidence of non-linearities in the $MPC * (1 + \epsilon)$ specification, expect for a small non-linearity at the very bottom and top. There is some evidence of nonlinearities for the elasticity, where the lowest and highest gas shares are more elastic, and individuals in the middle of the gas share distribution are the least elastic. The maximum difference is about 0.10 percentage points. So for an estimate of $\beta = MPC * (1 + \epsilon) = 0.8$, assuming the elasticity for those with the lowest gas share is about -0.25 and the elasticity from the middle of the distribution is -0.15, the MPC would be about 0.125 higher for the lowest group (1.07-0.941).

References

Kaplan, Greg and Giovanni L. Violante. 2014. "A Model of the Consumption Response to Fiscal Stimulus Payments." *Econometrica* 82(4): 1199-1239.

Mankiw, N. Gregory, and Matthew D. Shapiro. 1985. "Trends, Random Walks, and Tests of the Permanent Income Hypothesis." *Journal of Monetary Economics* 16: 165-174.

McKay, Alisdair, Emi Nakamura, and Jón Steinsson, 2016. "The Power of Forward Guidance Revisited." *American Economic Review* 106(10): 3133-3158.

Mian, Atif, and Amir Sufi. 2014. *House of Debt: How They (and You) Caused the Great Recession, and How We Can Prevent It from Happening*. Princeton University Press.

Appendix Table E.1. MPC by liquidity status.

Measure of Hand-to-mouth consumers (HtM)	Elasticity of demand for gasoline, ϵ			MPC		
	Horizon (weeks)			Horizon (weeks)		
	5	15	25	5	15	25
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Lagged HtM						
Non-HtM	-0.222	-0.182	-0.182	0.542	0.868	0.921
	(0.059)	(0.028)	(0.026)	(0.631)	(0.318)	(0.282)
	[0.003]	[0.002]	[0.003]	[0.054]	[0.040]	[0.040]
HtM	-0.271	-0.336	-0.361	0.509	0.492	0.596
	(0.068)	(0.032)	(0.033)	(0.701)	(0.261)	(0.344)
	[0.010]	[0.007]	[0.007]	[0.175]	[0.142]	[0.154]
P-value (Non-HtM=HtM)	0.119	0.000	0.000	0.930	0.093	0.219
	[0.000]	[0.000]	[0.000]	[0.858]	[0.009]	[0.036]
Panel B. Average HtM in 2013						
Non-HtM	-0.206	-0.158	-0.162	0.465	0.848	0.957
	(0.059)	(0.027)	(0.025)	(0.607)	(0.323)	(0.280)
	[0.003]	[0.003]	[0.003]	[0.058]	[0.044]	[0.044]
HtM	-0.314	-0.392	-0.412	1.111	1.090	1.215
	(0.071)	(0.040)	(0.038)	(0.584)	(0.276)	(0.321)
	[0.011]	[0.008]	[0.009]	[0.194]	[0.181]	[0.184]
P-value (Non-HtM=HtM)	0.026	0.000	0.000	0.185	0.469	0.374
	[0.000]	[0.000]	[0.000]	[0.002]	[0.211]	[0.190]
Panel C. Modal HtM in 2013						
Non-HtM	-0.211	-0.169	-0.174	0.515	0.885	0.992
	(0.059)	(0.027)	(0.025)	(0.598)	(0.319)	(0.279)
	[0.003]	[0.002]	[0.003]	[0.055]	[0.042]	[0.043]
HtM	-0.290	-0.330	-0.342	0.910	1.092	1.212
	(0.064)	(0.034)	(0.032)	(0.577)	(0.271)	(0.293)
	[0.009]	[0.007]	[0.007]	[0.153]	[0.134]	[0.135]
P-value (Non-HtM=HtM)	0.017	0.000	0.000	0.204	0.310	0.236
	[0.000]	[0.000]	[0.000]	[0.015]	[0.142]	[0.119]
Panel D. Extreme HtM in 2013						
Non-HtM	-0.220	-0.186	-0.192	0.588	0.969	1.071
	(0.058)	(0.027)	(0.025)	(0.597)	(0.316)	(0.278)
	[0.003]	[0.002]	[0.002]	[0.052]	[0.041]	[0.041]
HtM	-0.271	-0.339	-0.342	0.769	1.186	1.325
	(0.064)	(0.035)	(0.037)	(0.583)	(0.407)	(0.397)
	[0.020]	[0.015]	[0.016]	[0.317]	[0.288]	[0.285]
P-value (Non-HtM=HtM)	0.169	0.000	0.000	0.569	0.477	0.383
	[0.011]	[0.000]	[0.000]	[0.575]	[0.455]	[0.377]

Notes: the table reports estimates of MPC and ϵ based on equations (6)-(7) over k periods, where k is shown in the top row of the table. s_i^{gas} is the ratio of gasoline spending to non-gasoline spending for 2013 for consumer i. The title of each panel indicates how the presence/intensity of liquidity constraints is measured. Denote the dummy variable identifying hand-to-mouth consumers for a given month with D_{it}^*. Panel A uses a dummy variable equal to one if a consumer is liquidity constrained in period $t-k-1$ (recall that Δ_k operator calculates the growth rate between periods $t-k$ and t), i.e. $\widetilde{D}_{it} \equiv D_{i,t-k-1}^*$. For other panels, we construct three measures on the 2013 sample which is not used in the estimation of MPC and ϵ: the average value of D_{it}^* (this continuous variable provides a sense of frequency of liquidity constraints; we denote this measure with $\overline{D}_{i,2013}$), the modal value of D_{it}^* (most frequent value; we denote this measure with $\breve{D}_{i,2013}$), or the minimum value of $D_{i,2013}^*$ during the 2013 part of the sample. The latter measure, which we denote with $\widehat{D}_{i,2013}$ and refer to as "extreme," is equal to one only if a consumer is identified as hand-to-mouth in *every* month in 2013. Robust standard errors in parentheses are clustered by week and consumer. Standard errors reported in squared brackets are clustered at the consumer level. *P-value (Non-HtM=HtM)* is the p-value for the test of HtM and non-HtM responses being equal. See text for further details.

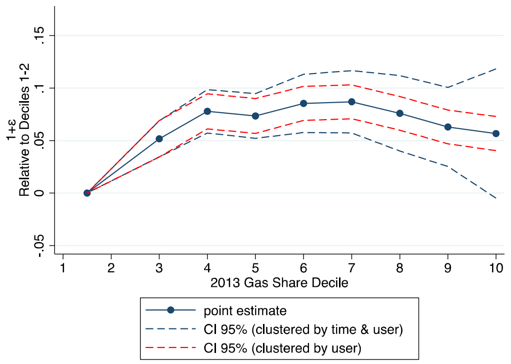
Appendix Table E.2. Deciles for the share of gasoline spending.

Decile	Mean of s_i within decile
1	0.004
2	0.017
3	0.027
4	0.037
5	0.047
6	0.058
7	0.071
8	0.088
9	0.117
10	0.204

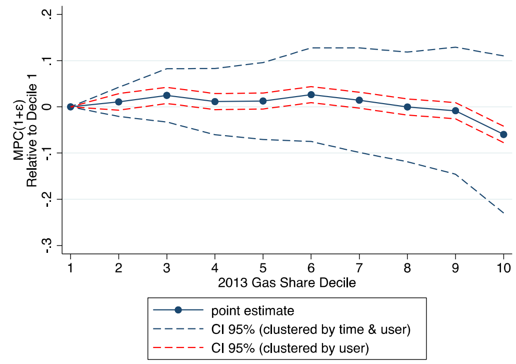
Appendix Figure E.1. Demand elasticity and MPC by declines of s_i.

5 Week Horizon

$(1+\epsilon)$

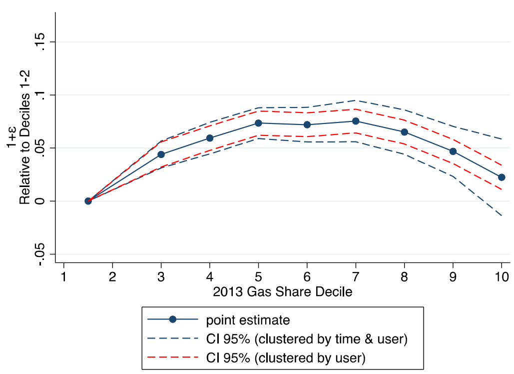


$MPC(1+\epsilon)$

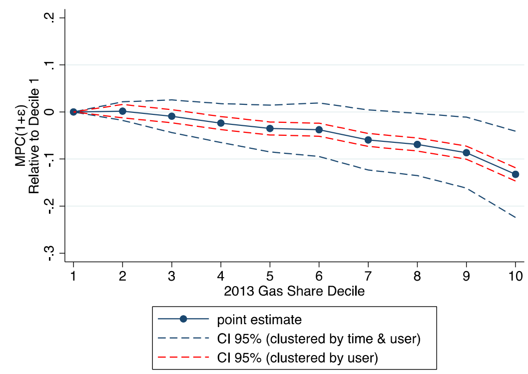


15 Week Horizon

$(1+\epsilon)$

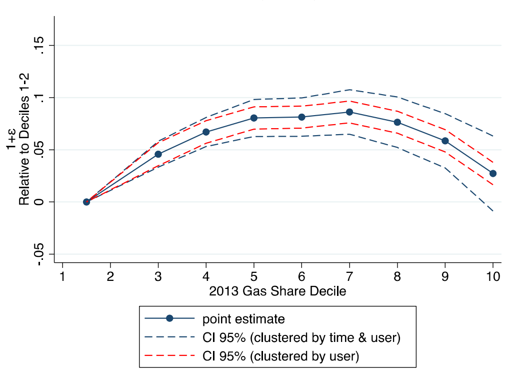


$MPC(1+\epsilon)$



25 Week Horizon

$(1+\epsilon)$



$MPC(1+\epsilon)$