THE RESPONSE OF CONSUMER SPENDING TO CHANGES IN GASOLINE PRICES*

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

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

First draft: August 2016
This draft: January 23, 2022

*This research is carried out in cooperation with a financial aggregation and bill-paying computer and smartphone application (the "app"). We are grateful to the executives and employees who have made this research possible. This research is supported by a grant from the Alfred P. Sloan Foundation. Shapiro acknowledges additional support from the Michigan node of the NSF-Census Research Network (NSF SES 1131500). Gorodnichenko thanks the NSF for financial support. We thank Arlene Wong and conference participants at NBER EFG and NBER SI (Consumption: Micro to Macro) for comments on an earlier draft of the paper.

[a]Claremont McKenna College
[b]University of California, Berkeley
[c]National Bureau of Economic Research
[d]University of Chicago
[e]University of Michigan
[f]Arizona State University

THE RESPONSE OF CONSUMER SPENDING TO CHANGES IN GASOLINE PRICES

This paper estimates how overall consumer spending responds to changes in gasoline prices. It uses the differential impact across consumers of the sharp drop in gasoline prices in 2014 for identification. This estimation strategy is implemented using comprehensive, high-frequency transaction-level data for a large panel of individuals. The average estimated marginal propensity to consume (MPC) out of unanticipated, permanent shocks to income is approximately one. This estimate accounts for the elasticity of demand for gasoline and potential slow adjustment to changes in prices. The high MPC implies that changes in gasoline prices have large aggregate effects.

Michael Gelman
Robert Day School of Economics and Finance
Claremont McKenna College
Claremont, CA 91711
mgelman@cmc.edu

Yuriy Gorodnichenko
Department of Economics
University of California, Berkeley
Berkeley, CA
and NBER
ygorodni@econ.berkeley.edu

Shachar Kariv
Department of Economics
University of California, Berkeley
Berkeley, CA
kariv@berkeley.edu

Dmitri Koustas
Harris Public Policy
University of Chicago
Chicago, IL
dkoustas@uchicago.edu

Matthew D. Shapiro
Department of Economics
and Survey Research Center
University of Michigan
Ann Arbor, MI 48109-1248
and NBER
shapiro@umich.edu

Dan Silverman
Department of Economics
Arizona State University
Tempe AZ
and NBER
Daniel.Silverman.1@asu.edu

Steven Tadelis
Haas School of Business
University of California, Berkeley
Berkeley, CA
and NBER
stadelis@berkeley.edu

I. Introduction

Few macroeconomic variables grab headlines as often and dramatically as do oil and gasoline prices. In 2014, policymakers, professional forecasters, consumers and businesses all wondered how the decline of oil prices from over \$100 per barrel in mid-2014 to less than \$50 per barrel in January 2015 would influence disposable incomes, employment, and inflation. A key component for understanding macroeconomic implications of this shock is the change in consumers' spending from the considerable resources freed up by lower gasoline prices (the average saving was more than \$1,000, or approximately 2 percent of total spending per household).[1] Estimating the quantitative impact of such changes is central to policy decisions. Yet, because of data limitations, a definitive estimate has proved elusive. Recently, big data have opened unprecedented opportunities to shed new light on the matter. This paper uses detailed transaction-level data provided by a personal financial management service over the 2013-2016 period to assess the spending response of consumers to changes in gasoline prices.

Specifically, we use this information to construct high-frequency measures of spending on gasoline and on non-gasoline purchases for a panel of more than half a million U.S. consumers. We use cross-consumer variation in the intensity of spending on gasoline interacted with the sharp decline in gasoline prices to identify and estimate the partial equilibrium marginal propensity to consume (MPC) out of savings generated by reduced gasoline prices. Given the low elasticity of demand for gasoline and the persistence of the oil price shock, one can think of this MPC as measuring the response of spending to a *permanent*, *unanticipated* income shock. Our baseline estimate of the MPC is approximately one. That is, consumers on average spend *all* their gasoline savings on non-gasoline items. There are lags in adjustment, so the strength of the response builds over a period of weeks and months.

Our results are informative along several dimensions. First, our estimate of the average MPC is largely consistent with the permanent income hypothesis (PIH), a theoretical framework that became a workhorse for analyses of consumption, and that has been challenged in previous studies. Second, our findings suggest that, ceteris paribus, falling oil prices can give a considerable boost to the U.S. economy via increased consumer spending (although other factors can offset output growth). Third, we document an important cross-sectional heterogeneity in that MPC

[1] According to the U.S. Consumer Expenditure Survey, average total household spending in 2014 was \$53,495 total, while the average household spending on gasoline was \$2,468.

declines with income. Hence, the average MPC should be interpreted with caution. Fourth, our analysis highlights the value of having high-frequency transaction data at the household level for precisely estimating consumer reactions to income and price shocks.

This paper is related to several strands of research. The first, surveyed in Jappelli and Pistaferri (2010), is focused on estimating consumption responses to income changes. Although the literature on the consumer responses to *anticipated*, *transitory* income shocks is abundant,[2] estimates for *unanticipated, highly persistent* income shocks are rare because identifying such shocks is particularly difficult. For example, income shocks due to job displacements (e.g., Stephens 2001) or health (e.g., Gertler and Gruber 2002) are likely combined with other changes in the lives of affected consumers which makes the identification of MPC challenging.[3] In our paper, we exploit a particularly clear-cut source of variation in household budgets (spending on gasoline) with several desirable properties. Specifically, we use a large, salient, unanticipated, permanent (more precisely, perceived by households to be permanent) shock. We examine spending responses at the weekly frequency—a key ingredient for matching the timing of changes in gasoline prices and the subsequent consumer spending responses—while, due to data limitations, most previous micro-level studies estimate responses at much lower frequencies. As we discuss below, the high-frequency dimension allows us to obtain highly informative estimates of the MPC.

The second strand to which we contribute studies the effects of oil prices on the economy (see Hamilton (2008), Kilian (2008), and Baumeister and Kilian (2016) for surveys). By and large, this literature uses macroeconomic time series to study aggregate reactions to oil price shocks. For example, Edelstein and Kilian (2009) and Känzig (2021) report estimated responses of consumer spending to oil price shocks. Baumeister and Kilian (2016) analyze aggregate data to shed light on

[2] A common finding in this line of research is that, in contrast to predictions of the PIH, consumers often spend only upon the realization of an income shock, rather than upon its announcement, although the size of this "excess sensitivity" depends on household characteristics. Baker (2018) and Kueng (2018) document this pattern using data similar to ours and Gelman et al. (2020) report it for the same data source that we use.

[3] An alternative strategy is to use statistical decompositions in the spirit of Jappelli and Pistafferi (2006) but these estimates of MPC may depend on the assumptions of the statistical models. Changes is taxes may provide a useful source of variation (see e.g., Neri et al. 2017), but it is often hard to identify the timing of these shocks (tax changes are typically announced well before the changes are implemented) and the persistence of shocks (tax changes could be reversed with a change in government). Another option is to use increases in the minimum wage (e.g., Aaronson et al. 2012) but since the minimum wage binds only for low-income households, interpretation of the findings is complicated by the liquidity constraints faced by low-income households.

the nature and macroeconomic consequences of the 2014 oil price shock. Our approach is complementary to this literature as we use micro-level data and cross-sectional variation in exposure to changes in gasoline prices to identify the consumption response and specifically estimate MPC from savings due to lower gasoline prices. Given our high-quality spending data, this approach allows us to obtain precise estimates as well as examine variation in the MPC at the micro level.

Finally, we contribute to the literature studying the interplay between consumer spending on gasoline and non-gasoline purchases at the micro level. Despite the importance of the MPC out of gasoline savings, research on the sensitivity of consumer non-gasoline spending to changes in the gasoline price has been scarce, in part due to data limitations. Available household consumption data tend to be low frequency, whereas consumer spending, gasoline prices, and consumer expectations can change rapidly. For example, the interview segment of the U.S. Consumer Expenditure Survey (CEX) asks households to recall their spending over the previous month. These data likely suffer from recall bias and other measurement errors that could attenuate estimates of households' sensitivity to changes in gasoline prices (see Committee on National Statistics 2013). The diary segment of the CEX has less recall error, but the panel dimension of the segment is short (14 days), making it difficult to estimate the consumer response to a change in prices. Because the CEX is widely used to study consumption, we do a detailed comparison of our approach using the app data with what can be learned from using the CEX. We find that analysis of the CEX produces much noisier estimates.

Given the limitations of the CEX, grocery store barcode data, such as from AC Nielsen, have become a popular source to measure higher-frequency spending. These data, however, cover only a limited category of goods. For example, gasoline spending by households is not collected in AC Nielsen, making it impossible to exploit heterogeneity in gasoline consumption across households. There are a few notable exceptions. Using loyalty cards, Hastings and Shapiro (2013) are able to match grocery barcode data to gasoline sold at a large grocery store retailer with gasoline stations on site. We show that households typically visit multiple gasoline station retailers in a month, suggesting limitations to focusing on consumer purchases at just one retailer. There is also some recent work using household data to identify a direct channel between gasoline prices and non-gasoline spending. Gicheva, Hastings and Villas-Boas (2010) use weekly grocery store data to examine the substitution to sale items as well as the response of total spending. They find that households are more likely to substitute towards sale items when gasoline prices are higher, but they must focus only on a subset of goods bought in grocery stores (cereal, yogurt, chicken and orange juice), making it difficult to extrapolate.

Perhaps the closest work to ours is a policy report produced by the J.P. Morgan Chase Institute (2015), which also uses big data to examine the response of consumers to the 2014 fall in gasoline prices and finds an average MPC of approximately 0.6. This report differs from our study in both its research design and its data. Most importantly, our data include a comprehensive view of spending, across many credit cards and banks. In contrast, the Chase report covers a vast number of consumers, but information on their spending is from Chase accounts only. Namely, any spending by consumers on non-Chase credit cards or checking accounts would be missed in the J.P. Morgan Chase Institute analysis, and measurement of household responses will therefore be incomplete. We confirm this by showing that an analysis based on accounts in one financial institution leads to a considerably attenuated estimate of the response of spending to changes in gasoline prices.

This paper proceeds as follows. Section II describes trends in gasoline prices, putting the recent experience into historical context. In Section III, we discuss the data, Section IV describes our empirical strategy, and Section V presents our results. Specifically, we report baseline estimates of the MPC and the elasticity of demand for gasoline. We contrast these estimates with the comparable estimates one can obtain from alternative data. In Section V we also explore robustness of the baseline estimates and potential heterogeneity of responses across consumers. Section VI concludes.

II. The 2014-2015 Change in Gasoline Prices: Unanticipated, Permanent. Large, and Exogenous

In this section, we briefly review recent dynamics in the prices of oil and gasoline and corresponding expectations of these prices. We emphasize two facts. First, households perceived the collapse of oil and gasoline prices in 2014-2015 as unanticipated and highly persistent. Second, the 2014-2015 price shock had a considerable component due to exogenous supply-side forces. These properties of the shock are important components of our identification strategy.

A. Unanticipated and Permanent

Because we focus on the micro-level consumer responses, we need to establish whether households anticipated the fall of gasoline prices in 2014 and what households believe about the persistence of shocks to gasoline prices. The Michigan Survey of Consumers has asked households about their expectations for changes in gasoline prices over the next one-year and five-year horizons. Panel A of Figure 1 plots the mean and median consumer expectations along with the actual price. While consumers expect a slightly higher price relative to the present price, the basic pattern is clear: the current price appears to be a good summary of expected future prices. In agreement with this observation, Anderson, Kellogg and Sallee (2012) fail to reject the null that consumer expectations for gasoline prices follow a random walk, which is consistent with consumers perceiving changes in gasoline prices as permanent. Panel B plots the forecasts errors for survey responses in the Michigan Survey of Consumers and documents that households were not anticipating large price changes in 2014-2015.[4] To give a reference point, Figure 1 also shows large movements in prices during the Great Recession (2007-2009) when commodity prices endogenously responded to aggregate economic conditions.

When put into historical context, the recent volatility in gasoline prices is large. Table 1 ranks the largest one-month percent changes in oil prices since 1947. When available, the change in gasoline prices over the same period is also shown.[5] The price drops in 2014-2015 are some of the largest changes in oil and gasoline prices in the last 60 years.[6] The 2014-2015 price drop is largely concentrated in just two months (December 2014 and January 2015) which provides us with a clear timing of the shock. Note that in 1986, gasoline prices and oil prices moved in opposite directions, indicating that the process generating gasoline prices can sometimes differ from oil.

---

[4] We report analogous figures for oil price futures in Appendix Figure C6. In contrast to households' expectations, the interpretation of movements in futures prices is more nuanced due to variation in liquidity and other factors, see Baumeister, Ellwanger and Kilian (2017) and Baumeister and Kilian (2016).

[5] Oil spot prices exist back to 1947, while the BLS maintains a gasoline price series for urban areas back to 1976. In our analysis, we use AAA daily gasoline prices retrieved from Bloomberg (3AGSREG). The series comes from a daily survey of 120,000 gasoline stations. These data almost perfectly track another series from the EIA which are point in-time estimates from a survey of 900 retail outlets as of 8am Monday.

[6] At the height of the COVID19 crisis in March and April 2020, oil prices fell by 42 percent in March and 43 percent in April. The price of gasoline declined by 8 percent in March and 17 percent in April.

B. Exogenous

Why did prices of oil and oil products such as gasoline fall so much in 2014-2015? In an early survey of the literature, Baffes et al. (2015) attributed the bulk of the decline to supply-side factors with the more minor demand-side explanations all coming from outside the United States. Specifically, this view emphasizes that key forces behind the decline were, first, OPEC's decision to abandon price support and, second, rapid expansion of oil supply from alternative sources (shale oil in the U.S., Canadian oil sands, etc.). Other work pointed to a smaller role of supply-side factors. For example, Baumeister and Hamilton (2019) assign approximately 40 percent of the decline to supply side and, using a different identification approach, Känzig (2021) finds a larger supply-side contribution. These estimates suggest that the dynamics of oil prices during 2014-2015 were not entirely driven by supply-side forces exogenous to the macroeconomic developments in the U.S. However, exogenous supply-side shocks accounted for a considerable share of variation thus making the 2014-2015 episode comparable to other events that are often used as illustrations of exogenous, supply-side shocks to oil prices. For example, Baumeister and Hamilton (2019) also assign approximately 40 percent of price changes during the 1990-1991 oil price spike due to supply-side shocks after the invasion of Kuwait. In contrast, Hamilton (2009) and others observe that the run up in oil and gasoline prices around 2007-2009 can be largely attributed to booming demand, stagnant production, and speculators, and the consequent decline of the prices during this period, to collapsed global demand (e.g., the Great Recession and Global Financial Crisis). In summary, we view the variation in oil prices during 2014-2015 as having a sufficiently large component that is exogenous to the U.S. consumers.

III. Data

Our analysis uses high-frequency data on spending from a financial aggregation and bill-paying computer and smartphone application (henceforth, the "app").[7] The app had approximately 1.4 million active users in the U.S. in 2013.[8] Users can link almost any financial account to the app,

---

[7] These data have previously been used to study the high-frequency responses of households to shocks such as the government shutdown (Gelman et al. 2020) and anticipated income, stratified by spending, income, and liquidity (Gelman et al. 2014).

[8] All data are de-identified prior to being made available to the project researchers. Analysis is carried out on data aggregated and normalized at the individual level. Only aggregated results are reported.

including bank accounts, credit card accounts, utility bills, and more. Each day, the app logs into the web portals for these accounts and obtains central elements of the user's financial data including balances, transaction records and descriptions, the price of credit and the fraction of available credit used. Using data for a similar service, Baker (2018) documents that over 90 percent of users link all their checking, savings, credit card, and mortgage accounts. Given the non-intrusive automatic data collection, attrition rates are moderate (approximately five percent per quarter).

We draw on the entire de-identified population of active users and data derived from their records from January 2013 until March 2016. The app does not collect demographic information directly and, thus, we are unable to study heterogeneity in responses across demographic groups or to use weights or similar methods to correct possible imbalances in the population of the app's users. However, for a subsample of users, the app employed a third-party that gathers both public and private sources of demographics, anonymizes them, and matches them back to the de-identified dataset. Table 1 in Gelman et al. (2014) (replicated in Appendix Table C1) compares the gender, age, education, and geographic distributions in a subset of the sample to the distributions in the U.S. Census American Community Survey (ACS), representative of the U.S. population in 2012. The app's user population is heterogeneous (including large numbers of users of different ages, education levels, and geographic location) and, along some demographic dimensions, contains proportions similar to those found in the US population. Consistent with this pattern, Baker (2018) observes that, as the online industry had matured, the differences between the population of a similar app's users and the U.S. population became small by 2013. That said, one should bear in mind that this large cross-section (and hence precise estimates) may come at the cost of potentially biased estimates to the extent that our data are not representative of the U.S. population in ways that correlate with spending behavior with respect to gasoline prices. Another limitation of our dataset is that it starts in 2013, which limits us in studying possible pre-trends in the data.

A. Identifying Spending Transactions

Not every debit reported by the app is spending. For example, a transfer of funds from one account to another is not. To avoid double counting, we exclude transfers across accounts, as well as credit card payments from checking accounts that are linked within the app. If an account is not linked, but we still observe a payment, we count this as spending when the payment is made. We identify transfers in several ways. First, we search if a payment from one account is matched to a receipt in another account within several days. Second, we examine transaction description strings to identify common flags like "transfer," "tfr," etc. To reduce the chance of double counting, we exclude the largest single transaction that exceeds $1,000 in a given week, as this kind of transaction is very heavily populated by transfers, credit card payments, and other non-spending payments (e.g., payments to the U.S. Internal Revenue Service). We include cash withdrawals from the counter and ATM in our measure of spending. To ensure that accounts in the app data are reasonably linked and active, we keep all users who were in the data for at least 8 weeks in 2013 and who did not have breaks in their transactions for more than two weeks. We also drop users with cards that we observe to have gone out of sync with the app. More details are provided in Appendix A.

B. Using Machine Learning to Classify Type of Spending

Our analysis requires classification of spending by type of goods. To do so, we address several challenges in using transactional data from bank accounts and credit cards. First, transactional data are at the level of a purchase at an outlet. For many purchases, a transaction will include many different goods. In the case of gasoline, purchases are carried out mainly at outlets that exclusively or mainly sell gasoline. Hence, gasoline purchases are relatively easy to identify in transactional data. Second, for the bulk of transactions in our data, we must classify the outlet from the text of the transaction description, rather than classifications provided by financial institutions. We therefore use a machine learning (ML) algorithm to classify spending based on transaction descriptions. In this section, we provide an outline of the classification routine, and compare our ML predictions in the data provided by the app with external data. As economic analysis increasingly uses naturally-occurring transactional data to replace designed survey data, applications of ML like the one we use will be increasingly important.

The ML algorithm constructs a set of rules for classifying transactions as gasoline or non-gasoline. This requires a training data set to build a classification model, and a testing data set not used in the training step to validate the model predictions. Two of the account providers in the data classify spending directly in the transaction description strings using merchant category codes

(MCCs). MCCs are four-digit codes used by credit card companies to classify spending and are also recognized by the U.S. Internal Revenue Service for tax reporting purposes. Our main MCC of interest is 5541, "Automated Fuel Dispensers." Purchases of gasoline could also fall into MCC code 5542, "Service Stations," which in practice covers gasoline stations with convenience stores.[9] We group transactions with these two codes together because distinguishing transactions as 5542 or 5541 without the MCC is nearly impossible with only the transaction descriptions.[10]

A downside of this approach is that transactions at a Service Station may either be for gasoline, for food and other items, or both. According to the National Association of Convenience Stores (NACS), which covers gasoline stations, purchases of non-gasoline items at gasoline stations with convenience stores (i.e., "Service Stations") account for about 30 percent of sales at "Service Stations." Although the app data do not permit us to differentiate gasoline and non-gasoline items at "Service Stations," we can use transaction data from "Automated Fuel Dispensers" (which do not have an associated convenience store), as well as external survey evidence, to separate purchases of non-gasoline items from purchases of gasoline. Specifically, according to the 2015 NACS Retail Fuels Report (NACS 2015), 35 percent of gasoline purchases are associated with going inside a gasoline station's store. Conditional on going inside the store, the most popular activities are to "pay for gasoline at the register" (42%), "buy a drink" (36%), "buy a snack" (33%), "buy cigarettes" (24%), and "buy lottery tickets" (22%). The last four items are likely to be associated with relatively small amounts of spending. This conjecture is consistent with the distribution of transactions for "Service Stations" and "Automated Fuel Dispensers" in the data we study. In particular, approximately 60 percent of transactions at "Service Stations" are less than \$10 while the corresponding share for "Automated Fuel Dispensers" is less than 10 percent. As we discuss below, the infrequent incidence of gasoline purchases totaling less than \$10 is also consistent with other data sources. Thus, we exclude Service Stations transactions less than \$10 to filter out purchases of non-gasoline items.

Using one of the two providers with MCC information (the one with more data), we train a Random Forest ML model to create binary classifications of transactions into those made at a gasoline station/service station and those that were made elsewhere. Figure 2 shows an example

---

[9] "Service Stations" do not include services such as auto repairs, motor oil change, etc.

[10] E.g., a transaction string with word "Chevron" or "Exxon" could be classified as either MCC 5541 or MCC 5542.

of a decision tree used to classify transactions into gasoline and non-gasoline spending. A tree is a series of rules that train the model to classify a purchase as gasoline or not. The rules minimize the decrease in accuracy when a particular model "feature," in our case transaction values and words in the transaction strings, is removed. In the Figure 2 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." Figure 2 also demonstrates how the decision tree combines transaction string keywords with transaction amounts. For example, "oil" is a very strong predictor of gasoline purchase, but it can be further refined by the transaction amount. The tree continues until all the data are classified.

We then use the second provider to validate the quality of our ML model.[11] The ML model is able to classify spending with approximately 90% accuracy in the testing data set, which is a high level of precision. Both Type I and Type II error rates are low. See Appendix Table B.1. More details on the procedure can be found in Appendix B.

We can also use the app data to investigate which gasoline stations consumers typically visit. The top ten chains of gasoline stations in the app data account for most of gasoline spending. On average, the app data suggest that the typical consumer does 66 percent of his or her gasoline spending in one chain and the rest of gasoline spending is spread over other chains. Thus, while for a given consumer there is a certain degree of concentration of gasoline purchases within a chain, an analysis focusing on only one gasoline retailer, such as in Gicheva, Hastings and Villas-Boas (2010) or Hastings and Shapiro (2013), particularly one not in the top ten chains, would miss a substantial amount of gasoline spending.

C. Comparison with the Consumer Expenditure Survey

[11] Card providers use slightly different transaction strings, and one may be concerned that training the model on a random subsample of data from both card providers, and testing it on another random subsample, can provide a distorted sense of how our ML model performs on data from other card providers. Thus, using a card from one account provider to train, and testing on an entirely different account provider, helps to assure that the ML model is valid outside of the estimation sample. Classification of transactions based on ML applied to both card providers yields very similar results.

We compare our measures of gasoline and non-gasoline spending with similar measures from the Consumer Expenditure Survey (CEX).[12] We use both the CEX Diary Survey and Interview Survey. In the diary survey, households record all spending in written diaries for 14 days. Therefore, this survey provides an estimate of daily gasoline spending that should be comparable to the daily totals we observe in the app. In Figure 3, we compare the distribution of spending in our data (solid lines) and in the diary survey (dashed line). We find that the distributions are very similar, with one notable exception: the distribution of gasoline purchases in the app data has more mass below $10 (solid gray line) than the CEX Diary data. As we discussed above, this difference is likely to be due to our inability to differentiate gasoline purchases and non-gasoline purchases at "Service Stations." In what follows, we restrict our ML predictions to be greater than $10 (solid black line).

The CEX Diary Survey provides a limited snapshot of households' gasoline and other spending. In particular, since a household on average only makes 1 gasoline purchase per week in the diary, we expect only to observe 2 gasoline purchases per household, which can be a noisy estimate of gasoline spending at the household level. Idiosyncratic factors in gasoline consumption that might push or pull a purchase from one week to the next could influence the measure of a household's gasoline purchases by 50% or more. In addition, because the survey period in the diary is so short, household fixed effects cannot be used to control for time-invariant household heterogeneity. Hence, while a diary survey could be a substitute for the app data in principle, the short sample of the CEX diary makes it a poor substitute in practice.[13]

The CEX Interview Survey provides a more complete measure of total spending, as well as a longer panel (4 quarters), from which we can make a comparison with estimates based on spending reported by the app at longer horizons. Panel A of Figure 4 reports the histogram (bin size is set to $1 intervals) of monthly spending on gasoline in the CEX Interview data for 2013-2014.[14] The distribution has clear spikes at multiples of $50 and $100 with the largest spikes at $0

[12] While the definition of the spending unit is different in the CEX ("household") and the app ("user"), Baker (2018) shows for a similar dataset that linked accounts generally cover the whole household.
[13] We have done a comparison of the CEX diary spending for January 2013 through December 2014. In a regression of log daily spending for days with positive spending on month time effects and day of week dummies, the month effects estimated in the CEX and app have a correlation of 0.77. (Finer than monthly comparison of the app and CEX is not possible because the CEX provides only the month and day of week, but not the date, of the diary entry.)
[14] The CEX Interview Survey question asks households to report their "Average monthly expense for gasoline."

and \$200. In contrast, the distribution of gasoline purchases in the app data has a spike at \$0 but the rest of the distribution exhibits considerably less bunching, particularly at large values like \$200 or \$400 that correspond with reporting \$50 or \$100 per week, respectively. In addition, the distribution of gasoline spending has a larger mass at smaller amounts in the app data than in the CEX Interview data. These differences are consistent with recall bias in the CEX Interview Survey data. As argued by Binder (2017), rounding in household surveys can reflect a natural uncertainty of households about how much they spent in this category.

Table 2 compares moments for gasoline and non-gasoline spending across the CEX and the app data. We find that the means are similar across data sources. For example, mean (median) biweekly gasoline spending in the CEX Diary Survey is \$84.72 (\$65.00), while the app counterpart is \$85.82 (\$53.44). Similarly, mean (median) non-gasoline spending is \$1,283.36 (\$790.56) in the CEX Diary Survey and \$1,605.59 (\$1,112.33) in the app data. The standard deviation (interquartile range) tends to be a bit larger in the app data than in the CEX, which reflects a thicker right tail of spending in the app data. This pattern is consistent with top-coding and under-representation of higher-income households in the CEX, a well-documented phenomenon (Sabelhaus et al. 2015). The moments in the CEX Interview Survey (quarterly frequency) are also in line with the moments in the app data. For example, mean (median) spending on gasoline is \$647 (\$540) in the CEX Interview Survey data and \$614 (\$457) in the app data, while the standard deviations (interquartile ranges) are \$531 (\$630) and \$591 (\$657) respectively. In each panel of Table 2, we also compare the distribution of the ratio of gasoline spending to non-gasoline spending, a central ingredient in our analysis. The moments for the ratio in the CEX and the app data are similar. For instance, the mean ratio is 0.08 for the CEX Interview Survey and 0.07 for the app data, while the standard deviation of the ratio is 0.07 for both the CEX Interview Survey and the app data.[15]

In summary, spending in the app data is similar to spending in the CEX data. Thus, although participation in the app is voluntary, app users have spending patterns similar to the population. In addition to reflecting survey recall bias and top-coding, some of the differences could reflect consumers buying gasoline on cards that are not linked to the app (such as credit cards specific to gasoline station chains), the ML procedure missing some gasoline stations, or

[15] Appendix Figure C1 shows the density of the gasoline to non-gasoline spending ratio for the CEX and app data.

gasoline spending done in cash that we could not identify. We will address these potential issues in our robustness tests.

IV. Empirical Strategy

The discourse on potential macroeconomic effects of a fall in gasoline prices often centers on the question of how savings from the fall in gasoline prices are used by consumers. Specifically, policymakers and academics are interested in the MPC from savings generated by reduced gasoline prices. For example, Janet Yellen (Dec 2014) compared the fall in gasoline prices to a tax cut: "[The decline in oil prices] is something that is certainly good for families, for households, it's putting more money in their pockets, having to spend less on gas and energy, so in that sense it's like a tax cut that boosts their spending power."[16] In line with this logic, we define the MPC out of changes in gasoline prices as:

$$dC_{it} \equiv -MPC * d(GasolineSpending_{it}) = -MPC * d(P_t Q_{it}) \qquad (1)$$

where i and t index consumers and time, C is spending on non-gasoline items, P is the price of gasoline, and Q is the quantity of consumed gasoline. Note that we define the *MPC* as an *increase* in spending (measured in dollars) in response to a dollar *decrease* in spending on gasoline after the price of gasoline declines.[17] Note that equation (1) is related to previous studies using tax rebates to measure MPC (e.g., Shapiro and Slemrod, 2003 and Johnson, Parker, and Souleles 2006): a cut in the marginal tax rate on personal income (akin to a change in P_t) increases after-tax earnings (akin to a change in $P_t Q_{it}$), which in turn are translated into a change in consumer spending via the MPC.

Equation (1) is a *definition*, not a behavioral relationship. Of course, Q_{it}, the quantity of gasoline purchased, and overall non-gasoline spending, C_{it}, are simultaneously determined, with simultaneity being an issue at the individual as well as aggregate level. Because Q_{it} is endogenous,

[16] "Highlights: Fed chief Yellen's news conference after FOMC meeting." *Reuters*. 12/12/2014. Available at: https://www.reuters.com/article/us-usa-fed-highlights/highlights-fed-chief-yellens-news-conference-after-fomc-meeting-idUSKBN0JV2O320141217

[17] Because the MPC may differ across groups of people, our notation and estimation refer to the average MPC.

we develop an econometric relationship that yields identification of the MPC based on the specific sources of variation of gasoline prices discussed in the previous sections.

At the aggregate level, one important determinant of gasoline spending is macroeconomic conditions. As discussed in Section II, the 2007-2008 collapse in gasoline prices has been linked to the collapse in global demand due to the financial crisis—demand for gasoline fell, driving down the price of gasoline while demand was falling for other goods as well. Individual-level shocks are another important source of simultaneity bias and threat to identification. Consider a family going on a road trip to Disneyland; this family will have higher gasoline spending (long road trip) and higher total consumption in that week due to spending at the park. Another example is a person who suffers an unemployment spell; this person will have lower gasoline spending (not driving to work) and lower other spending (a large negative income shock).

This discussion highlights that gasoline purchases and non-gasoline spending are affected by a variety of shocks. Explicitly modelling all possible shocks, some of which are expected in advance by households (unobservable to the econometrician), would be impossible. Fortunately, this is not required to properly identify the policy-relevant parameter—the sensitivity of non-gasoline spending to changes in gasoline spending induced by exogenous changes in the price of gasoline. This parameter may be interpreted as a partial derivative of non-gasoline spending with respect to the price of gasoline and thus could be mapped to a coefficient estimated in a regression, for which we only need to satisfy a weaker set of conditions. First, we need exogenous (to households), unanticipated shocks to gasoline prices. These shocks should be unrelated to the regression residual absorbing determinants of non-gasoline consumption unrelated to changes in gasoline prices. Second, we need to link non-gasoline spending to the price of gasoline (i.e., P_t), rather than purchases of gasoline ($P_t Q_{it}$).

As we established in Section II, shocks to gasoline prices in the period of our analysis were unanticipated to households, were perceived by households to be permanent, and they had a considerable exogenous component. To link the partial derivative of interest to a regression coefficient and to link it with cross-sectional variation in pre-determined propensity to spend on gasoline, we manipulate equation (1) as follows:

$$
\begin{aligned}
\frac{dC_{it}}{\bar{C}_i} &= d\log C_{it} = -MPC \times \frac{d(P_tQ_{it})}{\bar{C}_i} = -MPC \times \frac{d(P_tQ_{it})}{(\overline{PQ})_i} \times \frac{(\overline{PQ})_i}{\bar{C}_i} \\
&= -MPC \times \left[\frac{dP_t}{\bar{P}}\left(1 + \frac{dQ_{it}}{\bar{Q}_i} \times \frac{\bar{P}}{dP_t}\right)\right] \times s_i \\
&= -MPC \times (1+\epsilon) \times s_i \times d\log P_t \qquad (2)
\end{aligned}
$$

where bars denote pre-shock values, $s_i \equiv \frac{(\overline{PQ})_i}{\bar{C}_i}$ is the ratio of gasoline spending to non-gasoline spending,[18] and ϵ is the price elasticity of demand for gasoline (a negative number). Now the only source of time variation in the right-hand side of the equation is the price of gasoline. The identifying variation in equation (2) comes from time-series fluctuations in the price of gasoline interacted with the predetermined cross-sectional share of spending on gasoline.[19] The cross-section variation is essential for this paper since there is a single large episode of gasoline price movements in the sample period. One can also derive the specification from a utility maximization problem and link the MPC to structural parameters (see Appendix D). Thus, regressing log non-gasoline spending on the log of gasoline price multiplied by the ratio of gasoline spending to non-gasoline spending yields an estimate of $-MPC(1+\epsilon)$.

Note that we have an estimate of $-MPC$ scaled by $1+\epsilon$, but the scaling should be small if demand is inelastic. As discussed below, there is some variation in the literature on ϵ's estimated using household versus aggregate data. To ensure that a measure of ϵ is appropriate for our sample, we note:

$$
\begin{aligned}
d\log P_tQ_{it} &= d\log P_t + d\log Q_{it} \\
&= d\log P_t + d\log P_t \frac{d\log Q_{it}}{d\log P_t} = \left(1 + \frac{d\log Q_{it}}{d\log P_t}\right) \times d\log P_t \\
&= (1+\epsilon) \times d\log P_t. \qquad (3)
\end{aligned}
$$

[18] We calculate s_i as the ratio of consumer i's annual spending on gasoline to his/her annual spending on non-gasoline items in 2013. Using annual frequency in this instance helps to address seasonal variation in gasoline spending as well as considerable high frequency variation in the intensity of gasoline spending (e.g., trips to gasoline stations, spending per trip). Additionally, the use of 2013 data to calculate the share makes it pre-determined with respect to the shock to gasoline prices in the estimation period. In short, by using s_i for 2013, we approximate the response around the point where gasoline prices were high.

[19] Edelstein and Kilian (2009) and Baumeister and Kilian (2016) consider a similar specification at the aggregate level.

Similar to equation (2), the only source of time variation in the right-hand side of equation (3) is the price of gasoline. Thus, a regression of $d \log P_t Q_{it}$ on $d \log P_t$ yields an estimate of elasticity $(\widehat{1+\epsilon})$, which is the partial derivative of gasoline spending with respect to the price of gasoline, and the residual in this regression absorbs determinants of gasoline purchases unrelated to the changes in the price of gasoline.[20] The estimated $(1+\epsilon)$ and $-MPC(1+\epsilon)$ can be combined to obtain the MPC.[21]

In the derivation of equations (2) and (3) we deliberately did not specify the time horizon over which responses are computed, as these may vary with the horizon. For example, with lower prices, individuals may use their existing cars more intensively or may purchase less fuel-efficient cars. There may be delays in adjustment to changes in prices (e.g., search for a product). Households might take time to process the price change (Coibion and Gorodnichenko 2015). The very-short-run effects may also depend on whether a driver's tank is full or empty when the shock hits.

To obtain behavioral responses over different horizons, we build on the basic derivation above and estimate a multi-period long-differences model, where both the MPC and the price elasticity are allowed to vary with the horizon. Additionally, we introduce aggregate and idiosyncratic shocks to overall spending, and idiosyncratic shocks to gasoline spending. Hence,

$$\Delta_k \log C_{it} = \beta_k \times s_i \times \Delta_k \log P_t + \psi_t + \vartheta_{it} \tag{4}$$

$$\Delta_k \log P_t Q_{it} = \delta_k \Delta_k \log P_t + u_{it} \tag{5}$$

where $\beta_k = -MPC_k(1+\epsilon_k)$, $\delta_k = (1+\epsilon_k)$, $\Delta_k x_t = x_t - x_{t-k}$ is a *k*-period-difference operator, ψ_t is the time fixed effect, and ϑ_{it} and u_{it} are individual-level shocks to spending.[22] ϵ_k measures

[20] Because the dependent variable is spending on gasoline rather than volume of gasoline, elasticity ϵ estimated by this approach also includes substitution across types of gasoline (Hastings and Shapiro 2013).

[21] In this derivation, we implicitly assume that the change in the price of gasoline does not change prices of other goods. This assumption is reasonable given that energy prices account for a small cost share of typical goods produced in the economy. However, some commodities and services (especially energy services, fuels, public transportation) may be more sensitive to changes in the price of gasoline. We find (Appendix Figure C3) that while gasoline prices collapsed in 2014-2015, the retail prices of energy services, fuels, and public transportation showed little (if any) reaction. This weak response likely reflects the fact that prices of these commodities and services are highly regulated.

[22] Note that there are time effects only in equation (4). Since we maintain that changes in gasoline prices are exogenous over the time period, time effects are not needed for consistency of estimation of either (4) or (5). In (4), they may improve

the elasticity of demand over k periods, that is, $\epsilon_k \equiv dlog(Q_{it}/Q_{i,t-k})/dlog(P_t/P_{t-k})$. In a similar vein, $MPC_k \equiv -\beta_k/\delta_k$ measures the MPC from a dollar of savings due to lower gasoline prices over k periods. By varying k, we can recover the average response over k-periods so that we can remain agnostic about how quickly consumers respond to a change in gasoline prices. Given that our specification is in differences, we control for consumer time-invariant characteristics (gender, education, location, etc.) as well as for the level effect of s_i on non-gasoline spending, i.e., time-invariant characteristics are differenced out. To minimize adverse effects of extreme observations, we winsorize dependent variables $\Delta_\mathrm{k} \log C_{it}$ and $\Delta_\mathrm{k} \log P_t Q_{it}$ as well as s_i at the bottom and top one percent.

Because we are interested in the first-round effects of the fall in gasoline prices on consumer spending, we include the time fixed effects in specification (4). As a result, we obtain our estimate after controlling for common macroeconomic shocks and general equilibrium effects (e.g., changes in wages, labor supply, investment). Thus, consistent with the literature estimating MPC for income shocks (e.g., Shapiro and Slemrod 2003, Johnson et al. 2006, Parker et al. 2008, Jappelli and Pistaferri 2010), we estimate a partial equilibrium MPC.

We assume a common price of gasoline across consumers in this derivation. In fact, the co-movement of gasoline prices across locations is very strong (the first principal component accounts for 97 percent of variation in gasoline prices at the city level; Appendix Figure C2 illustrates this strong co-movement) and thus little is lost by using changes in aggregate gasoline prices. Furthermore, when computing s_i we use gasoline spending rather than gasoline prices and thus our measure of s_i takes into account geographical differences in levels of gasoline prices. We find nearly identical results when we use local gasoline prices.

Note that gasoline and oil prices are approximately random walks and thus $\Delta_\mathrm{k}\log P_t$ can be treated as an unanticipated, permanent shock. To the extent oil prices and, hence, gasoline prices are largely determined by global factors or domestic supply shocks, rather than domestic demand—which is our maintained assumption for our sample period—OLS yields consistent

efficiency by absorbing aggregate shocks to overall spending. We cannot include time effects in (5) because they would completely absorb the variation in gasoline prices. But again, note that the presence of an aggregate component in *u* does not make the estimates of δ biased under our maintained assumption that gasoline prices are exogenous to the U.S. economy in the estimation period. The standard errors account for cross-sectional and time dependence in the error term.

estimates of MPC and ϵ. Formally, we assume that the idiosyncratic shocks to spending are orthogonal to these movements in gasoline prices. Given the properties of the shock to gasoline prices in 2014-2015, the PIH model predicts that the response of spending from the resulting change in resources should be approximately equal to the change in resources ($MPC \approx 1$) and take place quickly.

The approach taken in specifications (4) and (5) has several additional advantages econometrically. First, as discussed in Griliches and Hausman (1986), using long differences helps to enhance signal-to-noise ratio in panel data settings. Second, specifications (4) and (5) allow straightforward statistical inference. Because our shock ($\Delta_k \log P_t$) is effectively national and we expect serial within-user correlation in spending, we use Driscoll and Kraay (1998) standard errors. This approach to constructing standard errors is much more conservative than the common practice of clustering standard errors only by a consumer, employer, or location (e.g., Johnson et al. 2006, Levin et al. 2017). To make our results comparable to previous studies, we also report standard errors clustered on users only. Note that we estimate specification (4) and (5) as a system so that we can use the delta method to compute standard errors for MPC from $\hat{\beta}_k/\hat{\delta}_k$. Third, although the variables are expressed in logs, equation (2) shows that we estimate a MPC rather than an elasticity and thus there is no need for additional manipulation of the estimate. This aspect is important in practice because the distribution of spending is highly skewed (in our data, the coefficient of skewness for weekly spending is approximately four) and specifications estimating MPC on levels of spending (rather than logs) are likely sensitive to what happens in the right tail of the spending distribution. Finally, because oil and gasoline prices change every day and the decline in the price of oil (and gasoline) was spread over time, there is no regular placebo test on a "no change" period or before-after comparison. However, these limitations are naturally addressed using regression analysis.

To summarize, our econometric framework identifies the MPC from changes in gasoline prices by interacting two sources of variation: i) a time-series shock to gasoline prices was large, was perceived by households to be permanent, and had a considerable exogenous component; ii) the pre-determined cross-sectional variation in the share of spending on gasoline. The econometric specification also accounts for the response of spending on gasoline to lower prices by allowing a non-zero elasticity of demand for gasoline and allowing for dynamic adjustment of gasoline spending to changes in gasoline prices.

V. Results

In this section, we report estimates of MPC and ϵ for different horizons, frequencies, and populations. We also compare estimates based on our app data to the estimates based on spending data from the CEX.

A. Sensitivity of Expenditure to Gasoline Prices

We start our analysis with the estimates of MPC and ϵ at weekly frequency for different response horizons. Panel A of Figure 5 shows $\hat{\epsilon} = \hat{\delta} - 1$, and 95 percent confidence bands, for $k = 0, \dots, 26$ weeks. Table 3, Row 1, gives the point estimates for selected horizons. The point estimates indicate that the elasticity of demand for gasoline stabilizes around week 15 at -0.16. When we use our conservative Driscoll-Kraay standard errors, confidence intervals are very wide at short horizons; estimates become quite precise at horizons of 12 weeks and longer. In contrast, the conventional practice of clustering standard errors by user yields tight confidence bands, but these likely understate sampling uncertainty in our estimates because there is considerable within-period dependence in the data.[23]

This estimate is broadly in line with previously reported estimates (see Brons et al. (2008) and Espey (1998) for surveys), although it is on the lower end of recent estimates which can reflect differences in the sample of households covered by the app and in the sample period that we study. Using aggregate data, the results in Hughes, Knittel and Sperling (2008) suggest that U.S. gasoline demand is significantly more inelastic today compared with the 1970s. Regressing monthly data on aggregate per capita consumption of gasoline on changes in gasoline prices, they estimate a short-run (monthly) price elasticity of -0.034 to -0.077 for the 2001 to 2006 period, compared with

[23] We have assumed a linear relationship across gasoline spending shares, s_i. Consumers with different s_i may differ in their responses to changes in gas prices. In Appendix E, we test for nonlinearities in our estimates of the MPC and elasticity of demand for gas by s_i deciles. We find only small differences across deciles of s_i, with the lowest and highest-usage consumers being slightly more elastic, by about 0.10 percentage points.

-0.21 to -0.34 for the 1975-1980 period. The U.S. Energy Information Administration (EIA 2014) also points to an elasticity close to zero, and also argues this elasticity has been trending downward over time.[24] In contrast to Hughes, Knittel and Sperling (2008), our findings suggest that gasoline spending could still be quite responsive to gasoline price changes. In general, our results lie in between the Hughes, Knittel and Sperling's estimates and previous estimates using household expenditure data to measure gasoline price elasticities. Puller and Greening (1999) and Nicol (2003) both use the CEX interview survey waves from the 1980s to the early 1990s to estimate the elasticity of demand. The approaches taken across these papers are very different. Nicol's (2003) approach is to estimate a structural demand system. Puller and Greening (1999), on the other hand, take advantage of the CEX modules about miles traveled that were only available in the 1980s, as well as vehicle information. Both papers find higher price elasticities of demand at the quarterly level, with estimates in Nicol (2003) ranging from -0.185 for a married couple with a mortgage and 1 child, to -0.85 for a renter with two children, suggesting substantial heterogeneity across households. Puller and Greening's (1999) baseline estimates are -0.34 and -0.47, depending on the specification. A more recent paper by Levin et al. (2017) uses city level price data and city level expenditure data obtained from Visa credit card expenditures. They estimate the elasticity of demand for gasoline to be closer to ours, but still higher, ranging from −0.27 to −0.35. Their data are less aggregate (MSA level) than the other studies, but more aggregate than ours because we observe individual level data. Also, we observe expenditures from all linked credit and debit cards and are not restricted only to Visa.

Panel C of Figure 5 shows our resulting estimate of the $\widehat{MPC} = \hat{\beta}/\hat{\delta}$ from system estimation, with point estimates at selected horizons in the first row of Table 3. At short time horizons (contemporaneous and up to 3 weeks), the estimates vary considerably from nearly 2 to 0.5 but the estimates are very imprecise when we use Driscoll-Kraay standard errors. Starting with the four-week horizon, we observe that $\widehat{MPC}$ steadily rises over time and becomes increasingly precise. After approximately 12 weeks, $\widehat{MPC}$ stabilizes between 0.8 and 1.0 with a standard error of 0.24. The estimates suggest that, over longer horizons, consumers spend nearly *all* their gasoline

[24] EIA (2014) reports, "The price elasticity of motor gasoline is currently estimated to be in the range of -0.02 to -0.04 in the short term, meaning it takes a 25% to 50% decrease in the price of gasoline to raise automobile travel 1%. In the mid 1990's, the price elasticity for gasoline was higher, around -0.08."

savings on non-gasoline items. The standard errors are somewhat smaller at monthly horizons (4-5 weeks) since the shock. While this pattern is not surprising given that $\hat{\beta}$ and $\hat{\delta}$ in equations (4) and (5) at long horizons have better signal-to-noise ratios, we suspect this is also because the residual variance in consumption tends to be lower at monthly frequency due to factors like frequency of shopping, recurring spending, and bills paid, while in other weeks, the consumption process has considerably more randomness (see Coibion, Gorodnichenko, and Koustas 2021). Similar to the case of ϵ, confidence bands are much tighter when we use standard errors clustered only by user.[25] Note that even with the conservative standard errors, our estimates are quite precise. For comparison, Känzig's (2021) response of consumption estimated on macroeconomic data is not statistically significant at the 10 percent level for any horizon.

There are not many estimates of the MPC derived from changes in gasoline prices. The JPMorgan Institute (2015) report examines the same time period that we do using similar data. It finds an MPC of 0.6, lower than our estimate. This finding likely arises from the use of data from a single financial institution rather than our more comprehensive data. This is an important advantage of the app data because many consumers have multiple accounts across financial institutions. The app's users have accounts on average in 2.6 different account providers (the median is 2). As a result, we have a more complete record of consumer spending. To illustrate the importance of this point, we rerun our specification focusing on a subgroup of consumers with accounts at the top three largest providers.[26] Specifically, we restrict the sample to accounts only at a specific provider so that we can mimic the data observed by a single provider. In rows (2), (4) and (6) of Table 3 we report estimates of ϵ and the MPC at horizons 5, 15 and 25 weeks for the case when we use *any* account at the provider. The MPC estimates based on data observed by a single provider are lower and have larger standard errors than the baseline, full-data MPC estimates reported in row (1). For example, the $\widehat{MPC}$ for Provider 1 (row 2) at the 25-week horizon is 0.462, which is approximately half of the baseline $\widehat{MPC}$ at 0.99. The Driscoll-Kraay standard error for

[25] Our measure of spending covers purchases of durable and non-durable goods. The drop in gasoline prices can stimulate consumers to spend more on cars (there is a modest increase in light weight vehicle sales in 2015; see Appendix Figure C4) so that the MPC may exceed one from new car owners. Unfortunately, MCC codes are too coarse for car dealers (e.g., these codes include repairs, maintenance, leasing) to identify car purchases. Aggregate time series indicate no materially important change in the purchases of electric vehicles around the gasoline price drop in 2014-2015 (see Appendix Figure C5).
[26] These providers cover 49.6 percent of accounts in the data and 55.0 percent of total spending.

the former estimate is 0.3, so that we cannot reject equality of the estimates as well as equality of the former estimate to zero. However, with the conventional practice of clustering standard errors only by user, one can reject equality of the estimates.

One may be concerned that having only one account with a provider may signal incomplete information because the user did not link all accounts with the app. To address this concern, we restrict the sample further to consider users that have at least one checking and one credit-card account with a given provider. In this case, one may hope that the provider is servicing "core" activities of the user. In rows (3), (5) and (7), we re-estimate our baseline specification with this restriction. With the exception of large provider #3, we find estimates largely similar to the case of any account, that is, the estimated sensitivity to changes in gasoline prices is attenuated, and in all cases is more imprecise relative to the baseline where we have accounts linked across multiple providers.

These results for the single-provider data are consistent with the view that consumers can specialize their card use. For example, one card (account) may be used for gasoline purchases while another card (account) may be used for other purchases. In these cases, because single-provider information systematically misses spending on other accounts, MPCs estimated on single-provider data could be attenuated severely. We conjecture that using loyalty cards of a single gasoline retailer may also lead to understated estimates of the MPC because loyalty cards are used only by 18 percent of consumers (NACS 2015).

B. Robustness

While our specification has important advantages, there are nevertheless several potential concerns. First, if s_i in specification (4) is systematically underestimated because a part of gasoline spending is missing from our data, for instance, due to gasoline retailer cards that are not linked to the app, then our estimate of the MPC will be mechanically higher. Second, suppose instead that we are misclassifying some spending, or that consumers buy a large portion of their gasoline in cash, so that this spending shows up in our dependent variable. Misclassifying gasoline spending as non-gasoline spending will generate a positive correlation between non-gasoline spending and the gasoline price. Third, while a random walk may be a good approximation for the dynamics of gasoline prices, one may be concerned that gasoline prices have a predictable component, so that estimated reaction mixes up responses to unanticipated and predictable elements of gasoline prices. Indeed, some changes in gasoline prices are anticipated due to seasonal factors.[27]

A practical implication of the first concern (i.e., cases where consumers use gasoline retailer cards that are not linked to the app) is that consumers with poorly linked accounts should have zero spending on gasoline. To evaluate if these cases could be quantitatively important for our estimates of MPC and ϵ, we estimate specifications (4) and (5) on the sample that excludes households with zero gasoline spending in 2013 (recall that the app data have a larger spike at zero than the counterpart in the CEX Interview Survey). Row (2) of Table 4 reports MPC estimates for this restricted sample at horizons $k = \{5,15,25\}$. We find that these estimates are very close to the baseline reported in row (1).

To address the second concern about cash spending, we note that, according to NACS (2015), less than a quarter of consumers typically pay for gasoline in cash and approximately 80 percent of consumers use credit and debit cards for purchases of gasoline. Furthermore, cash spending only shows up in the dependent variable, generating a positive correlation that will cause us to *underestimate* the MPC. In a robustness check, we exclude ATM and other cash withdrawals from the dependent variable. Row (3) of Table 4 shows that both the MPC and elasticity of demand estimated on these modified data are nearly identical to the baseline estimates. This finding is consistent with the intensity of using cash as means of payment being similar for gasoline and non-gasoline spending.

For the third concern relating to expected changes in gasoline prices, we turn to data from the futures market. In particular, we use changes in one-month-ahead futures for spot prices at New York Harbor (relative to last week's prediction for the month ahead) instead of the change in gasoline prices since last week. Specifically, let F_t^h denote the futures price at time t for month $t+h$. Then, in lieu of $\Delta_\mathrm{k}\log P_t$ in our baseline specification (4), we instead use $\Delta_\mathrm{k}\log\mathcal{F}_t \equiv \log F_t^1 - \log F_{t-k}^1$ for $k \in \{1, \dots ,25\}$. While the focus on one-month change is arguably justified given approximate random walk in gasoline prices, we also try the average change in the futures curves for gasoline prices over longer horizons (two years) to have a measure of changes in gasoline prices that are perceived as persistent: $\Delta_\mathrm{k}\widetilde{\log}\mathcal{F}_t \equiv \frac{1}{24}\sum_{h=1}^{24}\left(\log F_t^h - \log F_{t-k}^h\right)$. In either

[27] In the summer, many states require a summer blend of gasoline which is more expensive than a winter blend.

one-month change (row 4 of Table 4) or average change over two years (row 5), the results are very similar to our baseline.

C. Comparison with MPC using CEX

To evaluate the significance of using high-quality transaction-level data for estimating the sensitivity of consumers to income and price shocks, we estimate the sensitivity using conventional, survey-based data sources such as the Consumer Expenditure Survey (CEX). This survey provides comprehensive estimates of household consumption across all goods in the household's consumption basket and is the most commonly used household consumption survey. In this exercise, we focus on the interview component of the survey which allows us to mimic the econometric analysis of the app data.

In this survey, households are interviewed for 5 consecutive quarters and asked about their spending over the previous quarter. Note that the quarters are not calendar quarters; instead, households enter the survey in different months and are asked about their spending over the previous three months. The BLS only makes available the data from the last 4 interviews; therefore, we have a one-year panel of consumption data for a household. Given the panel design of the CEX Interview Survey, we can replicate aspects of our research design described above. Specifically, we calculate the ratio of gasoline spending to non-gasoline spending in the first interview. We then estimate the MPC in a similar regression over the next three quarters for households in the panel.[28]

In the first row of Table 5, we estimate our baseline specification for the app data at the quarterly frequency. In contrast to the weekly estimates, our estimate of the elasticity of gasoline spending is notably noisier and not statistically different from zero.[29] The estimates for the MPC at a 6-month horizon are slightly lower than the estimates based on the weekly frequency, although

[28] Our build of the CEX data follows Coibion et al. (2017).

[29] In general, we find that aggregation to lower frequencies lowers our elasticity estimate. On one hand, the probability of having no gas spending declines, so more households are identifying the elasticity estimate for each period. In addition, shopping behavior can matter at higher frequencies: suppose households are more likely to "fill up" when gas prices are low, but only put in a few gallons of gas "as needed" when gas prices are high. This results in more weekly transactions and fewer weeks with 0 spending when gas prices are high. We find some evidence of this: the probability of any gas purchase in a week is lower when gas prices are lower.

the Driscoll-Kraay standard errors do not allow us to reject the null of equality of our MPC estimate over time or across frequencies.

Note that in estimates from the app in row 1 we continue to use complete histories of consumer spending over 2013-2016 while the CEX tracks households only for four quarters. To assess the importance of having a long spending series at the consumer level, we "modify" the app data to bring it even closer to the CEX data. Specifically, for every month of our sample, we randomly draw a cohort of app users and track this cohort for only four consecutive quarters, thus mimicking the data structure of the CEX. Then, for a given cohort, we use the first quarter of the data to calculate s_i and use the remainder of the data to estimate ϵ and MPC. Results are reported in row 2 of Table 5. Generally, patterns observed in row 1 are amplified in row 2. In particular, the estimated MPC increases more strongly in the horizon when we track consumers for only four quarters relative to the complete 2013-2016 coverage.

Panel B of Table 5 presents estimates based on the CEX. To maximize the precision of CEX estimates, we begin by applying our approach to the CEX data covering 1980-2015. For this specification, we use BLS urban gasoline prices which provide a consistent series over this time period (see note for Table 1). The point estimates (row 3) indicate that non-gasoline spending *declines* in response to decreased gasoline prices. Standard errors are so large that we cannot reject the null of no response. The estimated elasticity of demand for gasoline is approximately -0.4, which is a double of the estimates based on the app data and is similar to some of the previous CEX-based estimates (e.g., Nicol, 2003).

One should be concerned that the underlying variation of gasoline prices is potentially different across datasets. The dramatic decline in gasoline prices in 2014-2015 had considerable supply-side and foreign-demand components, but it is less clear that one may be equally confident about the dominance of this source of variation over a longer sample period. Indeed, Barsky and Kilian (2004) and others argue that oil prices have often been demand-driven in the past. In this case, one may find a wrong-signed or a non-existent relationship between gasoline prices and non-gasoline spending. To address this identification challenge, we focus on instances when changes in oil prices were arguably determined by supply-side factors.

Specifically, we follow Hamilton (2009, 2011) and consider several episodes with large declines in oil prices: (i) the 1986 decline in oil prices (1985-1987 period); (ii) the 1990-1991 rise and fall in oil prices (1989-1992 period); (iii) the 2014-2015 decline on oil prices. Estimated MPCs and elasticities for each episode are reported in rows (4)-(6). The 1986 episode generates positive MPCs but the standard errors continue to be too high to reject the null of no response. The 2014-2015 episode generates similar, implausible large estimates of MPC, although the estimates are more precise.[30] The 1990-1992 episode yields negative MPCs with large standard errors.

In summary, the CEX-based point estimates are volatile and imprecise. The data are inherently noisy. Moreover, when limited to sample periods that have credibly exogenous variation in gasoline prices, the sample sizes are far too small to make precise inferences. Furthermore, these estimates do not appear to be particularly robust. These results are consistent with a variety of limitations of the CEX data such as small sample size, recall bias, and under-representation of high-income households. These results also illustrate advantages of using high-frequency (weekly) data relative to low-frequency (quarterly) data for estimating sensitivity of consumer spending to gasoline price shocks. The app's comprehensive, high frequency data, combined with a natural experiment—the collapse of oil and gasoline prices in 2014—help us resolve these issues and obtain precise, stable estimates of MPC and elasticity of demand for gasoline.

D. Spending Response by Income

Although we have little demographic information about the app's users, we can use transaction descriptions to gauge some user characteristics and hence examine micro-level heterogeneity in MPC. Specifically, we use payroll inflows, a stable source of income for most users, to construct a measure of permanent income at the user level and study how MPC varies along this dimension.[31] The standard theory predicts that MPC should not vary with permanent income, i.e., [changes in] consumption should be equal to [changes in] permanent income for all people. On the other hand, Straub (2019) argues that MPC can decrease in the level of permanent income if households have non-homothetic preferences over consumption across periods (permanently richer households save a large share of their income). Discriminating between these theories is difficult given challenges

[30] Alexander and Poirier (2020) use CEX data to study the response of consumer spending to the 2014-2015 oil price shock. Using a different empirical approach, they find an MPC that is greater than one.

[31] In appendix E, we also explore how MPC varies with liquidity status and the share of spending on gas. Consistent with PIH, we find that MPC is the same for liquidity-constrained and unconstrained households. We also find that MPC is similar for households with different gas spending shares.

in measuring consumer spending and identifying variation in permanent income. Fortunately, our data and the 2014-2015 oil price shock offer an opportunity to shed more light on this.

Using payroll deposits (after taxes and other deductions) for 2013, we group users into income terciles and estimate the following regressions:

$$\Delta_{\mathrm{k}} \log C_{it} = \beta_1 \times s_i \times \Delta_{\mathrm{k}} \log P_t + \sum_{j=2}^{3}[\beta_j \times s_i^{gas} \times \Delta_{\mathrm{k}} \log P_t \times 1\{Tercile = j\}] + \psi_t + \omega_t \times 1\{Tercile = 2\} + \lambda_t \times 1\{Tercile = 3\} + \vartheta_{it}, \quad (6)$$

$$\Delta_{\mathrm{k}} \log PQ_{it} = \alpha + \delta_1 \times \Delta_{\mathrm{k}} \log P_t + \sum_{j=2}^{3}(\delta_j \times \Delta_{\mathrm{k}} \log P_t \times 1\{Tercile = j\} + \xi_j \times 1\{Tercile = j\}) + u_{it}. \quad (7)$$

This specification is equivalent to running separate regressions by income tercile, i.e., we are focusing on variation within income groups. Results are reported in Table 6.

We find that lower-income households are the most responsive both in terms of their elasticity of demand and MPCs. In particular, estimates of the elasticity of demand for gasoline are significantly different across the income groups at all horizons. Lowest income households have an elasticity of around -0.25, while higher income households have a medium-run elasticity of around -0.08. In terms of the MPC, differences across income groups are small at 5- and 15-week horizons but they become statistically significant at the 25-week horizon. Specifically, we estimate an MPC of 1.02 for the lowest-income tercile, 0.74 for medium income, and 0.64 for the highest income households. Although we cannot reject the null of each estimate being equal to one, these results suggest that the average MPC should be interpreted with caution as the average masks important heterogeneity. These findings are consistent with the predictions in Straub (2019) and, hence, can contribute to our understanding of U.S. trends in macroeconomic aggregates (e.g., a decline in interest rates) and inequality.[32]

VI. Conclusion

How consumers respond to changes in gasoline prices is a central question for policymakers and researchers. We use big data from a personal financial management service to examine the

[32] Previous research studying the Alaska permanent fund found the opposite, that the MPC was increasing in income (Kueng 2018), however one important difference is that these payments are largely predictable.

dynamics of consumer spending during the 2014-2015 period when gasoline prices plummeted by 50 percent. Given the low elasticity of demand for gasoline, this major price reduction generated a large windfall for consumers equal to approximately 2 percent of total consumer spending. We document that on average, the MPC out of these savings is approximately one. Since the change in gasoline prices was unexpected and permanent, this estimate can be interpreted as capturing *MPC out of permanent income*, an object that has been most difficult to estimate with previously available data.

While estimating the macroeconomic effects of the change in oil prices is beyond the scope of this paper, this partial equilibrium estimate provides a first-step input for quantifying the effects on the aggregate economy, which depend on several factors. The aggregate effects of changes in gasoline prices potentially depend on general equilibrium effects and redistribution of resources in the economy. The aggregate response to a gasoline price shock may be a function of the sensitivity of, for example, sectoral wages and employment to energy price shocks (see Appendix D for a model). Depending on specific assumptions about utility and production functions, general equilibrium effects can amplify or attenuate the immediate effects that we estimate. Moreover, there are income effects arising from the ownership of energy resources both domestically and abroad that will have macroeconomic effects. Nevertheless, any offsetting macroeconomic effects, e.g., from changes in oil field production or from exports to foreign, oil-rich countries, do not obviate the interest in estimates of response of U.S. consumers to a very significant shock to their budget sets coming from gasoline prices.

We also show why previous attempts to estimate the MPC out of gasoline savings led to lower and/or more imprecise estimates due to data limitations (e.g., low frequency of data, incomplete coverage of consumer spending, short panel) in earlier studies. Our analysis highlights the substantial potential of big data from household financial accounts for enhancing national economic statistics, as well as estimates of key, policy-relevant macroeconomic parameters.

References

Aaronson, Daniel, Sumit Agarwal and Eric French. 2012. “The Spending and Debt Response to Minimum Wage Hikes,” *American Economic Review* 102(7): 3111-3139.

Alexander, Patrick, and Louis Poirier. 2020. "Did U.S. Consumers Respond to the 2014-2015 Oil Price Shock? Evidence from the Consumer Expenditure Survey," *Energy Journal* 41(1).

Anderson, Soren T., Ryan Kellogg and James M. Sallee. 2012. "What do consumers believe about future gasoline prices?" *Journal of Environmental Economics and Management* 66(3): 383-403.

Baffes, John, M. Ayhan Kose, Franziska Ohnsorge, and Marc Stocker, 2015. "The Great Plunge in Oil Prices: Causes, Consequences, and Policy Responses," World Bank Group Policy Research Note #15/01.

Baker, Scott, 2018. "Debt and the Consumption Response to Household Income Shocks," *Journal of Political Economy* 126(4): 1504-1557

Barsky, Robert B., and Lutz Kilian. 2004. "Oil and the Macroeconomy Since The 1970s," *Journal of Economic Perspectives* 18(4): 115-134.

Baumeister, Christiane, and James D. Hamilton. 2019. "Structural Interpretation of Vector Autoregressions with Incomplete Identification: Revisiting the Role of Oil Supply and Demand Shocks," *American Economic Review* 109(5): 1873-1910.

Baumeister, Christiane, and Lutz Kilian. 2016. "Lower Oil Prices and the U.S. Economy: Is This Time Different?," *Brookings Papers on Economic Activity* 47(2 (Fall)): 287-357.

Baumeister, Christiane, and Lutz Kilian. 2016. "Forty Years of Oil Price Fluctuations: Why the Price of Oil May Still Surprise Us," *Journal of Economic Perspectives* 30(1): 139-160.

Baumeister, Christiane, Reinhard Ellwanger, and Lutz Kilian. 2017. "Did the Renewable Fuel Standard Shift Market Expectations of the Price of Ethanol?" NBER Working Paper 23752.

Binder, Carola. 2017. "Measuring Uncertainty Based on Rounding: New Method and Application to Inflation Expectations." *Journal of Monetary Economics* 90: 1-12.

Brons, Martijn, Peter Nijkamp, Eric Pels, and Piet Rietveld. 2008. "A meta-analysis of the price elasticity of gasoline demand. A SUR approach," *Energy Economics* 30(5): 2105-2122.

Coibion, Olivier, and Yuriy Gorodnichenko. 2015. "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts," *American Economic Review* 105: 2644–2678.

Coibion, Olivier, Yuriy Gorodnichenko, Lorenz Kueng, and John Silvia. 2017. "Innocent Bystanders? Monetary Policy and Inequality in the U.S." *Journal of Monetary Economics* 88: 70-89.

Coibion, Olivier, Yuriy Gorodnichenko, and Dmitri Koustas. 2021. "Consumption Inequality and the Frequency of Purchases," *American Economic Journal: Macroeconomics.* 13(4): 449-482.

Committee on National Statistics. 2013. *Measuring What We Spend: Toward a New Consumer Expenditure Survey*, Don A. Dillman and Carol C. House, Editors. Washington, D.C.: National Academies Press.

Driscoll, John C., and Aart C. Kraay 1998. "Consistent Covariance Matrix Estimation With Spatially Dependent Panel Data," *Review of Economics and Statistics* 80(4): 549-560.

Edelstein, P., and Kilian, L. (2009). "How sensitive are consumer expenditures to retail energy prices?" *Journal of Monetary Economics* 56(6): 766–779.

Energy Information Administration (EIA). 2014. "Gasoline prices tend to have little effect on demand for car travel." Last accessed June 2, 2016. Available at: http://www.eia.gov/todayinenergy/detail.cfm?id=19191.

Espey, Molly. 1998. "Gasoline demand revisited: an international meta-analysis of elasticities," *Energy Economics* 20(3): 273-295.

Gelman, Michael, Shachar Kariv, Matthew D. Shapiro, Dan Silverman and Steven Tadelis. 2020. "How Individuals Respond to a Liquidity Shock: Evidence from the 2013 Government Shutdown." *Journal of Public Economics* 189:103917

Gelman, Michael, Shachar Kariv, Matthew D. Shapiro, Dan Silverman and Steven Tadelis. 2014. "Harnessing Naturally Occurring Data to Measure the Response of Spending to Income," *Science* 345(6193): 212-215.

Gertler, Paul, and Jonathan Gruber. 2002. "Insuring Consumption Against Illness." *American Economic Review* 92(1): 51-70.

Gicheva, Dora, Justine Hastings and Sofia Villas-Boas. 2010. "Investigating Income Effects in Scanner Data: Do Gasoline Prices Affect Grocery Purchases?" *American Economic Review: Papers & Proceedings* 100: 480-484.

Griliches, Zvi, and Jerry A. Hausman. 1986. "Errors in variables in panel data," *Journal of Econometrics* 31(1): 93-118.

Hamilton, James D. 1983. "Oil and the Macroeconomy since World War II." *Journal of Political Economy* 91(2): 228-248.

Hamilton, James D. 2009. "Causes and Consequences of the Oil Shock of 2007-08." *Brookings Papers on Economic Activity.* Spring 2009. 215-283.

Hamilton, James D. 2011. "Historical Oil Shocks," in *Routledge Handbook of Major Events in Economic History*, pp. 239-265, edited by Randall E. Parker and Robert Whaples, New York: Routledge Taylor and Francis Group, 2013.

Hamilton, James. 2008. "Oil and the Macroeconomy," in *New Palgrave Dictionary of Economics*, 2nd edition, edited by Steven Durlauf and Lawrence Blume, Palgrave McMillan Ltd.

Hastings, Justine S. and Jesse M. Shapiro. 2013. "Fungibility and Consumer Choice: Evidence from Commodity Price Shocks." *Quarterly Journal of Economics* 128(4): 1449-1498.

Hughes, Jonathan E., Christopher R. Knittel and Daniel Sperling. 2008. "Evidence of a Shift in the Short-Run Price Elasticity of Gasoline Demand." *Energy Journal* 29(1): 93-114.

J.P. Morgan Chase Institute. 2015. *How Falling Gas Prices Fuel the Consumer: Evidence from 25 Million People.* Available at https://www.jpmorganchase.com/corporate/institute/document/jpmc-institute-gas-report.pdf.

Jappelli, Tullio, and Luigi Pistaferri. 2006. "Intertemporal Choice and Consumption Mobility," *Journal of the European Economic Association* 4(1): 75-115.

Jappelli, Tullio, and Luigi Pistaferri. 2010. "The Consumption Response to Income Changes," *Annual Review of Economics* 2(1): 479-506.

Johnson, David S., Jonathan A. Parker and Nicholas S. Souleles. 2006. "Household Expenditure and the Income Tax Rebates of 2001." *American Economic Review* 96(5): 1589-1610.

Känzig, Diego R. 2021. "The Macroeconomic Effects of Oil Supply News: Evidence from OPEC Announcements," *American Economic Review* 111(4): 1092-1125.

Kilian, Lutz. 2008. "The Economic Effects of Energy Price Shocks." *Journal of Economic Literature* 46(4): 871-909.

Kilian, Lutz. 2009. "Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market." *American Economic Review* 99(3): 1053-69.

Kueng, Lorenz. 2018. "Excess Sensitivity of High-Income Consumers." *Quarterly Journal of Economics* 133: 1693-1751.

Levin, Laurence, Matthew S. Lewis and Frank A. Wolak. 2017. "High Frequency Evidence On The Demand For Gasoline." *American Economic Journal: Economic Policy* 9(3): 314-47.

National Association of Convenience Stores (NACS). 2015. *Retail Fuels Report*. Available at http://www.nacsonline.com/YourBusiness/FuelsReports/2015/Documents/2015-NACS-Fuels-Report_full.pdf .

Neri, Andrea, Concetta Rondinelli, and Filippo Scoccianti. 2017. "Household spending out of a tax rebate: Italian "€80 tax bonus"," Working Paper Series 2099, European Central Bank.

Nicol, C.J. 2003. "Elasticities of demand for gasoline in Canada and the United States." *Energy Economics* 25: 201-214.

Parker, Jonathan A., Nicholas S. Souleles, David S. Johnson and Robert McClelland. 2013. "Consumer Spending and the Economic Stimulus Payments of 2008." *American Economic Review* 103(6): 2530-2553.

Puller, Steven and Lorna A. Greening. 1999. "Household adjustment to gasoline price change: an analysis using 9 years of US survey data." *Energy Economics* 21: 37-52.

Sabelhaus, John, David Johnson, Stephen Ash, David Swanson, Thesia I. Garner, John Greenlees, and Steve Henderson. 2015. "Is the Consumer Expenditure Survey Representative by Income?" in Carroll, Crossley, and Sabelhaus, eds., *Improving the Measurement of Consumer Expenditures*, University of Chicago Press.

Shapiro, Matthew D., and Joel Slemrod. 2003. "Consumer Response to Tax Rebates." *American Economic Review* 93(1): 381-396.

Stephens, Melvin. 2001. "The Long-Run Consumption Effects Of Earnings Shocks," *Review of Economics and Statistics* 83(1): 28-36.

Straub, Ludwig. 2019. "Consumption, Savings, and the Distribution of Permanent Income." Manuscript. Available at https://scholar.harvard.edu/files/straub/files/cons_ineq_rates.pdf.