Understanding Household Consumption and Saving Behavior using Account Data

by

Michael L. Gelman

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in the University of Michigan
2017

Doctoral Committee:

Professor Matthew D. Shapiro, Chair
Professor John V. Leahy
Professor Martin C. Schmalz
Professor Melvin Stephens Jr.

Michael L. Gelman

mgelman@umich.edu

ORCID iD: 0000-0002-1714-7442

For my family and DC

ACKNOWLEDGEMENTS

First and foremost, I would like to thank Matthew Shapiro for his guidance and support throughout the dissertation writing process. I would also like to thank the members of my dissertation committee John V. Leahy, Martin C. Schmalz, and Melvin Stephens Jr., for their invaluable advice and feedback. I also benefited greatly from discussions with Daphne Chang, Michael Gideon, Gaurav Khanna, Minjoon Lee, Dhiren Patki, Dan Silverman, Ben Thompson, Mike Zabek, Fudong Zhang, and Xiaoqing Zhou. I acknowledge fellowship support from the Rackham Graduate School at the University of Michigan and the F. Thomas Juster Economic Behavior Research Award. I also thank my family for their support and understanding. Lastly, I am heavily indebted to Daphne for her endless patience and support throughout the past few years. I would not have made it this far without her.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

ABSTRACT

Understanding Household Consumption and Saving Behavior using Account Data

by

Michael Gelman

Chair: Matthew D. Shapiro

This dissertation seeks to better understand household consumption and saving behavior using account data from a personal finance app.

The first chapter examines the result that cash on hand is the most important source of variation in explaining heterogeneity in the marginal propensity to consume (MPC). While the standard hypothesis is that differences in financial circumstances caused by temporary income shocks explain this result, this paper finds that differences across persistent characteristics are just as important. To reach this finding, this paper develops a buffer stock model with discount factor heterogeneity and estimates it using a novel panel data set from a personal finance app that jointly measures spending, income, and liquid assets. In the model, within-individual variation in cash on hand results from temporary income shocks while across-individual variation in cash on hand results from differences in persistent characteristics. The panel nature of the data separately identifies temporary and persistent drivers of the MPC while previous studies using cross-sectional data typically confound these concepts. Simulations from the estimated model imply that ignoring heterogeneity in persistent

characteristics leads to underestimating the aggregate MPC.

The second chapter examines how individuals adjusted spending and saving in response to a temporary drop in income due to the 2013 U.S. government shutdown. The shutdown cut paychecks by 40% for affected employees, which was recovered within 2 weeks. Because it affected only the timing of payments, the shutdown provides a distinct experiment allowing estimates of the response to a liquidity shock holding income constant. Spending dropped sharply implying a naïve estimate of the marginal propensity to spend of 0.58. This estimate overstates how consumption responded. While many individuals had low liquidity, they used multiple strategies to smooth consumption including delay of recurring payments such as mortgages and credit card balances. This is joint work with Shachar Kariv, Matthew D. Shapiro, Dan Silverman, and Steven Tadelis.

The third chapter estimates how overall consumer spending responds to changes in gasoline prices. It uses the differential impact across consumers of the sudden, large drop in gasoline prices in 2014 for identification. This estimation strategy is implemented using comprehensive, daily transaction-level data for a large panel of individuals. The estimated marginal propensity to consume (MPC) is approximately one, a higher estimate than that found in less comprehensive or well-measured data. This estimate takes into account 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. This is joint work with Yuriy Gorodnichenko, Shachar Kariv, Dmitri Koustas, Matthew D. Shapiro, Dan Silverman, and Steven Tadelis.

The fourth chapter examines the response of food expenditures to the receipt of paychecks using financial account data from a personal finance app. Similar to previous studies, this paper finds that food expenditures increase during the week the paycheck is received. While the standard explanation for this result is temporary liquidity constraints, this paper argues otherwise. Intuitively, it's unlikely that individuals will be liquidity constrained during the weeks they receive their paycheck. Therefore, their decision to spend more during weeks in which they have more liquidity likely reflects preferences and not constraints. The intuition is formalized through specifying a buffer stock model of consumption. Model simulations show that indeed consumption behavior is not affected by liquidity during the week of the paycheck. The empirical results match the theoretical predictions and confirm that liquidity constraints cannot explain excess sensitivity.

CHAPTER I

What Drives Heterogeneity in the Marginal Propensity to Consume? Temporary Shocks vs Persistent Characteristics

1.1 Introduction

The marginal propensity to consume (MPC) out of income changes is of interest to both policymakers and academics. Studies analyzing the MPC have played a prominent role in government reports documenting and forecasting the macroeconomic effects of fiscal stimulus (Congressional Budget Office (2009), Council of Economic Advisers (2010)). Moreover, academics study the MPC out of various forms of changes in income to evaluate theoretical models of consumption (see Jappelli and Pistaferri (2010) for an excellent survey).

A key result in the empirical literature is that individuals with low financial resources (cash on hand) tend to have a higher MPC (See for example Parker et al. (2013), Jappelli and Pistaferri (2014), and Parker (2015)). Yet the literature is divided over the theoretical mechanisms that drive the negative correlation between the MPC and cash on hand. The lack of consensus stems from the fact that most studies analyzing the correlation between the MPC and cash on hand use cross-sectional data that confounds the various theoretical mechanisms. For example, a cross-sectional snapshot of cash on hand may be determined either by recent temporary shocks to income or persistent characteristics such as time preference. The first contribution of this paper is to overcome this identification obstacle by developing a novel panel data set that captures the spending response to multiple tax refunds over several years. The second contribution is to elucidate the theoretical mechanisms that drive MPC heterogeneity and to map these mechanisms to the empirical results by specifying a parsimonious buffer stock model with discount factor heterogeneity. The third contribution is to show through model simulations that ignoring heterogeneity in persistent characteristics leads to underestimating the aggregate MPC.

In general, there are a plethora of mechanisms that can explain the negative correlation between the MPC and cash on hand. In order to make the discussion manageable, I follow the dichotomy laid out in Parker (2015) between the two main classes of models used to explain MPC heterogeneity. One view is that temporary income shocks combined with precautionary savings or borrowing constraints play the main role. Some examples include the textbook buffer stock model with ex-ante identical individuals (Zeldes (1989), Deaton (1991), Carroll (1997)) and the wealthy hand-to-mouth model of Kaplan and Violante (2014). Another view is that persistent characteristics such as preferences or behavioral traits are the root cause. This may arise from simple impatience such as in Campbell and Mankiw (1989) and Krusell and Smith (1998). It may also arise from more complex mechanisms such as limited attention, problems of self-control, or propensity to plan as in Reis (2006), Angeletos et al. (2001), or Ameriks, Caplin and Leahy (2003). Simply put, the two views in the literature boil down to temporary circumstances versus persistent characteristics and hence I use the terms "circumstances view" and "characteristics view" to distinguish the two.

The main impediment to disentangling these two views is that circumstances and characteristics are not easily separately identified in existing datasets. Since circumstances vary over time while characteristics are constant, observing both within-person cash on hand and MPC over time is vital to identification. Most data sets, however, only allow researchers to estimate the cross-sectional relationship between the MPC and cash on hand. For example, the Consumer Expenditure Survey (CEX) has detailed enough data to identify the consumption response to income changes, but lacks a long enough panel structure to estimate multiple MPCs within an individual. Conversely, the Panel Study of Income Dynamics (PSID) has a long panel element, but lacks enough detail to isolate the source of income changes. Without a combination of a long panel and detailed consumption, income, and liquid balance data, it is difficult to disentangle circumstances from characteristics. Perhaps the study that comes closest to disentangling circumstances from characteristics is Sahm, Shapiro and Slemrod (2012). They directly ask individuals how two separate policy-induced income changes affected their spending behavior. Their results show that changes in within-individual financial conditions can explain differences in spending behavior. Unfortunately, they do not have precise liquidity measures.

The first contribution of this paper is to empirically decompose the fraction of MPC variance explained by within- and across-individual differences in cash on hand. The key data innovation is developing a novel panel dataset that includes joint spending, income, and liquid saving behavior from a personal finance app over several years. Using the detailed app data, I identify the receipt of several federal tax refunds within the same individual. I then estimate the monthly spending response using the high-frequency spending observations. Finally, I use the high-frequency liquid balance data to capture within- and across-individual variation in cash on hand. I find that within- and across-individual differences in cash on hand play roughly equal roles in explaining MPC variance. This is consistent with the results in Parker (2015) that show persistent characteristics such as time preferences are an important factor in explaining heterogeneity in the MPC.

The second contribution of this paper is to interpret the empirical results I find through the lens of a buffer stock saver model with discount factor heterogeneity.

This relatively parsimonious model is able to capture the role of both circumstances and characteristics. The role of circumstances is reflected in the model by temporary shocks to income which induce within-individual differences in cash on hand. The role of characteristics is reflected in the model by heterogeneity in the discount factor which induces across-individual differences in cash on hand. Holding the variance of temporary shocks constant, a higher dispersion in the discount factor will lead to a more prominent role of across-individual variation in explaining MPC variance. Using this logic, the mean and dispersion of the discount factor is estimated from the data using the method of simulated moments. The estimates are roughly in line with the literature and show that this procedure produces sensible results.

The third contribution of the paper is to use the estimated model to evaluate the implications for fiscal stimulus. I estimate the model separately under the circumstances and characteristics view. Under the characteristics view, heterogeneity in persistent characteristics leads to a higher aggregate MPC because the high MPC for impatient individuals outweighs the low MPC for patient individuals. This effect is amplified if temporary income shocks due to a recession are disproportionately concentrated on impatient individuals. The simulations show that under the characteristics view where persistent characteristics are important, the distribution of preferences will influence the aggregate MPC. Using the estimated parameters from the data used in this paper, ignoring these persistent characteristics leads to underestimating the aggregate MPC.

The rest of the paper is organized as follows. Section 4.4 lays out the theoretical framework I use to generate predictions about consumption and saving behavior under the two views which I will take to the data. Section 4.2 discusses the dataset and provides some descriptive statistics. Section 1.4 presents the empirical results used to evaluate which view is more consistent with the data. Section 1.5 estimates the parameters of the model via the method of simulated moments. Section 1.6 discusses policy implications and section 1.7 concludes.

1.2 Theoretical framework

This section describes the theoretical framework used to analyze individual decisions. It introduces a buffer stock model with discount factor heterogeneity and formally defines the circumstances versus the characteristics view of MPC heterogeneity. It then generates predictions about MPC heterogeneity which are taken to the data in later sections.

1.2.1 Model description

Individuals behave according to the standard "buffer-stock" saver model in the spirit of Zeldes (1989), Deaton (1991), and Carroll (1997). The main difference with previous studies is the introduction of preference heterogeneity via the discount factor signified by the i subscript on β.

Optimization problem Individual i solves the following utility maximization problem

$$\max_{\{C_{ij}\}_{j=t}^{\infty}} \mathbb{E}_t \left[\sum_{j=t}^{\infty} \beta_i^{j-t} \frac{{C_{ij}}^{1-\theta}}{1-\theta} \right] \tag{1.1}$$

subject to

$$A_{it+1} = (1+r)\left(A_{it} + Y_{it} - C_{it}\right) \tag{1.2}$$

$$A_{it+1} \geq b \tag{1.3}$$

$$Y_{it} = \bar{Y}_i(1-\rho) + \rho Y_{it-1} + \varepsilon_{it} \tag{1.4}$$

$$\varepsilon_{it} \overset{iid}{\sim} N(0, \sigma_Y^2) \tag{1.5}$$

where $\beta_i, r, C_{it}, A_{it}$ and Y_{it} represent the time discount factor, the interest rate, consumption, liquid assets, and income respectively.

Normalization Carroll (2004) showed that this problem can be rewritten by normalizing all variables by the level of permanent income. Following his notation, I define lowercase variables as uppercase variables divided through by the level of permanent income. Therefore $c_{it} = C_{it}/\bar{Y}_i, a_{it} = A_{it}/\bar{Y}_i$ and so on. This normalization is very useful because the same solution to the model can be used to jointly characterize the behavior of all individuals who share the same β_i and Y_{it} process while allowing the actual level of $\bar{Y}_i$ to differ.

Model Horizon An infinite horizon version of the model is chosen to abstract away from life cycle features. Carroll (2004) shows that the infinite horizon framework can be thought of as the limiting behavior of an individual when they are far away from their end of life. This assumption is reasonable for the population analyzed in this paper and will be discussed further in the data section. When buffer stock motives are strong enough, agents are more concerned with smoothing short term shocks rather than saving for retirement.

Income process Similarly to Zeldes (1989) and Deaton (1991), income follows an AR(1) processes. Because the time series of the data only span 4 years, permanent shocks are not well identified. To match the model, the subsequent empirical analysis will condition on individuals who have a fairly stable income process and therefore have not experienced any large permanent shocks in the data.

Solution The consumption problem specified above does not admit a closed form solution and is therefore solved computationally. I reformulate the individual's problem in terms of a functional equation and define cash on hand $x_{it} = a_{it} + y_{it}$ to simplify the state space. This variable represents the amount of resources available to the individual in the beginning of the period.

The individual then solves the optimization problem

$$V(x_{it}) = \max_{a_{it+1}} \{u(c_{it}) + \beta_i \mathbb{E}[V(x_{it+1})]\} \tag{1.6}$$

subject to

$$x_{it+1} = (1+r)\,(x_{it} - c_{it}) + y_{it+1} \tag{1.7}$$

and the previous constraints (4.4), (4.5), and (4.6).

Substituting in for c_{it} and x_{it+1} results in an equation in terms of x_{it}, a_{it+1}, and y_{it+1}

$$V(x_{it}) = \max_{a_{it+1}} \left\{ u\left(x_{it} - \frac{a_{it+1}}{1+r}\right) + \beta_i \mathbb{E}[V(a_{it+1} + y_{it+1})] \right\} \tag{1.8}$$

The individual maximizes utility by choosing next period saving (a_{it+1}) conditional on cash on hand (x_{it}). The model is solved using the method of endogenous gridpoints suggested in Carroll (2006). This solution method results in the value function $V(x_{it})$ and the policy function $a_{it+1}(x_{it})$ which maps the state variables x_{it} into the optimal control variable a_{it+1}. The consumption function is calculated using constraint (4.4) so that $c_{it}(x_{it}) = x_{it} - \frac{a_{it+1}}{1+r}$.

1.2.2 Circumstances and characteristics view

In order to understand the mechanisms that drive MPC heterogeneity, I adopt the dichotomy laid out in Parker (2015) between classes of models that can explain the relationship between cash on hand and the MPC. In the first class of models, temporary circumstances cause cash on hand to fluctuate. If individuals have concave consumption functions, low cash on hand leads to high MPCs and high cash on hand leads to low MPCs. Therefore, the MPC will depend on what circumstances individuals find themselves in and so I call this view the "circumstances view." Some examples include the textbook buffer stock model with ex-ante identical individuals (Zeldes (1989), Deaton (1991), Carroll (1997)) and the wealthy hand-to-mouth model of Kaplan and Violante (2014). In the second class of models, persistent characteristics drive the correlation between cash on hand and the MPC. This may arise from simple impatience such as in Campbell and Mankiw (1989) and Krusell and Smith (1998). It may also arise from more complex mechanisms such as limited attention, problems of self-control, or propensity to plan as in Reis (2006), Angeletos et al. (2001), or Ameriks, Caplin and Leahy (2003). Therefore, even though individuals may find themselves in good or bad circumstances, their average behavior over time will depend on differences across persistent characteristics such as the discount factor. I call this view the "characteristics view."

In the model described in the previous section, temporary shocks to income capture temporary circumstances while heterogeneity in the discount factor captures persistent characteristics. In general, "characteristics" may refer to a broad range of traits such as impatience, risk aversion, present bias, and inattention. I choose to parametrize characteristics as heterogeneity in the discount factor for two reasons.

The first reason is that recent studies suggest heterogeneity in the discount factor may be important for explaining the heterogeneity in the MPC. Parker (2015) shows that lack of smoothing is correlated not with temporary fluctuations but with persistent characteristics such as impatience.[1] He concludes that this behavior is consistent with models that exhibit heterogeneity in preference such as Campbell and Mankiw (1989), Krusell and Smith (1998), and Hurst (2003). Along a similar vein, Baugh, Ben-David and Park (2014) study the weekly response of spending to the receipt of a tax refund and find a strong immediate spending response which decays very rapidly. They argue that agents who are constrained but patient would exhibit a spike up in spending but would then smooth spending over the following weeks. Therefore they conclude that the spending response to tax refunds is consistent with some agents

[1]The measure is the answer to the question "In general, are you or other household members the sort of people who would rather spend your money and enjoy it today or save more for the future?" with a binary choice of 'spend now' and 'save for the future.'

who exhibit myopia.

The second reason I choose to model characteristics as heterogeneity in the discount factor is that for purposes of modeling consumption behavior, the MPC is largely a function of the curvature of the consumption function. Changes in the discount factor alter the curvature of the consumption function is similar ways to changes in risk aversion. Therefore, whether heterogeneity is introduced via the discount factor or risk aversion is not well identified from consumption behavior. The key is that introducing heterogeneity in the discount factor will capture persistent characteristics which are not correlated with high-frequency shocks to income.

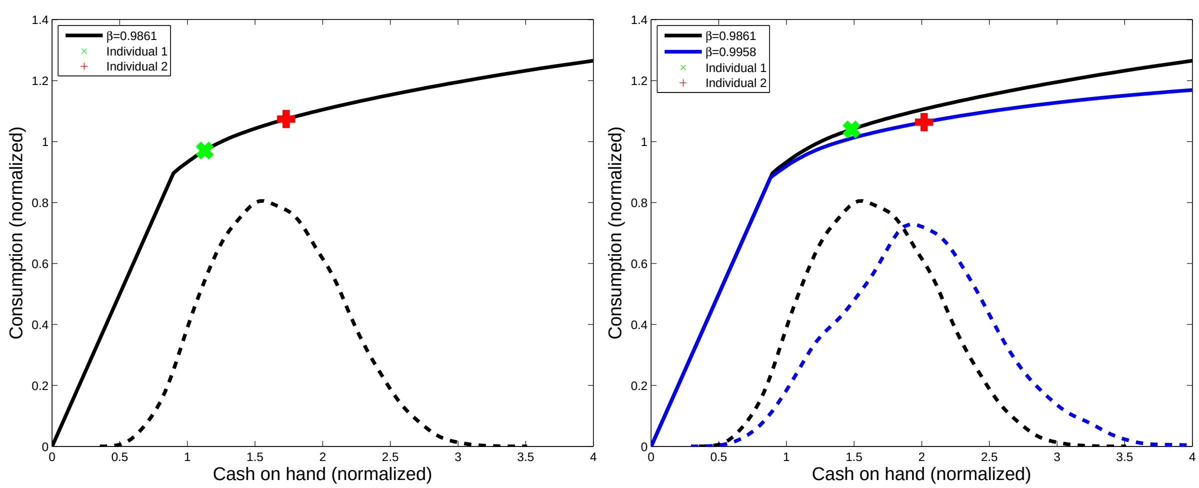
Under the circumstances view, MPC heterogeneity is driven entirely by temporary shocks to income and so $\beta_i = \bar{\beta}$. Under the characteristics view, MPC heterogeneity is driven both by temporary shocks to income and heterogeneity across individuals. This is captured by defining $\beta_i \sim U\ (\overline{\beta} - \Delta, \overline{\beta} + \Delta)$ as in Carroll et al. (2015) and Krueger, Mitman and Perri (2016).

Figure 1.1 provides a simple characterization of the sources of heterogeneity under the two views via the optimal consumption function and the distribution of cash on hand. The solid line represents the consumption function while the dotted line represents the distribution of cash on hand conditional on a particular discount factor. Panel (a) shows that under the circumstances view, heterogeneity is driven entirely by differences in cash on hand. Differences between individuals are represented by different points along the consumption function. For example, the individual represented by “x” may have received a negative shock and therefore exhibits lower cash on hand than the individual represented by “+”. Because the consumption function is concave, a lower cash on hand level is associated with lower consumption and a steeper slope (higher MPC). It is differences in circumstances that generates the correlation between the MPC and cash on hand.

Alternatively, panel (b) depicts heterogeneity under the characteristics view. The main difference is that individuals with different discount factors have different consumption functions and different distributions of cash on hand. For example, the individual represented by "+" has a higher discount factor relative to the individual represented by "x." The more patient individual has a flatter consumption function and a distribution of cash on hand that is shifted to the right. In the characteristics view, the discount factor jointly determines average MPC and average cash on hand. Impatient individuals will tend to have higher MPCs and lower cash on hand and vice versa. Contrary to the circumstances view, persistent characteristics now play a role in generating the correlation between the MPC and cash on hand.

Figure 1.1: Comparison of views

(a) Circumstances view (b) Characteristics view



Notes: Panel (a) and (b) plot the consumption function and distribution of cash on hand under the circumstances view and characteristics view respectively.

1.2.3 Target buffer stock behavior

A key mechanism to help distinguish between the two views is so called "target buffer stock" behavior. Under such behavior, individuals target a cash on hand to income ratio over time that is determined by their preferences and income uncertainty. While cash on hand will fluctuate due to temporary shocks to labor income, individuals will endogenously change their consumption behavior to achieve their target cash on hand. This implies that any snapshot of cash on hand at a point in time will reflect both recent temporary shocks and persistent characteristics. Because individuals react to temporary shocks by moving back towards their preferred buffer stock, taking a time average of cash on hand should isolate the level of cash on hand attributable to preferences.

Carroll (2004) defines the target buffer stock as the cash on hand value x^* such that $E[x^*] = x^*$. In other words, when cash on hand equals the target buffer stock, inviduals do not desire a different level of cash on hand. If cash on hand is not equal to the target buffer stock, individuals will alter their consumption behavior so that x_t converges back to x^*. Carroll (2004) then shows that for each individual, this value is unique and stable. This behavior can be understood by analyzing the well known second order approximation of the euler equation derived from the first order condition of the optimization problem represented by equations 4.2-4.6.

$$\underbrace{\Delta ln(c_{it+1})}_{\text{consumption growth}} \approx \overbrace{\frac{r-\delta_i}{\theta}}^{\text{impatience}} + \underbrace{\frac{\theta}{2}\sigma^2_{it+1}(x_{it})}_{\text{precautionary savings}} + \varepsilon_{it+1} \tag{1.9}$$

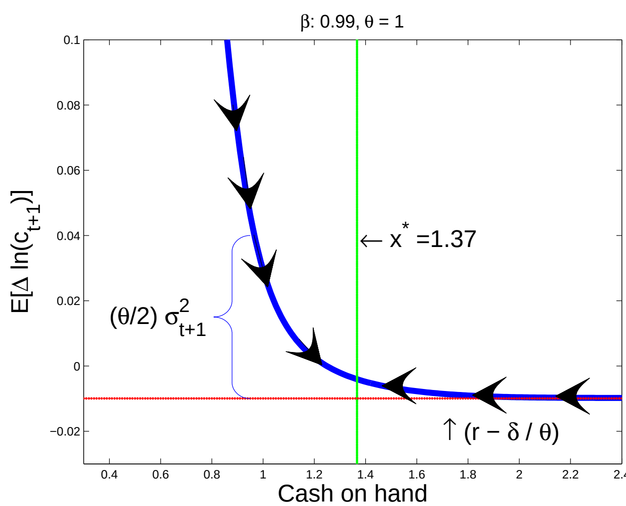
where c_{it} is normalized consumption, $\delta_i = \frac{1}{\beta_i} - 1$ is the discount rate, θ is the coefficient of relative risk aversion, σ^2_{it} is a measure of consumption growth volatility, r is the interest rate, and ε_{it} is a mean zero rational expectations error.

A buffer stock saver is influenced by two opposing factors. The first factor is that they are impatient and so weigh consumption today more than consumption tomorrow. This will tend to cause cash on hand to fall over time. Conversely, as pointed out in Kimball (1990), a positive third derivative of the utility function induces a precautionary savings motive which will tend to cause cash on hand to rise over time. Individual behavior will then depend on which motive is stronger.

These opposing factors are captured by the terms labeled "impatience" and "precautionary savings." The impatience term reflects the standard life cycle permanent income hypothesis (LC-PIH) motivation where consumption growth is a constant function of the interest rate, discount factor, and coefficient of relative risk aversion (or the elasticity of intertemporal substitution). Since this term is constant, the relative strength of each factor is driven by the non-constant precautionary savings term. The term $\sigma^2_{it+1}(x_{it})$ represents consumption growth volatility and is a function of cash on hand (x_{it}). Because this term is a complicated function of preferences and temporary shocks, it is hard to analytically derive the exact relationship. However, we do know that it is decreasing in x_{it}. The intuition is that when x_{it} is small, an individual is not able to smooth shocks very well leading to a wide range of possible consumption values in the next period depending on the realization of the labor income shock. This translates into high variability in consumption growth. Conversely, when x_{it} is high, an individual is easily able to smooth consumption in the face of income shocks so there will be little variation in consumption growth. In the limit, as $x_{it} \to \infty$, precautionary fears become irrelevant and an individual will behave according to the standard LC-PIH. The coefficient $\frac{\theta}{2}$ implies that consumption growth is an increasing function of the variance of consumption growth. Furthermore, the impact of uncertainty is increasing in risk aversion. Intuitively, this means that risk averse individuals will prefer not to put themselves in positions where they will face low levels of consumption. They achieve this by holding enough buffer stock to weather negative income shocks.

Figure 1.2 illustrates target buffer stock behavior by plotting expected consumption growth as a function of cash on hand. The vertical green line represents the target buffer stock level, and so behavior is determined by whether cash on hand is to the right or left of this value. When cash on hand is to the right of the target level, impatience dominates and cash on hand will fall back to the target level. More specifically, higher values of x_t will lead to lower values of $\sigma^2_{t+1}(x_t)$ and hence lower values of $\Delta ln(c_{t+1})$. As $x_t \to \infty$, $\Delta ln(c_{t+1})$ approaches $\frac{r-\delta}{\theta}$. Therefore, if cash on hand is too high, impatience will lead individuals to spend down cash on hand to finance consumption in the present period. Conversely, if cash on hand is to the left of the target level, the precautionary savings term dominates behavior. Lower values of x_t will lead to higher values of $\sigma^2_{t+1}(x_t)$ and $\Delta ln(c_{t+1})$. Intuitively, if cash on hand drops too low, the precautionary saving motive will prompt individuals to build back up their buffer stock. These opposing forces will constantly push cash on hand to its target level of x^*.

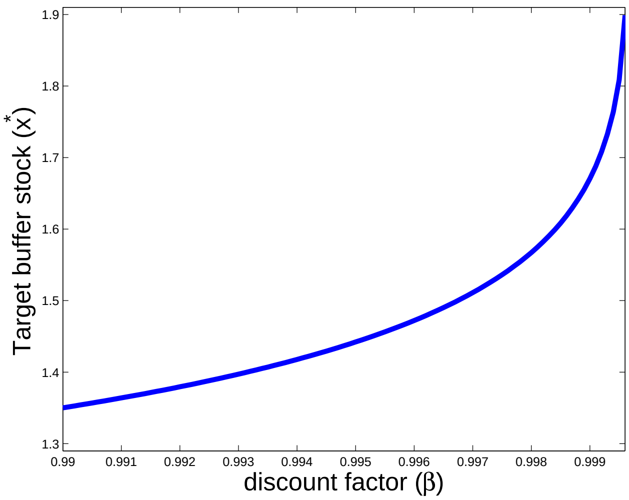
Figure 1.2: Target buffer stock behavior



Notes: The vertical line represents the stable target buffer stock level. The horizontal line represents the consumption growth rate in the absence of any precautionary savings motives. This figure is in the spirit of Figure Ia in Carroll (1997) but uses a different calibration.

Another important characteristic of target buffer stock x^* is that holding all else constant, it is a increasing function of the discount factor. While holding a buffer stock is helpful for protecting against income shocks, maintaining a high buffer stock comes at the expense of present consumption. Therefore, the more impatient individuals are, the more they will prefer to consume today instead of holding a large buffer stock. Figure 1.3 graphically demonstrates the positive relationship between x^* and β. This relationship will allow x^* to be interpreted as a proxy for the discount factor.

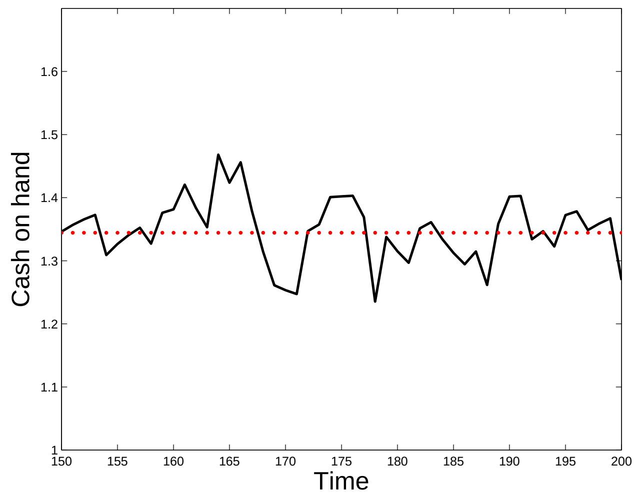
Figure 1.3: Target buffer stock and the discount factor



Notes: β refers to the monthly discount factor.

Lastly, Figure 1.4 shows the time series behavior of simulated cash on hand within an individual. The horizontal dashed line represents the target buffer stock level. As expected, temporary shocks cause cash on hand to deviate from the target value x^*. However, because the target buffer stock level is a stable equilibrium, individual consumption x_t will tend towards x^* over time. I utilize this behavior to decompose cash on hand into a circumstances and characteristics component as $x_t = (x_t - x^*) + x^*$. The next section will explore how these dynamics will aid in identifying the differential relationship of cash on hand and the MPC under the two views.

Figure 1.4: Cash on hand time series



Notes: The horizontal line represents the stable target buffer stock level.

1.2.4 Model simulation

Before analyzing the actual data, it's helpful to understand how consumption behavior differs under the circumstances and characteristics view. To this end, this section simulates the consumption response to income under the two views. In order to create a tight link with the data, I attempt to model the empirical environment that I observe within the dataset as closely as possible.

The dataset used in the empirical section includes transaction-level consumption, income, and cash on hand measures from a person finance app. I take advantage of the transaction-level granularity of the data to identify receipts of multiple tax refunds within individuals. These tax refund are then used in turn to calculate the MPC out of a change in income.

The simulation environment is chosen to match this empirical environment very closely. Therefore, I simulate the consumption reaction of 200 individuals to the receipt of a tax refund every 12 months over a period of 4 years. For each tax refund received, I calculate the MPC and cash on hand of each individual. I then explore how the relationship between the MPC and cash on hand differ under the two different views.

The main result is that the relationship between the MPC and cash on hand only differs when the panel structure of the data is used. Intuitively, cross-sectional snapshots will confound the role of circumstances and characteristics in driving MPC heterogeneity.

1.2.4.1 Calibration

The parameter values used to calibrate the model are listed in Table 4.4 below and represent monthly time periods. The utility function is specified as constant relative risk aversion (CRRA) with $\theta = 1$. The parameters $\overline{\beta}$ and Δ are set to the parameters estimated in the later part of the paper. The parameters ρ and σ_y are estimated using the income process observed in the dataset.[2] $refund_{it}$ represents the average tax refund to income ratio observed in the data set. The interest rate is set to the monthly rate on checking/savings accounts and the borrowing limit is set to zero.

Table 1.1: Parameter values

Parameter	Value	Notes	Description
$u(x)$	$\frac{x^{1-\theta}}{1-\theta}$	CRRA utility	utility function
θ	1	standard	coefficient of relative risk aversion
$\overline{\beta}$	0.9894		average discount factor
Δ	0.0103	0 for circumstance model	discount factor dispersion
ρ	0	estimated from dataset	income shock persistence
σ_y	0.20	estimated from dataset	S.D. of temporary shocks
$refund_{it}$	0.6	estimated from dataset	average normalized refund
r	0.01 / 12	monthly r on checking/saving	interest rate
b	0	no borrowing condition	borrowing limit

Notes: The parameters correspond to a monthly frequency.

[2]The estimate for $\hat{\rho} = 0.065$. Given how close it is to 0, I choose to set ρ to 0 in the simulation because it greatly reduces the complexity of model by allowing me to remove a state variable that normally needs to keep track of the previous value of income. The low estimate of $\hat{\rho}$ reflects the fact that the sample is selected on individuals who receive regular paychecks. This sample restriction is made to fit the model which doesn't have permanent shocks or periods of unemployment.

1.2.4.2 Variable definitions

The main variables used in the analysis are the MPC and cash on hand. This section provides definitions for these concepts.

Definition: The MPC at time t for individual i is defined as

$$MPC_{it} = \frac{\Delta C_{it}}{\Delta Y_{it}} = \frac{\sum_{j=t}^{t+2} c_{ij} - \sum_{j=t-1}^{t-3} c_{ij}}{refund_{it}} \tag{1.10}$$

Because each period in the model is one month, this value represents the quarterly change in consumption as a fraction of the tax refund. For periods in which a tax refund is not received, the MPC is undefined.

Definition: Pre-refund cash on hand at time t for individual i is defined as

$$coh_{it}^{PR} = \frac{\sum_{j=t-1}^{t-3} x_{ij}}{3} \tag{1.11}$$

This measure captures the average level of cash on hand three months prior to receiving the tax refund. It is meant to mimic the measures of liquidity captured in survey data commonly used in studies estimating the consumption response to income changes.

Definition: Average cash on hand for individual i

$$\overline{coh}_i = \frac{\sum_{j=t}^{T} x_{ij}}{T} \tag{1.12}$$

This measure is meant to capture the target level of buffer stock for individual i described in the previous section and is used as a proxy for the discount factor. This measure is not usually captured in survey data such as the Consumer Expenditure Survey because the panel dimension is relatively short.

1.2.4.3 The relationship between MPC and cash on hand

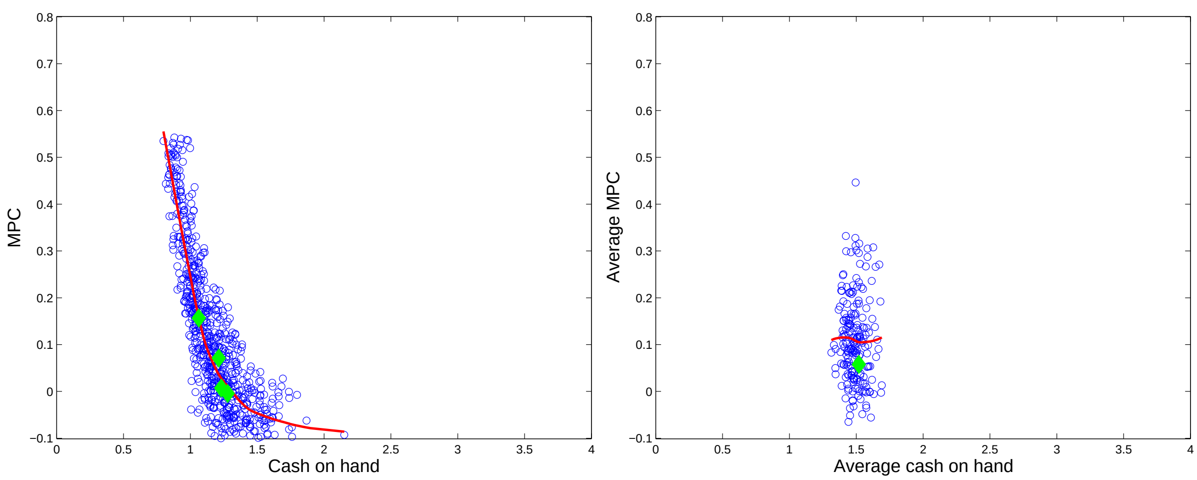
After simulating the data, I calculate MPC_{it}, coh_{it}^{PR}, and $\overline{coh_i}$ for each individual. Figure 1.5 shows the relationships between these variables under the assumptions of the circumstances view where $\beta_i = \bar{\beta}$.

Panel (a) presents a scatter plot of the MPC and pre-refund cash on hand overlaid with a local linear smoothed line. In this panel, each point represents an observation for individual i and time t. For example, the green diamonds represent all observations for a particular individual. Because each individual receives four refunds, there are four points. There is a clear negative relationship between MPC_{it} and coh_{it}^{PR}. This pattern is consistent with the concavity of the consumption function suggested by Carroll and Kimball (1996). Since the MPC is the slope of the consumption function, a concave consumption function will result in a high MPC when cash on hand is low and vice versa. Jappelli and Pistaferri (2014) also report a similar relationship when they explicitly ask individuals what their MPC would be out of a hypothetical income shock.

Panel (a) is analogous to plotting the relationship of the MPC and cash on hand in a pooled cross-section. As discussed earlier, a snapshot of cash on hand in time will reflect both circumstances as well as characteristics. In order to isolate the characteristics component of cash on hand, panel (b) presents a scatter plot of the average MPC and average cash on hand. Note that now each observation represents one individual. This is reflected in the fact that the four green diamonds in panel (a) are collapsed into one green diamond in panel (b). Once I collapse the data by average across time within an individual, the strong negative relationship between the MPC and cash on hand is no longer present. Under the circumstances view, the lack of heterogeneity in the discount factor leads to all individuals having the same target buffer stock level. Therefore, there should not be any systematic relationship between average cash on hand and any other individual level variable. The temporary shocks are beyond the control of the individual and so pre-refund cash on hand levels will influence the response to tax refunds. After the shocks have occurred, however, individuals will alter their behavior to return to their desired buffer stock level. Over a long enough horizon, this preference-driven behavior is the main determinant of the level of cash on hand. Under our parametrization, four years is a long enough time horizon for average cash on hand to reflect the theoretical target buffer stock level.

Figure 1.5: Relationship between MPC and cash on hand under the circumstances view

(a) Pooled cross-section (b) Average



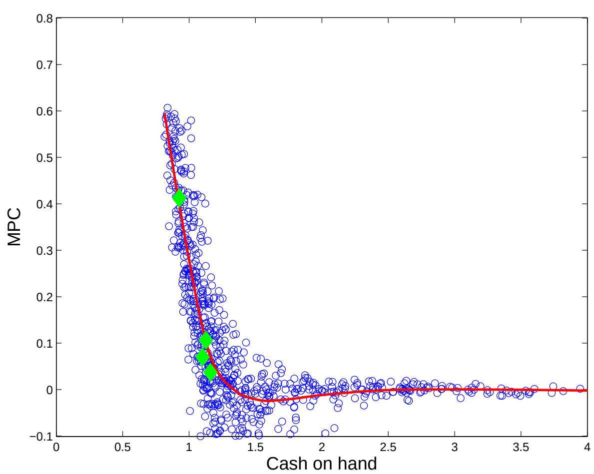
Notes: Panel (a) plots the relationship between pre-refund cash on hand and the MPC for individual i at time t using simulated data. Panel (b) plots the relationship between average cash on hand and the avearge MPC for individual i. In both plots, the solid red line represents a local-linear smoothed curve and the green diamond represents all observations for a randomly chosen individual. The first 100 periods of the simulations are discarded to allow individuals to reach steady-state.

Figure 1.6 repeats the exercise in Figure 1.5 under the assumptions of the characteristics view where $\beta_i \sim U\ (\overline{\beta} - \Delta, \overline{\beta} + \Delta)$. The results in panel (a) look similar across the two views. Once again, a strong negative relationship exists between MPC_{it} and coh_{it}^{PR}; however, it's not clear whether this is driven by the concavity of the consumption function or the differences in the discount factor across individuals. This formalizes the idea that observing the relationship between the MPC and cash on hand in the cross-section cannot identify which view is likely to be correct.

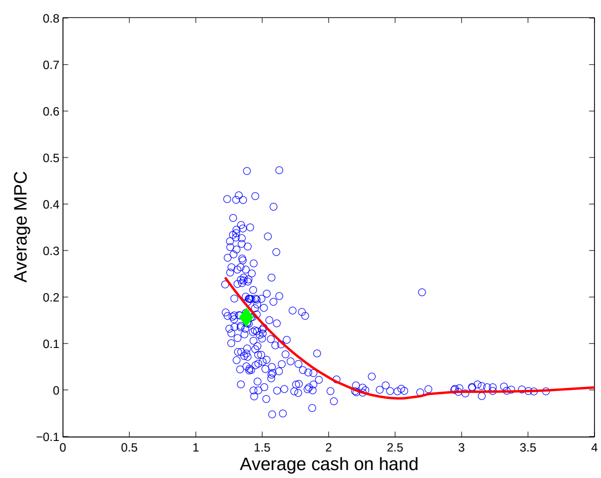
Once again, the problem stems from the fact that any snapshot of cash on hand is influenced both by recent changes to temporary circumstances as well as persistent characteristics. Plotting panel (b) under the characteristics view reveals that the relationship between $\overline{MPC}_i$ and $\overline{coh}_i$ exhibits a strong negative relationship. This result is driven by the fact that discount factors are allowed to vary across individuals. On average, impatient individuals with low discount factors will tend to hold low cash on hand and have high MPCs and vice versa. Even after averaging out the temporary shocks, these persistent characteristics drive the negative correlation between the average MPC and average cash on hand.

Figure 1.6: Relationship between MPC and cash on hand under the characteristics view

(a) Pooled cross-section



(b) Average



Notes: Panel (a) plots the relationship between pre-refund cash on hand and the MPC for individual i at time t using simulated data. Panel (b) plots the relationship between average cash on hand and the avearge MPC for individual i. In both plots, the solid red line represents a local-linear smoothed curve and the green diamond represents all observations for a randomly chosen individual. The first 100 periods of the simulations are discarded to allow individuals to reach steady-state.

In summary, estimating the cross-sectional relationship between the MPC and pre-refund cash on hand will lead to similar results under both views. A negative correlation is observed regardless of which view actually holds in the data. The views can only be distinguished by isolating the persistent characteristics component by calculating the average MPC and average cash on hand within individuals. The circumstances view implies a very weak relationship between the average MPC and average cash on hand while the characteristics view implies a strong negative relationship.

1.2.4.4 Variance decomposition

While the previous section helps to visualize the differences between the two views, it is also helpful to introduce a more quantitative measure that captures which view is more consistent with the data.

Regardless of which view is correct, the analysis in the previous section shows that MPC_{it} is a function of cash on hand. Furthermore, the section on buffer stock behavior showed that cash on hand can be decomposed into a circumstances and characteristics component. This decomposition can be used to determine which view is more likely to hold in the data. If the circumstances view is more likely, MPC_{it} should mainly be a function of changes in circumstances due to temporary labor income shocks. Alternatively, if the characteristics view is more likely, MPC_{it} should also be a function of characteristics such as the discount factor. To test this hypothesis, the MPC_{it} in specified in the following way.

$$MPC_{it} = \alpha + \underbrace{\gamma_1 \times \overline{coh_i}}_{\text{characteristics}} + \underbrace{\gamma_2 \times (coh_{it}^{PR} - \overline{coh_i}^{PR})}_{\text{circumstances}} + \varepsilon_{it} \tag{1.13}$$

where $\mathbb{E}[\varepsilon_{it}] = 0$. While the discount factor is not explicitly observed, the buffer stock model implies that average cash on hand is a function of the discount factor. Therefore $\overline{coh_i}$ is used to capture the characteristics component of cash on hand. The circumstances component of cash on hand is captured by using pre-refund cash on hand (coh_{it}^{PR}). Because the level of coh_{it}^{PR} is still related to the discount factor, it is demeaned by its average ($\overline{coh_i}^{PR}$) in order to extract the temporary component that is orthogonal to the individual level average.

Under this specification, the variance is easily decomposed because all the terms are uncorrelated with each other (see appendix section 1.8.1 for more details). The following equation applies the variance operator to both sides.

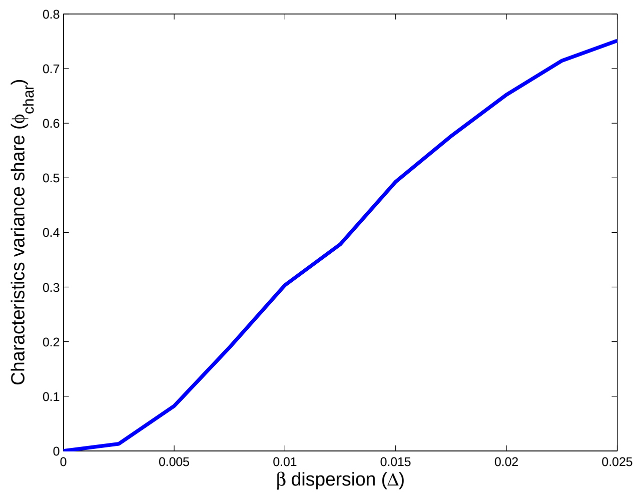
$$var(MPC_{it}) = var(\alpha) + var(\gamma_1 \times \overline{coh}_i) + var(\gamma_2 \times (coh_{it}^{PR} - \overline{coh}_i^{PR})) + var(\varepsilon_{it}) \quad (1.14)$$

Defining $var(\gamma_1 \times \overline{coh}_i) = \sigma_{char}^2$ and $var(\gamma_2 \times (coh_{it}^{PR} - \overline{coh}_i^{PR})) = \sigma_{circ}^2$, these terms capture the variance contribution of the characteristics and circumstances component of cash on hand respectively. Another way to think about this equation is that the characteristics component captures across-individual variation and the circumstances component captures within-individual variation. Under the circumstances view, σ_{circ}^2 should be very high relative to σ_{char}^2. This captures the idea that the variance in MPC_{it} is mostly driven by circumstances. Analogously, most of the variation in MPC_{it} should be driven by within-individual differences. Under the characteristics view, σ_{char}^2 is around the same size or larger than σ_{circ}^2. This captures the fact that variance in MPC_{it} is driven by both circumstances and characteristics. Stated differently, both within- and across- individual variation is important in explaining variation in MPC_{it} under the characteristics view.

Defining $\phi_{char} = \frac{\sigma_{char}^2}{\sigma_{char}^2 + \sigma_{circ}^2}$, this value represents the fraction of $var(MPC_{it})$ explained by cash on hand that is attributable to characteristics. Since ϕ_{char} is bound between 0 and 1, it can be used to determine which view is more likely. A value near 0 is consistent with the circumstances view while values away from 0 are more consistent with the characteristics view.

The characteristics share of variance (ϕ_{char}) can also be connected back to the model. Recall that under the circumstances view $\beta_i = \overline{\beta}$, while under the characteristics view $\beta_i \sim U\ (\overline{\beta} - \Delta, \overline{\beta} + \Delta)$. Higher values of the dispersion in the discount factor (Δ) lead to greater heterogeneity in average cash on hand levels. Holding the variance of temporary shocks constant, this should lead to a greater contribution of the characteristics component of cash on hand in explaining the MPC. Figure 1.7 shows this relationship by calculating ϕ_{char} under different values of Δ while holding all other parameters constant. As expected, ϕ_{char} is an increasing function of Δ.

Figure 1.7: Relationship between the dispersion of β and ϕ_{char}



Notes: This figure plots the relationship between the dispersion in the discount factor Δ against the characteristics component variance share (ϕ_{char}).

In summary, calculating ϕ_{char} in the data will identify which view is more consistent with the data. Furthermore, ϕ_{char} will later be used to help estimate Δ using the method of simulated moments.

1.3 Data

This section describes the data source, sample filters, variable definitions and descriptive statistics.

1.3.1 Data source

This paper utilizes a novel dataset derived from de-identified transactions and account data, aggregated and normalized at the individual level. The data are captured in the course of business by a personal finance app.[3][4] More specifically, the app offers financial aggregation and bill-paying services. Users can link almost any financial account to the app, 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. Prior to analysis, the data are stripped of personally identifying information such as name, address, or account number. The data have scrambled identifiers to allow observations to be linked across time and accounts.

We draw on the entire de-identified population of active users and data derived from their records from December 2012 until July 2016. For a subset of the data, we have made use of demographic information provided to the app by a third party. Table 4.1 compares the age, education, gender, and geographic distributions in the sample that matched with an email address to the distributions in the U.S. Census American Community Survey (ACS), representative of the U.S. population in 2012.

---

[3] These data have previously been used to study the high-frequency responses of households to shocks such as the government shutdown (Gelman et al., 2015) and anticipated income, stratified by spending, income and liquidity (Gelman et al., 2014).

[4] Similar account data has been used in Baugh, Ben-David and Park (2014), Baker (2015), Kuchler (2015), and Ganong and Noel (2016).

Table 1.2: App user demographics

Education	Not Completed College	Completed College	Completed Graduate School
ACS	66.62	24.02	9.36
App	70.42	23.76	5.83

Ages 25 and over. Sample size - ACS: 2,176,103 App: 28,057

Age	18-20	21-24	25-34	35-44	45-54	55-64	65+
ACS	5.85	7.28	17.44	17.24	18.78	16.00	17.41
App	0.59	5.26	37.85	30.06	15.00	7.76	3.48

Sample size - ACS: 2,436,714 App: 35,417

Gender	Male	Female
ACS	48.56	51.44
App	59.93	40.07

Sample size - ACS: 2,436,714 App: 59,072

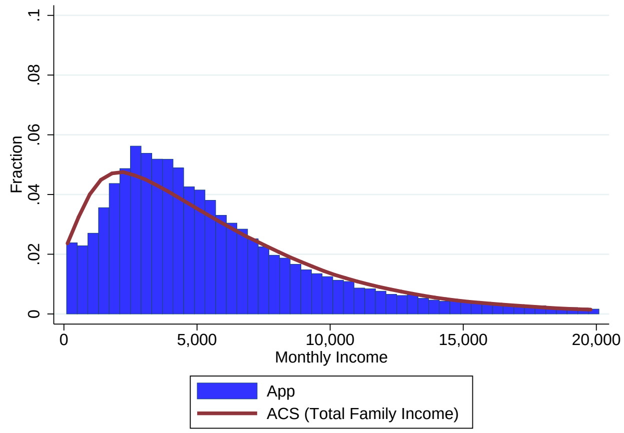
Region	Northeast	Midwest	South	West
ACS	17.77	21.45	37.36	23.43
App	20.61	14.62	36.66	28.11

Sample size - ACS: 2,441,532 App: 63,745

Source: Gelman et al. (2014).

Figure 4.2.1 compares the income distribution in the app to total family income in the ACS. Users who use the app are on average higher income than individuals surveys in the ACS.

Figure 1.8: Income comparison



Source: Gelman et al. (2014).

In summary, the app is not perfectly representative of the US population, but it is heterogeneous, including large numbers of users of different ages, education, income, and geographic location.

1.3.2 Sample filters

The sample is filtered on various characteristics to ensure that the analysis sample matches the model specified in the earlier sections.

First, the model assumes the researcher observes a comprehensive view of spending, income, and liquid assets. Therefore, I require data from individuals who add all (or most) of their accounts, generate a long time series of observations, and have positive income in each month. This reduces the sample size because there is a large amount of churn from users who try out the app but later decide not to continue using it. Moreover, there are some users that only want to track one or two credit cards without adding all their other accounts.

Second, the model is meant to abstract away from life cycle motives and large permanent shocks to income so that reactions stem from either temporary circumstances or persistent characteristics. Therefore, I condition on individuals who receive regular paychecks.

Lastly, since the MPC is estimated from the consumption reaction to tax refunds, I condition on individuals who received more than 1 tax refund in the sample.

In summary, I select users based on length of panel, number of accounts, connectedness of accounts, regular paycheck status, no missing income data, and whether they received more than 1 tax refund.

1.3.2.1 Defining account linkage

The analysis may be biased if all accounts that are used for receiving income and making expenditures are not observed. For example, an individual may have a checking account that is used to pay most bills and a credit card that it used when income is low. If credit card expenditures are not properly observed the MPC will be biased downwards.

In order to identify linked accounts, I use a method that calculates how many credit card balance payments are also observed in a checking account. I define the variable *linked* as the ratio of the number of credit card balance payments observed in all checking accounts that matches a particular payment that originated from all credit card accounts. For example, a typical individual will pay their credit card bill once a month. If they existed in the data for the whole year, they will have 12 credit card balance payments. If 10 of those credit card payments can be linked to a checking account the variable $linked = \frac{10}{12} \approx 0.83$.

One drawback to this approach is that it requires individuals to have a credit card account. To ensure that those without credit cards are still likely to have linked accounts, I also condition on individuals who have three or more accounts.

1.3.2.2 Defining regular paycheck

In order to identify regular paychecks, I start by using keywords that are commonly associated with these transactions (see appendix section 1.8.2 for more details). I condition on four statistics to ensure that these transactions represent regular paychecks.

1. Number of paychecks ≥ 5
2. Median paycheck amount $> \$200$
3. Median absolute deviation of days between paychecks is ≤ 5
4. Coefficient of variation of the paycheck amount ≤ 1

1.3.2.3 Sample size

Table 1.3 shows the evolution of the sample size from all users in the sample to those that survive the selection criteria. The criteria selects users who have a long time series ($\geq$ 40 months), a high linked account ratio ($\geq$ 0.8), a reasonable number of accounts linked ([3,15]), receive a regular paycheck, receive positive income in each month, and receive more than 1 tax refund. I choose to drop users that have over 15 accounts linked because these accounts typically represent business users. The final sample may seem small but this is due to fact that most individuals only try out the app for a short amount of time. Baker (2015) uses a similar sample selection criteria that results in a final sample that is also roughly 5% of the full sample.

Table 1.3: Sample Filters

	N	%
Full sample as of December 2012	883,529	100
Long time series (N $\geq$ 40)	341,841	39
Linked ratio $\geq$ 0.8	264,043	30
Linked accounts $\in$ [3,15]	197,530	22
Has regular paycheck	146,112	17
Has no months with zero income	77,052	9
Has > 1 tax refund	48,059	5

1.3.3 Variable definitions

Most survey data sets such as the consumer expenditure survey (CEX), panel study of income dynamics (PSID), and survey of consumer finances (SCF) are created with the explicit goal of facilitating academic research. The data set used in this study is naturally occurring and was not explicitly designed for use in academic studies. Constructing variables in this data set to match our models is not necessarily a trivial exercise. In order to study the relationship between the MPC out of tax refunds and cash on hand, the main variables I utilize are consumption, income, tax refunds, and liquid assets.

1.3.3.1 Consumption

The empirical analysis will focus on non-durable consumption because durable goods are not explicitly modeled. In particular, I attempt to match the composition of the widely used "strictly non-durable" definition from Lusardi (1996).

The raw data consists of individual transactions with characteristics such as amount, transaction type (debit or credit), and transaction description. While the type of spending (non-durable, durable) is not directly observed, I use a machine learning (ML) algorithm (see appendix section 4.8.1 for more details) to aid in categorization. The goal of the ML algorithm is to provide a mapping from transaction descriptions to spending categories. For example, any transaction with the keyword "McDonalds" should map into "Fast Food". A subset of these categories are then combined to create the consumption variable.

The finest level of categorization is derived from merchant category codes (MCCs) which are directly observable in two of the account providers in the data. 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. The ML algorithm works by using a subset of the data where the truth is known in order to create a mapping from transaction description to MCCs.

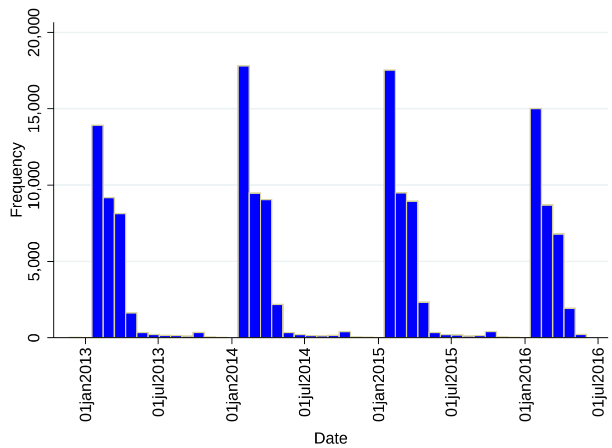
After training the ML algorithm on the data where the truth is known, the algorithm is then applied to the rest of the data set. I then define consumption as spending on restaurants, groceries, gasoline, entertainment, and services.

1.3.3.2 Tax refunds

In order to disentangle temporary circumstances from persistent characteristics, it's important to observe several MPCs across time within an individual. While many studies have analyzed the MPC out of tax rebates, one disadvantage of tax rebates is that they occur at a fairly low frequency. Since most people receive federal tax refunds in multiple years, this study utilizes the MPC out of tax rebates over time within individuals.

Federal tax refunds are identified by searching for identifying keywords in the transaction description (all tax refunds include the keywords "TAX", "TREAS", and "REF"). I exclude individuals that receive multiple tax refunds within the same year. Figure 1.9 shows the time series of the count of tax refunds observed in the data from December 2012 to July 2016. The figure shows that most tax refunds are received in February, March, April, and May.

Figure 1.9: Federal tax refund time series



1.3.3.3 Income

Income is important in determining the variance of temporary shocks as well as an input into cash on hand. Total income is defined as the sum of all inflows from checking and saving accounts minus incoming transfers.

In order to calibrate the income process in the model, I first estimate the time series properties of the income process using the equation below. To fit the model, I subtract out any tax refunds and normalize by average income. The equation specifies an AR(1) model in non-tax refund normalized income and controls for seasonality using monthly indicator variables.

$$y_{it} = \rho y_{it-1} + month_t + \varepsilon_{it} \tag{1.15}$$

Table 1.4 shows the results of estimating equation (4.5). The value of 0.065 indicates that there is a small amount of persistence in the income process. This is much lower than standard estimates because the sample conditions on individuals who receive a regular paycheck.

Table 1.4: Income process estimation

VARIABLES	(1) y_t
y_{t-1}	0.065***
	(0.001)
Constant	0.815***
	(0.001)
Observations	2,166,690
R-squared	0.07

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

I also estimate the variance of temporary shocks as $var(\varepsilon_{it}) = 0.041$. This is the value that is used throughout the analysis to calibrate the model.

1.3.3.4 Cash on hand and liquid assets

As discussed in section 4.4 (theoretical framework), cash on hand plays a crucial role in identifying changes in circumstances as well as providing a proxy for characteristics. Cash on hand is defined as $X_{it} = A_{it-1} + Y_{it}$ where A_{it-1} represents liquid balances for individual i in the previous period and Y_{it} represents income received in the current period.

Liquid balances (A) are defined as the sum of checking and saving account balances observed in the app. These balances are captured daily as the app takes a snapshot of the balance from each provider.

1.3.3.5 Normalization

To match the theoretical framework, the main variables in the empirical analysis are normalized by individual average income. The normalization is denoted with lower case variables and so $c_{it} = C_{it}/\bar{Y}_i, x_{it} = X_{it}/\bar{Y}_i$ and so on. The observed level of average income serves as a proxy for unobserved permanent income.

1.3.4 Summary statistics

This section provides summary statistics of the main variables used in the analysis. The mean and median values for spending and income are roughly in line with the data used in Baker (2015) from a different personal finance app.

Table 1.5: Summary statistics

Variable	Mean	p25	p50	p75
Spending	\$6,107	\$2,779	\$4,509	\$7,365
Income	\$6,290	\$3,375	\$5,035	\$7,642
Liquid balance	\$8,306	\$876	\$2,365	\$7,153
Tax refund	\$2,981	\$1,090	\$2,205	\$4,241

Notes: N=48,059

1.4 Empirical results

This section discusses the empirical results used to test whether the circumstances view or characteristics view is more consistent with the data. Using various different approaches, it finds that even after controlling for within-individual variation in cash on hand, across-individual cash on hand still explains a large portion of MPC heterogeneity. This finding is more consistent with the characteristics view rather than the circumstances view.

1.4.1 Tax refund impulse response function

As a preliminary step, I estimate the consumption impulse response function to receiving a tax refund. This analysis helps to confirm that the variables are constructed properly and behave according to economic theory. More specifically, I estimate the distributed lag of receiving a tax refund using the following specification.

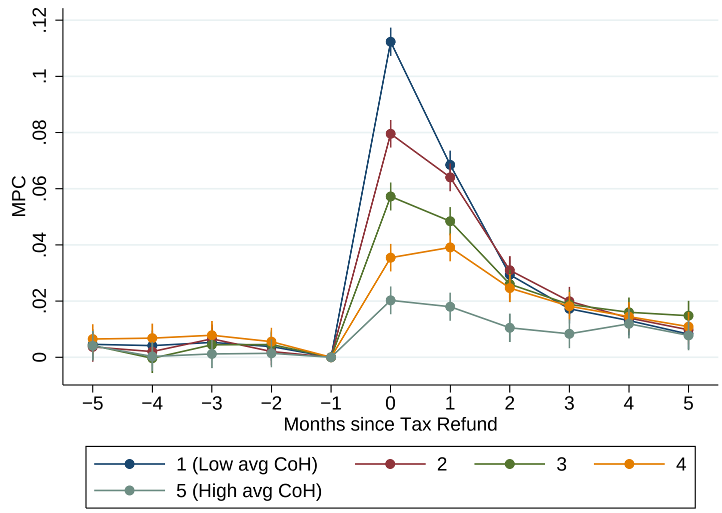
$$c_{it}^q = \alpha^q + \sum_{j=-6}^{6} MPC_j^q \times ref_{it-j}^q + \delta_t^q + \varepsilon_{it}^q,\ where\ q \in \{1,..,5\} \tag{1.16}$$

The q superscript represent quintiles of average cash on hand ($\overline{coh_i}$), c_{it} represents normalized consumption, ref_{it-j} represents the normalized tax refund, δ_{it} represents month fixed effects, and ε_{it} is the error term. Figure 1.10 below plots the MPC_j for each cash on hand quintile. The estimates show that there is little anticipatory response of consumption to receiving a tax refund and much of the response occurred within the first three months. The magnitude of the response is roughly in line with Souleles (1999) which examines the consumption response to income tax refunds in the CEX. Souleles (1999) does not calculate the MPC across cash on hand quintiles so I compare those results with the average response across all individuals in this paper. The average response is very similar to the third quintile in Figure 1.10 (see appendix Figure 1.8.1). A more recent paper by Baugh, Ben-David and Park (2014) studies the weekly response of spending to the arrival of tax refunds using similar account data. This paper does not explicitly calculate the MPC but finds that individual spending reacts strongly when the refund is received followed by a quick decay.

Splitting the sample up into quintiles of average cash on hand reveals the heterogeneous response in the data. Individuals in the lowest quintile of cash on hand tend to react much more strongly to the receipt of a tax refund relative to those in the highest quintile of cash on hand. This relationship is broadly consistent with most of the literature examining the consumption response to income changes (for example, many of the studies discussed in Jappelli and Pistaferri (2010)).

In summary, the estimated response of consumption to receiving a tax refund is similar in dynamics and magnitude to previous studies. This fact helps to confirm that both consumption, tax refunds, and cash on hand are identified properly in the data set.

Figure 1.10: Tax refund impulse response function



Notes: 1,445,560 observations from 48,059 individuals. The vertical bars on each coefficient represent 95% confidence intervals using heteroskedasticity robust errors clustered at the individual level.

1.4.2 The relationship between the MPC and cash on hand

In order to determine which view the data is more consistent with, this section analyzes the relationship between the MPC and cash on hand using two different levels of aggregation. The first level of aggregation is at the individual-refund level and the second level of aggregation is at the quantile level.

1.4.2.1 Individual-refund level analysis

I estimate the quarterly MPC out of tax refunds for individual i at time t is using the following specification.

$$c_{it} = \alpha_{ir} + MPC_{it} \times ref_{it} + \delta_t + \varepsilon_{it} \tag{1.17}$$

where i represents individual, t represents month, α_{ir} represents a dummy variable for each individual-refund year[5], ref_{it} represents the refund amount, δ_t represents time

[5]More specifically, this variable represents a series of dummy variables for each individual that takes a value of 1 for the three month windows before after a refund is received and 0 for all other

fixed effects, and ε_{it} is the error term.

The estimated MPC measures from this specification are then plotted against different concepts of cash on hand in Figure 1.11. Panel (a) plots the results of a smoothed local linear kernel regression of the relationship between the individual-refund level MPC (MPC_{it}) and pre-refund cash on hand (coh_{it}^{PR}). The MPC is falling rapidly as cash on hand increases until it starts to level out around a value of 1.6. This empirical relationship is consistent with the simulation results presented earlier in Figure 1.5 and 1.6. While previous studies have shown that a negative correlation exists between the MPC and cash on hand, this is the first paper to estimate the relationship using smooth kernel regressions with such a high level of precision. This high level of flexibility and precision provides novel evidence that the relationship between the MPC and cash on hand is consistent with a concave consumption function as argued by Carroll and Kimball (1996).

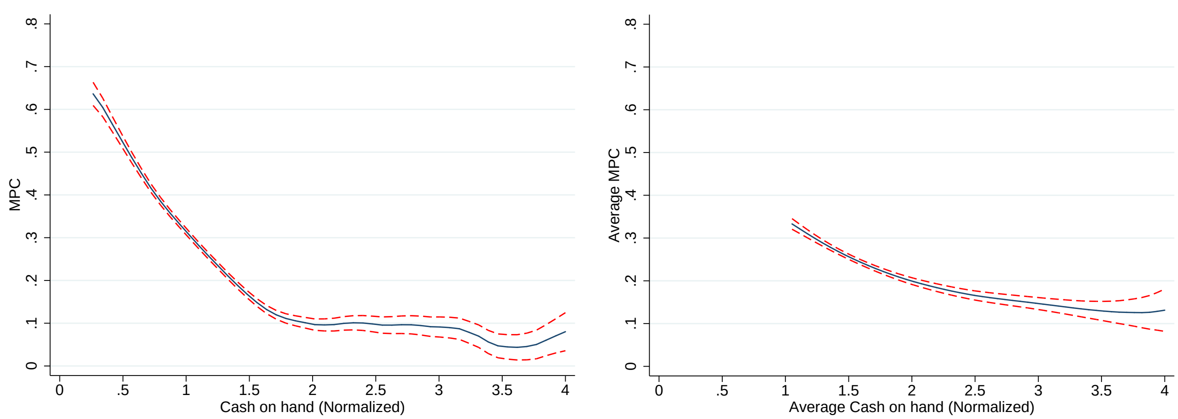
Panel (b) plots the relationship between the average MPC ($\overline{MPC}_i$) and average cash on hand ($\overline{coh}_i$). The results imply a statistically significant negative relationship between $\overline{MPC}_i$ and $\overline{coh}_i$. To my knowledge, this is the first paper to use panel data to estimate this relationship. This is important because averaging across time within individual isolates the role of persistent characteristics in driving the relationship between the MPC and cash on hand. The earlier simulation results showed that estimating the cross-sectional relationship between the MPC and cash on hand is not sufficient to separately disentangle the circumstances view from the characteristics view. This is made clear when comparing panel (a) in 1.5 and 1.6. The two views can only be disentangles by isolating the characteristics component by estimating the relationship between the average MPC and average cash on hand represented in panel (b) of 1.5 and 1.6. The significant negative relationship between $\overline{MPC}_i$ and

periods. This variable ensures that the MPC captures the change in consumption during the three months after receiving the refund relative to the three months prior to receiving the refund. This is the definition of the quarterly MPC.

$\overline{coh_i}$ imply that the characteristics view is more likely to hold in the data. Recall that under the characteristics view, differences in the discount factor across individuals generates a correlation between the average MPC and average cash on hand. Impatient individuals will tend to have higher average MPCs and lower average cash on hand and vice versa.

Figure 1.11: MPC and cash on hand

(a) MPC_{it} and coh_{it}^{PR} (b) $\overline{MPC_i}$ and $\overline{coh_i}$



Notes: 129,823 observations from 48,059 individuals in panel (a). 48,059 observations from 48,059 individuals in panel (b). The vertical bars on each coefficient represent 95% confidence intervals using heteroskedasticity robust errors clustered at the individual level. Variables are winsorized at the 5% level.

1.4.2.2 Quantile level estimates

This section estimates the MPC at the quantile level. More specifically, it estimates the MPC for each group defined by the interaction of coh_{it}^{PR} and $\overline{coh_i}$ quintiles. The econometric specification is

$$c_{it} = \alpha_{jk} + MPC_{jk} \times ref_{it} + \delta_t + \varepsilon_{it} \tag{1.18}$$

where i represents individual, t represents month, j refers to pre-refund cash on hand quintile, and k refers to average cash on hand quintile.

More concretely, MPC_{jk} represents the MPC for individuals with pre-refund cash on hand quintile j and average cash on hand quintile k. The average cash on hand quintile is an individual-level trait and so does not vary within i. On the other hand, j is allowed to vary within individual based on the level of cash on hand that is observed before the tax refund is received. To understand these concepts better, table 1.6 tabulates the median levels of each quintile.

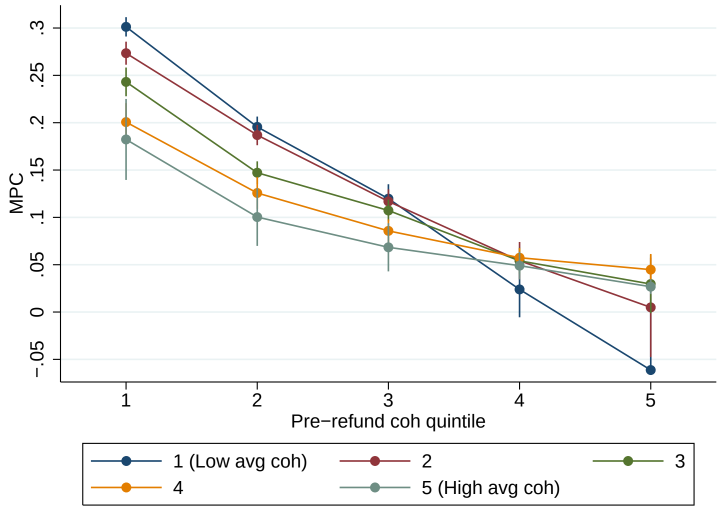
Table 1.6: Quintile sample statistics

	coh_{it}^{PR}		$\overline{coh}_i$	
Quintile	median	N	median	N
1	0.78	26,681	1.20	26,681
2	1.09	26,680	1.36	26,680
3	1.38	26,680	1.58	26,680
4	1.85	26,680	2.00	26,680
5	3.51	26,680	3.49	26,680
Total	1.38	133,401	1.59	133,401

Figure 1.12 plots the coefficients of MPC_{jk}. When the coh_{it}^{PR} quintile is low, the MPCs are ordered highest to lowest by the quintiles of $\overline{coh}_i$. For example, when coh_{it}^{PR} is 1, the point estimate is approximately 0.3 for individuals in the lowest $\overline{coh}_i$ quintile and approximately 0.18 for those in the highest $\overline{coh}_i$ quintile. The dispersion of the MPC within coh_{it}^{PR} falls as we move from the lowest to the highest quintile.[6]

[6]While this appears at odds with the large negative point estimate for those in the lowest $\overline{coh}_i$ quintile when the coh_{it}^{PR} is high, this estimate is extremely noisy and we cannot reject that the point estimate is different from the other $\overline{coh}_i$ quintiles estimates at the 5% level.

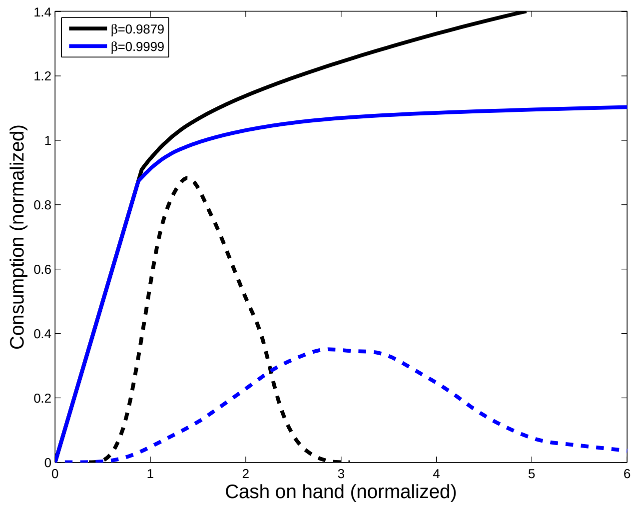
Figure 1.12: MPC by quintile interactions



This phenomenon is consistent with the heterogeneity in the discount factor laid out by the characteristics view. To illustrate, Figure 1.13 plots the consumption function and distribution of cash on hand for an impatient and patient individual. The solid black line represents the impatient individual and the solid blue line represents the patient individual. The dotted lines represent the kernel density estimates of the distribution of cash on hand for each individual. If the tax refund is received when individuals hold low cash on hand, the dispersion in the MPC will be relatively high because the consumption functions have very different slopes at this point. Under the circumstances view, all individuals have the same consumption function, so there would be no heterogeneity in the MPC conditional on pre-refund cash on hand.

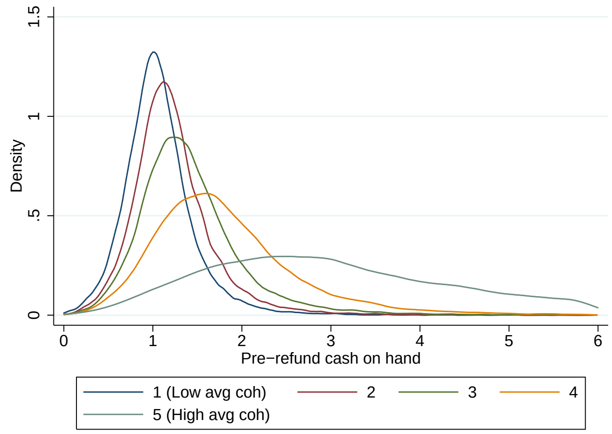
The distribution of pre-refund cash on hand is also consistent with the characteristics view. Figure 1.13 shows that in the simulated data, the cash on hand distribution of impatient individuals is more tightly centered around a lower mean. Conversely, the cash on hand distribution for patient individuals is more dispersed around a higher mean.

Figure 1.13: Theoretical consumption function and distribution of cash on hand



To check whether this same pattern of the distribution of coh_{it}^{PR} holds in the data, Figure 1.14 plots the empirical coh_{it}^{PR} distribution by $\overline{coh}_i$ quintiles. Consistent with the theory, individuals with low average cash on hand tend to have a tighter distribution of pre-refund cash on hand centered around a lower mean. Conversely, individuals with high average cash on hand tend to have a more disperse distribution of pre-refund cash on hand centered around a higher mean. This pattern explains the size of the confidence intervals for each estimate of MPC_{jk} in Figure 1.12. For individuals with low average cash on hand, estimates at the lower quintiles of pre-refund cash on hand are measured with relatively high precision. However, the estimates for pre-refund cash on hand quintiles 4 and 5 are rather imprecise because it is rare that these individuals hold such high levels of pre-refund cash on hand.

Figure 1.14: Empirical coh_{it}^{PR} distribution by $\overline{coh}_i$ quintiles



To summarize, this section estimated the relationship between the MPC and cash on hand at both the individual-refund and quantile level. Both levels of aggregation confirm that persistent characteristics play a role in explaining MPC heterogeneity above and beyond temporary circumstances. I interpret these findings as evidence in favor of the characteristics view. The analysis also provides novel evidence that the joint income, consumption, and saving behavior is consistent with the buffer stock model which includes heterogeneity in the discount factor.

1.4.3 Variance decomposition

This section decomposes the variance of the MPC that is attributable to cash on hand into circumstances and characteristics components. The analysis first starts by adapting the quintile level analysis in the previous section to isolate the circumstances and characteristics components of cash on hand. The MPCs for the adjusted quintile interactions are estimated and plotted to visualize the decomposition. Lastly, the point estimates of the share of the variance in the MPC explained by both circumstances and characteristics components of cash on hand are calculated.

1.4.3.1 Graphical analysis

The graphical analysis starts by adjusting the quintiles in the previous section to capture the effect of circumstances and characteristics. The previous section estimated the MPC using the interactions of quintiles of pre-refund (coh_{it}^{PR}) and average cash on hand ($\overline{coh_i}$). Previous sections showed that $\overline{coh_i}$ captures the characteristics component of cash on hand because it acts as a proxy for the discount factor in the buffer stock theory. coh_{it}^{PR} does not, however, isolate the circumstances component of cash on hand because it is also influenced by the discount factor. To isolate the circumstances component, demeaned pre-refund cash on hand is used. More precisely, the quintiles are based on $coh_{it}^{PR} - \overline{coh_i}^{PR}$ instead of coh_{it}^{PR}. Table 1.7 shows the mean of the demeaned pre-refund cash on hand quintiles and the median of the average cash on hand quintiles. As expected, $coh_{it}^{PR} - \overline{coh^{PR}}_i$ has a mean of 0 and is approximately normally distributed.

Table 1.7: Quintile sample statistics

	$coh_{it}^{PR} - \overline{coh^{PR}}_i$		$\overline{coh_i}$	
Quintile	mean	N	median	N
1	-0.77	26,681	1.20	26,681
2	-0.22	26,680	1.36	26,680
3	-0.02	26,680	1.58	26,680
4	0.17	26,680	2.00	26,680
5	0.84	26,680	3.49	26,680
Total	0.00	133,401	1.59	133,401

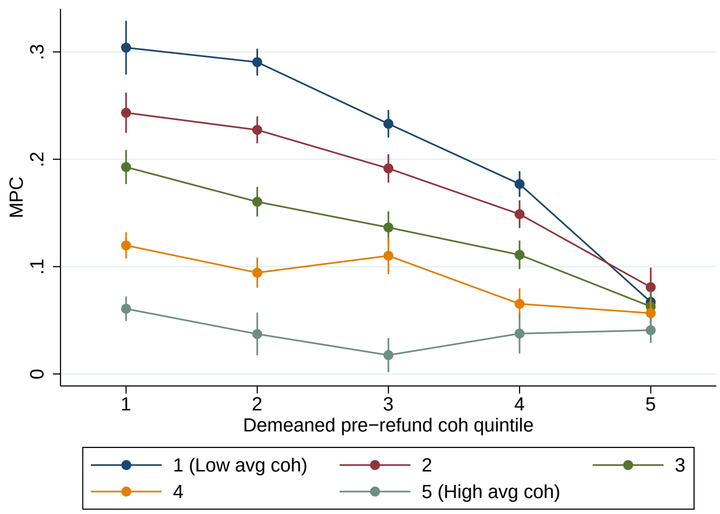
The MPC for each quintile interaction is estimated using the following specification

$$c_{it} = \alpha_{dk} + MPC_{dk} \times ref_{it} + \delta_t + \varepsilon_{it} \tag{1.19}$$

where i represents individual, t represents months, d represents demeaned pre-refund cash on hand quintiles , and k represents average cash on hand quintiles.

Figure 1.15 plots the coefficients of MPC_{dk}. The main difference with Figure 1.12 in the previous section is that now the demeaned pre-refund cash on hand quintiles represent different actual cash on hand levels. This isolates the circumstances component of cash on hand and also leads to a more even distribution of observations across the quintiles. This is reflected in the fact that the standard errors are fairly consistent across $coh_{it}^{PR} - \overline{coh^{PR}}_i$ quintiles relative to using the raw quintiles of coh_{it}^{PR}.

Figure 1.15: MPC by quintile interactions



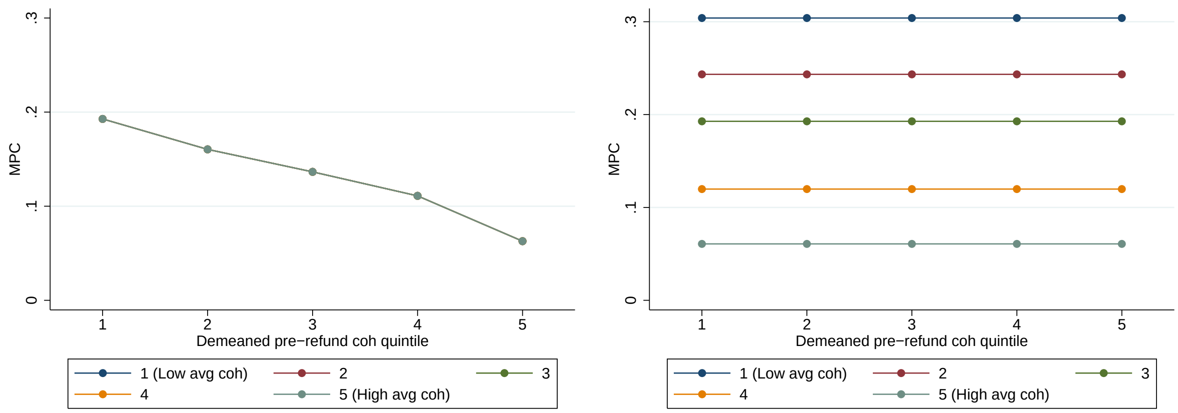
This figure can be thought of as decomposing the circumstances and characteristics components of cash on hand represented by within and across individual variation. For example, consider the top blue line which represents individuals with low average cash on hand. The MPC drops from about 0.3 to about 0.08 when moving from the first to the last quintile of $coh_{it}^{PR} - \overline{coh^{PR}}_i$. Because the blue line holds average cash on hand constant, this drop from 0.3 to 0.08 represents the change in MPC when cash on hand changes within a person due to a change in circumstances. Another pattern that emerges is that the MPC drops more for low average cash on hand individuals relative to high relative cash on hand individuals. This pattern is explained by once again referring to the simulated consumption functions in Figure 1.13. For impatient individuals (identified in the data via low average cash on hand), their pre-refund cash on hand distribution is tightly centered around a lower mean. The left tail of the distribution includes regions where the consumption function is very steep while the consumption function flattens out as cash on hand increases. This is consistent with the large change in MPC seen for the low average cash on hand individuals as cash on hand moves from the lowest to the highest quintile of $coh_{it}^{PR} - \overline{coh^{PR}}_i$. Conversely, patient individuals (identified in the data via high average cash on hand) have a more dispersed distribution around a larger mean. The cash on hand distribution rarely falls into areas where the consumption function is very steep. Therefore, there will be a less dramatic change in the size of the MPC as cash on hand moves from the lowest to the highest quintile of $coh_{it}^{PR} - \overline{coh^{PR}}_i$.

A change in the persistent characteristics component of cash on hand holding circumstances constant is represented by looking at how the MPC changes when holding the $coh_{it}^{PR} - \overline{coh^{PR}}_i$ quintile constant and moving across $\overline{coh}_i$ quintiles. For example, $coh_{it}^{PR} - \overline{coh^{PR}}_i$ quintile 3 represents the case where pre-refund cash on hand is close to the mean for each individual. At this quintile, the MPC ranges from about 0.23 for those with low $\overline{coh}_i$ and about 0.02 for those with high $\overline{coh}_i$. This distance represents across individual variation or variation that results from the persistent characteristics component of cash on hand which is a proxy for the discount factor.

To better understand how circumstances and characteristics influence the estimates in this section, Figure 1.16 shows how Figure 1.15 would look if only one source of variation was important.

Figure 1.16: Alternative scenarios

(a) Only circumstances matter (b) Only characteristics matter



For example, panel (a) shows the scenario in which circumstances drives all the variation in the MPC. In this case, the MPC will fall as demeaned pre-refund coh increases. However, there will is no variation across average cash on hand quintiles. Conversely, panel (b) shows the scenario in which characteristics drives all the variation in the MPC. In this case, the MPC does not change as demeaned pre-refund cash on hand quintiles change. All the variation is driven by across individual variation and so the result is horizontal parallel lines. The estimates plotted in Figure 1.15 represent a middle ground between the extremes in Figure 1.16. The next section builds upon this intuition and quantitatively estimates the contribution of the circumstances and characteristics component in explaining the variance of the MPC.

1.4.3.2 Regression analysis

The previous section estimated the MPC for each interaction of $coh_{it}^{PR} - \overline{coh}_i^{PR}$ and $\overline{coh}_i$ quintiles. This section uses the same variables to calculate the point estimates of the share of the variance in the MPC explained by both the circumstances and characteristics components of cash on hand. I use the same framework defined earlier in section 1.2.4.4 which approximates the relationship between the MPC and the different concepts of cash on hand as follows.

$$MPC_{it} = \alpha + \underbrace{\gamma_1 \times \overline{coh}_i}_{\text{characteristics}} + \underbrace{\gamma_2 \times (coh_{it}^{PR} - \overline{coh}_i^{PR})}_{\text{circumstances}} + \varepsilon_{it} \tag{1.20}$$

where $\mathbb{E}[\varepsilon_{it}] = 0$. Under this specification, the variance is easily decomposed because all the terms are uncorrelated with each other. The following equation applies the variance operator to both sides.

$$var(MPC_{it}) = var(\alpha) + var(\gamma_1 \times \overline{coh}_i) + var(\gamma_2 \times (coh_{it}^{PR} - \overline{coh}_i^{PR})) + var(\varepsilon_{it}) \tag{1.21}$$

Defining $var(\gamma_1 \times \overline{coh}_i) = \sigma^2_{char}$ and $var(\gamma_2 \times (coh_{it}^{PR} - \overline{coh}_i^{PR})) = \sigma^2_{circ}$, these terms capture the variance contribution of the persistent characteristics and circumstances component of cash on hand respectively. Defining $\phi_{char} = \frac{\sigma^2_{char}}{\sigma^2_{circ} + \sigma^2_{char}}$, this value represents the fraction of $var(MPC_{it})$ explained by cash on hand that is attributable to characteristics.

The results from estimating specification (1.20) are presented in table 1.8 below. The sign of the coefficients show that both $\overline{coh}_i$ and $(coh_{it}^{PR} - \overline{coh}_i^{PR})$ vary negatively with the MPC. This is consistent with economic theory, the earlier empirical analysis, and the empirical literature. Calculating the ratio $\phi_{char} = \frac{\sigma^2_{char}}{\sigma^2_{circ} + \sigma^2_{char}} = 0.46$ shows that about half of the variance of the MPC that is explained by cash on hand is driven by the characteristics component. This is in line with the graphical results in the previous section that showed both the characteristics and circumstances component play a role in explaining the variance of the MPC.

Table 1.8: Variance decomposition

VARIABLES	$\hat{\gamma}$	$\hat{Var}$	$\hat{\gamma}^2 \times \hat{Var}$	$Var\hat{S}hare$
$\overline{coh}_i$	-0.051***	1.040	0.0027	0.46***
	(0.003)			(0.033)
$(coh_{it}^{PR} - \overline{coh}_i^{PR})$	-0.093***	0.370	0.0032	0.54***
	(0.004)			(0.033)
Observations	129,823	129,823	129,823	129,823

For $\hat{\gamma}$, heteroskedasticity and cluster robust standard errors in parentheses.
For $Var\hat{S}hare$, cluster bootstrapped standard errors in parenthesis with 10,000 draws.
*** p<0.01, ** p<0.05, * p<0.1

There are some alternative ways to calculate the characteristics variance share using the framework laid out in this section. The first alternative is to define the circumstances share differently. In this section, the circumstances component is defined as pre-refund cash on hand subtracted by its mean $(coh_{it}^{PR} - \overline{coh}_i^{PR})$. The alternative methodology is to demean pre-refund cash on hand by total average cash on hand instead of just the average of pre-refund cash on hand $(coh_{it}^{PR} - \overline{coh}_i)$. The difference arises because the circumstances component only captures the state of liquidity preceding the receipt of a refund that occurs once a year. Demeaning using $\overline{coh}_i^{PR}$ represents deviations from average liquidity right before the refund is received while demeaning using $\overline{coh}_i$ also includes seasonal fluctuations in liquidity over the course of the year. Appendix 1.8.5 discusses the calculation of the characteristics variance share using this alternative method. While the estimate of the characteristics share is qualitatively different under the alternative specification, it is still large enough (0.42) to favor the characteristics view over the circumstances view.

The second alternative is to include higher order terms in specification (1.20). While including higher order terms helps to better capture the total variance of the MPC, it results in roughly the same characteristics variance share as not including the terms. Appendix 1.8.6 presents the analysis with the higher order terms included.

To summarize, this section used both graphical and regression analysis to show that the data is more consistent with the characteristics view rather than the circumstances view.

1.5 Structural estimation

This section connects the empirical results back to the model by estimating the model parameters via the method of simulated moments. The estimation proceeds in two steps. In the first step, I estimate and calibrate the parameters of the model that don't rely on the explicit solution of the model. In the second stage, I estimate the remaining parameters of the model that rely on the model solution conditional on the first stage estimates.

1.5.1 First stage estimation and calibration

I calibrate the coefficient of risk aversion (θ), the interest rate (r), and the borrowing limit (b) by setting them to reasonable values. As mentioned earlier, the discount factor (β) and θ are not easily separately identified so I choose to set $\theta = 1$ which allows me to compare β to other papers using similar methods such as Carroll et al. (2015) and Krueger, Mitman and Perri (2016).

The income process is governed by the level of persistent (ρ) and the standard deviation of income shocks (σ_y). I estimate these parameters directly from the data by using the panel nature of the income process. The estimation process was presented earlier in section 1.3.3.3. The average level of tax refunds in the data is also estimated directly from the data by taking the unconditional mean of normalized tax refunds.

The values of the first stage parameters are listed below in table 1.9.

Table 1.9: First stage parameter values

Parameter	Value	Source	Description
$u(x)$	$\frac{x^{1-\theta}}{1-\theta}$	CRRA utility	utility function
θ	1	standard	coefficient of relative risk aversion
ρ	0	income time series	income shock persistence
σ_y	0.20	income time series	S.D. of temporary shocks
$refund_{it}$	0.6	tax refund distribution	average normalized refund
r	0.01 / 12	external savings data	interest rate
b	0	no borrowing condition	borrowing limit

Notes: The parameters correspond to a monthly frequency.

1.5.2 Second stage estimation

In the second stage, I use the method of simulated moments to estimate the parameters that rely explicitly on the model. This estimation procedure is used because there is no simple analytic expression for the theoretical moments in the model.

More specifically, the average discount factor ($\overline{\beta}$) and the dispersion in the discount factor (Δ) are estimated by matching the fraction of $var(MPC_{it})$ explained by cash on hand that is attributable to characteristics (ϕ_{char}) and median cash on hand ($\widetilde{CoH_i}$). The parameters are exactly identified because I use two moments to estimates two parameters.

The parameter estimate $\hat{\Theta} = \{\hat{\overline{\beta}}, \hat{\Delta}\}$ is the solution to the criterion function

$$\hat{\Theta} = \arg\min_{\Theta}(m_{data} - m_{sim}(\Theta))(m_{data} - m_{sim}(\Theta))' \tag{1.22}$$

where $m = \{\phi_{char}, \widetilde{CoH_i}\}$, m_{data} represent moments calculated from the data, and $m_{sim}(\Theta)$ represent moments calculated from simulating the model under parameters Θ.

The parameter estimates and empirical moments are shown in table 1.10.

Table 1.10: Second stage parameter and moment estimates

Parameter	Value	Description
$\hat{\overline{\beta}}$	0.9894	average discount factor
$\hat{\Delta}$	0.0103	discount factor dispersion
Moment		
$\hat{\phi}_{char}$	0.4600	characteristics variance share
$\widetilde{\hat{CoH}}_i$	1.5900	median cash on hand

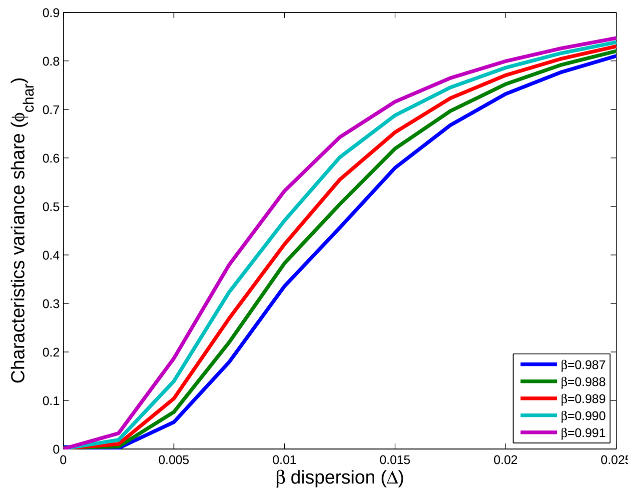
Notes: The parameters correspond to a monthly frequency.

The estimate of the average discount factor ($\overline{\beta}$) is mainly driven by the fact that the median level of cash on hand is 1.59 times monthly income. This moment represents the target buffer stock conditional on the first stage value of the variance of income and risk aversion. The estimate of discount factor dispersion (Δ) is mainly driven by the fact that roughly half of the variance in the MPC is driven by the characteristics component of cash on hand. This implies a fairly important role of across individual heterogeneity and is reflected in the 0.0103 value.

The estimated parameters are roughly in line with Carroll et al. (2015) who calibrate their model at the quarterly level by matching either liquid financial and retirement assets.

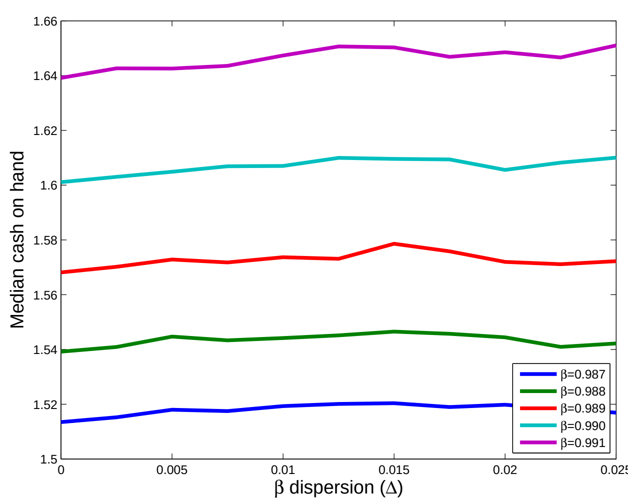
The method of simulated moments jointly estimates $\Theta = \{\overline{\beta}, \Delta\}$ by matching the moments $m = \{\phi_{char}, \widetilde{CoH}_i\}$. Plotting how the simulated moments $m_{sim}(\Theta)$ vary with Θ is helpful in developing intuition about identification. Figure 1.17 plots ϕ_{char} as a function of each parameter. Changes in Δ are represented on the x-axis and changes in $\overline{\beta}$ are represented by different colored lines. The figure shows that ϕ_{char} is much more sensitive to changes in Δ relative to changes in $\hat{\beta}$. ϕ_{char} can be thought of as measuring across individual variation in the MPC. Therefore, holding the variance of temporary shocks constant while increasing the dispersion of types of individuals will lead to a higher share of the variance being driven by persistent differences across individuals. The actual level of $\overline{\beta}$ does not influence this measure as much. Therefore, we can think of Δ being identified primarily through ϕ_{char}.

Figure 1.17: Characteristics component share



Similarly, Figure 1.18 plots $\widetilde{CoH}_i$ as a function of each parameter. The figure shows that $\widetilde{CoH}_i$ is much more sensitive to changes in $\overline{\beta}$ relative to changes in Δ. Earlier analysis showed that more patient individuals tend to hold a higher buffer stock. Therefore, the relationship between $\overline{\beta}$ is straightforward. While Δ does have some effect on $\widetilde{CoH}_i$ it is relatively minor. Therefore, we can think of $\overline{\beta}$ being identified primarily through $\widetilde{CoH}_i$.

Figure 1.18: Median cash on hand

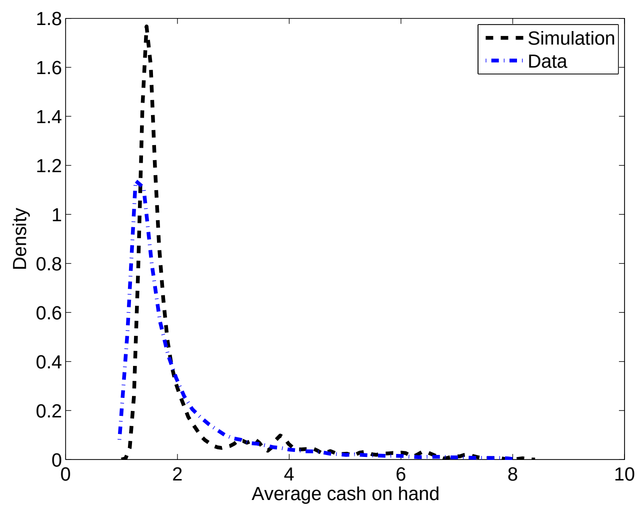


1.5.3 Fit of other variables

This section assesses the fit of variable that weren't explicitly targeted in the estimation procedure.

Cash on hand distribution Figure 1.19 compares the average cash on hand distribution in the model to the data. The fitted model is able to replicate the long right tail of the average cash on hand distribution in the cross-section. This partly explains why the estimates are similar to Carroll et al. (2015). Their paper estimates the dispersion of the discount factor by matching the shape of the liquid assets distribution. Therefore, introducing heterogeneity in the discount factor is important to explaining the relationship between the MPC and cash on hand as well as explaining inequality in the wealth distribution.

Figure 1.19: Average cash on hand distribution



MPC The aggregate MPC in the model is 0.19 compared to 0.14 in the data. While this paper doesn't focus on the aggregate MPC, it is reassuring to know that the model is able to capture it relatively well.

1.5.4 Alternative specification

This section shows the results of estimating alternative specifications. In one case, I estimate risk aversion while holding the time discount factor constant. In another case, I estimate the present-bias term in a quasi-hyperbolic discounting model.

1.5.4.1 Estimating risk aversion

An alternative way to specify the model is to fix time preference (β) and vary risk aversion (θ). I fix $\beta = 0.9894$ and allow θ to be distributed uniformly. Table 1.11 shows the estimated parameters under this alternative specification. The estimated θ falls within the range commonly seen in other studies.

Table 1.11: Alternative parameter estimates

Parameter	Value	Description
$\hat{\bar{\theta}}$	1.1353	average risk aversion
$\hat{\Delta}$	0.8081	risk aversion dispersion
Moment		
$\hat{\phi}_{char}$	0.4600	characteristics variance share
$\widetilde{\widehat{CoH}}_i$	1.5900	median cash on hand

Notes: The parameters correspond to a monthly frequency.

1.5.4.2 Estimating present-bias

Another way of interpreting characteristics is the level of present bias of individuals. This section estimates heterogeneity in the present-bias term of a quasi-hyperbolic discount model while holding the long run discount factor and risk aversion parameters constant. I use the buffer-stock consumption model with quasi-hyperbolic discounting from Harris and Laibson (2002).[7] The individual is now modeled as a sequence of autonomous temporal selves. These selves are indexed by the respective

[7] Without a commitment device, a quasi-hyperbolic individual behaves much the same as an impatient exponential discounter.

periods, t=0,1,2,.... Self t receives payoff

$$\mathbb{E}_t\left[U(C_{it})+\beta_i\sum_{j=1}^{\infty}\delta^j U(C_{it})\right] \tag{1.23}$$

where I allow for heterogeneity in β and a common value for δ. In order to match the highest level of patience under the exponential discounting model, I set $\delta = 0.9997$.

The corresponding Bellman equation is

$$W_{it}(X_{it}) = \max_{C_{it}} U(C_{it}) + \delta\mathbb{E}\left[(W_{it+1} - \epsilon_i U\circ g\circ W'_{it+1})((1+r)(X_{it} - C_{it} + y_{it+1})\right] \tag{1.24}$$

where $\epsilon_i = 1-\beta_i$ and $g = (U')^{-1}$

I solve the model numerically using an endogenous grid method suitably modified to account for the new terms introduced by quasi-hyperbolic preferences. This makes use of the quasi-hyperbolic euler equation which is

$$U'(C(X_{it}) \geq (1+r)\mathbb{E}_t\Big[\Big(C'(X_{it+1})\beta\delta + (1-C'(X_{it+1}))\delta\Big)U'(C(X_{it+1})\Big] \tag{1.25}$$

I estimate the parameters using the method of simulated moments and display the results in table 1.12.

Table 1.12: Alternative parameter estimates

Parameter	Value	Description
$\hat{\bar{\beta}}$	0.8943	average present bias factor
$\hat{\Delta}$	0.0495	present bias dispersion
Moment		
$\hat{\phi}_{char}$	0.4600	characteristics variance share
$\widetilde{\hat{CoH}_i}$	1.5900	median cash on hand

Notes: The parameters correspond to a monthly frequency.

1.6 Policy Implications

1.6.1 Tax rebate simulation

This section analyzes the consumption response to a tax rebate under the two different views. While tax refunds are modeled as anticipated changes to income, I model the tax rebate as an unanticipated shock. Since tax rebates are often issued in times of recession, I perform the simulation with and without aggregate shocks to income.[8]

Great recession shock I calibrate the magnitude of income shocks due to the great recession from the PSID. In order to match the model, I first split the sample up into different quintiles of cash on hand.[9] I use the average over the period 2001-2007 to create the quintiles. I then calculate income shocks by taking the log difference between labor income in 2009 relative to average income using the period 2001-2007. Table 1.13 shows the mean value for cash on hand and the income shock for each quintile.

[8]Since the model is not a general equilibrium model, the concept of aggregate shock is an income shock above and beyond the standard temporary shock.

[9]The corresponding PSID variable is total balance in saving and checking accounts.

Table 1.13: Quintile sample statistics

Quintile	cash on hand	income shock	N
1	0.005	-0.123	542
2	0.042	-0.124	612
3	0.114	-0.087	611
4	0.293	-0.077	587
5	15.72	-0.058	459
Total	0.00	-0.095	2811

The MPC is then calculated by simulating 10,000 individuals under each view and shocking all individuals with unexpected income equivalent to one month of income. In order to understand the role of heterogeneity, I also calculate the MPC at different quintiles of average cash on hand. Table 1.14 shows the results of the simulation.

The columns are divided into two sections based on whether the great recession shock is applied or not. Under each shock scenario the MPC is calculated under the circumstances and the characteristics view. For the great recession shock the characteristics view is further split into two cases. In one case I shock all individuals with the average income shock. In another case, I use the true distribution of income shocks across individuals. As expected, under the circumstances view there is no heterogeneity along the persistent characteristics dimension which is captured by the average cash on hand quintiles. Also as expected, under the characteristics view, individuals with low average cash on hand tend to have high MPCs while those with high average cash on hand tend to have lower MPCs. Under the estimated parameters, this leads to a larger aggregate MPC under the characteristics view relative to the circumstances view.

The simulation under the great recession shock shows similar patterns in heterogeneity but the gap between the aggregate MPC is slightly higher. By breaking out the response into an average shock and individual shock case, we see that the higher MPC is due to the fact that individuals with low average cash received a larger shock than those with high cash on hand.

Table 1.14: Tax rebate simulation

	No aggregate shock		Great recession shock		
	Circumstances	Characteristics	Circumstances	Characteristics (average shock)	Characteristics (individual shock)
Aggregate	0.32	0.37	0.51	0.55	0.58
$\overline{coh_i}$ Q1	0.32	0.63	0.50	0.85	0.96
$\overline{coh_i}$ Q2	0.32	0.47	0.51	0.68	0.77
$\overline{coh_i}$ Q3	0.32	0.34	0.51	0.54	0.54
$\overline{coh_i}$ Q4	0.32	0.27	0.52	0.46	0.44
$\overline{coh_i}$ Q5	0.31	0.13	0.51	0.24	0.19
$\overline{\beta}$	0.991	0.989	0.991	0.989	0.989
Δ	0.000	0.010	0.000	0.010	0.010

Notes: Simulated N=10,000

In summary, the tax rebate simulation shows that the aggregate MPC can differ under the two views. However, the are some important caveats. First, the sample used in this study is not necessarily representative of the U.S. population. Second, the uniform distribution assumption is made partially for convenience and simplicity. In future work, using a more representative sample and more realistic assumptions about the distribution of preferences will lead to more accurate estimates of the aggregate MPC.

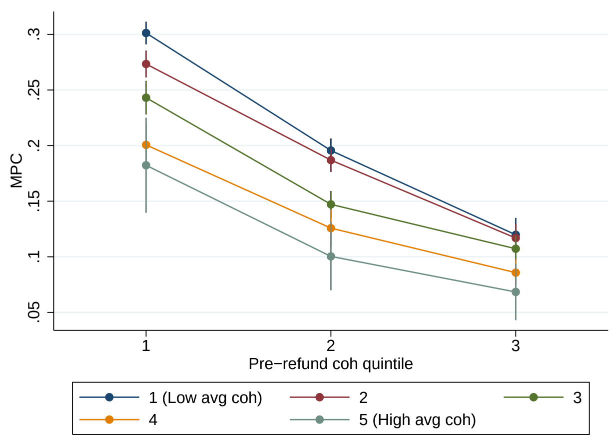
1.6.2 MPC targeting

Tax rebates are usually issued as part of an economic stimulus package to boost consumption and are loosely targeted based on income. For example, the 2008 tax rebates started to phase out for taxpayers with adjusted gross income greater than $75,000 for single individuals and $150,000 for joint filers.

The results from this paper imply that it is possible to target the MPC much more precisely. For example, Figure 1.20 shows that the MPC varies predictably based on interactions of pre-refund and average cash on hand quintiles. Individuals with low pre-refund and low average cash on hand tend to have the highest MPCs. Conversely individuals with high pre-refund and high average cash on hand tend to have the lowest MPCs. This implies that in order to maximize the MPC, individuals with low pre-refund and average cash on hand should receive higher rebates. In terms of the terminology used in this paper, the fiscal authority can choose to target circumstances or characteristics.

In order to target circumstances, the fiscal authority can calculate recent deviations in income from permanent income (proxied using a recent average). Characteristics can be estimated by calculating target buffer stock levels using interest income filed on recent tax returns.

Figure 1.20: MPC by pre-refund and average cash on hand quintile interactions



While targeting the MPC in a way suggested by the model may be politically or operationally unfeasible, the analysis highlights that understanding the sources of heterogeneity in the MPC can provide fiscal authorities more levers in tailoring stimulus packages.

1.7 Conclusion

This paper tests the two leading views in the literature on what theoretical mechanisms drive the negative correlation between the MPC and cash on hand. Under the circumstances view, individuals are ex-ante identical but differ in the circumstances they face. Under the characteristics view, the economy is populated by different types of individuals. These views are represented using a parsimonious buffer stock model with discount factor heterogeneity. Testing the two views is complicated by the fact that most data sources are not able to disentangle the effect of circumstances and characteristics. This paper overcomes these challenges by by using a novel panel dataset on joint spending, income, and liquid saving behavior from a personal finance app. Identification is achieved by comparing the MPC within individuals over time relative to the MPC across individuals.

The empirical results show that conditional on cash on hand levels in a certain year, average cash on hand levels explain a significant amount of MPC heterogeneity. Stated in terms of the model, even conditional on temporary circumstances, persistent characteristics are important in explaining MPC heterogeneity. Furthermore, a variance decomposition shows that persistent characteristics explain roughly half of the variance in the MPC while temporary circumstances explain the other half. This evidence shows that the characteristics view is much more consistent with the data than the circumstance view.

Lastly, the dispersion of the discount factor is estimated using the simulated method of moments and is roughly in line with other studies. Using the estimated parameters, the spending response to a tax rebate is simulated under the two views. The results show that ignoring heterogeneity in persistent characteristics will under predict the aggregate MPC. Therefore, the simulations show that which view obtains in the data can have important implications for policy relevant outcomes such as the aggregate MPC. Future research should focus on estimating the parameters under a more representative sample and relaxing the assumptions on the preference distribution. This will lead to more realistic and reliable estimates of the aggregate MPC out of fiscal stimulus.

Bibliography

Ameriks, John, Andrew Caplin, and John Leahy (2003) "Wealth Accumulation and the Propensity to Plan," *The Quarterly Journal of Economics*, Vol. 118, No. 3, pp. 1007–1047.

Angeletos, George-Marios, David Laibson, Andrea Repetto, Jeremy Tobacman, and Stephen Weinberg (2001) "The Hyperbolic Consumption Model: Calibration, Simulation, and Empirical Evaluation," *Journal of Economic Perspectives*, Vol. 15, No. 3, pp. 47–68.

Baker, Scott R (2015) "Debt and the consumption response to household income shocks," *Available at SSRN*.

Baugh, Brian, Itzhak Ben-David, and Hoonsuk Park (2014) "Disentangling Financial Constraints, Precautionary Savings, and Myopia: Household Behavior Surrounding Federal Tax Returns,"Technical report, National Bureau of Economic Research.

Campbell, John Y. and N. Gregory Mankiw (1989) "Consumption, Income and Interest Rates: Reinterpreting the Time Series Evidence," NBER Chapters, National Bureau of Economic Research, Inc.

Carroll, Christopher (2004) "Theoretical Foundations of Buffer Stock Saving," Working Paper 10867, National Bureau of Economic Research.

Carroll, Christopher D. (1997) "Buffer-Stock Saving and the Life Cycle/Permanent

Income Hypothesis," *The Quarterly Journal of Economics*, Vol. 112, No. 1, pp. 1–55.

——— (2006) "The method of endogenous gridpoints for solving dynamic stochastic optimization problems," *Economics Letters*, Vol. 91, No. 3, pp. 312–320.

Carroll, Christopher D. and Miles S. Kimball (1996) "On the Concavity of the Consumption Function," *Econometrica*, Vol. 64, No. 4, pp. 981–992.

Carroll, Christopher D., Jiri Slacalek, Kiichi Tokuoka, and Matthew White (2015) "The distribution of wealth and the marginal propensity to consume," Working Paper Series 1655, European Central Bank.

Congressional Budget Office (2009) "Did the 2008 Tax Rebates Stimulate Short-Term Growth?." Economic and Budget Issue Brief.

Council of Economic Advisers (2010) "The Economic Impact of the American Recovery and Reinvestment Act." Third Quarterly Report, Executive Office of the President.

Deaton, Angus (1991) "Saving and Liquidity Constraints," *Econometrica*, Vol. 59, No. 5, pp. 1221–1248.

Ganong, Peter and Pascal Noel (2016) "How Does Unemployment Affect Consumer Spending?."

Gelman, Michael (2017) "What Drives Heterogeneity in the Marginal Propensity to Consume? Temporary Shocks vs Persistent Characteristics."

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*, Vol. 345, No. 6193, pp. 212–215.

——— (2015) "How Individuals Smooth Spending: Evidence from the 2013 Government Shutdown Using Account Data," Working Paper 21025, National Bureau of Economic Research.

Hall, Robert E (1978) "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence," *Journal of Political Economy*, Vol. 86, No. 6, pp. 971–87.

Harris, Christopher J. and David Laibson (2002) *Hyperbolic Discounting and Consumption*, pp. 258–298. Eighth World Congress.

Hurst, Erik (2003) "Grasshoppers, Ants, and Pre-Retirement Wealth: A Test of Permanent Income," Working Paper 10098, National Bureau of Economic Research.

Jappelli, Tullio and Luigi Pistaferri (2010) "The Consumption Response to Income Changes," *Annual Review of Economics*, Vol. 2, No. 1, pp. 479–506.

——— (2014) "Fiscal Policy and MPC Heterogeneity," *American Economic Journal: Macroeconomics*, Vol. 6, No. 4, pp. 107–136.

Kaplan, Greg and Giovanni L Violante (2014) "A model of the consumption response to fiscal stimulus payments," *Econometrica*, Vol. 82, No. 4, pp. 1199–1239.

Kimball, Miles S. (1990) "Precautionary Saving in the Small and in the Large," *Econometrica*, Vol. 58, No. 1, pp. 53–73.

Krueger, Dirk, Kurt Mitman, and Fabrizio Perri (2016) "Macroeconomics and Household Heterogeneity," Working Paper 22319, National Bureau of Economic Research.

Krusell, Per and Jr. Smith, Anthony A. (1998) "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, Vol. 106, No. 5, pp. 867–896.

Kuchler, Theresa (2015) "Sticking to your plan: Hyperbolic discounting and credit card debt paydown," *Available at SSRN 2629158*.

Laibson, David (1997) "Golden Eggs and Hyperbolic Discounting," *The Quarterly Journal of Economics*, Vol. 112, No. 2, pp. 443–478.

Lusardi, Annamaria (1996) "Permanent Income, Current Income, and Consumption: Evidence From Two Panel Data Sets," *Journal of Business & Economic Statistics*, Vol. 14, No. 1, pp. 81–90.

Parker, Jonathan A. (2015) "Why Don't Households Smooth Consumption? Evidence from a 25 million dollar experiment,"Technical report, Manuscript, MIT.

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*, Vol. 103, No. 6, pp. 2530–53.

Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay (2011) "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, Vol. 12, pp. 2825–2830.

Reis, Ricardo (2006) "Inattentive consumers," *Journal of Monetary Economics*, Vol. 53, No. 8, pp. 1761–1800.

Sahm, Claudia R., Matthew D. Shapiro, and Joel Slemrod (2012) "Check in the Mail or More in the Paycheck: Does the Effectiveness of Fiscal Stimulus Depend on How It Is Delivered?" *American Economic Journal: Economic Policy*, Vol. 4, No. 3, pp. 216–250.

Shapiro, Jesse M. (2005) "Is there a daily discount rate? Evidence from the food stamp nutrition cycle," *Journal of Public Economics*, Vol. 89, No. 23, pp. 303–325.

Souleles, Nicholas S. (1999) "The Response of Household Consumption to Income Tax Refunds," *The American Economic Review*, Vol. 89, No. 4, pp. 947–958.

Stephens, Melvin (2003) ""3rd of tha Month": Do Social Security Recipients Smooth Consumption between Checks?" *The American Economic Review*, Vol. 93, No. 1, pp. 406–422.

——— (2006) "Paycheque Receipt and the Timing of Consumption," *The Economic Journal*, Vol. 116, No. 513, pp. 680–701.

Zeldes, Stephen P. (1989) "Optimal Consumption with Stochastic Income: Deviations from Certainty Equivalence," *The Quarterly Journal of Economics*, Vol. 104, No. 2, pp. 275–298.

1.8 Appendix

1.8.1 Proof for variance decomposition

In equation 1.13, α, γ_1, γ_2 are assumed to be constant while ε_{it} is assumed to be uncorrelated with any of the other regressors. Therefore, the only terms that can plausibly generate non-zero covariance terms are $\overline{coh}_i$ and $(coh_{it}^{PR} - \overline{coh}_i^{PR})$. The theoretical covariance is calculated using the standard covariance formula as follows:

$$\mathbb{E}[(\overline{coh}_i - \mathbb{E}[\overline{coh}_i])(coh_{it}^{PR} - \overline{coh}_i^{PR} - \mathbb{E}[coh_{it}^{PR} - \overline{coh}_i^{PR}]) = \tag{1.26}$$

$$\mathbb{E}[\mathbb{E}[(\overline{coh}_i - \mathbb{E}[\overline{coh}_i])(coh_{it}^{PR} - \overline{coh}_i^{PR} - \mathbb{E}[coh_{it}^{PR} - \overline{coh}_i^{PR})|i]] = \tag{1.27}$$

$$\mathbb{E}[(\overline{coh}_i - \mathbb{E}[\overline{coh}_i])\mathbb{E}[(coh_{it}^{PR} - \overline{coh}_i^{PR} - \mathbb{E}[coh_{it}^{PR} - \overline{coh}_i^{PR})|i]] = \tag{1.28}$$

$$\mathbb{E}[(\overline{coh}_i - \mathbb{E}[\overline{coh}_i])0] = \tag{1.29}$$

$$0 \tag{1.30}$$

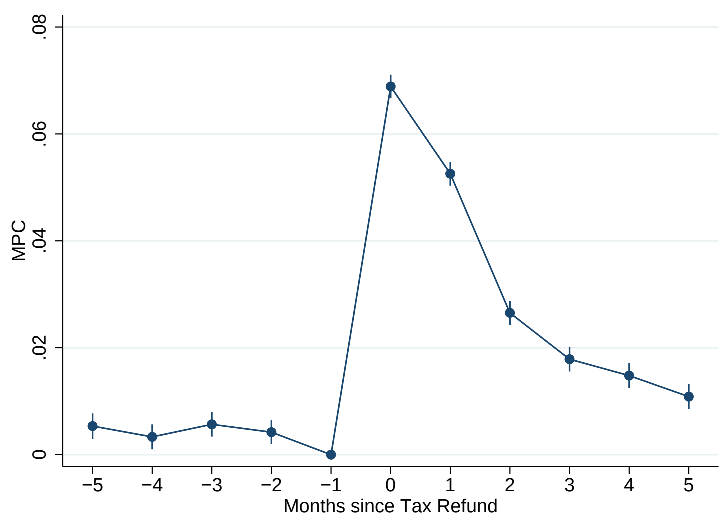
where the second line uses the law of iterated expectations and the third line uses the fact that $\overline{coh}_i - \mathbb{E}[\overline{coh}_i]$ is a constant once we condition on individual i. Intuitively, any variable that is invariant at the individual level should not be correlated with any variable that is demeaned at the individual level.

1.8.2 Identifying paychecks

Keywords used to identify paychecks are "dir dep","dirde p","salary","treas xxx fed","fed sal","payroll","ayroll","payrll","payrl","payrol","pr payment","adp","dfas-cleveland","dfas-in" and DON'T include the keywords "ing direct","refund","direct deposit advance","dir dep adv."

1.8.3 Tax refund impulse response function

Figure 1.8.1: Tax refund impulse response function



Notes: 1,445,560 observations from 48,059 individuals. The vertical bars on each coefficient represent 95% confidence intervals using heteroskedasticity robust errors clustered at the individual level.

1.8.4 Machine learning algorithm

Most transactions in the data do not contain direct information on spending category types. However, category types can be inferred from existing transaction data.

In general, the mapping is not easy to construct. If a transaction is made at "McDonalds," it's easy to surmise that the category is "Fast Food Restaurants." However, it is much harder to identify smaller establishments such as "Bob's store." "Bob's store" may not uniquely identify an establishment in the data and it would take many hours of work to look up exactly what types of goods these smaller establishments sell. Luckily, the merchant category code (MCC) is observed for two account providers in the data. 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. If an individual uses an account provider that provides MCC information "Bob's store" will map into a spending category type.

The mapping from transaction data to MCC can be represented as $Y = f(X)$ where Y represents a vector of MCC codes and X represents a vector of transactions data. The data is partitioned into two sets based on whether Y is known or not.[10] The sets are also commonly referred to as training and prediction sets. The strategy is to then estimate the mapping $\hat{f}(\cdot)$ from (Y_1, X_1) and predict $\hat{Y}_0 = \hat{f}(X_0)$.

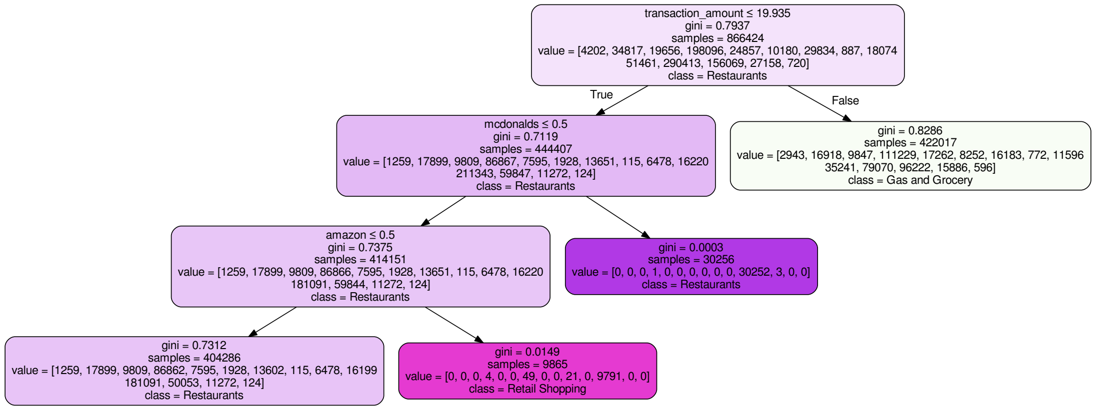
One option for the mapping is to use the multinomial logit model since the dependent variable is a categorical variable with no cardinal meaning. However, this approach is not well suited to textual data because each word would need its own dummy variable. Furthermore, interactions may be important for classifying spending categories. For example "jack in the box" refers to a fast food chain while "jack s surf shop" refers to a retail store. Including a dummy for each word can lead to about 300,000 variables. Including interaction terms will cause the number of variables to grow exponentially and will typically be unfeasible to estimate.

In order to handle the textual nature of the data I use a machine learning algorithm called random forest. A random forest model is composed of many decision trees that map transaction data to MCCs. This mapping is created by splitting the sample up

[10] Y_0 represents the set where Y is not known and Y_1 represents the set where Y is known.

into nodes depending on the features of the data. For example, for transactions that have the keyword “McDonalds” and transaction amounts less that $20, the majority of the transactions are associated with a MCC that represents fast food. To better understand how the decision tree works, Figure 4.8.11 shows an example. The top node represents the state of the data before any splits have been made. The first row “transaction_amount $\leq$ 19.935” represents the splitting criteria of the first node. The second row is the Gini measure which is explained below. The third row show that there are 866,424 total transactions to be classified in the sample. The fourth row “value=[4202,34817,...,27158,720]” shows the number of transactions in each spending category. The last row represents the majority class in this node. Because “Restaurants” has the highest number of transactions, assigning a random transaction to this category minimizes the categorization error without knowing any information about the transaction. At each node in the tree, the sample is split based on a feature. For example, the first split will be based on whether the transaction amount is $\leq$ 19.935. The left node represents all the transactions for which the statement is true and vice versa. Transactions $\leq$ 19.935 are more likely to be “Restraunts” spending while transactions $>$ 19.934 are more likely to be “Gas and Grocery.” In our example, the sample is split further to the left of the tree. Transactions with the string “mcdonalds” are virtually guaranteed to be “Restaurant” spending. A further split shows that the string “amazon” is almost perfectly correlated with the category “Retail Shopping.” How does the algorithm decide which features to split the sample on? The basic intuition is that the algorithm should split the sample based on features that lead to the largest disparities in the different groups. For example, transactions that have the word “mcdonalds” will tend to split the sample into fast food and non-fast food transactions so it is a good feature to split on. Conversely, “bob” is not a very good feature to split on because it can represent a multitude of different types of spending depending on what the other features are.

Figure 1.8.2: Decision tree example



I state the procedure more formally by adapting the notation used in (Pedregosa et al., 2011). Define the possible features as vectors $X_i \in R^n$ and the spending categories as vector $y \in R^l$. Let the data at node m be presented by Q. For each candidate split $\theta = (j, t_m)$ consisting of a feature j and threshold t_m, partition the data into $Q_{left}(\theta)$ and $Q_{right}(\theta)$ subsets so that

$$Q_{left}(\theta) = (X, y)|x_j \leq t_m \tag{1.31}$$

$$Q_{right}(\theta) = Q \setminus Q_{left}(\theta) \tag{1.32}$$

The goal is then to split the data at each node in the starkest way possible. A popular quantitative measure of this idea is called the Gini criteria and is represented by

$$H(X_m) = \sum_k p_{mk}(1 - p_{mk}) \tag{1.33}$$

where $p_{mk} = 1/N_m \sum_{x_i \in R_m} \mathbb{I}(y_i = k)$ represents the proportion of category k observations in node m.

If there are only two categories, the function is is minimized at 0 when the transactions are perfectly split into the two categories[11] and maximized when the transactions are evenly split between the two categories.[12]

Therefore, the algorithm should choose the feature to split on that minimizes the Gini measure at node m

$$\theta^* = argmin_\theta \frac{n_{left}}{N_m} H(Q_{left}(\theta)) + \frac{n_{right}}{N_m} H(Q_{right}(\theta)) \tag{1.34}$$

The algorithm acts recursively so the same procedure is performed on $Q_{left}(\theta^*)$ and $Q_{right}(\theta^*)$ until a user-provided stopping criteria is reached. The final outcome is a decision rule $\hat{f}(\cdot)$ that maps features in the transaction data to spending categories.

This example shows that decision trees are much more effective in mapping high dimensional data that includes text to spending categories. However, fitting just one tree might lead to over-fitting. Therefore, a random forest fits many trees by bootstrapping the samples of the original data and also randomly selecting the features used in the decision tree. With the proliferation of processing power, each tree can be fit in parallel and the final decision rule is based on all the decision trees. The most common rule is take the majority decision of all the trees that are fit.

1.8.5 Demeaning by $\overline{coh_i}$ instead of $\overline{coh_i}^{PR}$

This section compares the results for the characteristics variance share when demeaning coh_{it}^{PR} by $\overline{coh_i}$ instead of $\overline{coh_i}^{PR}$. The difference is that $\overline{coh_i}$ is no longer the mean of coh_{it}^{PR} so the demeaned term will not have mean 0. This also means that the demeaned term will not be orthogonal to $\overline{coh_i}$.

Taking the variance of both sides under the alternative demeaned term results in the following equation that now involves a covariance term.

[11]because 0*1 + 1*0 = 0.

[12]because 0.5*0.5 + 0.5*0.5 = 0.5.

$$var(MPC_{it}) = var(\alpha) + \gamma_1^2 var(\overline{coh}_i) + \gamma_2^2 var(coh_{it}^{PR} - \overline{coh}_i) + 2\gamma_1\gamma_2 cov(\overline{coh}_i, coh_{it}^{PR} - \overline{coh}_i) + var(\varepsilon_{it}) \quad (1.35)$$

Table 1.15 shows the variance-covariance matrix of the MPC and coh terms.

Table 1.15: Variance-covariance matrix

	MPC_{it}	$\overline{coh}_i$	$(coh_{it}^{PR} - \overline{coh}_i^{PR})$	$(coh_{it}^{PR} - \overline{coh}_i)$
MPC_{it}	0.817			
$\overline{coh}_i$	-0.053	1.040		
$(coh_{it}^{PR} - \overline{coh}_i^{PR})$	-0.034	0.000	0.370	
$(coh_{it}^{PR} - \overline{coh}_i)$	-0.046	-0.130	0.368	0.530

In order calculate the variance shares, I first estimate the $\hat{\gamma}$ terms. Table 1.16 shows the results when the different methods of demeaning are used.

Table 1.16: $\hat{\gamma}$ estimation

	(1)	(2)
VARIABLES	MPC_{it}	MPC_{it}
$\overline{coh}_i$	-0.05***	-0.06***
	(0.00)	(0.00)
$(coh_{it}^{PR} - \overline{coh}_i^{PR})$	-0.09***	
	(0.00)	
$(coh_{it}^{PR} - \overline{coh}_i)$		-0.10***
		(0.00)
Observations	129,823	129,823

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Using the input from table 1.15 and 1.16, table 1.17 shows the results of the variance decomposition for each method of demeaning. The column "share" calculates the share of each component among the terms that represent cash on hand.

Table 1.17: Variance decomposition

	$\overline{coh}_i^{PR}$		$\overline{coh}_i$	
	value	share	value	share
$var(MPC_{it})$	0.8169	-	0.8169	-
$\gamma_1^2 var(coh_i)$	0.0027	0.46	0.0042	0.53
$\gamma_2^2 var(coh_{it}^{PR} - \overline{coh}_i^{PR})$	0.0032	0.54	-	-
$\gamma_2^2 var(coh_{it}^{PR} - \overline{coh}_i)$	-	-	0.0055	0.69
$2\gamma_1\gamma_2 cov(\overline{coh}_i, (coh_{it}^{PR} - \overline{coh}_i))$	-	-	-0.0017	-0.21
$var(\varepsilon)$	0.8110	-	0.8089	-

Under the baseline specification, the characteristics variance share is 0.46. If we assume the covariance term is split evenly between the characteristics and circumstances component in the alternative specification, the characteristics variance share is 0.42.

1.8.6 Including higher order terms

This section calculates the characteristics variance share after including higher order terms for the circumstances and characteristics components.

$$\begin{aligned}MPC_{it} = \alpha + &\underbrace{\gamma_1 \times \overline{coh}_i + \gamma_2 \overline{coh}_i^2 + \dots}_{\text{characteristics}} + \\ &\underbrace{\gamma_{n+1} \times (coh_{it}^{PR} - \overline{coh}_i^{PR}) + \gamma_{n+2} \times (coh_{it}^{PR} - \overline{coh}_i^{PR})^2 + \dots}_{\text{circumstances}} + \\ &\varepsilon_{it} \quad (1.36)\end{aligned}$$

where $\mathbb{E}[\varepsilon_{it}] = 0$ and n is the number of terms included for each component. The variance of both sides is calculated as

$$var(MPC_{it}) = var(\alpha) + var\Big(\gamma_1 \times \overline{coh}_i + \gamma_2 \overline{coh}_i^2 + \dots\Big) +$$
$$var\Big(\gamma_{n+1} \times (coh_{it}^{PR} - \overline{coh}_i^{PR}) + \gamma_{n+2} \times (coh_{it}^{PR} - \overline{coh}_i^{PR})^2 + \dots\Big) + var(\varepsilon_{it}) \quad (1.37)$$

I define the variance contribution as $\sigma^2_{char} = var\Big(\gamma_1 \times \overline{coh}_i + \gamma_2 \overline{coh}_i^2 + \dots\Big)$ and $\sigma^2_{circ} = var\Big(\gamma_{n+1} \times (coh_{it}^{PR} - \overline{coh}_i^{PR}) + \gamma_{n+2} \times (coh_{it}^{PR} - \overline{coh}_i^{PR})^2 + \dots\Big)$. These terms capture the variance contribution of the characteristics and circumstances component of cash on hand respectively. Defining $\phi_{char} = \frac{\sigma^2_{char}}{\sigma^2_{circ} + \sigma^2_{char}}$, this value represents the fraction of $var(MPC_{it})$ explained by cash on hand that is attributable to characteristics.

The polynomial approximation stops improving after the sixth degree polynomial. Table 1.18 shows the coefficient estimates for the cash on hand terms.

Table 1.18: Excess sensitivity estimates

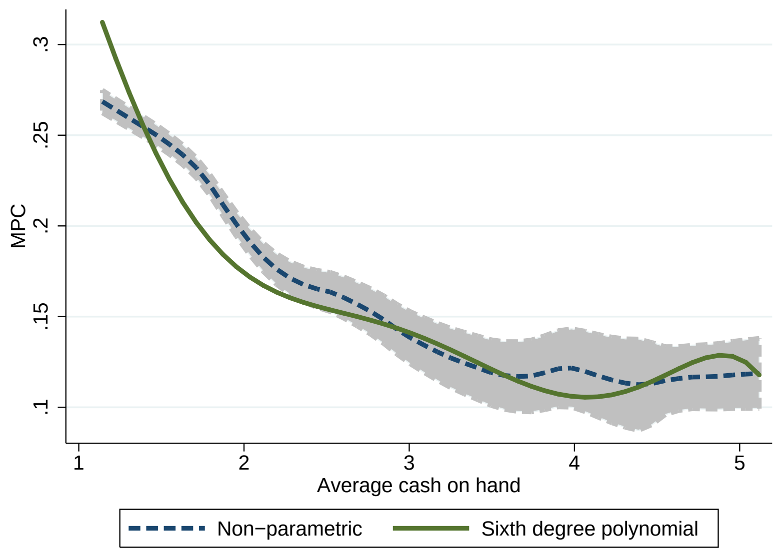
VARIABLES	(1) $\hat{\gamma}$
$\overline{coh}_i$	1.097
	(2.446)
$\overline{coh}_i^2$	-1.757
	(2.454)
$\overline{coh}_i^3$	1.076
	(1.248)
$\overline{coh}_i^4$	-0.322
	(0.340)
$\overline{coh}_i^5$	0.047
	(0.047)
$\overline{coh}_i^6$	-0.003
	(0.003)
$(coh_{it}^{PR} - \overline{coh}_i^{PR})$	-0.193***
	(0.008)
$(coh_{it}^{PR} - \overline{coh}_i^{PR})^2$	-0.004
	(0.011)
$(coh_{it}^{PR} - \overline{coh}_i^{PR})^3$	0.075***
	(0.006)
$(coh_{it}^{PR} - \overline{coh}_i^{PR})^4$	-0.005
	(0.005)
$(coh_{it}^{PR} - \overline{coh}_i^{PR})^5$	-0.008***
	(0.001)
$(coh_{it}^{PR} - \overline{coh}_i^{PR})^6$	0.001**
	(0.000)
Observations	129,823

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

With up to the sixth-order terms included, $\sigma^2_{char} = 0.0038, \sigma^2_{circ} = 0.0046$, and $\phi_{char} = 0.45$.

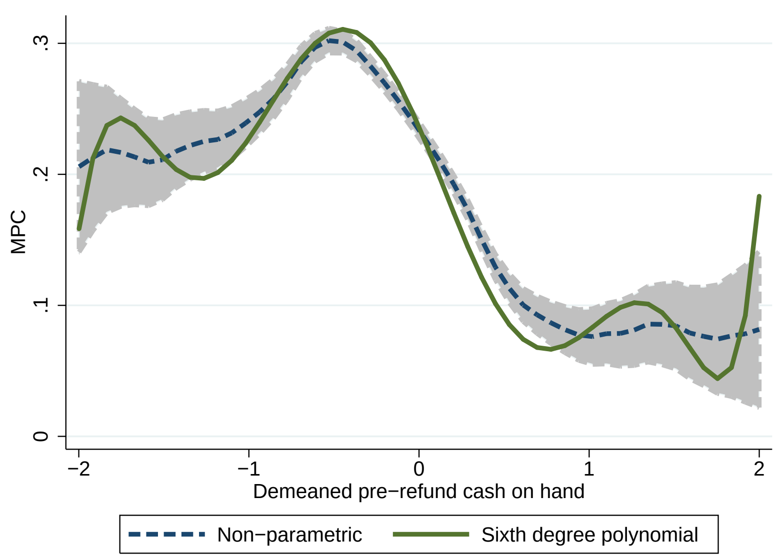
The following figures show the relationship between the MPC and each component of cash on hand under a zero degree kernel smoother and the sixth degree polynomial function.

Figure 1.8.3: Characteristics component



Notes: The non-parametric figure is created using a kernel smoothing method using only the mean. The gray area represents a 95% confident interval.

Figure 1.8.4: Circumstances component



Notes: The non-parametric figure is created using a kernel smoothing method using only the mean. The gray area represents a 95% confident interval.

CHAPTER II

How Individuals Smooth Spending: Evidence from the 2013 Government Shutdown Using Account Data

2.1 Introduction

How consumers respond to changes in income is a central concern of economic analysis and is key for policy evaluation. This paper uses the October 2013 U.S. Federal Government shutdown and a newly developed dataset of financial account records to examine how consumers with different levels of liquidity, income, and spending respond to a short-lived and entirely reversed drop in income. For affected government employees, the shutdown caused a sharp decline in income that was recovered within two weeks. The new dataset, derived from the de-identified account records of more than 1 million individuals living in the United States, provides a granular and integrated view of how individuals in different economic circumstances adjusted spending, saving, and debt in response to the shock.

The most important findings are, first, that many workers routinely have very low levels of liquidity, especially in the days just before their regular paycheck arrives. Second, and consistent with low liquidity, spending by affected workers declined sharply in response to the drop in income caused by the shutdown – though the drop lasted at most two weeks and was then offset by an equal increase. Third, the granularity and integration of the data reveal the means used by affected workers to smooth consumption—if not spending—most notably their delay of recurring expenses such as mortgage payments and credit card balances. Last, though many workers found very low-cost ways to weather the shock, some with low liquidity who were already relying on credit card debt accumulated still more credit card debt.

Prior studies that measure the response of individuals to changes in income have faced two challenges. First, the optimal reaction to an income change depends both on whether the change is anticipated, and on its persistence; but standard data sources make it difficult to identify shocks to expected income and the longevity of these shocks. Second, analysis and policy prescriptions often require a comprehensive view of the heterogeneous responses to an income change. Existing data typically capture only some dimensions with sufficient resolution. They may measure total spending with precision, but not savings or debt; or they measure spending and debt well, but do not measure income with similar accuracy.

We overcome the challenge of identifying income shocks and their persistence by using the 2013 U.S. Federal Government shutdown, which produced a significant, temporary, and easily identified negative shock to the incomes of a large number of employees. We address the challenge of measuring a household's full range of responses to this shock by exploiting a new dataset derived from the integrated transactions and balance data of more than 1 million individuals in the U.S.[1]

More specifically, the data allow us to distinguish Federal government employees subject to the shutdown. They are distinguished by the transaction description associated with direct deposit of their paychecks to their bank accounts. Knowing who was subject to the income shock, we can examine their responses in terms of spending and other variables before, during, and after the government shutdown. These responses are estimated by a difference-in-difference approach, where the outcomes of

[1] The data are captured in the course of business by a mobile banking app. While newly developed, this dataset has already proved useful for studying the high-frequency response of spending to regular, anticipated income by levels of spending, income, and liquidity (Gelman et al. 2014). The related literature section below discusses other studies that use similar types of account data.

affected government workers are compared with those of a control group consisting of workers that have the same biweekly pay schedule as the Federal government, but who were not subject to the shutdown. The control group is mainly non-Federal workers, though also includes some Federal workers not subject to the shutdown.

The pay of a typical affected worker was 40% below normal during the shutdown because the government was closed from October 1 to October 16, 2013, thus including the last four days of the previous ten-day pay period. By the next pay period, however, government operations had resumed and workers were reimbursed fully for the income lost during the shutdown. The transaction data clearly show this pattern for affected workers. This event combined with the distinctive features of the data, which link income, spending, and liquid assets at a high frequency for each individual, provides an unusual opportunity to study the response to a relatively sizeable shock that affected just the timing of income for individuals across the income distribution, without any net effect on their lifetime incomes. See the related literature section below for a discussion of the distinctions of this study.

An important fact revealed by the balance records is that many affected employees maintained low levels of liquid assets (checking and saving account balances), especially in the days just before their regular paychecks arrive. Prior to the shutdown, the median worker in the data held an average liquid assets balance sufficient to cover just eight days of average spending. Moreover, liquidity exhibits systematic changes over the pay-cycle. Just before payday, the median level of liquidity is only five days of average spending. Indeed, a substantial fraction of this population barely lives paycheck-to-paycheck. On the day before their paycheck arrives, the bottom third of the liquidity distribution has, on average, a liquid asset balance of zero.

Given such low levels of liquidity, it is perhaps unsurprising that the transaction records show a sharp drop in total spending by affected workers during the week of missing paycheck income. Weekly spending declined by roughly half the reduction in income and then recovered roughly equally over the two pay periods following the end of the shutdown. Econometric analysis reveals a marginal propensity to spend of about 0.58 as a response to the income shock. Most individuals reversed this drop in spending immediately after they receive the paychecks that reimbursed them for their lost income.

On its face, it is troubling that so many affected workers maintained such low liquidity and exhibited such a sharp spending response to an unexpected but brief delay in income. It suggests either that benchmark theories founded on a taste for smoothing consumption are badly specified; that households are inadequately buffered against even very temporary shocks; or that the financial markets that make smoothing possible are functioning poorly.

Further examination of the data reveals, however, that even consumers with low liquidity can smooth consumption better than spending using low-cost methods to shift the timing of payments for committed forms of expenditure. More detailed analysis shows that affected workers delayed mortgage payments, in particular; and many individuals shifted credit card balance payments. At the same time, the data show no increase in spending on credit cards; average debt only increased due to delays in debt payments. Hence, while they responded to the temporary shock by reducing spending, a large part of their reaction was to delay recurring payments that impose little to no penalty. This shows how consumers make use of short-term margins of adjustment that are mostly overlooked in the literature on methods of smoothing consumption in response to at least temporary income shocks. As such, it also reveals a potentially important welfare benefit of, especially, mortgages with low interest rates. Mortgages can function as a (cheap) line of credit that can help smooth even large, if brief, shocks to income at relatively low cost.

While the data show that many affected workers were able to use perhaps unconventional means to smooth consumption, if not spending, for some with low liquidity these methods were either inadequate or unavailable. This group, who was carrying some credit card debt already, emerged from the shutdown with still more debt owing to failure to make payments rather than new borrowing.

The remainder of the paper proceeds as follows. Section 2 describes the paper's relationship to prior studies of individual responses to income shocks. Section 3 provides key facts about the circumstances surrounding the shutdown. Section 4 describes the data and our research design. It establishes that many workers regularly have low liquidity prior to receiving their paycheck. Section 5 estimates the average response of spending and liquid assets to the shock. Section 6 considers heterogeneity in these responses across the liquidity distribution and examines the consequences for credit card debt.

2.2 Related Literature

The literature concerned with individual responses to income shocks is large. Jappelli and Pistaferri (2010) offer an insightful review. Relative to that large literature, a principal distinction of our paper is the integrated, administrative data that allow accurate observation of liquidity, of the income shock itself, and of several forms of response to the shock. These data thus provide measures of important constraints and outcomes that allow improved inference from the heterogeneous and multi-dimensional reactions to this change in income.

Prior studies of income shocks have mostly relied on the self-reports of survey respondents to provide information either about the shock or about the response of spending and savings and debt.[2] Carroll, Crossley, and Sabelhaus (2013), Dillman and House (2013), Einav and Levin (2014), and others, have called for increased use of administrative records to augment survey research. So far, however, the administra-

[2]See, for example, Souleles (1999), Browning and Crossley (2001), Shapiro and Slemrod (2003), and Johnson et al. (2006).

tive records available for research have typically represented just a slice of economic activity, either providing information about spending at just one retailer, or about the use of a few credit cards, or about just one form of spending.[3] Other approaches blend survey and administrative data. For example, Broda and Parker (2014) and Parker (2015) use consumer-based scanner data to study the response of spending to an income shock; and use surveys for measuring income.[4] This paper is different from these studies of purely administrative data and from those of blended data sources: The integrated data we use provide an accurate, high-resolution, and high frequency picture of liquidity before the shock, and both the spending and net saving responses to the shock.

By using integrated account records, this paper is part of a new and still small literature that includes Baker (2014), Kuchler (2014), Gelman et al. (2014), and Baker and Yannelis (2015). Baker (2014) uses account records from an online banking app, links them to external data on employers, and instruments for individuals' income changes with news about their employers. Because they are persistent, theory suggests that some of these income shocks (from layoffs or plant closings, e.g.) should have different implications for spending from the one caused by the government shutdown. Nevertheless, like the present paper, Baker (2014) finds evidence of the importance of liquidity (more than debt) for the spending response to an income shock. The present paper is distinct in its study of the methods by which those with very little liquidity smooth consumption through a temporary income shock.

Kuchler (2014) studies integrated account records from an online financial management service that elicits from its customers plans for paying down credit card debt. Kuchler (2014) uses those plans, along with the spending responses to income

[3]See, for example, Gross and Souleles (2002), Agarwal, Liu, and Souleles (2007), and Aguiar and Hurst (2005).

[4]Einav, Leibtag, and Nevo (2010) discuss the challenges that even scanner technologies like Nielsen Consumer Panel (formerly Homescan), and Feenstra and Shapiro (2003) discuss the challenges of using store-based scanner data to measure expenditure and prices.

changes, to evaluate a model of present-biased time preferences.

Gelman et al. (2014) use a small subsample of the same data we use in this paper to study the spending response to the arrival of predictable (paycheck and Social Security benefit) income. That paper did not examine other outcomes besides spending. Finally, in a complementary study completed shortly after ours, Baker and Yannelis (2015) use data from the same banking app used in Baker (2014) to describe the response of affected government workers to the 2013 shutdown. Baker and Yannelis (2015) focuses on income and spending, but does not integrate those outcomes with financial positions. Their analysis confirms that the spending and income response to the government shutdown is identifiable in these data sets. From these initial facts, their paper analyzes time allocation and home production. Our paper focuses on the precarious liquidity position individuals find themselves in near the end of the paycheck cycle and the different channels through which individuals smoothed their consumption.

The shutdown also is a distinctive shock. The shock is large, negative, and proportional to income. These features stand in contrast to shocks arising from government stimulus payments, which are positive and often weakly related to income. See, for example, Shapiro and Slemrod (1995, 2003), Johnson et al. (2006), Parker et al. (2013), Agarwal et al. (2007), Bertrand and Morse (2009), Broda and Parker (2014), Parker (2015), and Agarwal and Qian (2014). The government shutdown caused a 40% drop in anticipated paycheck income for individuals across a wide range of the income distribution. Thus, unlike the stimulus payments, the shutdown represented a sizeable, albeit temporary, shock even to high-income households.

2.3 The 2013 U.S. Government Shutdown

2.3.1 Background

The U.S. government was shut down from October 1 to October 16, 2013 because Congress did not pass legislation to appropriate funds for fiscal year 2014. While Federal government shutdowns have historical precedent, it was difficult to anticipate whether this shutdown would occur and how long it would last.[5] The shutdown was preceded by a series of legislative battles surrounding the Affordable Care Act (ACA), also known as Obamacare. Key events and their timing are described in Figure 1.

Opponents of the ACA in the House of Representatives sought to tie FY 2014 appropriations to defunding the ACA. They used the threat of a shutdown as a lever in their negotiations and thus generated considerable uncertainty about whether a shutdown would occur. Just days before the deadline to appropriate funding and avoid a shutdown, there was substantial uncertainty over what would happen. A YouGov/Huffington Post survey conducted on September 28-29, 2013 showed that 44% of U.S. adults thought Congress would reach a deal to avoid a shutdown while 26% thought they would not, and 30% were unsure. A similar survey taken after the shutdown began on October 2-3, 2013 showed substantial uncertainty over its expected duration. 7% thought the shutdown would last less than a week, 31% thought one or two weeks, 19% thought three or four weeks, and 10% thought the shutdown would last more than a month. 33% were unsure of how long it would last.[6] For most federal employees, therefore, the shutdown and its duration were likely difficult to anticipate at the outset. While it was not a complete surprise, it was a shock to many that the shutdown did indeed occur. On the other hand, as we will discuss in the next subsection, the shutdown was essentially resolved contemporaneously with

[5]There have been 12 shutdowns since 1980 with an average length of 4 days. The longest previous shutdown lasted for 21 days in 1995-1996. See Mataconis (2011).

[6]Each survey was based on 1,000 U.S. adults. See YouGov/Huffington Post (2013a, b).

the receipt of the paycheck affected by the shutdown. Hence, there was no reason based on permanent income to respond to the drop in income.

2.3.2 Impact on Federal Employees

Our analysis focuses on the consequences of the shutdown for a group of the approximately 2.1 million federal government employees. The funding gap that caused the shutdown meant that most federal employees could not be paid until funding legislation was passed. The 1.3 million employees deemed necessary to protect life and property were required to work. They were not, however, paid during the shutdown for work that they did during the shutdown. The 800,000 "non-essential" employees were simply furloughed without pay.[7] In previous shutdowns, employees were paid retroactively (whether or not they were furloughed). Of course, it was not entirely clear what would happen in 2013. On October 5, however, the House passed a bill to provide back pay to all federal employees after the resolution of the shutdown. While not definitive, this legislation was strong reassurance that the precedent of retroactive pay would be respected, as in fact it was when the shutdown concluded. After the October 5 Congressional action, most of the remaining income risk to employees was due to the uncertain duration of the shutdown and to potential cost-cutting measures that could be part of a deal on the budget.

Unlike most private sector workers, Federal workers are routinely paid with a lag of about a week, so the October 5 House vote came before reduced paychecks were issued. For most government employees, the relevant pay periods are September 22 - October 5, 2013 and October 6 - October 19, 2013. Because the shutdown started in the latter part of the first relevant pay period, employees did not receive payment for 5 days of the 14 day pay period. For most employees on a Monday to Friday work

[7]Some federal employees were paid through funds not tied to the legislation in question and were not affected. The Pentagon recalled its approximately 350,000 employees on October 5, reducing the number of furloughed employees to 450,000.

schedule, this would lead to 4 unpaid days out of 10 working days, so they would receive 40% less than typical pay. The actual fraction varies with hours and days worked and because of taxes and other payments or debits. Since the government shutdown ended before the next pay date, employees who only received a partial paycheck were fully reimbursed in their next paycheck.

Federal government employees are a distinctive subset of the workforce. According to a Congressional Budget Office report (CBO 2012), however, federal employees represent a wide variety of skills and experiences in more than 700 occupations. Compared to private sector employees, they tend to be older, more educated, and more concentrated in professional occupations. Table 1 below reproduces Summary Table 1 in the CBO report. Overall, total compensation is slightly higher for federal employees. Breaking down the compensation difference by educational attainment shows that federal employees are compensated relatively more at low levels of education while the opposite holds for the higher end of the education distribution. In the next section, we make similar comparisons based on Federal versus non-Federal employees in our data. The analysis must be interpreted, however, with the caution that Federal employees may not have identical behavioral responses as the general population. We return to this issue in the discussion of the results.

2.4 Data and Design

2.4.1 Data

The source of the data analyzed here is a financial aggregation and bill-paying computer and smartphone application that had approximately 1.5 million active users in the U.S. in 2013.[8] Users can link almost any financial account to the app, including

[8]We gratefully acknowledge the partnership with the financial services application that makes this work possible. 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.

bank accounts, credit card accounts, utility bills, and more. Each day, the application logs into the web portals for these accounts and obtains key elements of the user's financial data including balances, transaction records and descriptions, the price of credit and the fraction of available credit used.

We draw on the entire de-identified population of active users and data derived from their records from late 2012 until October 2014. The data are de-identified and the analysis is performed on normalized and aggregated user-level data as described in the Appendix. The firm does not collect demographic information directly and instead uses a third party business that gathers both public and private sources of demographics, anonymizes them, and matches them back to the de-identified dataset. Appendix Table A (replicated from Table 1 of Gelman et al. 2014) 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) that is representative of the U.S. population in 2012. The app's user population is not representative of the U.S. population, but it is heterogeneous, including large numbers of users of different ages, education levels, and geographic locations.

We identify paychecks using the transaction description of checking account deposits. Among these paychecks, we identify Federal employees by further details in the transaction description. The appendix describes details for identifying paychecks in general and Federal paychecks in particular. It also discusses the extent to which we are capturing the expected number of Federal employees in the data. The number of federal employees and their distribution across agencies paying them are in line with what one would expect if these employees enroll in the app at roughly the same frequency as the general population.

2.4.2 Design: Treatment and Controls

Much of the following analysis uses a difference-in-differences approach to study how Federal employees reacted to the effects of the government shutdown. The treatment group consists of Federal employees whose paycheck income we observe changing as a result of the shutdown. The control group consists of employees that have the same biweekly pay schedule as the Federal government who were not subject to the shutdown (see the Appendix for more details). The control group is mainly non-Federal employees, but also includes some Federal employees not subject to the shutdown.[9] Table 2 shows summary statistics from the app's data for these groups of employees. As in the CBO study cited above, Federal employees in our sample have higher incomes. They also have higher spending, higher liquid balances, and higher credit card balances.

We use the control group of employees not subject to the shutdown to account for a number of factors that might affect income and spending during the shutdown: these include aggregate shocks and seasonality in income and spending. Additionally, interactions of pay date, spending, and day of week are important (see Gelman et al. 2014). Requiring the treatment and control to have the same pay dates and pay date schedule (biweekly on the Federal schedule) is a straightforward and important way to control for these substantial, but subtle effects.

There is, of course, substantial variability in economic circumstance across individuals both within and across treatment and controls. We normalize many variables by average daily spending, or where relevant by average account balances) at the

[9] Employees not subject to the shutdown include military, some civilian Defense Department, Post Office, and other employees paid by funds not involved in the shutdown. An alternative strategy would use just the Federal workers not affected by the shutdown as the control. We do not adopt this strategy because we believe a priori that it is less suitable because the workers exempted from the shutdown are in non-random agencies and occupations. This selectivity makes them potentially less suitable as a control. For completeness, however, the Appendix includes key results using the unaffected Federal workers as the control. The findings are quite similar though less precise because of smaller sample size.

individual level. This normalization is a simple, and given the limited covariates in the data, a practical way to pool individuals with very different levels of income and spending. In particular, it serves to equalize the differences in income levels between treatment and control seen in Table 2.

By showing a wide span of data before and after the shutdown, Figure 2 provides strong evidence of the adequacy of the control group and the effectiveness of using average daily spending as the normalization. Figure 2 shows that the employees not subject to the shutdown have nearly identical movement in spending except during the weeks surrounding the shutdown. Thus the controls appear effective at capturing aggregate shocks, seasonality, payday interactions, etc. In particular, note the regular, biweekly pattern of fluctuations in spending. It arises largely from the timing of spending following receipt of the bi-weekly paychecks. There are also subtler beginning-of-month effects—also related to timing of spending. In subsequent figures we use a narrower window to highlight the effects of the shutdown.

Gelman et al. (2014) shows that much of the sensitivity of spending to receipt of paycheck, like that seen in Figure 2, arises from reasonable choices of individuals to time recurring payments—such as mortgage payments, rent, or other recurring bills—immediately after receipt of paychecks. Figure 2 makes clear that the control group does a good job of capturing this feature of the data and therefore eliminating ordinary paycheck effects from the analysis. The first vertical line in Figure 2 indicates the week in which employees affected by the shutdown were paid roughly 40% less than their average paycheck. There is a large gap between the treatment and control group during this week. Similarly, the second vertical line indicates the week of the first paycheck after the shutdown. The rebound in spending is discernable for two weeks. The figure thus demonstrates that the control group represents a valid counterfactual for spending that occurred in the absence of the government shutdown.

2.4.3 Liquidity Before the Shutdown

To understand how affected employees responded to the shutdown, it is useful to examine first how they and others like them managed their liquid assets prior to the shock. Analysis of liquid asset balances before the shutdown shows that, while some workers were well buffered, many were ill-prepared to use liquid assets to smooth even a brief income shock.

We define liquid assets as the balance on all checking and savings accounts. The measure of liquidity is based on daily snapshots of account balances. Hence, they are measures of the stock of liquid assets independent from the transactional data used to measure spending and income. Having such high-frequency data makes it possible to observe distinctive, new evidence on liquidity and how it interacts with shocks. Figure 3 shows median liquidity over the pay-cycle, by terciles of the distribution of liquidity. Liquidity is expressed as a ratio of checking and savings account balances to average daily total spending. The results are for the period prior to the shutdown and aggregate over both treatment and control groups.[10,11]

While the optimal level of liquidity is not clear, the figure shows the top third of the liquidity distribution is well-positioned to handle the income shock due to the shutdown. The median of this group could maintain more than a month of average spending with their checking and savings account balances, even in the days just before their paycheck arrives.

[10]The distributions for treatment and control are similar. For example, in the control group the median liquidity ratio for the first, second, and third, terciles of the liquidity distribution is, 2.9, 7.9, and 32.1, respectively. The analogous numbers for the treatment group are 3.3, 8.1, and 32.0.

[11]Liquidity peaks two days after a payday. The balance data are based on funds available, so liquidity should lag payday according to the banks funds-availability policy. There is at least one-day lag built into the data because the balances are scraped during the day, so will reflect a paycheck posted the previous day. Appendix Figure A4 shows that the two-day delay in the peak of liquidity is due to funds availability, not delays in posting based on interactions of day of payday and delays in posting of transactions over the week-end. (As discussed in the Appendix, even within the government bi-weekly pay schedule, there is some heterogeneity in day of week of the payday.) Additionally, liquidity is, of course, net of inflows and outflows. Recurring payments made just after the receipt of paycheck will therefore lead daily balances to understate gross liquidity right after the receipt of the paycheck.

The lower two-thirds of the liquidity distribution has a substantially smaller cushion. Over the entire pay-cycle, the middle tercile has median liquid assets equal to 7.9 days of average spending. Liquidity drops to only 5 days of average spending in the days just before their paycheck arrives. Thus, even in the middle of the liquidity distribution many would be hard pressed to use liquid savings to smooth a temporary loss of 4 days pay. The bottom third of this population is especially ill-prepared. Prior to the shutdown, the median of this group consistently arrives at payday with precisely zero liquid balances. (Balances can be negative owing to overdrafts.) These balance data thus reveal how, even among those with steady employment, large fractions of consumers do not have the liquid assets to absorb a large, but brief, shock to income.

2.5 Responses to the Shutdown

Having established that many (affected) workers had little liquidity prior to the shutdown, we now examine how their income, and various form of spending responded to the shutdown. Our method is to estimate the difference-in-difference, between treatment and control, for various outcomes using the equation,

$$y_{i,t} = \sum_{k=1}^{T} \delta_k \times Week_{i,k} + \sum_{k=1}^{T} \beta_k \times Week_{i,k} \times Shut_i + \Gamma' X_{i,t} + \epsilon_{i,t}, \qquad (2.1)$$

where y represents the outcome variable (total spending, non-recurring spending, income, debt, savings, etc.), i indexes individuals ($i \in \{1, ..., N\}$), and t indexes time ($t \in \{1, ..., T\}$). $Week_{i,k}$ is a complete set of indicator variables for each individual-week in the sample, $Shut_i$ is a binary variable equal to 1 if individual i is in the treatment group and 0 otherwise, and $X_{i,t}$ represents controls to absorb the predictable variation arising from bi-weekly pay week patterns.[12] The β_k coefficients capture the

[12]Specifically, $X_{i,t}$ contains dummies for paycheck week, treatment, and their interaction. This specification allows the response of treatment and control to ordinary paychecks to differ. These

average weekly difference in the outcome variables of the treatment group relative to the control group. Standard errors in all regression analyses are clustered at the individual level and adjusted for conditional heteroskedasticity.

2.5.1 Paycheck Income and Total Spending

We begin with an examination of how income, as measured in these data, was affected by the shutdown. External reports indicate that the paycheck income of affected employees should have dropped by 40% on average. The analysis of paycheck income here can thus be viewed, in part, as testing the ability of these data to accurately measure that drop. Once that ability is confirmed, we move to an evaluation of the spending responses.

Recall that we normalize each variable of interest, measured at the individual level, by the individual's average daily spending computed over the entire sample period. The unit of analysis in our figures is therefore days of average spending. Figure 4 plots the estimated β_k from equation (1) where y is normalized paycheck income. We plot three months before and after the government shutdown to highlight the effect of the event. The first vertical line (dashed-blue) represents the week that the shutdown began and the second vertical line (solid-red) represents the week in which pay dropped due to the shutdown, and the third when pay was restored.

Panel A of Figure 4 shows, as expected, a drop in income equal to approximately 4 days of average daily total spending during the first paycheck period after the shutdown.[13] This drop quickly recovers during the first paycheck period after the shutdown ends, as all users are reimbursed for their lost income. Some users received their reimbursement paychecks earlier than usual, so the recovery is spread across

controls are only necessary in the estimates for paycheck income.

[13] The biweekly paychecks dropped by 40 percent on average. For the sample, paycheck income is roughly 70 percent of total spending on average because there are other sources of income. So a drop of paycheck income corresponding to 4 days of average daily spending is about what one would expect (4 days $\approx 0.4 \times 0.7 \times 14$ days).

two weeks. The results confirm that the treatment group is indeed subject to the temporary loss and subsequent recovery of income that was caused by the government shutdown, and that the account data allow an accurate measure of those income changes.

Panel B of Figure 4 plots the results on total spending, showing the estimated β_k where y is normalized total spending. On average, total spending drops by about 2 days of spending in the week the reduced paycheck was received. Hence, the drop in spending upon impact is about half the drop in income. That implies a propensity to spend of about one-half—much higher than most theories would predict for a drop of income that was widely expected to be made up in the relatively near future. In the inter-paycheck week, spending is about normal. In the second week after the paycheck affected by the shutdown, spending rebounds with the recovery spread mainly over that week and the next one.

To ease interpretation we convert the patterns observed in Figure 4 into an estimate of the marginal propensity to spend (which we call the MPC as conventional). Let τ be the week of the reduced paycheck during the shutdown. The variable $s_{i,\tau-k}$ denotes total spending for individual i in the k weeks surrounding that week. To estimate the MPC, we consider the relationship,

$$s_{i,\tau-k} = \alpha_k + \beta_k^{MPC}(Paycheck_{i,\tau} - Paycheck_{i,\tau-2}) + \epsilon_{i,\tau-k}, \tag{2.2}$$

where $(Paycheck_{i,\tau} - Paycheck_{i,\tau-2})$ is the change in paycheck income. Both $s_{i,\tau-k}$ and $Paycheck_{i,\tau}$ are normalized by individual-level average daily spending as discussed above. We present estimates for the one and two week anticipation of the drop in pay ($k = 1$ and $k = 2$), the contemporaneous MPC ($k = 0$), and one lagged MPC ($k = -1$). We do not consider further lags because the effect of the lost pay is confounded by the effect of the reimbursed pay beginning at time $\tau + 2$.

There are multiple approaches to estimating equation (2). The explanatory variable is the change in paycheck. We are interested in isolating the effect on spending due to the exogenous drop in pay for employees affected by the shutdown. While in concept this treatment represents a 40 percent drop in income for the affected employees and 0 for the controls, there are idiosyncratic movements in income unrelated to the shutdown. First, not all employees affected by the shutdown had exactly a 40 percent drop in pay because of differences in work schedule or overtime during the pay period. Second, there are idiosyncratic movements in pay in the control group. Therefore, to estimate the effect of the shutdown using equation (2) we use an instrumental variables approach where the instrument is a dummy variable $Shut_i$. The IV estimate is numerically equivalent to the difference-in-difference estimator.[14]

Table 3 shows the estimates of the MPC. These estimates confirm that the total spending of government employees reacted strongly to their drop in income and that this reaction was focused largely during the week that their reduced paycheck arrived. The estimate of the average MPC is 0.58 in this week, with much smaller coefficients in the two weeks just prior. Thus, at the margin, about half of the lost income was reflected in reduced spending.

2.5.2 Spending and Payments by Type

Analyzing different categories of spending offers further insight into the response of these users to the income drop. We separate spending into non-recurring and recurring components. Recurring spending is identified using patterns in both the amount

[14]Estimating equation (2) by least squares should produce a substantially attenuated estimate relative to the true effect of the shutdown if there is idiosyncratic movement in income among the control group, some of which results in changes in spending. In addition, if the behavioral response to the shutdown differs across individuals in ways related to variation in the change in paycheck caused by the shutdown (e.g., because employees with overtime pay might have systematically different MPCs), the difference between the OLS and IV estimates would also reflect treatment heterogeneity. This heterogeneity could lead the OLS estimate to be either larger or smaller than the IV estimate, depending on the correlation between of the size of the shutdown-induced shock and the MPC. The OLS estimate of the MPC for the week the reduced paycheck arrived is 0.123, with a standard error of 0.004.

and transaction description of each individual transaction.[15] It identifies spending that, due to its regularity, is very likely to be a committed form of expenditure (see Grossman and Laroque (1990), Chetty and Szeidl (2007), and Postlewaite, et al. (2008)). Non-recurring spending is total spending minus recurring spending. These measures thus use the amount and timing of spending rather than an a priori categorization based on goods and services. This approach to categorization is made possible by the distinctive features of the data infrastructure.

Figure 5 presents estimates of the β_k from equation (1) where the outcome variable y takes on different spending, payment, or transfer categories. For each graph, the data are normalized by individual-level averages for the series being plotted. In the top two panels we can compare the normalized response of recurring and non-recurring spending and see important heterogeneity in the spending response by category. The results on total spending (Figure 4) showed an asymmetry in the spending response before and after the income shock; total spending dropped roughly by 2 days of average spending during the three weeks after the shutdown began and only rose by 1.6 days of average spending during the three weeks after the shutdown ended. The reaction of recurring spending drives much of that asymmetry; it dropped by 2.6 days of average recurring spending and rose only by 0.84 days once the lost income was recovered. Non-recurring spending exhibits the opposite tendency: it dropped by 1.8 days of average non-recurring spending and rose by 2.0 days. Thus, recurring spending drops more and does not recover as strongly as non-recurring spending.

To better understand this pattern of recurring expenditure and its significance we focus on a particular, and especially important, type of recurring spending—mortgage payments. Panel C of Figure 5 shows that, while the mortgage spending data is noisier

[15] We identify recurring spending using two techniques. First, we define a payment as recurring if it takes the same amount at a regular periodicity. This definition captures payments such as rent or mortgages. Second, we also use transaction fields to identify payments that are made to the same payee at regular intervals, but not necessarily in the same amount. This definition captures payments such as phone or utility bills that are recurring, but in different amounts. See appendix for further details. Gelman et al. (2014) uses only the first technique to define recurring payments.

than the other categories, there is a significant drop during the shutdown and this decline fully recovers in the weeks when the employees' missing income was repaid. In this way, we see that some users manage the shock by putting off mortgage payments until the shutdown ends. Indeed, many of those affected by the shutdown changed from paying their mortgage early in October to later in the month as shown in Figure 6. The irregular pattern of payment week of mortgage reflects interaction of the bi-weekly paycheck schedule with the calendar month. The key finding of this figure is that the deficit in payments of the treatment group in the second week of October is largely offset by the surplus of payments in the last two weeks of October.

Panel D of Figure 5 shows the response of account transfers to the income shock.[16] During the paycheck week affected by the shutdown, transfers fell and rebounded when the pay was reimbursed two weeks later. This finding implies a margin of adjustment, reducing transfers out of linked accounts, during the affected week. One might have expected the opposite, i.e., an inflow of liquidity from unlinked asset accounts to make up for the shortfall in pay. That kind of buffering is not present on average in these data.

Similar behavior is seen in the management of credit card accounts. Another relatively low-cost way to manage cash holdings is to postpone credit card balance payments. Panel E of Figure 5 shows there was a sharp drop in credit card balance payments during the shutdown, which was reversed once the shutdown ended. For users who pay their bill early, this is an easy and cost-free way to finance their current spending. Even if users are using revolving debt, the cost of putting off payments may be small if they pay off the balance right away after the shutdown ends. We examine credit card balances in greater detail in the next section.

Indeed, as we see in Panel F of Figure 5, there was no average reaction of credit

[16] These are transactions explicitly labeled as "transfer," etc. For linked accounts, they should net out (though it is possible that a transfer into and out of linked accounts could show up in different weeks). Hence, these transfers are (largely) to and from accounts (such as money market funds) that are not linked.

card spending to the shutdown. Thus, we find no evidence that affected employees sought to fund more of their expenditure with credit cards but instead floated, temporarily, more of their prior expenditure by postponing payments on credit card balances. Affected individuals who had ample capacity to borrow in order to smooth spending, by charging extra amounts to credit cards, had other means of smoothing, e.g., liquid checking account balances or the postponement of mortgage payments. On the other hand, those who one might think would use credit cards for smoothing spending because they had little cash on hand did not—either because they were constrained by credit limits or preferred to avoid additional borrowing. In the next section we will examine the consequences for credit card balances of these postponed balance payments, and later probe the heterogeneous responses of individuals by their level of liquidity.

This analysis of different categories of spending reveals that users affected by the shutdown reduce spending more heavily on recurring spending and payments compared to non-recurring spending. It is important to note that this behavior appears to represent, in many cases, a temporal shifting of payments and neither a drop in eventual spending over a longer horizon or a proportionate drop in contemporaneous consumption. These results thus provide evidence of the instruments that individuals use to smooth temporary shocks to income that has not been documented before. The drop in non-recurring spending shows, however, that this method of cash management is not perfect; it does not entirely smooth spending categories that better reflect consumption.

Spending could have fallen in part because employees stayed home and engaged in home production instead of frequenting establishments that they encounter during their work-day. Recall, however, that many employees affected by the shutdown were not, in fact, furloughed. They worked but did not get paid for that work on the regular schedule. In addition, Figure 7 shows that categories of expenditure that are quite close to consumption, such as a fast food and coffee shops spending index, show a sharp drop during the week starting October 10 when employees were out of work. Given that a cup of coffee or fast food meal is non-durable, one would not expect these categories to rebound after the shutdown. Yet, there is significant rebound after the shutdown. We interpret this spending as resulting from going for coffee, etc., with co-workers after the shutdown, perhaps to trade war stories.[17] Hence, in a sense, a cup of coffee is not entering the utility function as an additively separate non-durable.

2.5.3 Response of Liquid Assets

For users who have built up a liquid asset buffer, they may draw down on these reserves to help smooth income shocks. Figure 8 shows the estimated β_k from equation (1) where $y_{i,t}$ is the weekly average liquid balance normalized by its individual-level average (Panel A) or normalized by individual-level average daily total spending (Panel B). Because of the heterogeneity in balances, normalizing by average liquid balances leads to more precise results. Normalizing by total spending is less precise but allows for a more meaningful interpretation because it is in the same units as Figure 4. Consistent with the spending analysis, relative savings for the treatment group rises in anticipation of the temporary drop in paycheck income. There is a steep drop in the average balance the week of the lower paycheck as a result of the shutdown. The drop in liquidity is, however, substantially attenuated relative to the drop in income because of the drop in payments that is documented in the previous section. The recovery of the lost income causes a large spike in the balances, which is mostly run off during the following weeks. Figure 8B shows that liquid balances fell by around 2 day of average daily total spending. Therefore, on average, users reduced spending by about 2 days and drew down about 2 days of liquidity to fund their consumption

[17] Interestingly, the rebound is highest for the most liquid individuals (figure not reported) who are also higher income. This finding supports our notion that the rebound in coffee shop and fast food arises from post-shutdown socializing.

when faced with a roughly 4 days drop in income. These need not add up because of transfers from non-linked accounts and because of changes in credit card payments, though they do add up roughly at the aggregate level. In the next section, we explore the heterogeneity in responses as a function of liquid asset positions where specific groups of individuals do use other margins of adjustment than liquid assets.

2.6 Liquidity and the Heterogeneity in Response to the Income Shock

The preceding results capture average effects of the shutdown. There are important reasons to think, however, that different employees will react differently to this income shock, depending on their financial circumstances. Although all may have a desire to smooth their spending in response to a temporary shock, some may not have the means to do so.

In this section we examine the heterogeneity in the response along the critical dimension of liquidity. For those with substantial liquid balances relative to typical spending, it should be relatively easy to smooth through the shutdown. Section 4 showed, however, that many workers in these data have little liquidity, especially in the days just before their regular paycheck arrives. For those (barely) living paycheck-to-paycheck, even this brief drop in income may pose significant difficulties.

We investigate the impact of the shutdown for those with varying levels of liquidity by first further quantifying the buffer of liquid assets that different groups of workers had. Second, we return to each of the spending categories examined above and compare how different segments of the liquidity distribution responded to the income shock. Last, we study how the precise timing of the shock, relative to credit card due dates, influenced credit card balances coming out of the shutdown.

2.6.1 Liquidity and Spending

As before, we define the liquidity ratio for each user as the average daily balance of checking and savings accounts to the user's average daily spending until the government shutdown started on October 1, 2013 and then divide users into three terciles. Table 4 shows characteristics of each tercile. Users in the highest tercile have on average 54 days of daily spending on hand while the lowest tercile only has about 3 days. This indicates that a drop in income equivalent to 4 days of spending should have significantly greater effects for the lowest tercile compared with the highest tercile.

Figure 9 plots the estimates of β_ks from equation (1), for various forms of spending, by terciles of liquidity. The results are consistent with liquidity playing a major role in the lack of smoothing. Users with little buffer of liquid savings are more likely to have problems making large and recurring payments such as rent, mortgage, and credit card balances. In terms of average daily expenditure, spending for these recurring payments drops the most for low liquidity users. In contrast, the drop in non-recurring spending is similar across all liquidity groups. Like those with more liquidity, however, low liquidity users refrained from using additional credit card spending to smooth the income drop.

2.6.2 Liquidity and Credit Debt

The preceding results indicate that the sharp declines in recurring spending (especially mortgages) and credit card balance payments induced by the shutdown were particularly important strategies for those with lower levels of liquidity. The granularity of the data shows, however, that fine differences in timing are consequential when liquidity levels are so low.

To examine how individuals manage credit card payments and balances, we carry out the analysis at the level of the individual credit card account, rather than aggregating across accounts as in the previous section. The account-level analysis allows us to examine the role of payment due dates in the response to the shutdown. These due dates may represent significant requirements for liquidity. That they are staggered and unlikely to be systematically related to the timing of the shutdown provides another means for identifying behavioral responses that exploits the high resolution of the data infrastructure.

In this analysis, however, attention is restricted to the accounts of "revolvers." We focus, that is, on accounts held by those who, at some point during the study period (including the period of the shutdown), incurred interest charges on at least one of their credit cards, indicating that they carried some revolving credit card debt. This represents 63% of the treatment group and 63% of controls; and 70% of these workers fall in the lower two-thirds of the liquidity distribution. The complement of the revolver group is the "transactors." Members of this group routinely pay their entire credit card balance, and have a distinct monthly pattern of balances that reflects their credit card spending over the billing cycle and regular payment of the balance at the end of the cycle. Only 44% of transactors fall in the lower two thirds of the liquidity distribution. Including transactors would obscure the results for those who carry credit card debt.[18]

Figure 10 shows the response of credit card balances, at the account level, to the loss of income due to the shutdown. The estimates again present the difference-in-difference between accounts held by revolvers in the treatment group and those held by revolvers in the control group. These estimates are specified in terms of days since the account's August 2013 statement date instead of calendar time in order to show the effect of statement due dates. In Figure 10, Days 0 through 30 on the horizontal axis correspond to payment due dates in late August or in September 2013. (Payments are due typically 25 days after the statement date.) The different panels of

[18] We investigated those who shifted from being transactors to revolvers at the time of the shutdown. This group was so small (17%) that it did not yield interesting results. Given that transactors tend to have high liquidity (15.9 median ratio vs 7.7 for revolvers), the lack of such transitions is not surprising.

Figure 10 show alternative cuts of the data that we will explain next. Focus, however, on the first 25 days since the August statement date, i.e., due dates that occur in advance of the shutdown. Regardless of cut on the data, the difference-in-difference between treatment and control is essentially zero.

Panels A and B divide the sample of accounts into two groups based on the credit card statement date and, in particular, whether the statement date places them "at risk" for having to make a payment during the government shutdown. Panel A shows the accounts with statement dates on September 16-30, 2013. Panel B shows accounts that have statement dates on September 1-15. For those in the treatment group, the accounts with September 16-30 due dates (Panel A) are at risk. Based on our analysis of liquidity over the paycheck cycle (Figure 3) it is likely that the mid-October paycheck that is diminished by the shutdown would have been a primary source of liquidity for making the payment on these accounts that come due during that pay period. Indeed, Panel A reveals this effect. Control and treatment accounts start to diverge about a week to 10 days into the October billing cycle (days 35-37). By the time the November statement arrives (days 58-60), a significant gap emerges; relative to controls, treatment account balances are now significantly above average. They return to average in a month, presumably as affected workers use retroactive pay to make balance payments. Panel B, those who made their payments before the shutdown, shows no such effect (the hump starting at day 30 is prior to the shutdown and is not statistically significant.)

The high-resolution analysis made possible with the data infrastructure reveals that, when liquidity is so low, small differences in timing can matter. Workers whose usual credit card payment date fell before the shutdown adjusted on other margins; their balances did not rise. For others, the shutdown hit just as they would have normally made their credit card payment; they deferred credit card payments and their balances were elevated for a billing cycle or two before returning to normal levels.

These findings for credit cards reinforce the findings for mortgage payments found in the previous section and Figure 6. For those who typically made payments on mortgages early in the month, that is, prior to the receiving the paycheck reduced by the shutdown, there is little effect of the shutdown on mortgage payments. For those who make payments in the second half of the month, they can and often did postpone the mortgage payments as a way to respond to the shock to liquidity.

2.7 Conclusion

Living paycheck-to-paycheck lets workers consume at higher levels, but would seem to leave them quite vulnerable to income shocks. The results of this paper reveal how workers use financial assets and markets, sometimes in unconventional ways, to reduce that vulnerability and adjust to shocks when they do occur. The findings indicate that to the extent a large but brief shock to income is a primary risk, a lack of liquid assets as a buffer is not necessarily a sign of myopia or unfounded optimism. Rather, the reactions to the 2013 government shutdown studied in this paper indicate that workers can defer debt payments and thus maintain consumption (at low cost) despite limited liquid assets. They may face higher costs to access less liquid assets. Such illiquidity may be optimal even if it leads to short- or medium-run liquidity constraints (see Kaplan and Violante 2013). This paper shows that the majority of households have such liquidity constraints, yet they have mechanisms for coping with shocks to income so as to mitigate the consequences of such illiquidity.

This paper provides direct evidence on the importance of deferring debt payments, especially mortgages, as an instrument for consumption smoothing. Mortgages function for many as a primary line of credit. By deferring a mortgage payment, they can continue to consume housing, while waiting for an income loss to be recovered. For changing the timing of mortgage payments within the month due, there is no cost.

As discussed above, that is the pattern for the bulk of deferred mortgage payments. Moreover, the cost of paying one month late can also be low. Many mortgages allow a grace period after the official due date, in which not even late charges are incurred, or charge a fee that is 4-6 percent of the late payment. Being late by a month adds only modestly to the total mortgage when interest rates are low, and many mortgage service companies will not report a late payment to credit agencies until it is at least 30 days overdue. Even if there are penalties or costs, late payment of mortgage is a source of credit that is available without the burden of applying for credit.

Thus, this paper's findings indicate that policies that encourage homeownership and low-interest mortgages may have under-appreciated welfare benefits to those mortgage holders. Our results suggest expansion of mortgage availability not only finances housing, but has the added effect of making it easier to smooth through shocks to income. As in Herkenhoff and Ohanian (2013), who show how skipping mortgage payments can function as a form of unemployment insurance, the results here reveal how the ability to defer mortgage payments can be an important source of consumption insurance in the face of large, temporary income shocks.

The timing of credit card balance payments provides another source of managing liquidity to buffer consumption against a temporary decline in income. For those with low levels of liquid assets, deferring or reducing credit card payments is a convenient and relatively low-cost way to address a temporary income shortfall. Among credit card borrowers who had payment due dates during the pay period with the reduced paycheck, we see significant deferral of payments. Their credit card balances rose, and stayed elevated for a billing cycle or two before returning to normal.

The distinctive findings of this paper derive high-frequency data on transactions and balances that provide new and distinctive evidence on consumer behavior. The precision and resolution of these data allow insights into behavior that are obscured by conventional data sources.

Bibliography

Agarwal, Sumit, Chunlin Liu, and Nicholas S. Souleles (2007) "The Reaction of Consumer Spending and Debt to Tax Rebates: Evidence from Consumer Credit Data," *Journal of Political Economy*, 115(6): 986-1019.

Agarwal, Sumit, and Wenlan Qian (2014) "Consumption and Debt Response to Unanticipated Income Shocks: Evidence from a Natural Experiment in Singapore," *American Economic Review*, 104(12): 4205-30.

Aguiar, Mark and Erik Hurst (2005) "Consumption versus Expenditure," *Journal of Political Economy*, 113(5): 919-948.

Baker, Scott (2014) "Debt and the Consumption Response to Household Income Shocks," manuscript, Kellogg School of Management.

Baker, Scott, and Constantine Yannelis (2015) "Income Changes and Consumption: Evidence from the 2013 Federal Government Shutdown." SSRN Working paper, March 2015.

Bertrand, Marianne, and Adair Morse (2009) "Indebted Households and Tax Rebates," *American Economic Review*, 99(2): 418-23.

Broda, Christian and Jonathan A. Parker (2014) "The Economic Stimulus Payments of 2008 and the Aggregate Demand for Consumption," *Journal of Monetary Economics*, 68(S): S20-S3.

Carroll, Christopher, Thomas Crossley, and John Sabelhaus (2013) *Improving the Measurement of Consumer Expenditures*. University of Chicago Press.

Chetty, Raj and Adam Szeidl (2007) "Consumption Commitments and Risk Preferences," *Quarterly Journal of Economics*, 122(2): 831-877.

Congressional Budget Office (2012) "Comparing the Compensation of Federal and Private-Sector Employees."

Dillman, Don A. and Carol C. House eds. (2013) *Measuring What We Spend: Toward a New Consumer Expenditure Survey.* The National Academies Press.

Einav, Liran, Ephraim Leibtag, and Aviv Nevo (2010) "Recording discrepancies in Nielsen Homescan data: Are they present and do they matter?" *Quantitative Marketing and Economics*, 8(2): 207-239.

Einav, Liran and Jonathan D. Levin (2014) "Economics in The Age of Big Data," *Science*, 346(6210).

Feenstra, Robert C. and Matthew D. Shapiro eds. (2003) *Scanner Data and Price Indexes.* Chicago: University of Chicago Press.

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.

Gross, David B. and Nicholas S. Souleles (2002) "Do Liquidity Constraints and Interest Rates Matter for Consumer Behavior? Evidence from Credit Card Data," *The Quarterly Journal of Economics*, 117(1): 149-185.

Grossman, Sanford. J. and Guy Laroque (1990) "Asset Pricing and Optimal Portfolio Choice in the Presence of Illiquid Durable Consumption Goods," *Econometrica*, 58(1): 25-51.

Herkenhoff, Kyle F., and Lee E. Ohanian (2013) "Foreclosure Delay and U.S. Unemployment," Working paper, University of Minnesota Department of Economics.

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

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

Kuchler, Theresa (2014) "Sticking to Your Plan: Hyperbolic Discounting and Credit Card Debt Paydown," manuscript, NYU Stern School of Business.

Mataconis, Doug (2011) "A Brief History Of Federal Government Shutdowns" Outside the Beltway, 08 April 2011. Web. 10 Feb 2015.

Parker, Jonathan (2015) "Why Don't Households Smooth Consumption? Evidence from a 25 Million Dollar Experiment," unpublished paper, MIT.

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): 2530-53.

Postlewaite, Andrew, Larry Samuelson, and Dan Silverman (2008) "Consumption Commitments and Employment Contracts," *Review of Economic Studies*, 75(2): 559-578.

Shapiro, Matthew D., and Joel Slemrod (1995) "Consumer Response to the Timing of Income: Evidence from a Change in Tax Withholding," *American Economic Review*, 85(1): 274-283.

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

Souleles, Nicholas S. (1999) "The Response of Household Consumption to Income Tax Refunds," *American Economic Review*, 89(4): 947-958.

YouGov/Huffington Post (2013a) "Do you think President Obama and Republicans in Congress will reach a deal to avoid a government shutdown, or that they will not reach a deal and the government will shut down?" Question 4 of 6, September 28 – September 29, 2013.

YouGov/Huffington Post (2013b) "How long do you think the government shutdown will last before President Obama and Congress reach a new budget deal?" Question 5 of 5, October 2 – October 3, 2013.

Figure 2.7.1: Government shutdown timeline

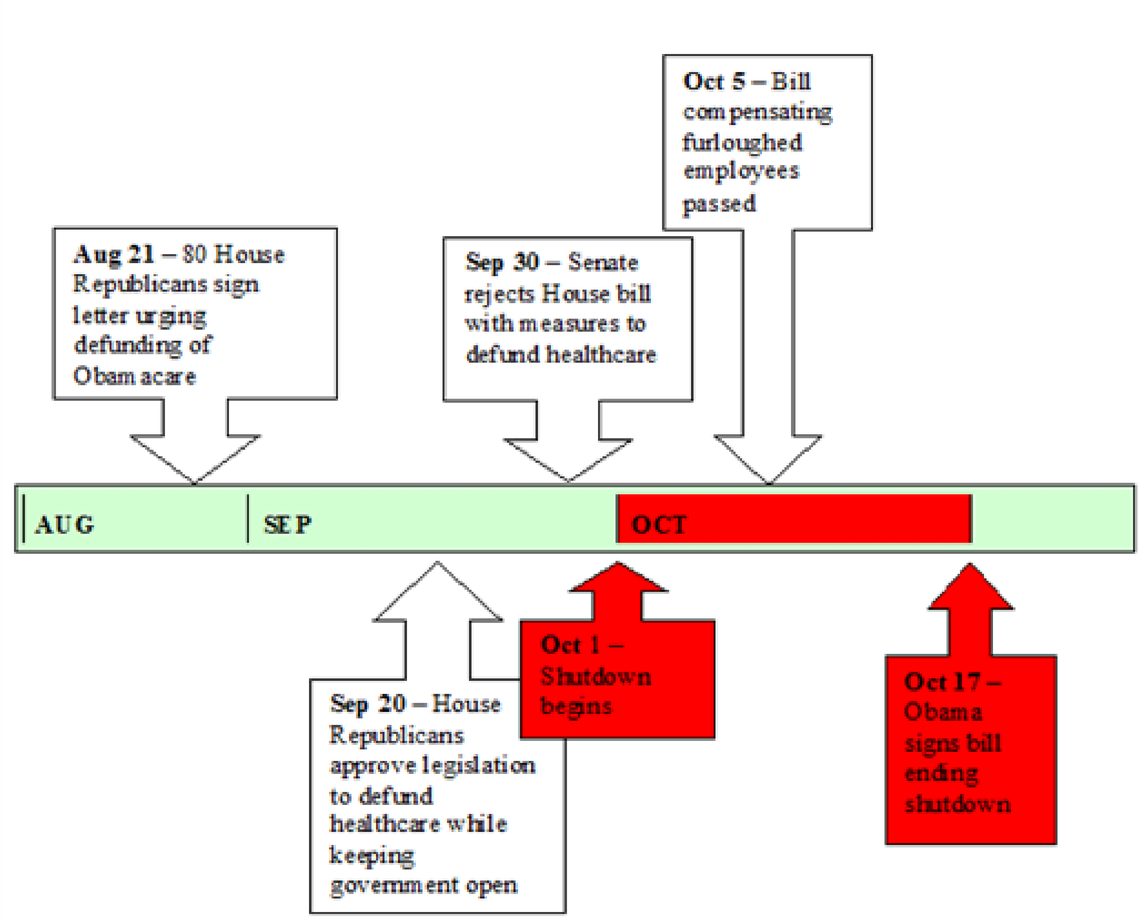


Figure 2.7.2: Time series of spending

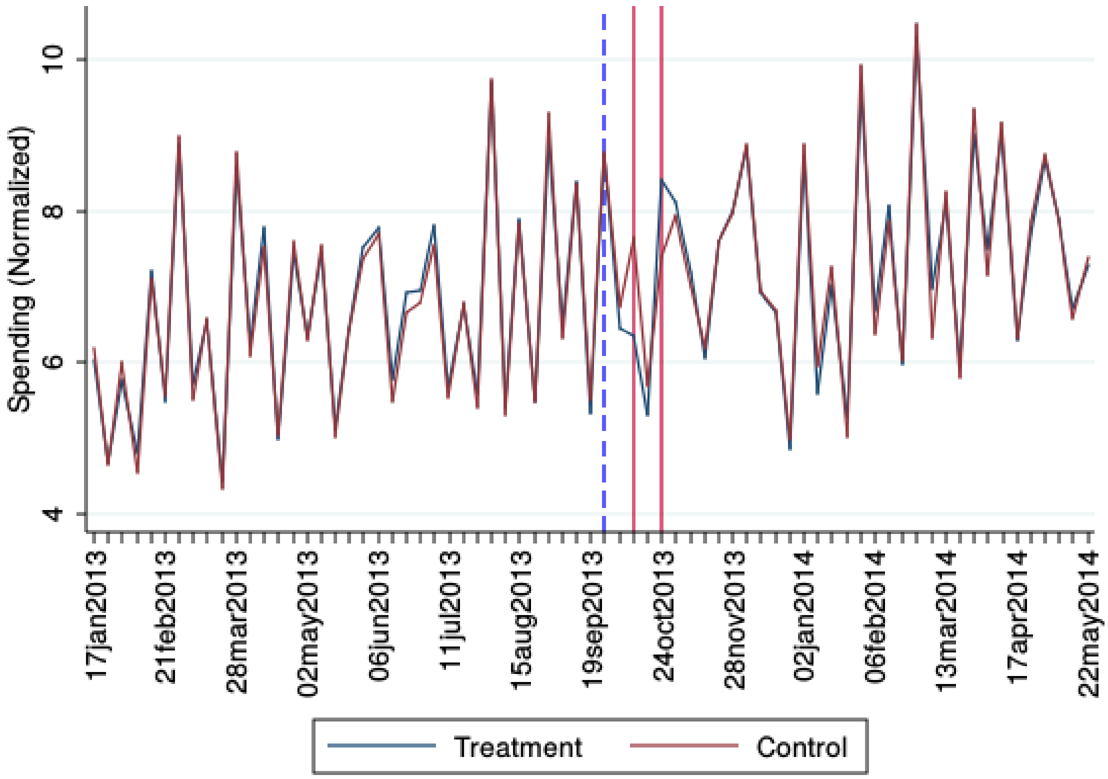


Figure 2.7.3: Pre-shutdown median liquidity over the paycheck cycle

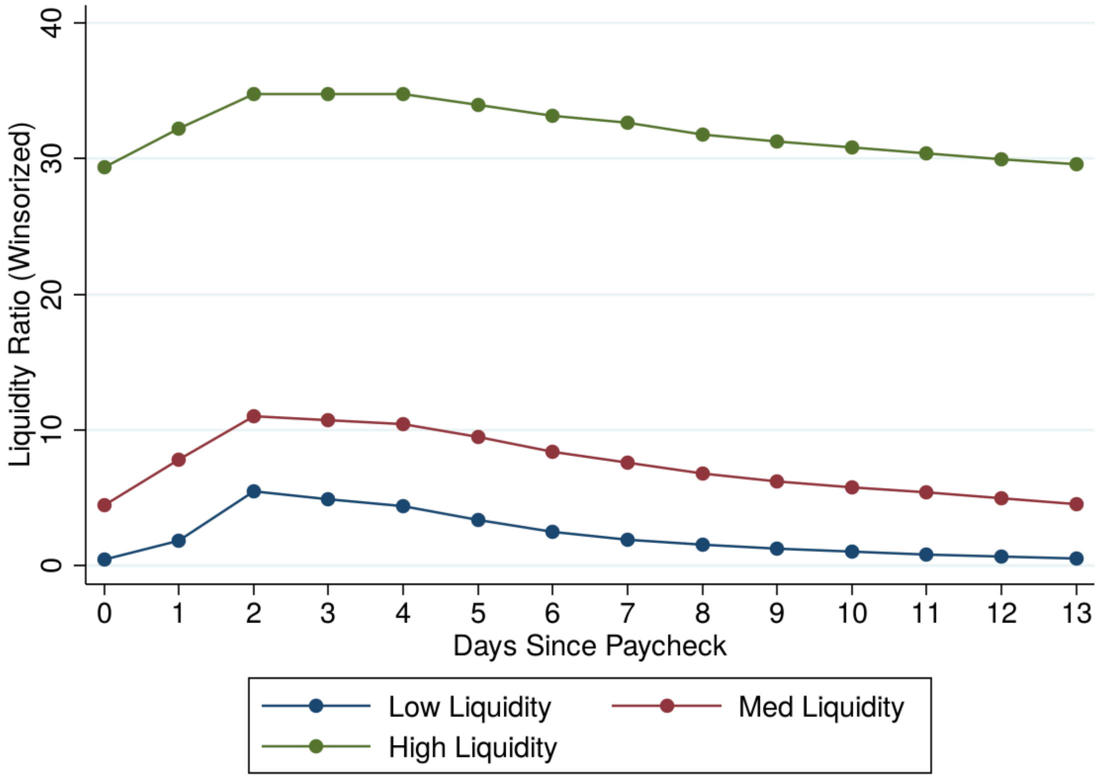


Figure 2.7.4: Estimated response of normalized paycheck income and normalized total spenidng to government shutdown

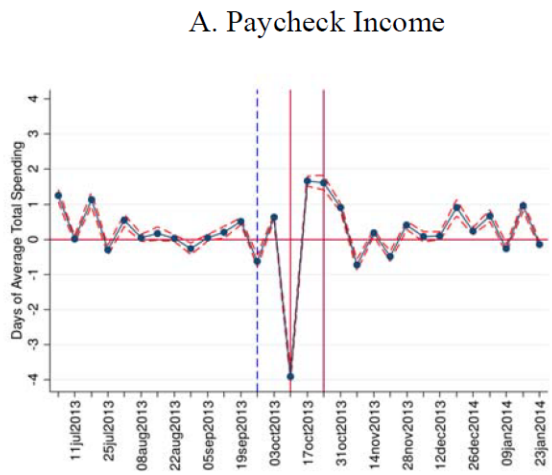


Figure 2.7.5: Estimated response of spending categories to government shutdown

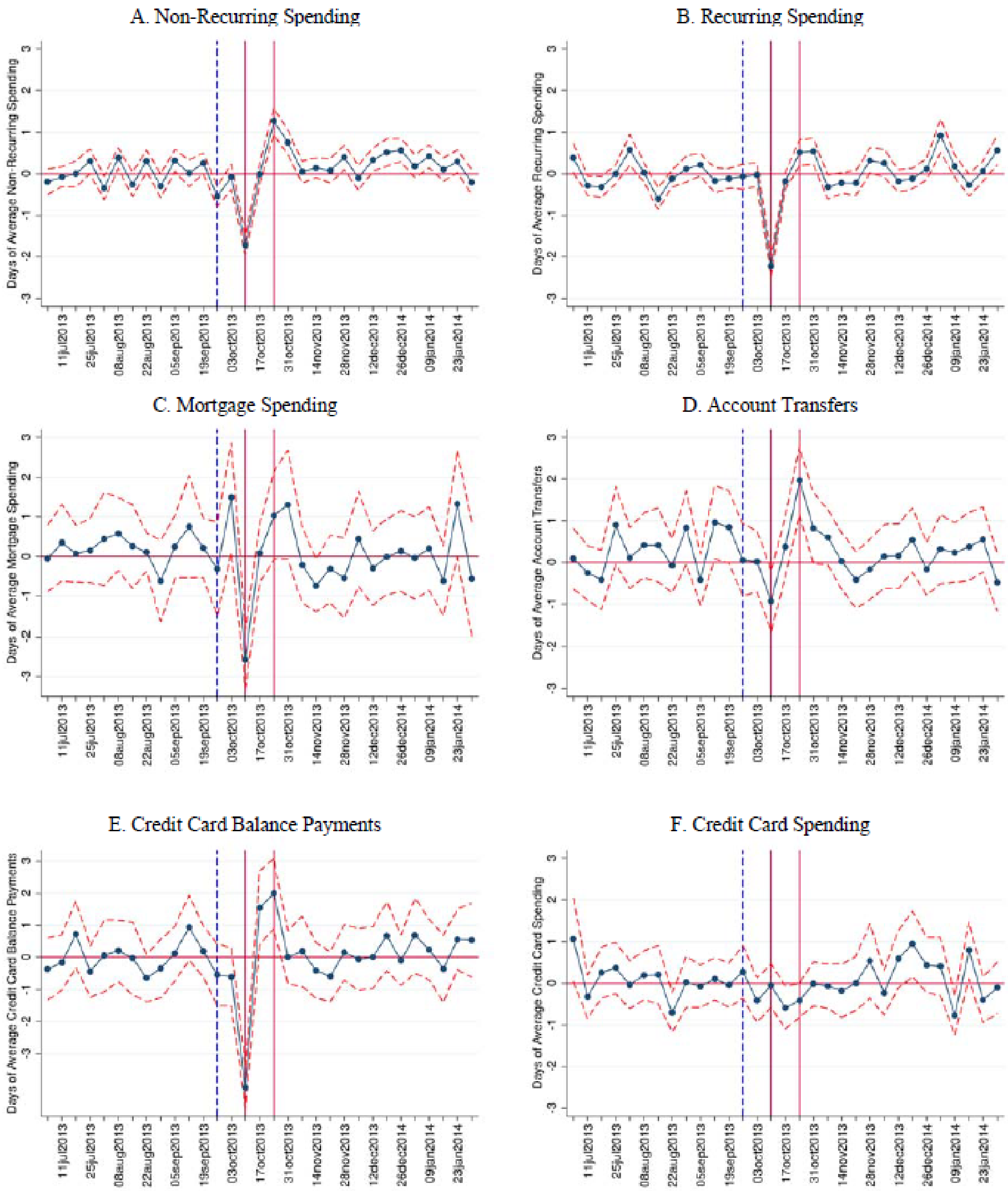


Figure 2.7.6: Distribution of week mortgage is paid

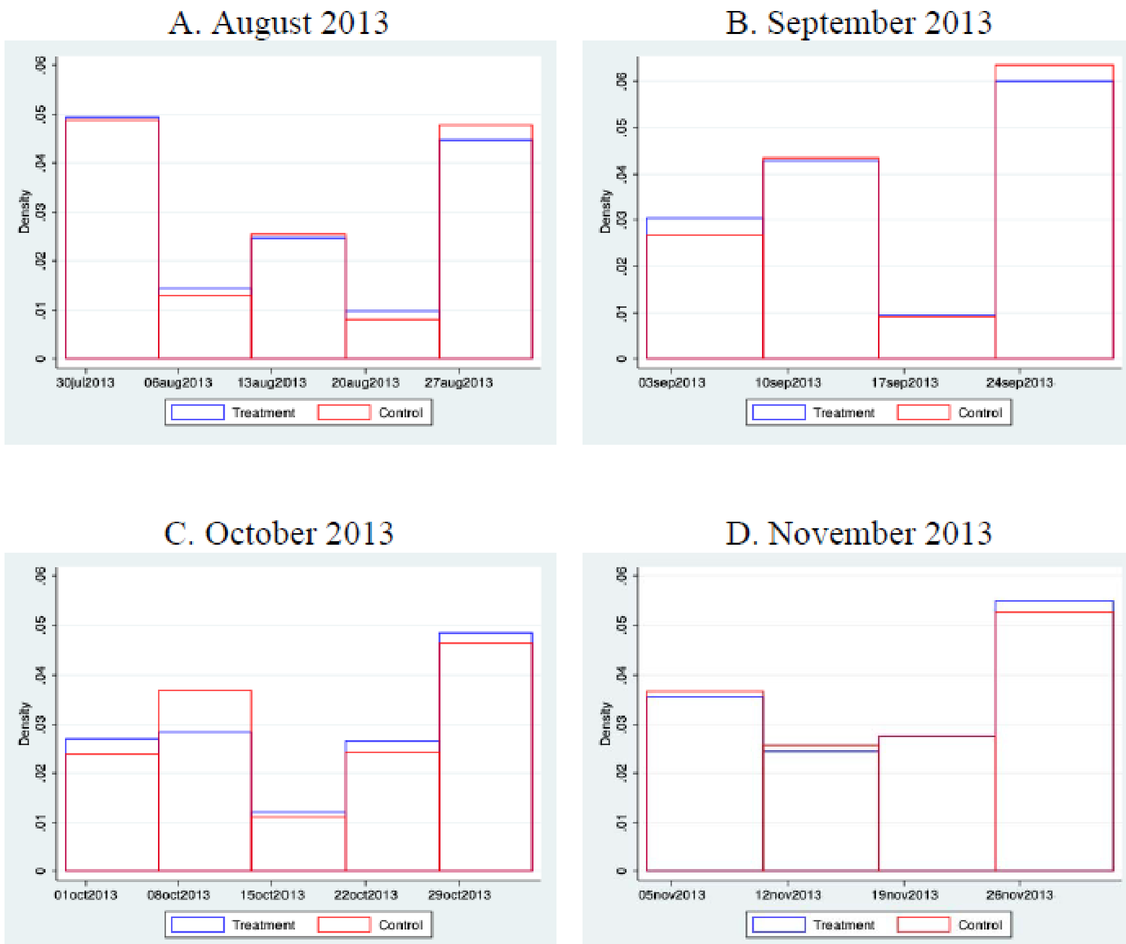


Figure 2.7.7: Estimated response of coffee shop and fast food spending to government shutdown

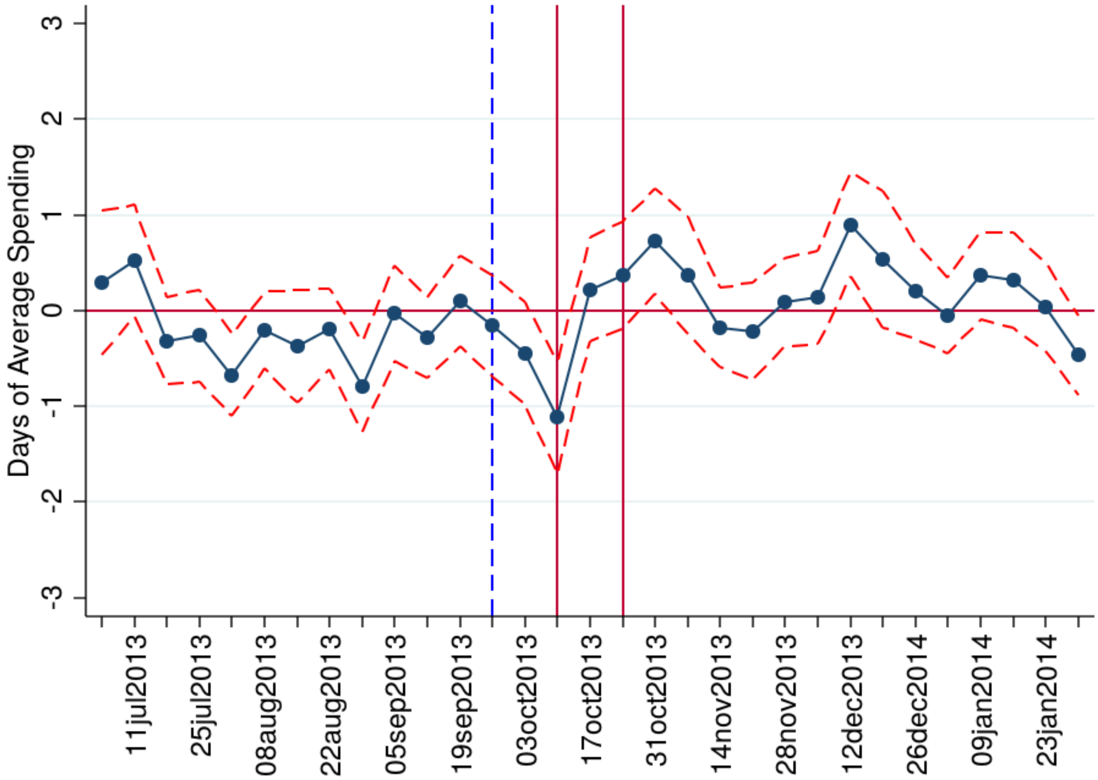


Figure 2.7.8: Estimated response liquid assets to government shutdown

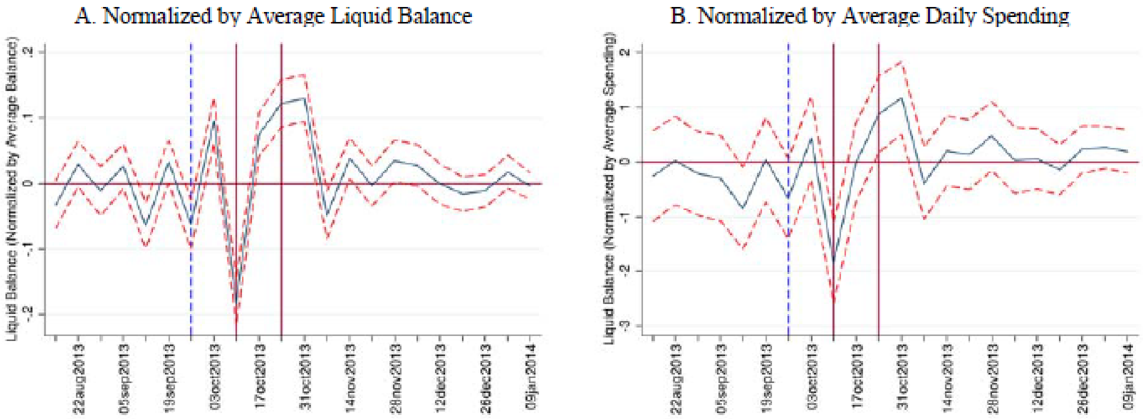


Figure 2.7.9: Estimated response of spending categories to government shutdown by liquidity tercile

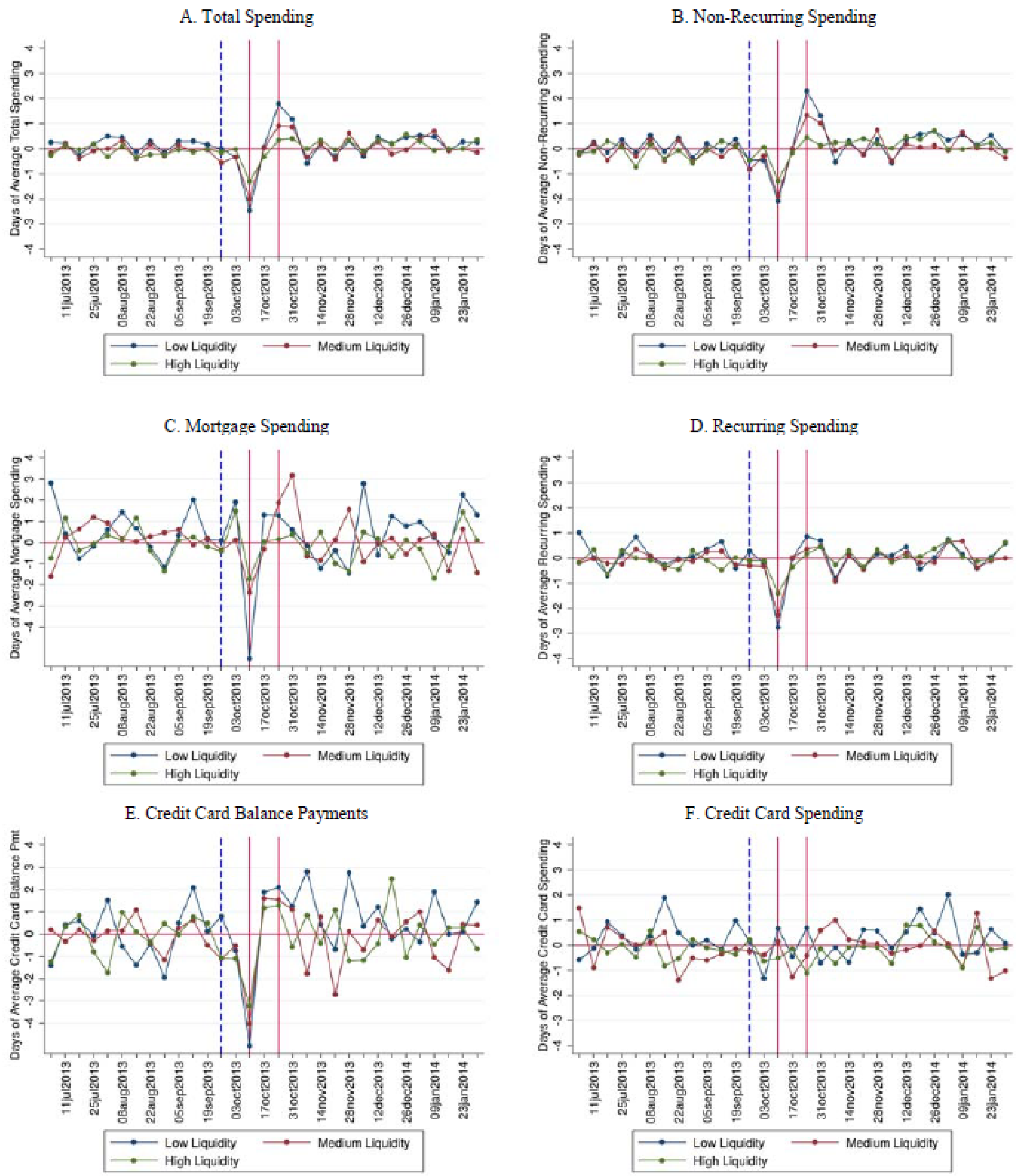


Figure 2.7.10: Estimated response of credit card debt to government shutdown

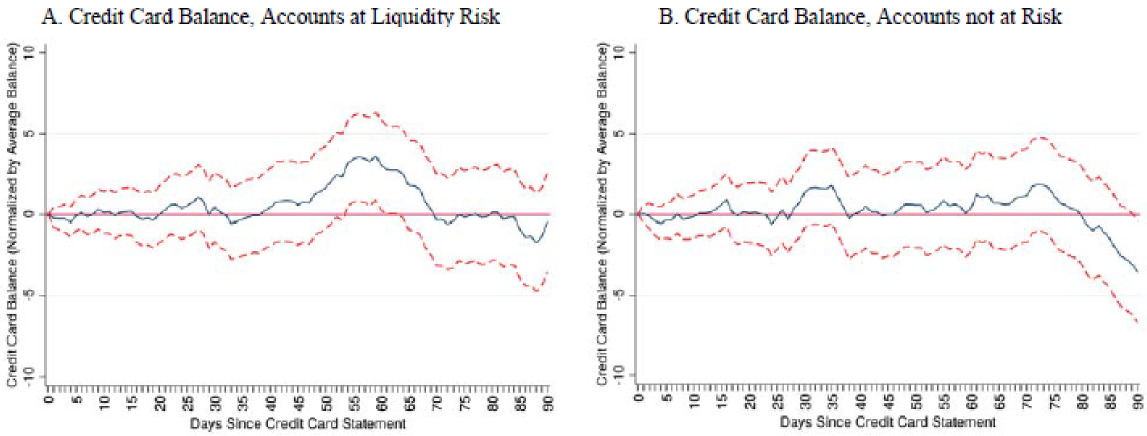


Table 2.1: Average hourly compensation of federal employees relative to that of private-sector employees, by level of educational attainment

	Difference in 2010 Dollars per Hour			Percentage Difference		
	Wages	Benefits	Total Compensation	Wages	Benefits	Total Compensation
High School Diploma or Less	$4	$7	$10	21%	72%	36%
Bachelor's Degree	-	$7	$8	-	46%	15%
Professional Degree or Doctorate	-$15	-	-$16	-23%	-	-18%

Table 2.2: Employee characteristics

	N	Average Weekly Income	Standard Deviation of Weekly Income	Average Weekly Spending	Average Normalized Liquid Balance (days)	Average Credit Card Debt
All Federal Employees	6,792	$1,728	$1,415	$1,855	27	$3,673
Affected by the Shutdown	3,804	$1,727	$1,326	$1,861	26	$3,785
Not affected by the Shutdown	2,988	$1,729	$1,521	$1,849	29	$3,529
Non-Federal Employees	91,692	$1,261	$1,360	$1,362	23	$2,461

Table 2.3: MPC estimates

Lag(k)	2	1	0	-1
$\hat{\beta}_{MPC}$	0.0653	0.0213	0.5765	0.0248
	(0.0308)	(0.0335)	(0.0271)	(0.0248)
Observations	98,476	98,476	98,476	98,476
SEE	7.107	6.780	6.602	6.263

Table 2.4: Liquidity ratio

	Treatment Group		Control Group	
Liquidity Ratio Tercile	Mean (Days)	N	Mean (Days)	N
1	2.9	851	2.8	25,105
2	8.4	1131	8.4	24,824
3	54.2	1201	54.7	24,754

CHAPTER III

The Response of Consumer Spending to Changes in Gasoline Prices

3.1 Introduction

Few macroeconomic variables grab headlines as often and dramatically as do oil 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 consumers' spending from resources freed up by lower gasoline prices. Estimating the quantitative impact of this element is central to policy decisions. Yet, because of data limitations a definitive estimate has proved elusive. Recently, "big data" has opened unprecedented opportunities to shed new light on the matter. This paper uses detailed transaction-level data provided by a personal financial management service to assess the spending response of consumers to changes in gasoline prices over the 2013-2016 period.

Specifically, we use this information to construct high-frequency measures of spending on gasoline and non-gasoline items for a panel of more than one million U.S. consumers. We use cross-consumer variation in the intensity of spending on gasoline interacted with the large, exogenous, and permanent decline in gasoline prices to identify and estimate the partial equilibrium marginal propensity to consume (MPC) out of savings generated by reduced gasoline prices. Our baseline estimate of the

MPC is approximately one. That is, consumers on average spend *all* of 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 useful and informative in several dimensions. First, our estimate of the 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, and also consistent with the PIH, we show that consumers' liquidity was not important for the strength of the consumer spending response to gasoline price shocks. Fourth, our analysis highlights the importance of having high-frequency transaction data at the household level for estimating consumer reactions to income and price shocks.

This paper is related to several strands of research. The first strand, surveyed in Jappelli and Pistaferi (2010), is focused on estimating consumption responses to income changes. Typically, studies in this area examine if and how consumers react to anticipated income shocks, which in many cases are transitory. As standard in this literature, our estimate provides the partial equilibrium response of household spending to a shock, not the general equilibrium outcome for aggregates. A common finding in this strand 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. For example, Kaplan and Violante (2014) and many others document that consumers with low liquidity holdings respond more strongly to anticipated income shocks. At the same time, estimating spending responses to unanticipated, highly persistent income shocks has been challenging, because identifying such shocks is particularly difficult.

We contribute to this literature in several ways. First, we exploit a particularly clear-cut source of variation in household budgets (spending on gasoline) with a number of desirable properties. Specifically, we use a large, salient, unanticipated, permanent (or perceived to be permanent) shock. Second, we examine spending responses at the weekly frequency while, due to data limitations, the vast majority of previous studies estimate responses at much lower frequencies. As we discuss below, the high-frequency dimension allows us to obtain crisp estimates of the MPC and thus provide a more informative input for policy making.

The second strand to which we contribute studies the effects of oil prices on the economy. In surveys of this literature, Hamilton (2008) and Kilian (2008) emphasize that oil price shocks can influence aggregate outcomes via multiple channels (e.g., consumer spending, changes in expectations) but disruption of consumers' (and firms') spending on goods other than energy is likely to be a key mechanism for amplification and propagation of the shocks. Indeed, given the low elasticity of demand for gasoline, changes in gasoline prices can materially affect non-gasoline spending budgets for a broad array of consumers. As a result, a decrease in gasoline prices can generate considerable savings for consumers which could be put aside (e.g., to pay down debt or save) or used to spend on items such as food, clothing, furniture, etc.

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. One reason for the dearth of research on the matter has been 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 based on the CEX produces noisy and attenuated estimates.

Grocery store barcode data, such as from AC Nielsen, have become a popular alternative 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. As a result, most estimates of MPC tend to be based on time series variation in aggregate series (see e.g. Edelstein and Kilian 2009).

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 select number of grocery stores with a gasoline station on site. Perhaps not surprisingly, we show that households typically visit multiple gasoline station retailers in a month, suggesting limitations to focusing on consumer purchases at just one gasoline 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, yet they 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 important, 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 limited to Chase accounts. If, for example, consumers use a non-Chase credit card or checking account, any spending on that account would be missed in the J.P. Morgan Chase Institute analysis, and measurement of household responses may therefore be importantly incomplete. In this paper, we confirm this by showing that an analysis based on accounts in one financial institution leads to a significantly 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.

3.2 Recent Changes in Gasoline Prices: Unanticipated, Permanent and Exogenous

In this section, we briefly review recent dynamics in the prices of oil and gasoline and corresponding expectations of future prices. We document that the collapse of oil and gasoline prices in 2014-2015 was highly persistent, unanticipated, and exogenous to demand conditions in the United States. These properties of the shock are important components of our identification strategy.

3.2.1 Unanticipated and Permanent

In Panel A of Figure 1, the solid black line shows the spot price of gasoline at New York Harbor, an important import and export destination for gasoline. The New York Harbor price is on average 70 cents lower than average retail prices, although the two series track each other very closely. The dashed line shows the one-year-ahead futures price for that date. The futures price tracks the spot price closely, suggesting the market largely treats gasoline price as a random walk—i.e., the best prediction for one-year-ahead price is simply the current price.

Panel B shows the difference between the realized and predicted spot price. The behavior of one-year-ahead forecast errors indicates that financial markets anticipated neither the run-up nor the collapse of gasoline prices in 2007-2009. Likewise, the dramatic decline in gasoline prices in 2014-2015 was not anticipated.

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 C of Figure 1 plots the mean and median consumer expectations along with the actual price and the mean one-year-ahead prediction in the Survey of Professional Forecasters. While consumers expect a slightly higher price relative to the present price than professional forecasters, the basic pattern is the same as in Panel A: the current price appears to be a good summary of expected future prices. Consistent with this observation, Anderson, Kellogg and Sallee (2012) fail to reject the null of a random walk in consumer expectations for gasoline prices. Thus, consumers perceive changes in gasoline prices as permanent. Also similar to the financial markets, consumers were not anticipating large price changes in 2007-2009 or 2014-2015 (Panel D).

Figure 1 shows large movements in prices during the Great Recession (shaded). Unlike the recent episode that is the subject of this paper, we would *not* use it to identify the MPC because this fluctuation in commodity prices in the Great Recession surely represents an endogenous response 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.[1] The price drops in 2014-2015 are some of the largest changes in oil and gasoline prices in the last 60 years. Note that in 1986, gasoline prices and oil prices actually moved in opposite directions, indicating that the process generating gasoline prices can sometimes differ from oil.

3.2.2 Exogenous

Why did prices of oil and oil products such as gasoline fall so much in 2014-2015? While many factors could have contributed to the dramatic decline in the prices, the consensus view, summarized in Baffes et al. (2015), attributes a bulk of the decline to supply-side factors. 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.). Consistent with this view, other commodity prices had modest declines during this period, which would not have happened if the decline in oil prices was driven by global demand factors. Observers note that the collapse of oil prices in 2014-2015 is similar in many ways to the collapse in 1985-1986, when more non-OPEC oil supply came from Mexico, the North Sea and other sources, and OPEC also decided to abandon price support. In short, available evidence suggests that the 2014-2015 decline in oil prices is a shock that was supply-driven and exogenous

[1] 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.

to U.S. demand conditions. In contrast, for the 2007-2009 episode, 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).

3.3 Data

Our analysis uses high-frequency data on spending from a financial aggregation and bill-paying computer and smartphone application (henceforth, the "app").[2] The app had approximately 1.4 million active users in the U.S. in 2013.[3] Users can link almost any financial account to the app, 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.

We draw on the entire de-identified population of active users and data derived from their records from January 2013 until February 2016. The app does not collect demographic information directly and instead uses a third party business 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 C2) 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

[2]These data have previously been used to study the high-frequency responses of households to shocks such as the government shutdown (Gelman et al. 2016) and anticipated income, stratified by spending, income and liquidity (Gelman et al. 2014).

[3]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.

user population is not representative of the U.S. population, but it is heterogeneous, including large numbers of users of different ages, education levels, and geographic location.

3.3.1 Identifying Spending Transactions

Not every transaction reported by the app is spending. For example, a transfer of funds from one account to another is not spending. 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. More details are provided in Appendix A.

3.3.2 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 in this study will be increasingly important.

We use an ML algorithm to construct a set of rules for classifying the data 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.[4] Because distinguishing gasoline purchases classified as 5542 or 5541 is nearly impossible with the information in transaction descriptions,[5] we group transactions with these two codes together.

A downside of this approach is that transactions at the convenience store associated with a gasoline station can be classified as a purchase of gasoline; that is, buying a food item in a gasoline station's convenience store can be classified as a purchase

[4]To be clear, "Service Stations" do not include services such as auto repairs, motor oil change, etc.

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

of gasoline. 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), only 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 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 transactions less than $10 to filter out purchases of non-gasoline items at "Service Stations."

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 of decision trees 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 on 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.[6] The ML model is able to classify spending with approximately 90% accuracy in the second provider not used to train the model, 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.

[6] 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.

3.3.3 Comparison with the Consumer Expenditure Survey

We compare our measures of gasoline and non-gasoline spending with similar measures from the Consumer Expenditure Survey (CEX). 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.[7]

[7] We have done a comparison of the CEX diary spending for January 2013 through December 2014 (the last time period that the CEX is available at this point). 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.)

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.[8] The distribution has clear spikes at multiples of $50 and $100 with the largest spikes at $0 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 (2015), 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 biweekly gasoline spending in the CEX Diary Survey is $84.72, while the app counterpart is $81.68. Similarly, non-gasoline spending is $1,283 in the CEX Diary Survey and $1,469 in the app data. The standard deviation, however, tends to be larger in the app data than in the CEX, which reflects a thick right tail of spending in the app data. This pattern is consistent with under-representation of higher-income households in the CEX, a well-documented phenomenon (Sabelhaus et al. 2015). Medians and inter-quartile ranges, measures of central tendency and variations that are less sensitive to the tails, are smaller in the app data than the CEX Diary Survey's counterparts. The moments in the CEX Interview Survey (quarterly frequency) are generally closer to the moments in the app data. For example, mean

[8]The CEX Interview Survey question asks households to report their "Average monthly expense for gasoline."

(median) spending on gasoline is \$647 (\$540) in the CEX Interview Survey data and \$628 (\$475) in the app data, while the standard deviations (interquartile ranges) are \$531 (\$630) and \$588 (\$660) 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 extremely 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.[9]

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. Some of the differences could reflect survey recall bias, 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 gasoline spending done in cash that we could not identify.[10] We will address these potential issues in our robustness tests.

3.4 Empirical Strategy

The discourse on potential macroeconomic effects of a fall in gasoline prices 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 marginal propensity to consume (MPC) from savings generated by reduced gasoline prices.[11] Define *MPC* as

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

[10] According to the 2015 NACS Retail Fuels Report (NACS 2015), approximately 80 percent of gasoline purchases are made with debit or credit card.

[11] 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."

$$dC_{\text{it}} = -MPC * d(P_t Q_{\text{it}})$$

where i and t index consumers and time, C is spending of 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 in response to a *decrease* in the price of gasoline.[12]

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. In this section, 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 aggregate economic 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, at the same time demand was falling for other goods. 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. Yet another example is a person who suffers an unemployment spell; this worker 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

[12]The MPC is likely different across groups of people, but our notation and estimation refers to the average MPC.

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 this, we only need to satisfy a weaker set of conditions. First, we need exogenous, 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 ($P_t Q_{\text{it}}$).

As we established in Section II, shocks to gasoline prices in the period of our analysis were unanticipated, exogenous, and permanent so that we have an exogenous source of variation. 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:

$$\begin{aligned}
\frac{dC_{\text{it}}}{C_i} &= d\log C_{\text{it}} = -MPC \times \frac{d(P_t Q_{\text{it}})}{C_i} = -MPC \times \frac{d(P_t Q_{\text{it}})}{(\text{PQ})_i} \times \frac{(\text{PQ})_i}{C_i} \\
&= -MPC \times \frac{d(P_t Q_{\text{it}})}{(\text{PQ})_i} \times s_i \\
&= -MPC \times \frac{Q_i dP_t + P dQ_{\text{it}}}{(\text{PQ})_i} \times s_i \\
&= -(MPC \times s_i \times \frac{dP_t}{P} + MPC \times s_i \times \frac{dQ_{\text{it}}}{Q_i}) \\
&= -MPC \times s_i \times d\log P_t - MPC \times s_i \times \left(\frac{dQ_{\text{it}}}{Q_i} \times \frac{P}{dP_t}\right) \times \frac{dP_t}{P} \\
&= -MPC \times s_i \times d\log P_t - MPC \times s_i \times \epsilon \times d\log P_t \\
&= -MPC \times (1+\epsilon) \times s_i \times d\log P_t \ (2)
\end{aligned}$$

where bars denote steady-state values, $s_i \equiv \frac{(\text{PQ})_i}{C_i}$ is the ratio of gasoline spending to non-gasoline spending,[13] and ϵ is the price elasticity of demand for gasoline (a

[13]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

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.[14] The cross-section variation is essential for this paper since there is 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, when we regress log non-gasoline spending on the log of gasoline price multiplied by the ratio of gasoline spending to non-gasoline spending, we get $-MPC\,(1+\epsilon)$.

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:

$$d\log P_t Q_{\text{it}} = d\log P_t + d\log Q_{\text{it}}$$
$$= d\log P_t + d\log P_t \tfrac{d\log Q_{\text{it}}}{d\log P_t} = \left(1+\tfrac{d\log Q_{\text{it}}}{d\log P_t}\right)\times d\log P_t$$
$$= (1+\epsilon)\times d\log P_t. \ (3)$$

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 $d\log P_t Q_{\text{it}}$ on $d\log P_t$ yields $(1 \hat{+} \epsilon)$, 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.[15] The estimated $(1+\epsilon)$ and $-MPC\,(1+\epsilon)$ can be combined to obtain the MPC.

In the derivation of equations (2) and (3) we deliberately did not specify the time

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. Given that gasoline prices are approximately random walks, we take s_i values in 2013 as stochastic steady-state values.

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

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

horizon over which sensitivities 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). It might take time to notice 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,

$\Delta_k \log C_{\text{it}} = \beta_k \times s_i \times \Delta_k \log P_t + \psi_t + \varepsilon_{\text{it}}$ (4)

$\Delta_k \log \text{PQ}_{\text{it}} = \delta_k \log P_t + u_{\text{it}}$ (5)

where $\beta = -MPC\,(1+\epsilon)$, $\delta = (1+\epsilon)$, $\Delta_k x_t = x_t - x_{t-k}$ is a k-period-difference operator, ψ_t is the time fixed effect, and ϵ_{it} and u_{it} are individual-level shocks to spending.[16] By varying k, we can recover the average impulse response over k-periods so that we can remain agnostic about how quickly consumers respond to a change in gasoline prices.[17] Given that our specification is in differences, we control for consumer time-invariant characteristics. 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, our estimate obtains after controlling for common

[16] Note that there are time effects only in equation (4). Since we have argued 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 efficiency by absorbing aggregate shocks to overall spending. It is not possible to 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 biased under our maintained assumption that gasoline prices are exogenous to the U.S. economy in the estimation period. (The standard errors account for residual aggregate shocks.)

[17] For example, if $\log C_{\text{it}} = \sum_{s=0}^{\infty} \psi_s \text{shock}_{t-s} + u_t$ and u_t summarizes variation orthogonal to the shock series of interest, then the impulse response is $\{\psi_s\}_{s=0}^{\infty}$ and the long-difference regression recovers $\beta_k = k^{-1} \sum_{s=0}^{k-1} \psi_s$.

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 Pistaferi 2010), we estimate a partial equilibrium MPC.

Note that gasoline and oil prices are approximately random walks and thus $\log P_t$ 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 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 national and we expect serial within-user correlation in spending, we cluster standard errors on two dimensions: time and person. This simplicity is particularly convenient in our case because we estimate equations (4) and (5) as a system to recover MPC from estimates of $-MPC\,(1+\epsilon)$ and $(1+\epsilon)$.

To summarize, our econometric framework identifies the MPC from changes in gasoline prices by interacting two sources of variation: a large, exogenous, and permanent change in gasoline prices, with the pre-determined 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 lagged adjustment of gasoline spending to changes in gasoline prices.

3.5 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.

3.5.1 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 $\hat{\epsilon}$ and 95 percent confidence bands for $k = 0, \ldots,$ 26 weeks. Table 3, Row 1, gives the point estimates for selected horizons. The point estimates indicate that the elasticity of demand for gasoline is increasing in the horizon (i.e., over time, consumers have greater elasticity of demand): estimated elasticity changes from -0.20 at the horizon of 15 weeks to -0.24 at the horizon of 25 weeks. Confidence intervals are very wide at short horizons; estimates become quite precise at horizons of 12 weeks and longer.

This estimate is broadly in line with previously reported estimates. 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 -0.21 to -0.34 for the 1975-1980 period. The Environmental Energy Institute (EIA 2014) also points to an elasticity close to zero, and also argues this elasticity has been trending downward over time.[18]

[18] 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.”

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 of these 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. Paul and Greening's (1999) baseline estimates are -0.34 and -0.47, depending on the specification. A more recent paper by Levin, Lewis and Wolak (2016) 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 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 B of Figure 5 shows the dynamics of $\hat{MPC}$ and 95 percent confidence bands over the same horizons 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. Starting with the four-week horizon, we observe that $\hat{MPC}$ steadily rises over time and becomes increasingly precise. After approximately 12 weeks, $\hat{MPC}$ stabilizes between 0.8 and 1.0 with a standard error of 0.3. The estimates suggest that, over longer horizons, consumers spend nearly *all* their gasoline savings on non-gasoline items. The standard errors are somewhat smaller at monthly horizons (4-5 weeks) since the shock. We suspect this is because the residual variance in consumption tends to be lower at monthly frequency due to factors like recurring spending, bill pay, and rules of thumb/behavioral reasons (shopping once per month), while in other weeks, the consumption process has considerably more randomness.[19]

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.[20] 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 0.515, which is approximately half of the baseline $\widehat{MPC}$ at 0.963, but the standard error for the former estimate is 0.387, so that we cannot reject equality of the estimates as well as

[19] Although there is some regional variation in the *level* of gasoline prices, the *comovement* of gasoline prices is very strong and thus little is lost by using 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 gasoline prices. We find nearly identical estimates when we use local gasoline prices.

[20] These providers cover 49.6 percent of accounts in the data and 55.0 percent of total spending.

equality of the former estimate to zero.

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. We find estimates largely similar to the case of any account, that is, the estimated sensitivity to changes in gasoline prices is attenuated and 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.

3.5.2 Robustness

While our specification has several 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 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.[21]

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 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. We find (row 3) 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 $\Delta_k \log P_t$ in our baseline specification (4), we instead use $\Delta_k \log \mathcal{F}_t \equiv \log F_t^1 - \log F_{t-k}^1$ for $k \in \{1, \ldots, 25\}$. While the focus on one-month change is arguably justified

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

given approximate random walk in gasoline prices, we also try the average change in the yield curves for gasoline prices over longer horizons (two years) to have a measure of changes in gasoline prices that are perceived as persistent: $\Delta_k \tilde{\log} \mathcal{F}_t \equiv \frac{1}{24}\sum_{h=1}^{24}\left(\log F_t^h - \log F_{t-k}^h\right)$. In either one-month change (row 4 of Table 4) or average change over two years (row 5), the results are very similar to our baseline.

3.5.3 Comparison with MPC using CEX

To appreciate the significance of using high-quality transaction-level data for estimating the sensitivity of consumers to income and price shocks, we estimated 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.[22] For this specification, we use BLS urban gasoline prices which provide a consistent series over this time period (see note for Table 1).

[22]Our build of the CEX data follows Coibion et al. (2012).

In the first row of Table 5, we estimate our baseline specification for the app data at the quarterly frequency: the estimates are slightly different from the estimates based on the weekly frequency, though much less precise. The standard errors are so large that we cannot reject the null of equality of the estimates over time or across frequencies.

Panel B of Table 5 presents estimates based on the CEX. To maximize the precision of CEX estimates, we apply our approach to the CEX data covering 1980-2014. 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 was largely determined by supply-side and foreign-demand factors, 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 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. The 1990-1992 episode yields negative MPCs with large standard errors. In what follows, we investigate to what extent the research design of the CEX contributes to these estimates.

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 elasticity of demand for gasoline is even lower at shorter horizons and even greater at the longer horizons. In a similar spirit, the estimated MPC increases more strongly in the horizon when we track consumers for only four quarters relative to the complete 2013-2016 coverage.

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, robust 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.

3.5.4 Heterogeneity in Responses

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 consumers with liquidity constraints ("hand-to-mouth" consumers, HtM) should exhibit a larger MPC to *transitory*, *anticipated* income shocks than consumers for whom these constraints are not binding (non-HtM consumers). 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. Given that the shock to the price of gasoline in our analysis was unanticipated and perceived as permanent, 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 "hand-to-mouth" consumers requires information on consumers' liquidity holdings *just before* they receive pay checks.[23] Because the Survey of Consumer Finances (SCF), the dataset used in Kaplan and Violante (2014), reports *average* 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 hand-to-mouth (HtM) and non-hand-to-mouth (non-HtM) consumers. As a result, MPC estimated in Kaplan and Violante (2014)

[23]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.

could provide a lower bound for the true 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. A consumer is classified as HtM in a given month if, 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 in the lower end of the range reported in Kaplan and Violante (2014).[24]

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

$$\Delta_k \log C_{\text{it}} = \beta_1 \times s_i \times \Delta_k \log P_t + \beta_2 \times s_i^{\text{gas}} \times \Delta_k \log P_t \times D_{\text{it}}$$

$$+\mu_0 \times D_{\text{it}} + \mu_1 \times s_i \times D_{\text{it}} + \psi_t + \omega_t \times D_{\text{it}} + \varepsilon_{\text{it}} \quad (6)$$

$$\Delta_k \log \text{PQ}_{\text{it}} = \delta_1 \times \log P_t + \delta_2 \times \log P_t \times D_{\text{it}} + \xi \times D_{\text{it}} + u_{\text{it}} \quad (7)$$

where D_{it} is a variable measuring the presence/intensity of liquidity constraints identifying HtM consumers, and $\omega_t \times D_{\text{it}}$ 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 $\tilde{D}_{\text{it}} \equiv D^*_{i,t-k-1}$ where D^*_{it} is a dummy variable equal to one if consumer i

[24] 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.

at time t satisfies Kaplan-Violante's 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 hand-to-mouth 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 $D_{i,2013}$), the modal value of $D^*_{i,t}$ (most frequent value;[25] 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 $\hat{D}_{i,2013}$, 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 Table 6 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 cannot reject the null of equal MPCs across the groups for any horizon or definition of HtM status. Furthermore, we cannot reject that estimated MPCs are equal to one. These results are consistent with theoretical predictions and thus lend more credibility to our baseline estimate of MPC equal to one.

3.6 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 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

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

to approximately 2 percent of total consumer spending. (Average total household spending in 2014 was $53,495 total, while the average household spending on gasoline was $2,468.)

We document that the marginal propensity to consume out of these savings is approximately one, which is consistent with the predictions of the permanent income hypothesis, given that the change in gasoline prices was unexpected and permanent. This partial equilibrium estimate provides a first-step input for quantifying the effects on the aggregate economy, which depend on a number of 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 first-round 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. While estimating the aggregate effects of the change in oil prices is beyond the scope of this paper, it is clear that a persistent increase of spending of 2 percent on the part of households purchasing gasoline is a large macroeconomic shock.

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 enormous potential of big data for enhancing national economic statistics, as well as estimates of key, policy-relevant macroeconomic parameters.

Bibliography

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.

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

Binder, Carola. 2015. "Measuring Uncertainty Based on Rounding: New Method and Application to Inflation Expectations."

Coibion, Olivier, Yuriy Gorodnichenko, Lorenz Kueng, and John Silvia. 2012. "Innocent Bystanders? Monetary Policy and Inequality in the U.S." NBER Working Paper No. 18170.

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

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.

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

Gelman, Michael, Shachar Kariv, Matthew D. Shapiro, Dan Silverman and Steven Tadelis. 2016. “How Individuals Smooth Spending: Evidence from the 2013 Government Shutdown Using Account Data.” Working Paper.

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.

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." *The Quarterly Hournal 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.*

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.

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

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.

Levin, Laurence, Matthew S. Lewis and Frank A. Wolak. 2016. "High Frequency Evidence On The Demand For Gasoline," NBER Working Paper no. 22345

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 Jn 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.

National Association of Convenience Stores (NACS). 2015. *Retail Fuels Report.*

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.

Figure 3.6.1: Gasoline prices and expectations

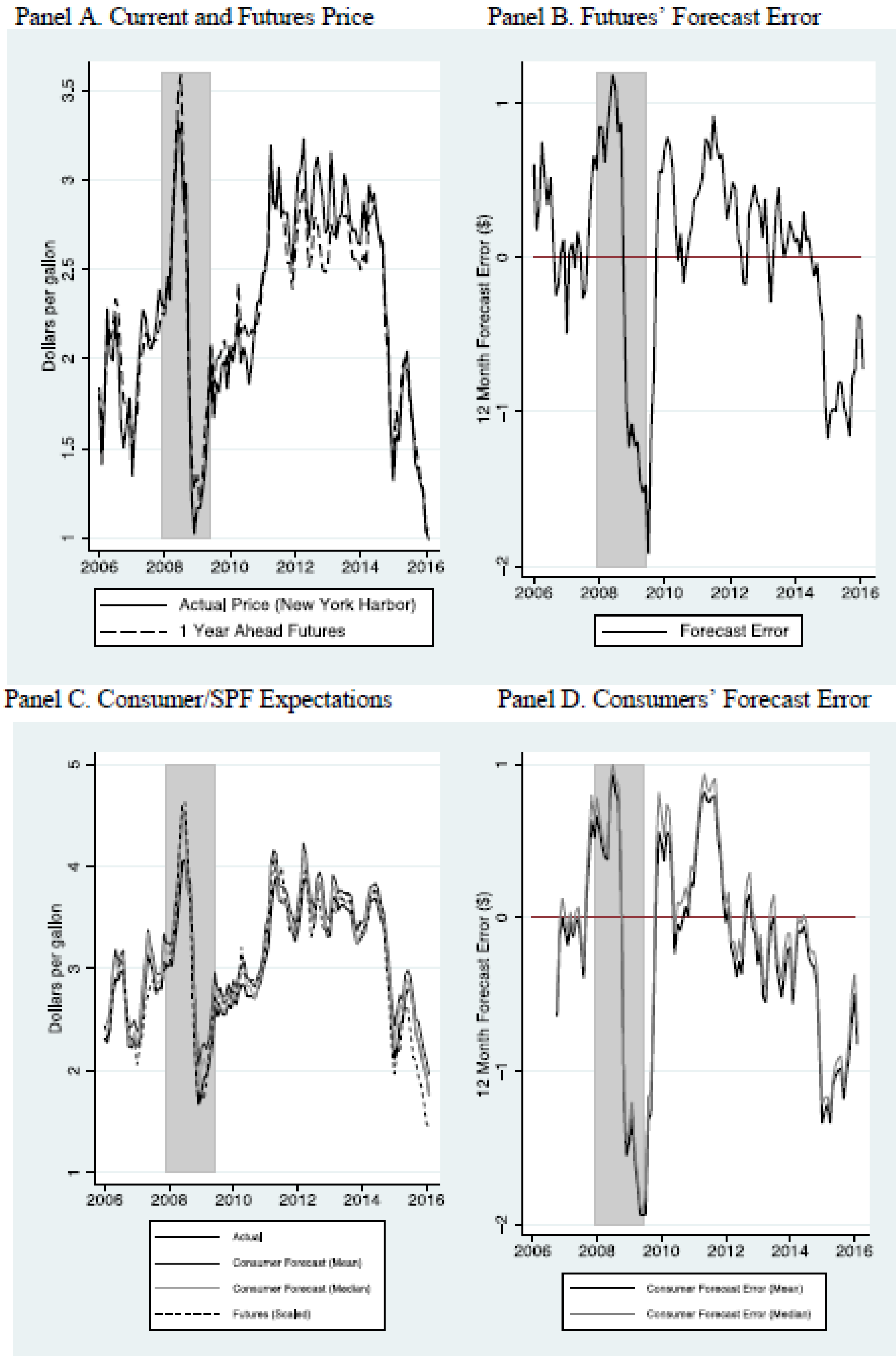


Figure 3.6.2: An example of machine learning decision tree

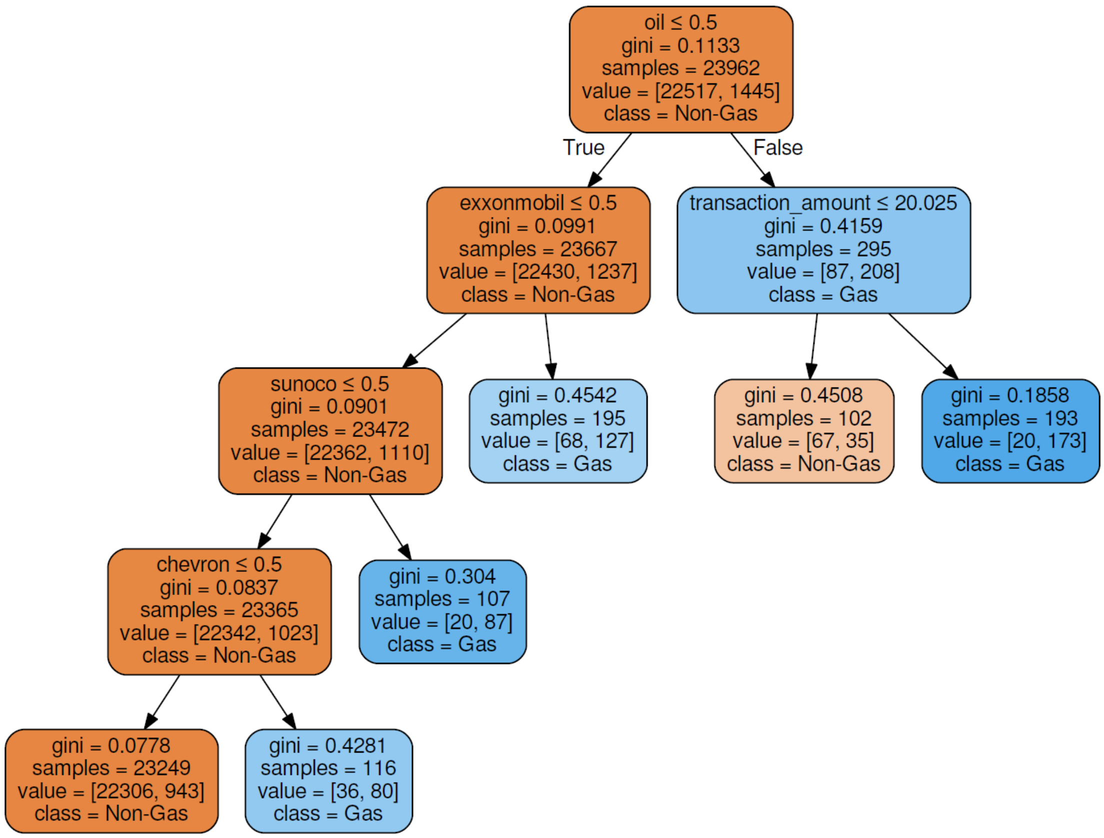


Figure 3.6.3: Distribution of log gasoline spending: CEX Diary vs App

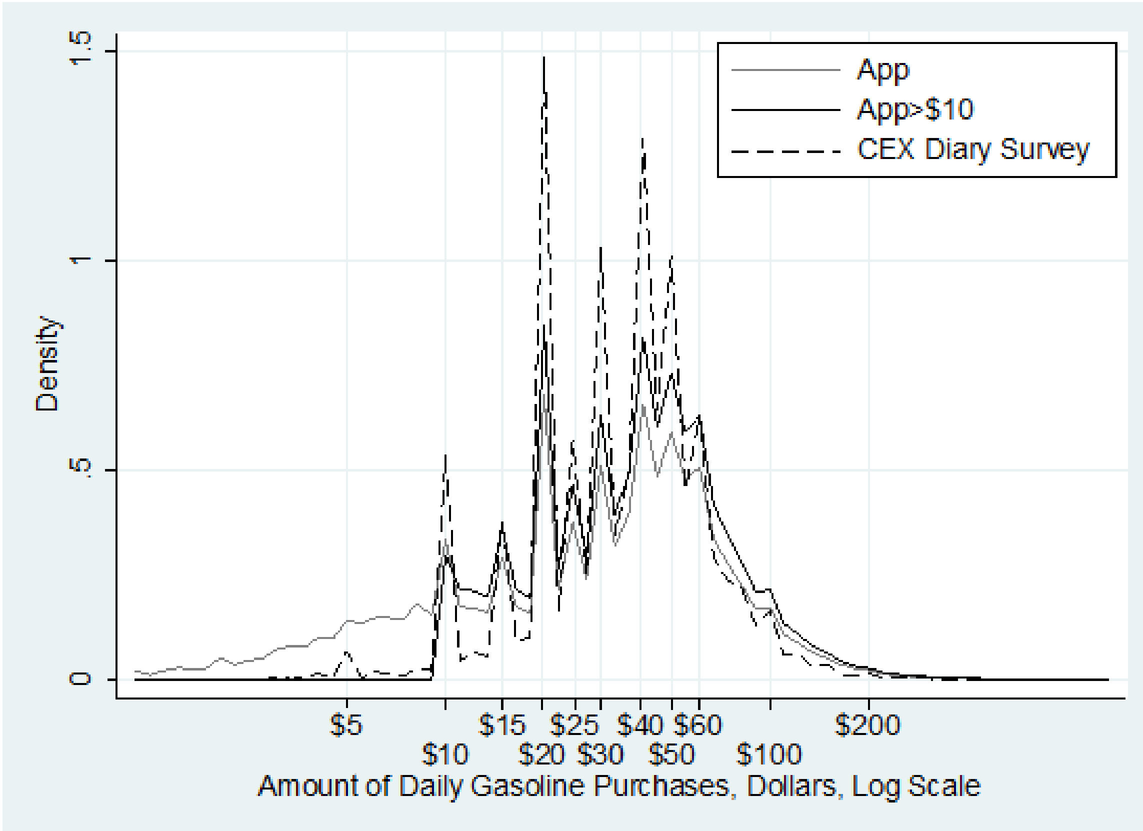


Figure 3.6.4: Reported gasoline spending (monthly)

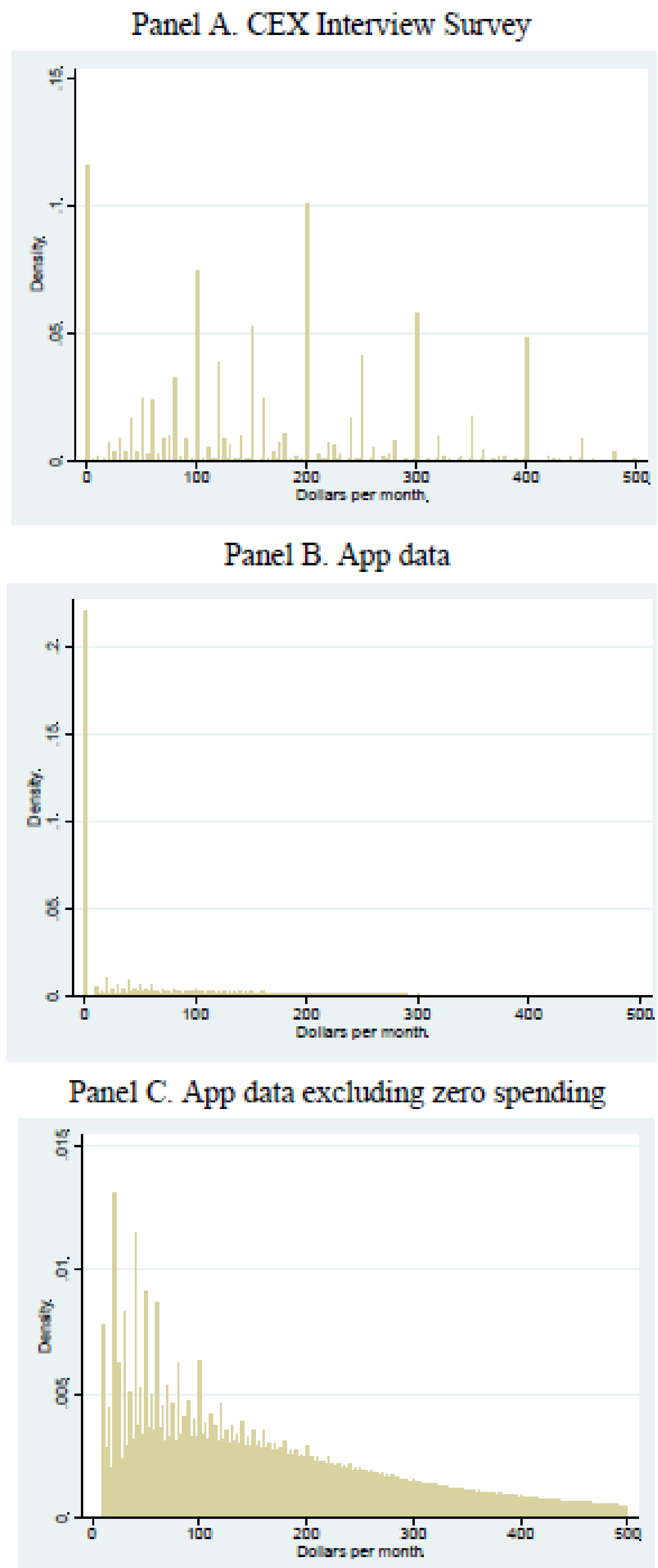


Figure 3.6.5: Dynamic response to a change in gasoline price

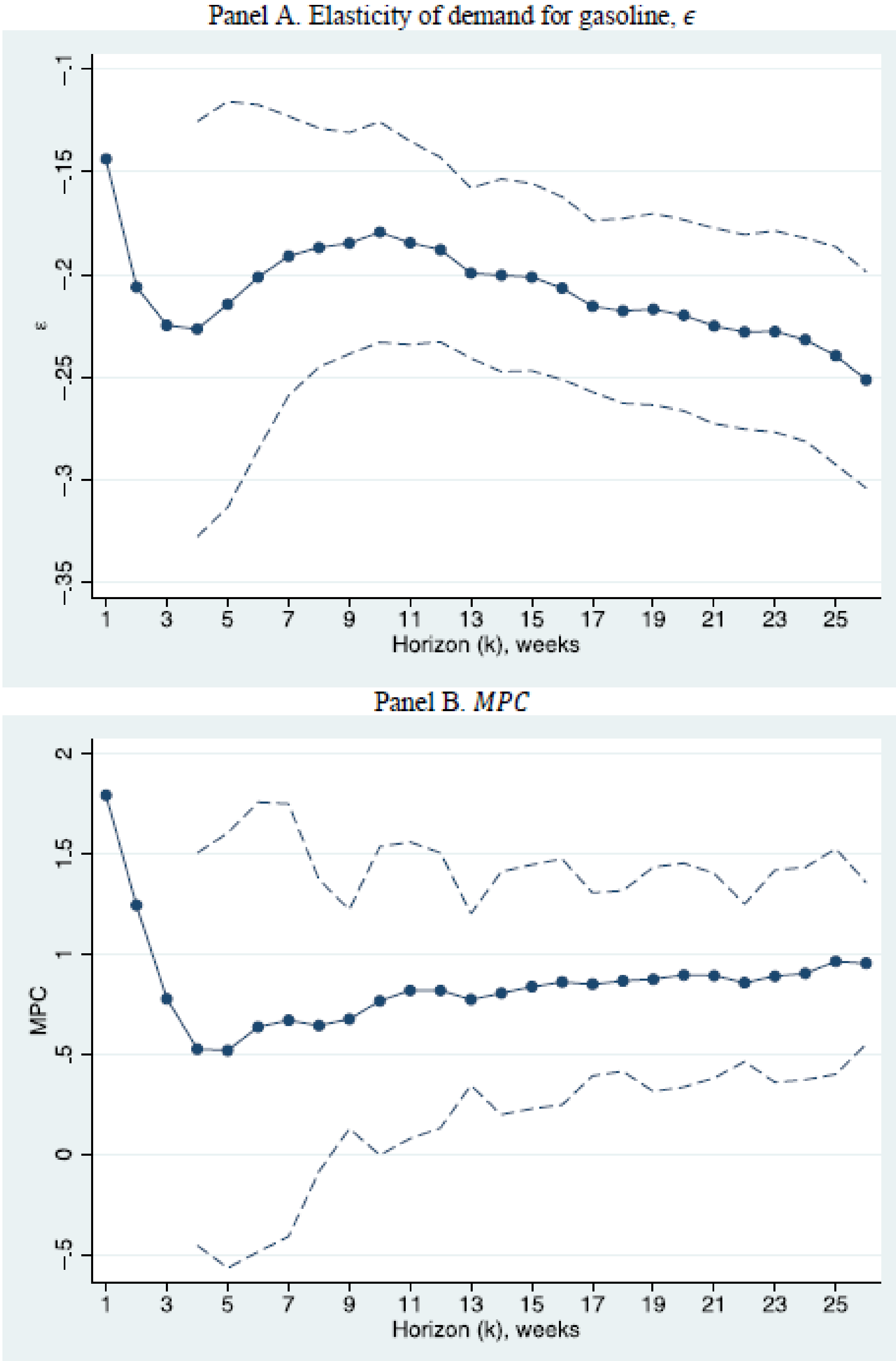


Table 3.1: Largest monthly changes in oil and gasoline prices

Largest Decreases			Largest Increases		
Date	Percent Change		Date	Percent Change	
	Oil	Gas		Oil	Gas
1986:2	-33	-6	1974:1	135	
2008:12	-28	-21	1990:8	47	10
2008:10	-26	-14	1986:8	30	-5
2008:11	-25	-32	1948:1	24	
2014:12	-22	-11	1990:9	23	9
2015:1	-20	-18	2009:3	23	1

Notes: Table shows the month-to-month percent change in West Texas Intermediate spot oil prices (FRED series OILPRICE and MCOILWTICO) and the corresponding change in average monthly regular gasoline prices, when available, from January 1946 – February 2016. For gasoline prices, the table use the BLS U.S. city average (BLS series APU000074714), since it is available further back in time than other available gasoline price data.

Table 3.2: Comparison of spending in the CEX and app data, 2013

Frequency and type of spending	Moment			
	Mean	St. Dev.	Median	Interquartile range
Panel A. Biweekly				
Spending on gas, dollars				
CEX Diary Survey	84.72	84.72	83.42	65.00
App	81.68	203.75	30.99	79.66
Spending on non-gasoline items, dollars				
CEX Diary Survey	1,283.36	1,470.93	790.56	1,380.66
App	1,468.65	3,617.73	557.38	975.32
Ratio of gasoline to non-gasoline spending				
CEX Diary Survey	0.15	0.25	0.06	0.14
App	0.12	0.25	0.04	0.12
Panel B. Quarterly				
Spending on gasoline, dollars				
CEX Interview Survey	646.63	530.87	540.00	630.00
App	627.94	588.24	475.33	660.18
Spending on non-gasoline items, dollars				
CEX Interview Survey	10,143.78	8,141.67	7,728.70	7,406.49
App	11,264.85	11,391.42	8,392.24	8,605.46
Ratio of gasoline to non-gasoline spending				
CEX Interview Survey	0.08	0.07	0.06	0.08
App	0.07	0.07	0.05	0.07

Notes: Means and standard deviation are from the distribution winsorized at the 1% level. The variables from the CEX use population sample weights. For Panel A, the ratio for a consumer/household is calculated as average value of the sum all gasoline spending during a biweekly period in 2013 divided by total non-gasoline spending in the corresponding biweekly period in 2013. For Panel B, the ratio for a consumer/household is calculated as the sum of all gasoline spending in a quarter, divided by total non-gasoline spending in that quarter.

Table 3.3: Estimated elasticity of demand and MPC: Baseline and estimates for single financial providers

Accounts	Sample	Row	Elasticity of demand for gasoline, ϵ			MPC		
			Horizon (weeks)			Horizon (weeks)		
			5	15	25	5	15	25
			(4)	(5)	(6)	(7)	(8)	(9)
Baseline	All	1	-0.215	-0.201	-0.240	0.518	0.837	0.963
			(0.050)	(0.023)	(0.027)	(0.554)	(0.311)	(0.287)
Large Provider #1	Any Account	2	-0.256	-0.265	-0.314	0.125	0.586	0.515
			(0.048)	(0.024)	(0.029)	(0.835)	(0.482)	(0.387)
	Core	3	-0.290	-0.261	-0.293	0.199	0.666	0.661
			(0.053)	(0.025)	(0.027)	(0.928)	(0.526)	(0.427)
Large Provider #2	Any Account	4	-0.242	-0.260	-0.367	0.064	0.321	0.642
			(0.069)	(0.038)	(0.054)	(0.425)	(0.218)	(0.266)
	Core	5	-0.223	-0.237	-0.326	-0.057	0.429	0.616
			(0.065)	(0.034)	(0.051)	(0.487)	(0.306)	(0.321)
Large Provider #3	Any Account	6	-0.183	-0.179	-0.246	0.281	0.561	0.639
			(0.051)	(0.034)	(0.045)	(0.736)	(0.431)	(0.371)
	Core	7	-0.161	-0.170	-0.241	0.204	0.575	0.734
			(0.052)	(0.035)	(0.046)	(0.595)	(0.366)	(0.349)

Notes: The table reports estimates of elasticity of demand for gasoline ϵ and marginal propensity to consume (MPC) based on specifications (4) and (5) for horizons 5, 15, and 25 weeks. Row 1 presents the baseline estimates based on the full sample. In the rest of the table, the sample is restricted to a single provider indicated in the left column. In other words, 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), the table report estimates for the case when we use any account of a provider. In rows (3), (5) and (7), the table report estimates based on "core accounts"; that is, to be part of the estimation sample, a user has to have at least one checking and one credit-card account with a given provider and have at least one transaction per month on each account. In all specifications, robust standard errors are clustered at both the consumer and week level. See text for further details.

Table 3.4: Robustness of MPC estimate

Sample	Row	Elasticity of demand for gasoline, ϵ			*MPC*		
		Horizon (weeks)			Horizon (weeks)		
		5	15	25	5	15	25
		(1)	(2)	(3)	(4)	(5)	(6)
Baseline	1	-0.215	-0.201	-0.240	0.518	0.837	0.963
		(0.050)	(0.023)	(0.027)	(0.554)	(0.311)	(0.287)
Exclude zero gasoline spending in 2013	2	-0.214	-0.201	-0.239	0.541	0.865	1.021
		(0.050)	(0.023)	(0.027)	(0.563)	(0.312)	(0.293)
Exclude ATM withdrawals	3	-0.213	-0.198	-0.235	0.401	0.870	1.062
		(0.050)	(0.023)	(0.027)	(0.754)	(0.443)	(0.391)
Change in one-month-ahead gasoline futures	4	-	-	-	0.204	0.855	0.907
					(0.440)	(0.234)	(0.216)
Average change in the yield curve of gasoline futures	5	-	-	-	0.388	1.033	1.110
					(0.531)	(0.281)	(0.259)

Notes: the table reports estimates of elasticity of demand for gasoline ϵ and marginal propensity to consume (MPC) based on specifications (4) and (5) for horizons 5, 15, and 25 weeks. Row 1 presents the baseline estimates based on the full sample. The estimation sample in row 2 excludes consumers with zero spending on gasoline in 2013. In row 3, we exclude ATM withdrawals and other cash withdrawals in calculation of the growth rate of non-gasoline spending. In rows 4 and 5, we replace actual changes in gasoline prices with changes in futures prices of gasoline in specification (4); specification (5) is estimated as in the baseline, so ϵ is the same as in row 1. Specifically, let F_t^h denote the futures price made at time t for period $t+h$. Then, in lieu of $\Delta_k \log P_t$ in our baseline specifications (4), we instead use $\Delta_k \log \mathcal{F}_t \equiv \log F_t^1 - \log F_{t-k}^1$ for $k \in \{1, \dots, 25\}$ in row 4 and the average change in the yield curves for gasoline prices over longer horizons (two years) $\Delta_k \widetilde{\log} \mathcal{F}_t \equiv \frac{1}{24}\sum_{h=1}^{24}(\log F_t^h - \log F_{t-k}^h)$ in row 5. In all specifications, robust standard errors are clustered at both the consumer and week level. See text for further details.

Table 3.5: Elasticity of demand for gasoline and MPC: Consumer Expenditure Survey (CEX) versus App

Data and Sample	Row	Elasticity of demand for gasoline, ϵ			*MPC*		
		Horizon (quarters)			Horizon (quarters)		
		1	2	3	1	2	3
		(1)	(2)	(3)	(4)	(5)	(6)
Panel A: App data (quarterly)							
Baseline	1	-0.084	-0.126	-0.383	1.000	1.050	2.320
		(0.118)	(0.110)	(0.157)	(0.478)	(0.627)	(0.797)
CEX sample design	2	0.005	-0.111	-0.432	0.732	1.574	4.942
		(0.073)	(0.066)	(0.057)	(0.548)	(0.287)	(1.017)
Panel B: CEX							
1980-2014	3	-0.429	-0.330	-0.352	-0.917	-0.526	-1.929
		(0.041)	(0.027)	(0.027)	(0.659)	(0.638)	(1.265)
1985-1987	4	-0.478	-0.396	-0.449	9.624	5.074	4.116
		(0.168)	(0.145)	(0.090)	(5.959)	(4.057)	(4.031)
1990-1992	5	-0.636	-0.562	-0.512	-6.420	-3.005	-4.917
		(0.136)	(0.142)	(0.106)	(4.775)	(4.516)	(6.450)
2014-2015	6	-0.431	-0.408	-0.447	1.384	2.020	6.895
		(0.084)	(0.043)	(0.074)	(1.407)	(1.568)	(2.306)

Notes: the table reports estimates of elasticity of demand for gasoline ϵ and marginal propensity to consume (MPC) based on specifications (4) and (5) for horizons 1, 2, and 3 quarters. The CEX estimates use the ratio of gasoline spending to non-gasoline spending calculated in the first interview, and exclude this period from estimation. For the baseline estimates in Row 1, we use the same 2013 ratio of gasoline spending to non-gasoline spending as in the baseline estimates, and aggregate the spending and gasoline prices to the quarterly level. In row 2, we replicate the CEX sampling scheme, randomly selecting a start month for a user and keeping only the data for the 12 month period that follows it (if a full 12 months of data follow). We similarly use the non-gasoline consumption calculated in the first quarter, and exclude this period from the estimation. In all specifications, robust standard errors are clustered at both the consumer and week level. See text for further details.

Table 3.6: 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	-0.191	-0.155	-0.179	0.569	0.894	0.660
	(0.057)	(0.025)	(0.026)	(0.749)	(0.438)	(0.356)
HtM	-0.239	-0.240	-0.293	0.547	1.044	0.896
	(0.058)	(0.030)	(0.031)	(0.829)	(0.472)	(0.471)
P-value (Non-HtM=HtM)	0.052	0.000	0.000	0.915	0.883	0.658
Panel B. Average HtM in 2013						
Non-HtM	-0.186	-0.164	-0.199	0.618	0.776	0.783
	(0.054)	(0.024)	(0.027)	(0.619)	(0.311)	(0.299)
HtM	-0.274	-0.317	-0.373	0.742	1.336	1.383
	(0.055)	(0.030)	(0.035)	(0.838)	(0.537)	(0.428)
P-value (Non-HtM=HtM)	0.007	0.000	0.000	0.932	0.073	0.077
Panel C. Modal HtM in 2013						
Non-HtM	-0.191	-0.172	-0.209	0.597	0.899	0.893
	(0.053)	(0.024)	(0.027)	(0.651)	(0.358)	(0.320)
HtM	-0.255	-0.287	-0.340	0.896	0.851	0.965
	(0.053)	(0.028)	(0.032)	(0.546)	(0.270)	(0.291)
P-value (Non-HtM=HtM)	0.012	0.000	0.000	0.579	0.410	0.641
Panel D. Extreme HtM in 2013						
Non-HtM	-0.195	-0.182	-0.221	0.613	0.924	0.933
	(0.053)	(0.024)	(0.028)	(0.639)	(0.355)	(0.319)
HtM	-0.284	-0.318	-0.365	1.029	0.912	1.064
	(0.056)	(0.029)	(0.033)	(0.571)	(0.297)	(0.310)
P-value (Non-HtM=HtM)	0.002	0.000	0.000	0.470	0.472	0.753

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 $\widetilde{D}_{i,2013}$), 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 are clustered by week and consumer and are reported in parentheses. *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.

CHAPTER IV

The Self-Constrained Hand to Mouth

4.1 Introduction

Ever since Hall's seminal work on testing the Life-cycle/permanent-income hypothesis (LC-PIH) (Hall, 1978), many studies have documented the fact that consumption responds to the arrival of predictable income (excess sensitivity). Many of these studies show that the strength of the consumption response varies by some measure of liquidity constraints such as income, liquid wealth, age, or occupation. These empirical results have led researchers to conclude that excess sensitivity is caused by temporary liquidity constraints.

This paper challenges this notion by arguing that individuals who receive regular paychecks are unlikely to be liquidity constrained during the week in which they are paid. This intuition is formalized by specifying a parsimonious buffer stock model of consumption with realistic paycheck dynamics. Model simulations show that in the week the paycheck is received, consumption behavior is unlikely to be affected by liquidity levels and so behavior is driven purely by preferences. By using a novel dataset on high frequency joint expenditure and liquid savings behavior, I show that indeed expenditure behavior on pay weeks is not affected by how much liquidity an individual holds. This simple buffer stock model can explain both patterns in the level of expenditures as well as the joint behavior of expenditure growth and liquidity levels. The main contribution of the paper is to show that the correlation between low average liquidity and excess sensitivity is not necessarily a sign of temporary liquidity constraints. The alternative explanation is that individual preferences determine both excess sensitivity and low average liquidity, thus generating the correlation seen in the data. We can then interpret excess sensitivity not as a failure of the LC-PIH, but as optimal behavior that reflects preferences.

The idea that excess sensitivity is caused by preferences and not temporary liquidity constraints is not new. There are a few papers such as Laibson (1997) and Shapiro (2005) which argue that quasi-hyperbolic discounting can explain the high frequency responses to changes in income. However, this is the first paper to show empirically that indeed individuals aren't liquidity constrained during the week that they receive their paychecks. This paper is is also related to Gelman (2017) which uses the same data set and also attempts to disentangle preferences and constraints. The main difference is that this paper uses high-frequency weekly data and focuses on the response to paychecks while Gelman (2017) focuses more on monthly data and examines the response to receiving a tax refund.

4.2 Data

This section describes the data source, sample filters, variable definitions and descriptive statistics.

4.2.1 Data source

This paper utilizes a novel dataset derived from de-identified transactions and account data, aggregated and normalized at the individual level. The data are captured in the course of business by a personal finance app.[1] More specifically, the app offers financial aggregation and bill-paying services. Users can link almost any

[1] These data have previously been used to study the high-frequency responses of households to shocks such as the government shutdown (Gelman et al., 2015) and anticipated income, stratified by spending, income and liquidity (Gelman et al., 2014). Similar account data from other apps have been used in Baugh, Ben-David and Park (2014), Baker (2015), Kuchler (2015), and Ganong and Noel (2016).

financial account to the app, 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. Prior to analysis, the data are stripped of personally identifying information such as name, address, or account number. The data have scrambled identifiers to allow observations to be linked across time and accounts.

We draw on the entire de-identified population of active users and data derived from their records from December 2012 until July 2016. For a subset of the data, we have made use of demographic information provided to the app by a third party. Table 4.1 compares the age, education, gender, and geographic distributions in the sample that matched with an email address to the distributions in the U.S. Census American Community Survey (ACS), representative of the U.S. population in 2012.

Table 4.1: App user demographics

Education	Not Completed College	Completed College	Completed Graduate School
ACS	66.62	24.02	9.36
App	70.42	23.76	5.83

Ages 25 and over. Sample size - ACS: 2,176,103 App: 28,057

Age	18-20	21-24	25-34	35-44	45-54	55-64	65+
ACS	5.85	7.28	17.44	17.24	18.78	16.00	17.41
App	0.59	5.26	37.85	30.06	15.00	7.76	3.48

Sample size - ACS: 2,436,714 App: 35,417

Gender	Male	Female
ACS	48.56	51.44
App	59.93	40.07

Sample size - ACS: 2,436,714 App: 59,072

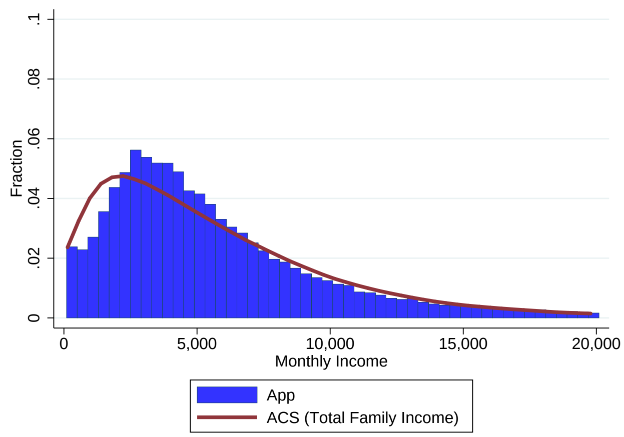
Region	Northeast	Midwest	South	West
ACS	17.77	21.45	37.36	23.43
App	20.61	14.62	36.66	28.11

Sample size - ACS: 2,441,532 App: 63,745

Source: Gelman et al. (2014).

Figure 4.2.1 compares the income distribution in the app to total family income in the ACS. Users who use the app are on average higher income than individuals surveys in the ACS.

Figure 4.2.1: Income comparison



Source: Gelman et al. (2014).

In summary, the app is not perfectly representative of the US population, but it is heterogeneous, including large numbers of users of different ages, education, income, and geographic location.

4.2.2 Defining the sample

The sample is filtered on various characteristics to mitigate measurement error. I filter users based on length of panel, number of accounts, connectedness of accounts, regular paycheck status, and no missing income data.

4.2.2.1 Defining account linkage

If all accounts that are used for receiving income and making expenditures are not observed, we may mistake mismeasurement for excess sensitivity. For example, an individual may have a checking account that is used to pay most bills and a credit card that it used when income is low. If credit card expenditures are not properly observed, it may look like expenditures is lower the week after a paycheck is received relative to the week in which the paycheck is received.

In order to identify linked accounts, I use a method that calculates how many credit card balance payments are also observed in a checking account. I define the variable *linked* as the ratio of the number of credit card balance payments observed in all checking accounts that matches a particular payment that originated from all credit card accounts. For example, a typical individual will pay their credit card bill once a month. If they existed in the data for the whole year, they will have 12 credit card balance payments. If 10 of those credit card payments can be linked to a checking account the variable $linked = \frac{10}{12} \approx 0.83$.

One drawback to this approach is that it requires individuals to have a credit card account. To ensure that those without credit cards are still likely to have linked accounts, I also condition on individuals who have three or more accounts.

4.2.2.2 Identifying regular paychecks

In order to identify regular paychecks, I start by using keywords that are commonly associated with these transactions.[2] I condition on four statistics to ensure that these transactions represent regular paychecks.

1. Number of paychecks ≥ 5
2. Median paycheck amount $> \$200$
3. Median absolute deviation of days between paychecks is ≤ 5
4. Coefficient of variation of the paycheck amount ≤ 1

[2]Keywords used to identify paychecks are "dir dep","dirde p","salary","treas xxx fed","fed sal","payroll","ayroll","payrll","payrl","payrol","pr payment","adp","dfas-cleveland","dfas-in" and DON'T include the keywords "ing direct","refund","direct deposit advance","dir dep adv."

5. Weekly or bi-weekly payroll schedule

For bi-weekly paychecks there are two possible payment schedules. I define these bi-weekly payroll patterns by "odd" or "even." Although this is an arbitrary definition, the main role of this variable is to create two mutually exclusive groups. My definition of week starts on Thursday and week 0 is December 6, 2012. Therefore "even" weeks are the weeks starting Dec 20, 2012, Jan 3, 2012, etc. I define a payroll schedule for a particular individual as "even" if 90% of paychecks are received on an even week. The odd week schedule is defined similarly.

4.2.3 Variable definitions

Most survey data sets such as the consumer expenditure survey (CEX), panel study of income dynamics (PSID), and survey of consumer finances (SCF) are created with the explicit goal of facilitating academic research. The data set used in this study is naturally occurring and was not explicitly designed for use in academic studies. Constructing variables in this data set to match our models is not necessarily a trivial exercise. In order to study the expenditure response to receiving a paycheck, the main variables I utilize are expenditure, paycheck income, and liquid assets.

4.2.3.1 Expenditures

The empirical analysis will focus on food expenditures befcause they are highly nondurable. In particular, I attempt to follow the widely used "strictly non-durable" definition from Lusardi (1996).

The raw data consists of individual transactions with characteristics such as amount, transaction type (debit or credit), and transaction description. While the type of expenditure (food, non-food) is not directly observed, I use a machine learning (ML) algorithm (see Appendix 4.8.1 for more details) to aid in categorization. The goal of the ML algorithm is to provide a mapping from transaction descriptions to expenditure categories. For example, any transaction with the keyword "McDonalds" should map into "Fast Food." A subset of these categories are then combined to create the expenditure variable.

The finest level of categorization is derived from merchant category codes (MCCs) which are directly observable in two of the account providers in the data. MCCs are four digit codes used by credit card companies to classify expenditures and are also recognized by the U.S. Internal Revenue Service for tax reporting purposes. The ML algorithm works by using a subset of the data where the truth is known in order to create a mapping from transaction description to MCCs.

After training the ML algorithm on the data where the truth is known, the algorithm is then applied to the rest of the data set. I then define expenditure as expenditures on fast food and restaurants.

4.2.3.2 Cash on hand and liquid assets

Cash on hand is defined as $X_{it} = A_{it-1} + Y_{it}$ where A_{it-1} represents liquid balances for individual i in the previous period and Y_{it} represents income received in the current period.

Liquid balances (A) are defined as the sum of checking and saving account balances observed in the app. These balances are captured daily as the app takes a snapshot of the balance from each provider.

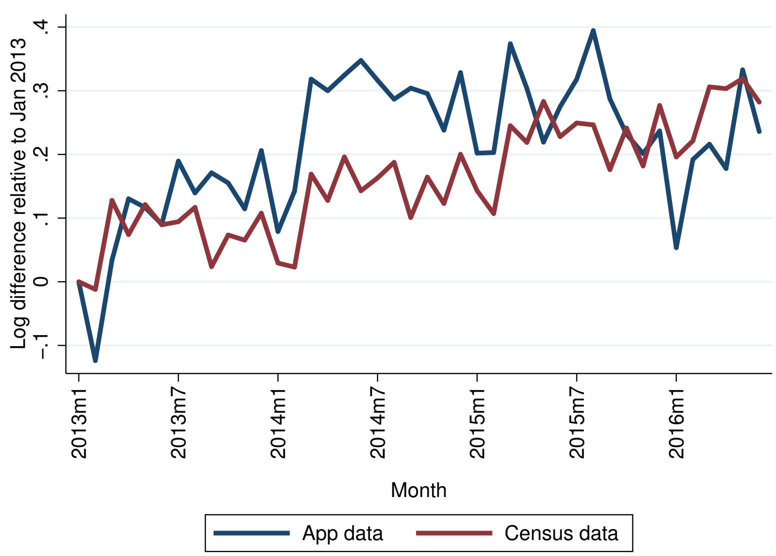
4.3 The expenditure response to paycheck arrival

This section documents the expenditure response to the arrival of a bi-weekly paycheck. By using two different bi-weekly schedules, I show that the expenditure response seen in the data is due to the receipt of a paycheck and not confounded with other events such as first of the month effects.

4.3.1 Time series figures

When analyzing high frequency excess sensitivity, it's important to focus on non-durable expenditures to make sure expenditures line up with consumption as much as possible. As discussed in the previous section, I use fast food and restaurant expenditures to test excess sensitivity of expenditure. Figure 4.3.2 compares this expenditure measure to a comparable expenditures series from the Census Bureau.[3] Because the app data and the Census data are in different units, I plot the log difference relative to Jan 2013 on the y-axis. While the app data is more volatile than the Census data, they both exhibit an upward trend over the time period.

Figure 4.3.2: Monthly food expenditures

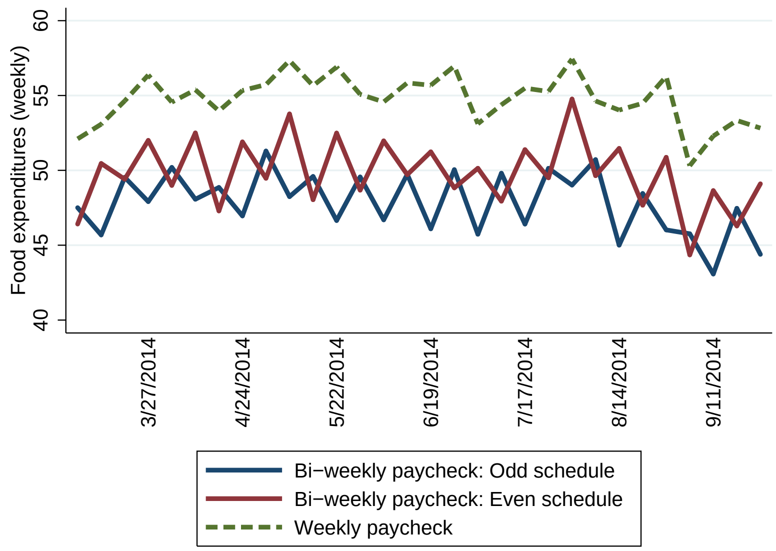


Using the high frequency nature of the data, Figure 4.3.3 plots weekly food expenditures for bi-weekly and weekly paycheck receivers. For bi-weekly paycheck receivers, I further distinguish between "odd" and "even" pay schedules. It's clear from the figure that there is a strong bi-weekly pattern in food expenditures. Furthermore, the opposing bi-weekly pay schedules make it clear that the spikes are associated with

[3] I combine the series "7221: Full service restaurants" and "7222: Limited service eating places" from the U.S. Census Bureau Monthly Retail Trade and Food Services report.

paycheck receipt and not other recurring events like the first of the month. The weekly paycheck series is much smoother but still follows the overall trend seen in the bi-weekly paycheck schedules.

Figure 4.3.3: Weekly food expenditures



4.3.2 Excess sensitivity of food expenditures

The time series for bi-weekly paycheck receivers in Figure 4.3.3 indicate that expenditures rise in sync with weeks in which individuals are paid. The time series for weekly paycheck receivers reveal that expenditures rise in some weeks even for those that receive a paycheck every week. In order to estimate the rise in expenditures from receiving a paycheck while controlling for seasonal expenditure fluctuation, I estimate the following specification.

$$log(Food_{it}) = \alpha_i + \beta_1 Even_t + \beta_2 Payweek_{it}^{Even} + \beta_3 Payweek_{it}^{Odd} + \varepsilon_{it} \qquad (4.1)$$

where $Even_t$ is an indicator variable for whether week t is an even week, $Payweek_{it}^{Even}$ and $Payweek_{it}^{Odd}$ are indicator variables for whether individual i receives bi-weekly paychecks on week t on the even and odd schedule respectively, and α_i represents an individual fixed effect. β_2 and β_3 capture the growth rate of food expenditures on payweeks for those on the bi-weekly even and odd schedule respectively. β_1 captures the growth rate of food expenditures on even weeks. Including the weekly paid individuals helps to control for these seasonal trends that aren't necessarily associated with receiving a paycheck like first of the month effects or holidays that tend to fall on even weeks.

Table 4.2 shows the coefficient estimates from estimating specification (4.1). The estimate of 0.012 on $Even_{it}$ represents the fact that food expenditures grow by 1% on average during even weeks. The coefficients on $Payweek_{it}^{Even}$ and $Payweek_{it}^{Odd}$ are nearly identical and we cannot reject the null hypothesis that the magnitudes are the same. These estimates imply that food expenditures grow by an additional 5.5% on weeks in which bi-weekly individuals are paid after controlling for general seasonal trends. The magnitude of these estimates are in line with Stephens (2003), Shapiro (2005), Stephens (2006), and Kuchler (2015). The granularity of the data allow for much more accurate measurement of receipt of paychecks which results in more precise estimates relative the the previous studies.

Table 4.2: Excess sensitivity estimates

	(1)
VARIABLES	$ln(Food_{it})$
$Even_t$	0.012***
	(0.002)
$Payweek_{it}^{Even}$	0.055***
	(0.003)
$Payweek_{it}^{Odd}$	0.054***
	(0.003)
Observations	3,193,752
R-squared	0.276

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The standard explanation for the excess sensitivity seen in table 4.2 is that individuals are temporary liquidity constrained. Following the literature, table 4.3 re-estimates equation (4.1) for three different terciles of 2013 average liquidity. The estimation only uses data from 2014 and onward to ensure that there is no mechanical correlation with the measure used to split the sample. In line with the previous literature, individuals that have lower levels of liquidity tend to react more strongly to the receipt of a paycheck relative to those with higher levels of liquidity. For example, food expenditures increase by 10% on average during weeks in which a paycheck is received for individual with low average levels of liquidity relative to 2% for individuals with high levels of liquidity. The coefficient on $Even_t$ is fairly similar across liquidity terciles. This is consistent with the view that that $Even_t$ captures aggregate trends that are common to all individuals.

Table 4.3: Excess sensitivity estimates by liquidity tercile

	(1)	(2)	(3)
VARIABLES	Low avg liquidity	Medium avg liquidity	High avg liquidity
$Even_t$	0.009***	0.010***	0.014***
	(0.003)	(0.003)	(0.003)
$Payweek_{it}^{Even}$	0.100***	0.043***	0.021***
	(0.005)	(0.005)	(0.005)
$Payweek_{it}^{Odd}$	0.099***	0.039***	0.017***
	(0.006)	(0.005)	(0.005)
Observations	748,692	754,908	701,221
R-squared	0.292	0.305	0.306

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This section has documented the presence of excess sensitivity of food expenditures to the receipt of a paycheck using financial account data from a personal finance app. The estimates are in line with the previous literature and provide more precise estimates than previous studies. The main goal of this section is to set the stage to further investigate whether the standard explanation that liquidity constraints explain excess sensitivity of expenditure to paychecks is correct. The next section introduces a theoretical model of consumption which will allow us to more formally test the standard explanation.

4.4 Buffer stock model of consumption

This section describes the model used to analyze consumption decisions. Individuals behave according to the standard "buffer-stock" saver model in the spirit of Zeldes (1989), Deaton (1991), and Carroll (1997).

Optimization problem An individual solves the following utility maximization problem

$$\max_{\{C_j\}_{j=t}^{\infty}} \mathbb{E}_t \left[\sum_{j=t}^{\infty} \beta^{j-t} \frac{{C_j}^{1-\theta}}{1-\theta} \right] \tag{4.2}$$

subject to

$$A_{t+1} = (1+r)\left(A_t + Y_t - C_t\right) \tag{4.3}$$

$$A_{t+1} \geq b \tag{4.4}$$

$$Y_t = \bar{Y} + \varepsilon_t \tag{4.5}$$

$$\varepsilon_t \overset{iid}{\sim} N(\mu_y, \sigma_y^2) \tag{4.6}$$

where β, r, C_t, A_t and Y_t represent the time discount factor, the interest rate, consumption, liquid assets, and income respectively. Each period t represents a bi-weekly pay period. Y_t is further broken down into a constant term $\bar{Y}$ which represents a recurring paycheck and a stochastic term ε_t that represents non-paycheck income.

Income process I model the income process to match individuals who receive bi-weekly paychecks. Therefore, individuals receive a paycheck every other period. Overall, paycheck income comprises 70% of total income.

Solution The consumption problem specified above does not admit a closed form solution and is therefore solved computationally. I reformulate the individual's problem in terms of a functional equation and define cash on hand $x_t = a_t + y_t$ to simplify the state space. This variable represents the amount of resources available to the individual in the beginning of the period.

The individual then solves the optimization problem

$$V(x_t) = \max_{a_{t+1}} \{u(c_t) + \beta \mathbb{E}[V(x_{t+1})]\} \tag{4.7}$$

subject to

$$x_{t+1} = (1 + r)(x_t - c_t) + y_{t+1} \tag{4.8}$$

and the previous constraints (4.4), (4.5), and (4.6).

Substituting in for c_t and x_{t+1} results in an equation in terms of x_t, a_{t+1}, and y_{t+1}

$$V(x_t) = \max_{a_{t+1}} \left\{ u\left(x_t - \frac{a_{t+1}}{1+r}\right) + \beta \mathbb{E}[V(a_{t+1} + y_{t+1})] \right\} \tag{4.9}$$

The individual maximizes utility by choosing next period saving (a_{t+1}) conditional on cash on hand (x_t). The model is solved using value function iteration which results in the value function $V(x_t)$ and the policy function $a_{t+1}(x_t)$ which maps the state variable x_t into the optimal control variable a_{t+1}. The consumption function is calculated using constraint (4.4) so that $c_t(x_t) = x_t - \frac{a_{t+1}}{1+r}$.

4.5 Model analysis

The buffer stock model introduced in the previous section can help us understand the cause of the excess sensitivity observed in section 4.3.2. In this section, I test whether the model can generate similar patterns as seen in the data. Furthermore, I explore which parameters are important for explaining the observed data. The parameter values used to calibrate the model are listed in Table 4.4 and represent weekly time periods. The utility function is specified as constant relative risk aversion (CRRA) with $\theta = 1$.

Table 4.4: Parameter values

Parameter	Value	Notes	Description
$u(x)$	$\frac{x^{1-\theta}}{1-\theta}$	CRRA utility	utility function
θ	1	standard	coefficient of relative risk aversion
μ_y	0.30	non-paycheck income share of 30%	
σ_y	0.10	estimated from dataset	S.D. of temporary shocks
$\bar{y}$	1.4	paycheck income share of 70%	
r	0.01 / 52	weekly r on checking/saving	interest rate
b	0	no borrowing condition	borrowing limit

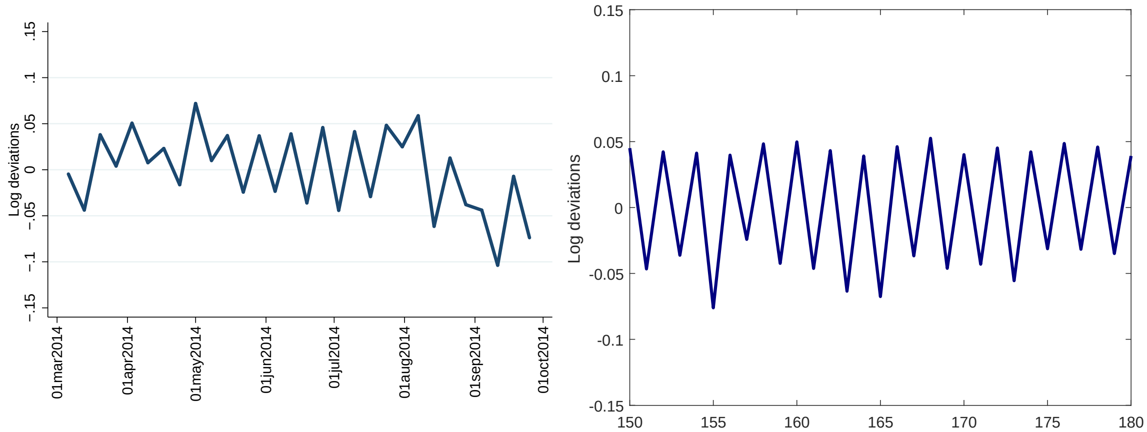
Notes: The parameters correspond to a weekly frequency.

4.5.1 Understanding excess sensitivity

As seen in figure 4.3.3, one important feature of the data when observed at a weekly aggregation is the consistent spike in expenditures during the paycheck week with a subsequent drop in the non-paycheck week. Figure 4.5.4 panel (a) below plots weekly log deviations of food expenditures to their average from March to October of 2014. Panel (b) plots a random subsample of simulated time series in the buffer-stock model. By modeling the receipt of a bi-weekly paycheck, the model can easily explain the spikes in expenditures upon paycheck receipt.

Figure 4.5.4: Weekly time series for bi-weekly paycheck receivers

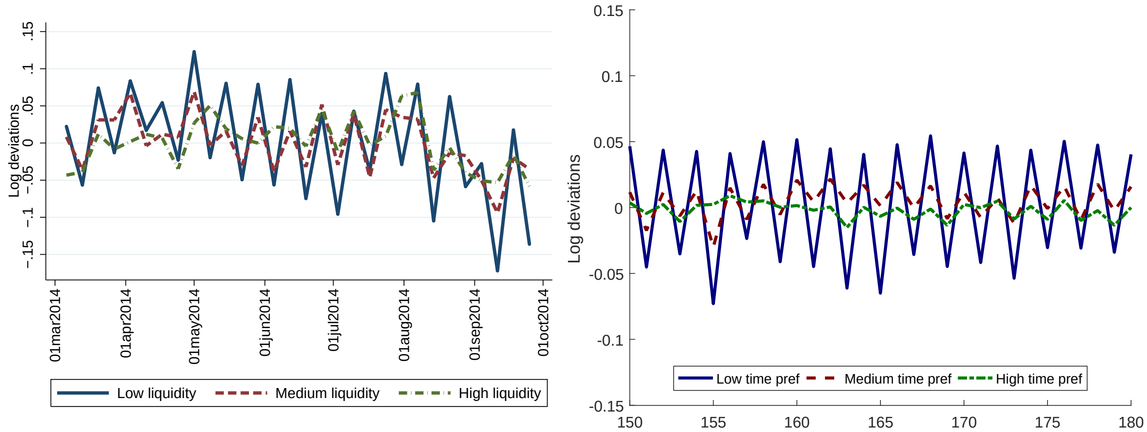
(a) Data (b) Model



The model further allows us to investigate what causes these spikes in expenditures. In this particular model, the time discount factor is the most important parameter that influences the spike in expenditures. This is seen in figure 4.5.5 panel (b) where I simulate the model for different time preference parameters. For patient individuals with high time preference, the time series is relatively smooth. Conversely, impatient individuals with low time preference exhibit much larger spikes. In the data, splitting up individuals into average liquidity terciles as in panel (a) leads to differences in the peaks and troughs of log deviations. Individuals with low average liquidity tend to react more strongly to the receipt of a paycheck relative to individuals with high average liquidity. Most studies see this evidence and conclude that temporary liquidity constraints explains excess sensitivity. However, in the model, temporary liquidity constraints cannot explain excess sensitivity because individuals are rarely constrained during the week in which they receive their paycheck. It is during the week in which they are paid that individuals make the decision on how to allocate expenditures between this week and next week. The week after the paycheck is received is simply a reaction to the decisions made during the paycheck week. The next section will make this more clear by more formally exploring how expenditure growth is determined in the paycheck week and the non-paycheck week.

Figure 4.5.5: Weekly time series for bi-weekly paycheck receivers (heterogeneity)

(a) Data (b) Model



4.5.2 Excess sensitivity and liquidity constraints

The excess sensitivity documented in the previous sections can be interpreted as positive consumption growth in weeks in which a paycheck is not received and negative consumption growth in weeks in which a paycheck is received. In order to understand excess sensitivity, it's important to understand what influences consumption growth. Luckily, the model provides a key equation that can help make this clear. The key equation can be derived from the optimality conditions of the consumption optimization problem specified in section 4.4. The second order approximation of the optimality condition is commonly known as the consumption euler equation and is written below as

$$\underbrace{\Delta ln(c_{t+1})}_{\text{consumption growth}} \approx \overbrace{\frac{r-\delta}{\theta}}^{\text{impatience}} + \underbrace{\frac{\theta}{2}\sigma^2_{t+1}(x_t)}_{\text{precautionary savings}} + \underbrace{\lambda_t(x_t)}_{\text{liquidity constraints}} + \varepsilon_{t+1} \tag{4.10}$$

where c_t is consumption, $\delta = \frac{1}{\beta} - 1$ is the discount rate, θ is the coefficient of relative risk aversion, σ^2_t is a measure of consumption growth volatility, r is the interest rate, and ε_t is a mean zero rational expectations error.

The equation shows that consumption growth is influenced by three terms. The first term is constant and represents desired consumption growth in the absence of any precautionary savings or liquidity constraints. It is driven by the difference between the interest rate and the time discount rate scaled by the intertemporal elasticity of substitution.

The second term represents precautionary savings motives. As explained in Kimball (1990), a positive third derivative of the utility function induces a precautionary savings motive which will tend to cause individuals to save for tomorrow in favor of consuming today. This term will tend to increase consumption growth by lowering consumption today.

Lastly, the third term represents liquidity constraints. If the constraint is binding, this term will also increase consumption growth because individuals cannot increase consumption today relative to their desired amount.

In general, it is difficult to derive analytical expressions for the precautionary savings and liquidity constraint terms. However, we do know that they are functions of cash on hand x_t. Variation in x_t is driven by both uncertainty income as well as predictable changes that arise from different consumption levels in paycheck and non-paycheck weeks. For the liquidity constraint term, there is a value of x_t for which the constraint will begin to bind and so it is a increasing function of x_t. Similarly, the precautionary savings motive is an increasing function of x_t. The intuition is that when x_t is small, an individual is not able to smooth shocks very well leading to a wide range of possible consumption values in the next period depending on the realization of the labor income shock. This translates into high variability in consumption growth. Conversely, when x_t is high, an individual is easily able to smooth consumption in the face of income shocks so there will be little variation in consumption growth. In the limit, as $x_t \to \infty$, liquidity constraints will be unlikely to bind and precautionary fears become irrelevant. In that case, consumption growth will be dominated by the impatience term.

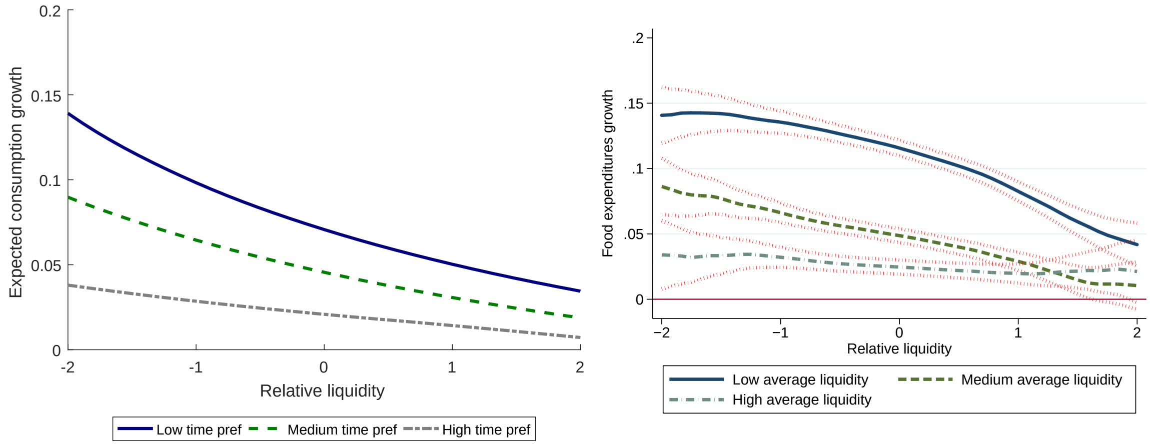
In order to better understand these mechanisms, panel (a) of figure 4.5.6 plots expected consumption growth from the model on the y-axis against relative liquidity for weeks in which the paycheck is not received on the x-axis. Relative liquidity is defined as the log difference of liquidity in time t from it's average. In general, expected consumption growth is positive because consumption tends to be lower in the non-paycheck week relative to the paycheck week. Furthermore, expected consumption growth increases as relative liquidity falls. Because the impatience term is not a function of liquidity, we can interpret the joint movement of consumption growth and liquidity as being driven by precautionary savings and liquidity constraints.

Typically, the theoretical relationship plotted in panel (a) is hard to estimate empirically. There are few datasets where liquidity is observed at such a high frequency jointly with expenditure growth and the timing of paycheck arrival. Utilizing these unique features of the financial app data, panel (b) of figure 4.5.6 estimates the empirical analogue to panel (a) by using realized food expenditures growth. More specifically, panel (b) plots a smoothed local linear relationship between food expenditure growth and log deviations from average liquidity in the week in which the paycheck is not received. This relationship is estimated for each tercile of average liquidity. Similar to the theoretical predictions, food expenditure growth is increasing as relative liquidity falls. During weeks in which individuals do not receive their bi-weekly paycheck, individuals are likely to be very sensitive to changes in liquidity and therefore, will have to lower their food expenditures when liquidity is low. Lastly, individuals with low average liquidity tend to have higher levels of food expenditure growth in non-pay weeks and are more sensitive to changes in relative liquidity. The interpretation under the buffer stock model is that low levels of time preference will jointly produce higher expenditure growth, higher sensitivity to relative liquidity, and low levels of average liquidity.

Figure 4.5.6: consumption/expenditure growth and relative liquidity (non-pay week)

(a) Model

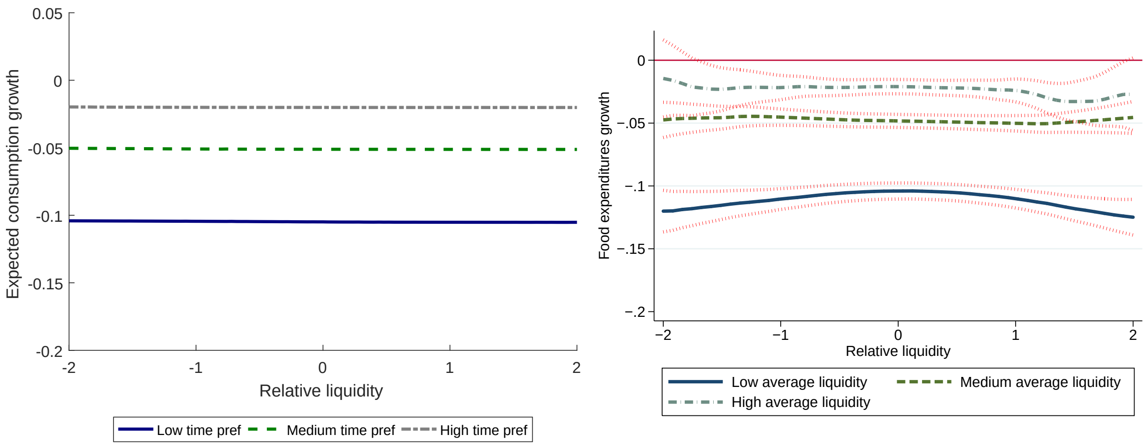
(b) Data



It's intuitive that liquidity constraints play an important role during the week in which inidivdiuals are not paid. However, it's harder to make the case that liquidity constraints are important during pay weeks. Figure 4.5.7 panel (a) confirms this intuition by plotting expected consumption growth against relative liquidity in weeks in which individuals are paid. In stark contrast to non-pay weeks, expected consumption growth is relatively flat. We can interpret this flatness as the absence of the precautionary savings and liquidity constraint terms in the euler equation. In the absence of these terms, equation (4.10) implies that impatience will determine expected expenditure growth. This is reflected in the fact that individuals with low time preference have lower rates of expenditure growth relative to individuals with high time preference. Panel (b) plots the empirical analogue to the theoretically derived relationships. Consistent with the model, food expenditure growth is much less sensitive to liquidity during pay weeks relative to non-pay weeks. Furthermore, individuals with low average liquidity tend to have lower levels of expenditure growth relative to those with high average liquidity.

Figure 4.5.7: consumption/expenditure growth and relative liquidity (pay week)

(a) Model (b) Data



The empirical relationship between food expenditure growth and relative liquidity is summarized in the table below. The table lists the estimated coefficients from the specification

$$\Delta ln(food_{it+1}) = \alpha_i \times payweek_{it} + \beta_2 \times liq^{pay}_{it-1} + \beta_3 \times liq^{nopay}_{it-1} + \varepsilon_{it+1} \qquad (4.11)$$

where $\alpha_i \times payweek_{it}$ represents individual fixed effects for both paycheck and non-paycheck weeks, and liq^{pay}_{it-1} and liq^{nopay}_{it-1} represent $t-1$ log liquidity in the payweek and non-pay week respectively for individual i. I use $t-1$ liquidity because I want to measure the resources individual have when they enter period t. The individual fixed effects for both paycheck and non-paycheck weeks allow us to interpret liquidity as the percent change in the previous week relative to the pay and non-pay week. This relative measure is important because the liquidity levels are different in pay and non-pay weeks. Equation 4.11 is then estimated for each liquidity tercile. Intuitively, the coefficients from the econometric specification estimate the slope of the linear relationship captured in panel (b) of figures 4.5.6 and 4.5.7.

Table 4.5: Relationship between expenditure growth and relative liquidity

VARIABLES	(1) Low avg liquidity	(2) Medium avg liquidity	(3) High avg liquidity
Pay week	0.001	-0.003	-0.002
	(0.002)	(0.003)	(0.004)
Non pay week	-0.037***	-0.026***	-0.009**
	(0.003)	(0.003)	(0.004)
Observations	363,714	416,502	383,654
R-squared	0.056	0.036	0.033

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The results suggest that relative liquidity matters in non-pay weeks but not for pay weeks. This is in line with the model as well as the intuition that individuals are very unlikely to be constrained in the week in which they receive their paycheck and so should not respond much to their liquidity levels at the beginning of the week.

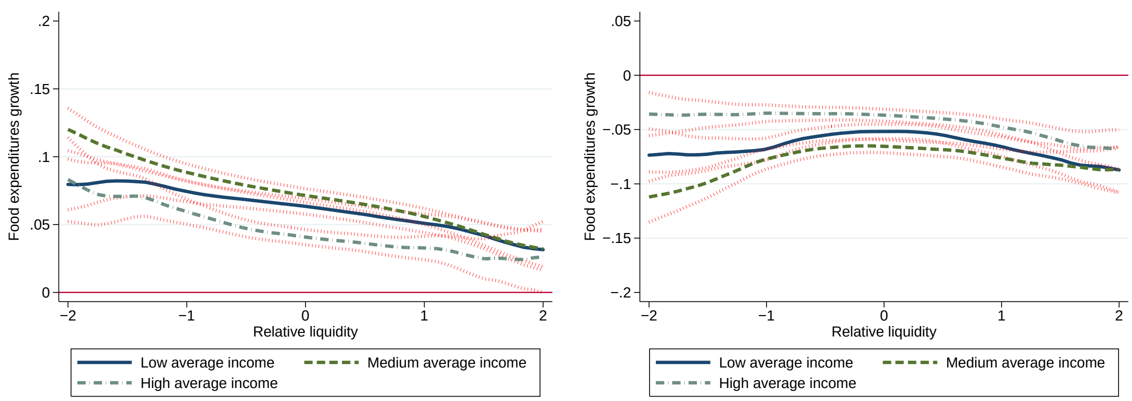
In previous studies, researchers have often observed that average liquidity levels are strong predictors of how individuals respond to paychecks. The analysis in this section makes it clear that we shouldn't interpret these results as evidence that temporary liquidity constraints explain excess sensitivity. Instead, the results are more consistent with a model in which time preferences jointly generate excess sensitivity as well as lower levels of average liquidity. In this simple buffer stock model, excess sensitivity reflects preferences and not constraints.

4.5.3 Excess sensitivity and income

If the explanation in the previous section is true, average liquidity can be thought of as a proxy for preferences. Conversely, paycheck income in the model is exogenous and so does not reflect preferences. To test this assumption, figure 4.5.8 estimates the relationship between food expenditure growth and relative liquidity for different terciles of paycheck income. The results show that paycheck income terciles do not differentiate between different levels of food expenditure growth as well as liquidity terciles. Furthermore, the ordering of the relationships by tercile doesn't generally match the model predictions.

Figure 4.5.8: Expenditure growth and relative liquidity

(a) Non-pay week (b) Pay week



4.6 Testing the implications of liquidity constraints using tax refunds

The main result in the paper is that expenditure growth is relatively unaffected by liquidity in the pay week while it is significantly affected by liquidity during the non-pay week. One way to see this mechanism in action is by looking at the response to receiving a tax refund. More specifically, if individuals are liquidity constrained during the non-pay week, we should observe a stronger reaction to the refund if it is received during a non-pay week relative to a pay week.

4.6.1 Expenditure growth

This section estimates the effect of receiving a tax refund on expenditure growth. It also tests whether the effect is different depending on whether the week in which the tax refund is received is a pay period or a non-pay period. The econometric specification is

$$\Delta ln(food_{it+1}) = \alpha_i + \beta_1 \times refund_{it} + \beta_2 \times payweek_{it} + \beta_3 \times refund_{it} \times payweek_{it} + \varepsilon_{it+1} \tag{4.12}$$

where $refund_{it}$ and $payweek_{it}$ are indicator variables for whether a refund or a paycheck was received for person i in week t and α_i is an individual-level fixed effect.

Table 4.6 shows the results from estimating equation 4.12 for each liquidity tercile. The coefficient on $payweek_{it}$ shows that expenditure growth is negative in weeks in which a paycheck is received. This is in line with the excess sensitivity captured in earlier results. The coefficient on $refund_{it}$ shows that for individuals with low and medium average liquidity, expenditure growth is negative in weeks in which a a tax refund is received. This indicates that individuals increase expenditures when they receive a tax refund. The positive coefficients on $refund_{it} \times payweek_{it}$ show that expenditure growth is less negative when the refund is received during weeks in which the paycheck is also received. This is consistent with the notion that individuals are more liquidity constrained in weeks in which they don't receive a paycheck. More specifically, for individuals with low average liquidity, expenditure growth is 12% lower during weeks in which a refund is received and a paycheck is not received. If the refund is received in the same week that the paycheck is received, expenditure growth is only 3% lower relative to weeks in which the refund is not received.

Table 4.6: Coefficient estimates

VARIABLES	(1) Low avg liquidity	(2) Medium avg liquidity	(3) High avg liquidity
$refund_{it}$	-0.118***	-0.056***	-0.010
	(0.023)	(0.021)	(0.023)
$payweek_{it}$	-0.221***	-0.094***	-0.047***
	(0.004)	(0.003)	(0.004)
$refund_{it} \times payweek_{it}$	0.084***	0.012	0.006
	(0.032)	(0.029)	(0.032)
Observations	375,965	419,802	385,760
R-squared	0.013	0.005	0.004

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

4.6.2 Expenditure growth by relative liquidity

This section takes a closer look at how receiving a tax refund affects the relationship between expenditure growth and relative liquidity. Figure 4.6.9 shows the relationship between expenditure growth and relative liquidity during weeks in which the paycheck is not received. As seen earlier, in weeks in which the paycheck is not received, expenditure growth has a strong negative relationship with relative liquidity. However, on weeks in which the tax refund is received, that strong negative relationship no longer holds. One way to interpret this is that individuals are usually very cash starved during weeks in which they don't receive their paycheck because they choose to consume more during weeks in which they receive their paychecks. Receiving a tax refund relaxes the liquidity constraints that usually bind. Due to the constraints being relaxed, expenditure growth is no longer affected by the amount of liquidity individuals carry over from the previous period.

Figure 4.6.9: Expenditure growth and relative liquidity (non-pay week)

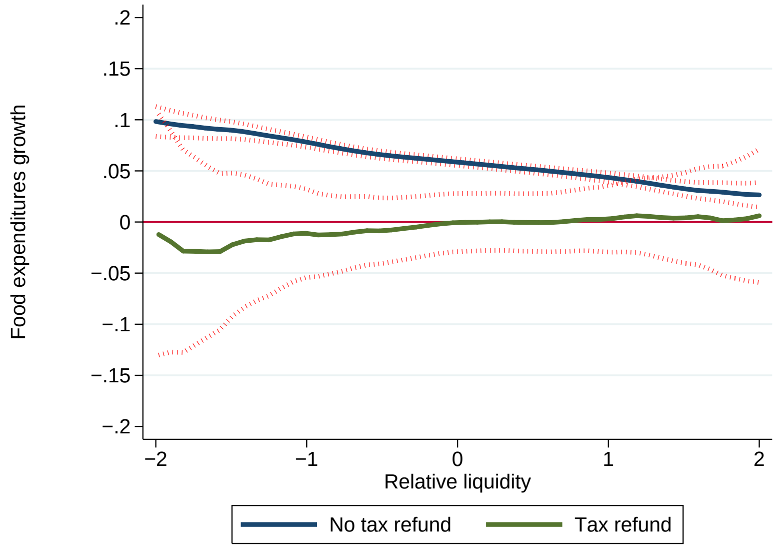
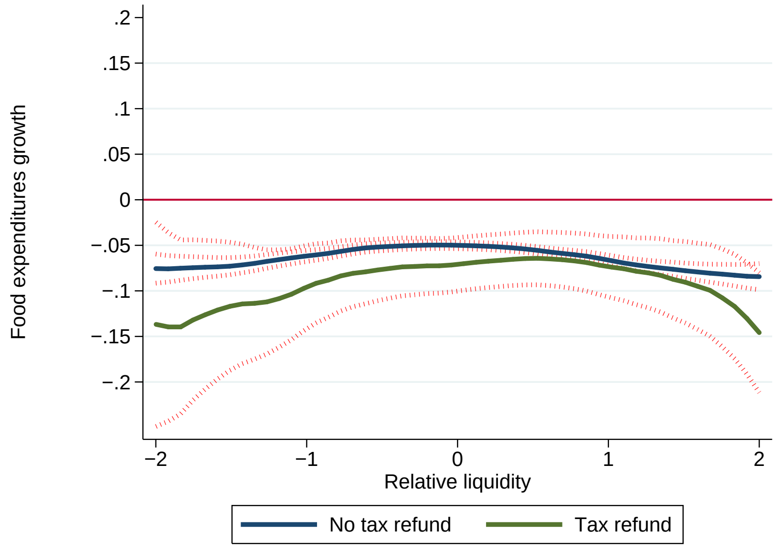


Figure 4.6.10 performs the same analysis but for weeks in which a paycheck is received. As seen in the previous sections, the relationship between expenditure growth and relative liquidity is much weaker during weeks in which the paycheck is received. Furthermore, because an individual typically has so much liquidity during pay weeks, the relationship does not appear to be very different in weeks in which a tax refund is also received.

Figure 4.6.10: Expenditure growth and relative liquidity (pay week)



Because tax refunds are only received once a year, the results conditioning on weeks in which a tax refund is received are much less precise. In order to more formally analyze how receiving a refund affects the relationship between expenditure growth rate and relative liquidity, I estimate the following econometric specification

$$\begin{aligned}\Delta ln(food_{it+1}) = \alpha_i + \alpha_i \times payweek_{it} + \beta_1 \times liq^{pay}_{it-1} + \beta_2 \times liq^{pay}_{it-1} \times refund_{it} + \\ \beta_3 \times liq^{nopay}_{it-1} + \beta_4 \times liq^{nopay}_{it-1} \times refund_{it} + \\ \beta_5 \times refund_{it} + \beta_6 \times refund_{it} \times payweek_{it} + \varepsilon_{it+1} \quad (4.13)\end{aligned}$$

where liq^{pay}_{it-1} and liq^{nopay}_{it-1} capture the log of liquidity in the previous period when the current period is a pay week or non-pay week respectively. The specification aims to capture the differential marginal effect of relative liquidity on expenditure growth in weeks in which a tax refund is received. The negative coefficient on liq^{nopay}_{it-1} replicates the earlier result that relative liquidity is an important determinant of expenditure growth in non-pay weeks. Furthermore, the small and statistically insignificant result on liq^{pay}_{it-1} replicates the earlier result that relative liquidity is *not* an important determinant of expenditure growth in *pay* weeks. The new results of interest are the coefficients on $liq^{pay}_{it-1} \times refund_{it}$ and $liq^{nopay}_{it-1} \times refund_{it}$. They represent the additional effect on liquidity on expenditure growth on weeks in which the tax refund is received. The small and statistically insignificant coefficient on $liq^{pay}_{it-1} \times refund_{it}$ confirms that since liquidity is already high on pay weeks, receiving additional liquidity in the form of a tax refund does not have much of an effect. The positive and statistically significant coefficient on $liq^{nopay}_{it-1} \times refund_{it}$ confirms that since individuals tend to be liquidity constrained during non-pay weeks, receiving extra liquidity cancels out the negative relationship between relative liquidity and expenditure growth during non-pay weeks. To test this idea more formally, I calculate $\beta_3 + \beta_4 = 0.0116$ with a p-value of 0.201. Therefore, the econometric specification confirms the results in figure 4.6.9 that liquidity no longer affects expenditure growth in non-pay weeks after the tax refund relieves liquidity constraints.

Table 4.7: Relationship between expenditure growth and relative liquidity

VARIABLES	(1) $\Delta ln(food_{it+1})$
liq^{pay}_{it-1}	0.000
	(0.002)
$liq^{pay}_{it-1} \times refund_{it}$	0.011
	(0.008)
liq^{nopay}_{it-1}	-0.026***
	(0.002)
$liq^{nopay}_{it-1} \times refund_{it}$	0.037***
	(0.009)
$refund_{it}$	-0.349***
	(0.073)
$refund_{it} \times payweek_{it}$	0.234**
	(0.096)
Observations	1,394,974
R-squared	0.037

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

To summarize, this section tested the implications of the effects of liquidity on expenditure growth during pay and non-pay weeks. The main analysis suggests that liquidity only affects expenditure growth in non-pay weeks because this is when liquidity is low. It tests this implication by studying a case in which liquidity is increased in the form of a tax refund. Similarly to what the theory and empirics suggest, receiving a tax refund has different effects whether it is received on a pay week or non-pay week. In general, expenditure growth is negative in weeks in which a tax refund is received as individuals increase expenditure relative to weeks in which a tax refund is not received. However, the analysis in this section shows that the impact of receiving a tax refund is greater in non-pay weeks. The analysis then proceeds by studying the effect of receiving a tax refund on the relationship between expenditure growth and relative liquidity. The analysis shows that in weeks in which the tax refund is received, liquidity no longer affects expenditure growth in the non-pay week. These results are consistent with the interpretation that individuals are liquidity constrained during the non-pay week. The receipt of the tax refund allows us to test this assumption and confirms that indeed when liquidity constraints are relaxed, relative liquidity no longer affects expenditure growth.

4.7 Conclusion

This paper has re-examined whether excess sensitivity of expenditure to the receipt of a paycheck is caused by temporary liquidity constraints. The main argument against such an interpretation is that individuals who receive paychecks are unlikely to be liquidity constrained in the weeks in which they receive their paychecks. Therefore, their expenditure reaction to a paycheck represents behavior that is driven by preferences rather than constraints. To formalize this intuition, I specify a parsimonious buffer stock model of consumption with realistic paycheck dynamics. Model simulations show that during the week in which a paycheck is received, consumption growth is not affected by changes in liquidity. I then test this assumption in the data and show that indeed liquidity does not affect expenditure growth in the week in which the paycheck is received.

Under the specified model, the spike up in expenditures during the pay week is driven by the fact that some individuals are impatient and prefer to consume more when they have resources. Indeed, in the data, excess sensitivity is strongest for those with low average liquidity. This is consistent with the model as impatient individuals react more strongly to paychecks while at the same time holding less liquidity on average.

In the model, impatient individuals intentionally leave less liquidity for themselves next period thus making them vulnerable to shocks in weeks in which a paycheck is not received. I further test this assumption by showing how an influx of liquidity affects expenditure behavior. In pay weeks, individuals are already awash with liquidity so they should not react much to extra liquidity. Conversely, in non-pay weeks, individuals have left themselves fewer resources and so should react strongly to liquidity. Using the extra liquidity provided by the receipt of a tax refund, I find that expenditure behavior once again matches the predictions of the model.

Both the model and the empirical results imply that excess sensitivity is not caused by temporary liquidity constraints. Instead, excess sensitivity is an optimal outcome for impatient individuals that face high frequency fluctuations in income.

Bibliography

Ameriks, John, Andrew Caplin, and John Leahy (2003) "Wealth Accumulation and the Propensity to Plan," *The Quarterly Journal of Economics*, Vol. 118, No. 3, pp. 1007–1047.

Angeletos, George-Marios, David Laibson, Andrea Repetto, Jeremy Tobacman, and Stephen Weinberg (2001) "The Hyperbolic Consumption Model: Calibration, Simulation, and Empirical Evaluation," *Journal of Economic Perspectives*, Vol. 15, No. 3, pp. 47–68.

Baker, Scott R (2015) "Debt and the consumption response to household income shocks," *Available at SSRN.*

Baugh, Brian, Itzhak Ben-David, and Hoonsuk Park (2014) "Disentangling Financial Constraints, Precautionary Savings, and Myopia: Household Behavior Surrounding Federal Tax Returns,"Technical report, National Bureau of Economic Research.

Campbell, John Y. and N. Gregory Mankiw (1989) "Consumption, Income and Interest Rates: Reinterpreting the Time Series Evidence," NBER Chapters, National Bureau of Economic Research, Inc.

Carroll, Christopher (2004) "Theoretical Foundations of Buffer Stock Saving," Working Paper 10867, National Bureau of Economic Research.

Carroll, Christopher D. (1997) "Buffer-Stock Saving and the Life Cycle/Permanent

Income Hypothesis," *The Quarterly Journal of Economics*, Vol. 112, No. 1, pp. 1–55.

——— (2006) "The method of endogenous gridpoints for solving dynamic stochastic optimization problems," *Economics Letters*, Vol. 91, No. 3, pp. 312–320.

Carroll, Christopher D. and Miles S. Kimball (1996) "On the Concavity of the Consumption Function," *Econometrica*, Vol. 64, No. 4, pp. 981–992.

Carroll, Christopher D., Jiri Slacalek, Kiichi Tokuoka, and Matthew White (2015) "The distribution of wealth and the marginal propensity to consume," Working Paper Series 1655, European Central Bank.

Congressional Budget Office (2009) "Did the 2008 Tax Rebates Stimulate Short-Term Growth?." Economic and Budget Issue Brief.

Council of Economic Advisers (2010) "The Economic Impact of the American Recovery and Reinvestment Act." Third Quarterly Report, Executive Office of the President.

Deaton, Angus (1991) "Saving and Liquidity Constraints," *Econometrica*, Vol. 59, No. 5, pp. 1221–1248.

Ganong, Peter and Pascal Noel (2016) "How Does Unemployment Affect Consumer Spending?."

Gelman, Michael (2017) "What Drives Heterogeneity in the Marginal Propensity to Consume? Temporary Shocks vs Persistent Characteristics."

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*, Vol. 345, No. 6193, pp. 212–215.

——— (2015) “How Individuals Smooth Spending: Evidence from the 2013 Government Shutdown Using Account Data,” Working Paper 21025, National Bureau of Economic Research.

Hall, Robert E (1978) “Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence,” *Journal of Political Economy*, Vol. 86, No. 6, pp. 971–87.

Harris, Christopher J. and David Laibson (2002) *Hyperbolic Discounting and Consumption*, pp. 258–298. Eighth World Congress.

Hurst, Erik (2003) “Grasshoppers, Ants, and Pre-Retirement Wealth: A Test of Permanent Income,” Working Paper 10098, National Bureau of Economic Research.

Jappelli, Tullio and Luigi Pistaferri (2010) “The Consumption Response to Income Changes,” *Annual Review of Economics*, Vol. 2, No. 1, pp. 479–506.

——— (2014) “Fiscal Policy and MPC Heterogeneity,” *American Economic Journal: Macroeconomics*, Vol. 6, No. 4, pp. 107–136.

Kaplan, Greg and Giovanni L Violante (2014) “A model of the consumption response to fiscal stimulus payments,” *Econometrica*, Vol. 82, No. 4, pp. 1199–1239.

Kimball, Miles S. (1990) “Precautionary Saving in the Small and in the Large,” *Econometrica*, Vol. 58, No. 1, pp. 53–73.

Krueger, Dirk, Kurt Mitman, and Fabrizio Perri (2016) “Macroeconomics and Household Heterogeneity,” Working Paper 22319, National Bureau of Economic Research.

Krusell, Per and Jr. Smith, Anthony A. (1998) “Income and Wealth Heterogeneity in the Macroeconomy,” *Journal of Political Economy*, Vol. 106, No. 5, pp. 867–896.

Kuchler, Theresa (2015) “Sticking to your plan: Hyperbolic discounting and credit card debt paydown,” *Available at SSRN 2629158*.

Laibson, David (1997) “Golden Eggs and Hyperbolic Discounting,” *The Quarterly Journal of Economics*, Vol. 112, No. 2, pp. 443–478.

Lusardi, Annamaria (1996) “Permanent Income, Current Income, and Consumption: Evidence From Two Panel Data Sets,” *Journal of Business & Economic Statistics*, Vol. 14, No. 1, pp. 81–90.

Parker, Jonathan A. (2015) “Why Don't Households Smooth Consumption? Evidence from a 25 million dollar experiment,” Technical report, Manuscript, MIT.

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*, Vol. 103, No. 6, pp. 2530–53.

Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay (2011) “Scikit-learn: Machine Learning in Python,” *Journal of Machine Learning Research*, Vol. 12, pp. 2825–2830.

Reis, Ricardo (2006) “Inattentive consumers,” *Journal of Monetary Economics*, Vol. 53, No. 8, pp. 1761–1800.

Sahm, Claudia R., Matthew D. Shapiro, and Joel Slemrod (2012) “Check in the Mail or More in the Paycheck: Does the Effectiveness of Fiscal Stimulus Depend on How It Is Delivered?” *American Economic Journal: Economic Policy*, Vol. 4, No. 3, pp. 216–250.

Shapiro, Jesse M. (2005) “Is there a daily discount rate? Evidence from the food stamp nutrition cycle,” *Journal of Public Economics*, Vol. 89, No. 23, pp. 303–325.

Souleles, Nicholas S. (1999) "The Response of Household Consumption to Income Tax Refunds," *The American Economic Review*, Vol. 89, No. 4, pp. 947–958.

Stephens, Melvin (2003) ""3rd of tha Month": Do Social Security Recipients Smooth Consumption between Checks?" *The American Economic Review*, Vol. 93, No. 1, pp. 406–422.

——— (2006) "Paycheque Receipt and the Timing of Consumption," *The Economic Journal*, Vol. 116, No. 513, pp. 680–701.

Zeldes, Stephen P. (1989) "Optimal Consumption with Stochastic Income: Deviations from Certainty Equivalence," *The Quarterly Journal of Economics*, Vol. 104, No. 2, pp. 275–298.

4.8 Appendix

4.8.1 Machine learning algorithm

Most transactions in the data do not contain direct information on expenditure category types. However, category types can be inferred from existing transaction data. In general, the mapping is not easy to construct. If a transaction is made at "McDonalds," it's easy to surmise that the category is "Fast Food Restaurants." However, it is much harder to identify smaller establishments such as "Bob's store." "Bob's store" may not uniquely identify an establishment in the data and it would take many hours of work to look up exactly what types of goods these smaller establishments sell. Luckily, the merchant category code (MCC) is observed for two account providers in the data. 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. If an individual uses an account provider that provides MCC information "Bob's store" will map into a expenditure category type.

The mapping from transaction data to MCC can be represented as $Y = f(X)$ where Y represents a vector of MCC codes and X represents a vector of transactions data. The data is partitioned into two sets based on whether Y is known or not.[4] The sets are also commonly referred to as training and prediction sets. The strategy is to then estimate the mapping $\hat{f}(\cdot)$ from (Y_1, X_1) and predict $\hat{Y}_0 = \hat{f}(X_0)$.

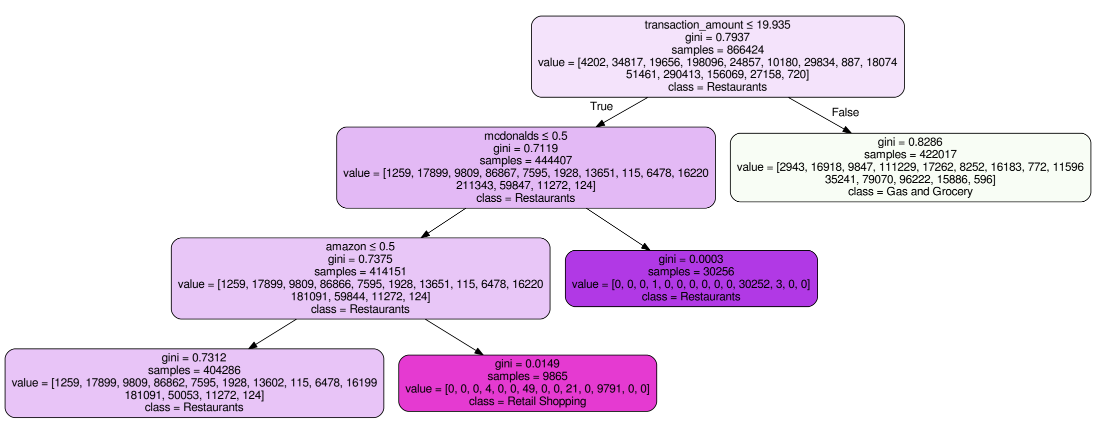
One option for the mapping is to use the multinomial logit model since the dependent variable is a categorical variable with no cardinal meaning. However, this approach is not well suited to textual data because each word would need its own dummy variable. Furthermore, interactions may be important for classifying expenditure categories. For example “jack in the box” refers to a fast food chain while “jack s surf shop” refers to a retail store. Including a dummy for each word can lead to about 300,000 variables. Including interaction terms will cause the number of variables to grow exponentially and will typically be unfeasible to estimate.

In order to handle the textual nature of the data I use a machine learning algorithm called random forest. A random forest model is composed of many decision trees that map transaction data to MCCs. This mapping is created by splitting the sample up into nodes depending on the features of the data. For example, for transactions that have the keyword “McDonalds” and transaction amounts less that \$20, the majority of the transactions are associated with a MCC that represents fast food. To better understand how the decision tree works, Figure 4.8.11 shows an example. The top node represents the state of the data before any splits have been made. The first row “transaction_amount $\leq$ 19.935” represents the splitting criteria of the first node. The second row is the Gini measure which is explained below. The third row show that there are 866,424 total transactions to be classified in the sample. The fourth row “value=[4202,34817,...,27158,720]” shows the number of transactions in each expenditure category. The last row represents the majority class in this node. Because

[4] Y_0 represents the set where Y is not known and Y_1 represents the set where Y is known.

“Restaurants” has the highest number of transactions, assigning a random transaction to this category minimizes the categorization error without knowing any information about the transaction. At each node in the tree, the sample is split based on a feature. For example, the first split will be based on whether the transaction amount is ≤ 19.935. The left node represents all the transactions for which the statement is true and vice versa. Transactions ≤ 19.935 are more likely to be “Restraunts” expenditure while transactions > 19.934 are more likely to be “Gas and Grocery.” In our example, the sample is split further to the left of the tree. Transactions with the string “mcdonalds” are virtually guaranteed to be “Restaurant” expenditure. A further split shows that the string “amazon” is almost perfectly correlated with the category “Retail Shopping.” How does the algorithm decide which features to split the sample on? The basic intuition is that the algorithm should split the sample based on features that lead to the largest disparities in the different groups. For example, transactions that have the word “mcdonalds” will tend to split the sample into fast food and non-fast food transactions so it is a good feature to split on. Conversely, “bob” is not a very good feature to split on because it can represent a multitude of different types of expenditure depending on what the other features are.

Figure 4.8.11: Decision tree example



I state the procedure more formally by adapting the notation used in (Pedregosa et al., 2011). Define the possible features as vectors $X_i \in R^n$ and the expenditure categories as vector $y \in R^l$. Let the data at node m be presented by Q. For each candidate split $\theta = (j, t_m)$ consisting of a feature j and threshold t_m, partition the data into $Q_{left}(\theta)$ and $Q_{right}(\theta)$ subsets so that

$$Q_{left}(\theta) = (X, y)|x_j \leq t_m \tag{4.14}$$

$$Q_{right}(\theta) = Q \setminus Q_{left}(\theta) \tag{4.15}$$

The goal is then to split the data at each node in the starkest way possible. A popular quantitative measure of this idea is called the Gini criteria and is represented by

$$H(X_m) = \sum_k p_{mk}(1 - p_{mk}) \tag{4.16}$$

where $p_{mk} = 1/N_m \sum_{x_i \in R_m} \mathbb{I}(y_i = k)$ represents the proportion of category k observations in node m.

If there are only two categories, the function is is minimized at 0 when the transactions are perfectly split into the two categories[5] and maximized when the transactions are evenly split between the two categories.[6]

Therefore, the algorithm should choose the feature to split on that minimizes the Gini measure at node m

$$\theta^* = argmin_\theta \frac{n_{left}}{N_m} H(Q_{left}(\theta)) + \frac{n_{right}}{N_m} H(Q_{right}(\theta)) \tag{4.17}$$

The algorithm acts recursively so the same procedure is performed on $Q_{left}(\theta^*)$ and $Q_{right}(\theta^*)$ until a user-provided stopping criteria is reached. The final outcome is a decision rule $\hat{f}(\cdot)$ that maps features in the transaction data to expenditure

[5]because 0*1 + 1*0 = 0.
[6]because 0.5*0.5 + 0.5*0.5 = 0.5.

categories.

This example shows that decision trees are much more effective in mapping high dimensional data that includes text to expenditure categories. However, fitting just one tree might lead to over-fitting. Therefore, a random forest fits many trees by bootstrapping the samples of the original data and also randomly selecting the features used in the decision tree. With the proliferation of processing power, each tree can be fit in parallel and the final decision rule is based on all the decision trees. The most common rule is take the majority decision of all the trees that are fit.