Harnessing Naturally-Occurring Data to Measure the Response of Spending to Income

Michael Gelman,[1] Shachar Kariv,[2] Matthew D. Shapiro,[1,3*]
Dan Silverman,[3,4] Steven Tadelis[3,5]

[1]Department of Economics, University of Michigan, Ann Arbor, MI, 48109, USA
[2]Department of Economics, University of California, Berkeley, Berkeley, CA, 94720, USA
[3]National Bureau of Economic Research (NBER), Cambridge, MA, 02138, USA
[4]Department of Economics, Arizona State University, Tempe, AZ, 85287, USA
[5]Haas School of Business, University of California, Berkeley, Berkeley, CA, 94720, USA

[*]To whom correspondence should be addressed; E-mail: shapiro@umich.edu.

This paper presents a new data infrastructure for measuring economic activity. The infrastructure records transactions and account balances yielding measurements with scope and accuracy that have little precedent in economics. The data are drawn from a diverse population that overrepresents males and younger adults, but contains large numbers of underrepresented groups. The data infrastructure permits evaluation of a benchmark theory in economics that predicts that individuals should use a combination of cash management, saving, and borrowing to make the timing of income irrelevant for the timing of spending. As with previous studies and in contrast with the predictions of the theory, there is a response of spending to the arrival of anticipated income. The data also show, however, that this apparent excess sensitivity of spending results largely from the coincident timing of regular income and regular

spending. The remaining excess sensitivity is concentrated among individuals with less liquidity.

Introduction

Economic researchers and policymakers have long sought high quality measures of individual income, spending, and assets from large and heterogeneous samples. For example, when policymakers consider whether and how to stimulate the economy they need to know how individuals will react to changes in their income. Will individuals spend differently? Will they save at a different rate, or reduce their debt, and when? There are many obstacles to obtaining reliable answers to these important questions. One obstacle is that existing data sources on individual income and spending have substantial limits in terms of accuracy, scope, and frequency.

This paper advances the measurement of income and spending with new high-frequency data derived from the actual transactions and account balances of individuals. It uses these measures to evaluate the predictions of a benchmark economic theory that states that the timing of anticipated income should not matter for spending. Like previous research, it finds that there is a response of spending to the arrival of anticipated income. The data show that, on average, an individual's total spending rises substantially above average daily spending the day a paycheck or Social Security check arrives, and remains high for at least the next four days. The data also allow construction of variables that show, however, that this apparent excess sensitivity of spending results in large part from the coincident timing of regular income and regular spending. The remaining excess sensitivity is concentrated among individuals who are likely to be liquidity constrained.

Traditionally, researchers have used surveys such as the Consumer Expenditure Survey (CEX) to measure individual economic activity. Such surveys are expensive to implement, require considerable effort from participants, and are therefore fielded infrequently with modestsized samples. Researchers have recently turned to administrative records, which are accurate and can be frequently refreshed, to augment survey research. So far, however, the administrative records have typically represented just a slice of economic activities. They have not provided simultaneous information about various sources of income and forms of spending.

The data described here result from transactions that are captured in the course of business by Check (www.check.me), a financial aggregation and service application. The resulting income data are accurate and comprehensive in that they capture income from several sources, and can be linked to similarly accurate and comprehensive information on spending. These raw data present important technical and conceptual challenges. The paper describes protocols necessary for turning them into a dataset with several features useful for research and policy analysis.

Data Description

Description of Users in Data Set

Check had approximately 1.5 million active users in the U.S. in 2012. They can link almost any financial account to the app, including bank accounts, credit cards, utility bills, and more. The application logs into the web portals for these accounts daily and obtains the user's primary financial data. The data are organized so users can obtain a comprehensive view of their financial situation.

The data we analyze are derived from a sample of approximately 75,000 Check users, selected at random from the pool of U.S.-based users who had at least one bank or credit card account, and covers 300 consecutive days spanning 2012 and 2013. The data are de-identified and the analysis is perfomed on normalized and aggregated user-level data as described in the text and Supplementary Online Material (SOM). Check does not collect demographic information directly and instead uses a third party that gathers both publicly and privately provided demographic data, anonymizes them, and matches them back to the data. Table 1 compares the gender, age, education, and geographic distributions in the Check 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 1: Check vs. ACS Demographics

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

Note: The sample size for Check is 59,072, 35,417, 28,057, and 63,745 for gender, age, education, and region respectively. The sample size for ACS is 2,441,532 for gender, age, region and 2,158,014 for education.

Table 1 shows that the data overrepresent males and those age 25-44. Education levels are broadly similar to those of the U.S. population and the geographic distribution of Check users is reasonably consistent with that of the U.S. population. Overall, the sample contains large numbers of even the most underrepresented sociodemographic groups. For example, the sample contains about 3,000 individuals age 65 and older. At a point in time, the CEX contains information on approximately 1,100 individuals age 65 and older.[1]

Measuring Income and Spending

Summary statistics for the raw data are provided in Tables S1 and S2 of the SOM. The data allow us to calculate total income and to identify separately paychecks and Social Security payments using the description fields of transactions. Similarly, we measure total spending and sub-categories of spending. We identify recurring and non-recurring income and spending by looking for transactions that occur at regular periodicity and have regular amounts.

Income

We derive two measures of income: The first sums all transactions that represent credits to a user's non-credit-card accounts, excluding transfers from one account to another. The second isolates only those transactions that credit paychecks and Social Security payments, using a list of keywords commonly found in the description field. Fig. 1 shows the distribution of average monthly income measures at the user level.

Total monthly income depicted in Fig. 1(a) has a median of $4,800, and a mean of $8,923. The long and heavy right tail reflects income inequality, and also includes large one-off transactions from asset sales. Paycheck and Social Security income, shown in Fig. 1(b), is less skewed with a median of $2,900 and a mean of $3,951. The figure also displays a kernel density estimate of the distribution of monthly incomes reported in the U.S. Census Bureau's ACS. The income concepts in the ACS and Check data have important differences. Fig. 1(a) shows the distribution of ACS monthly pre-tax household income. The Check data shown in Fig. 1(a) are net of any (tax) withholding and may be aggregated from either individual or household income.

[1] We note, however, the willingness to provide login credentials may select on personal characteristics or increased need for financial organization. The extent of this selection could be assessed with surveys of Check users, the results of which could be compared to existing surveys of representative populations. Alternatively, random samples of the population could be encouraged to link their financial accounts to the app, and the transaction and balances of this population could be compared with that of Check users.

Despite these differences, the ACS and Check distributions are qualitatively similar. Fig. 1(b) shows the ACS distribution of wages, salaries and the sum of wages and salaries and Social Security payments, which are more closely aligned with their analogue in the Check data. The shape of the ACS distribution is again similar to Check's.

Spending

For credit card accounts, we identify spending as transactions that post debits to the account. Non-credit card accounts are similar, but a sum of their debits will overstate spending because some may represent credit card payments or transfers between accounts. Consequently, spending measures exclude debits we can identify as such payments or transfers either by amount or by transaction description.

We consider three measures of spending: (1) total spending, calculated using the method just described, (2) non-recurring spending, and (3) spending on fast food and coffee shops. See Fig. S1 in the SOM for the distribution of these average weekly spending measures at the user level. Non-recurring spending subtracts from total spending both ATM cash withdrawals and expenditures of at least $30 that recur, in the exact same amount (to the cent), at regular frequencies such as weekly or monthly. It isolates spending that, due to its irregularity, is not easily timed to match the arrival of income. This measure thus uses the amount and timing of spending rather than an a priori categorization based on goods and services, an approach made possible by the distinctive features of the data infrastructure. The fast food and coffee shop measure is identified using keywords from the transaction descriptions. This measure isolates an especially discretionary, non-durable, and highly divisible form of spending, which we use in the analysis of the spending response to anticipated income.

The Spending Response to Anticipated Income

Quantifying Excess Sensitivity

A benchmark theory indicates that the anticipated arrival of a payment should not affect the timing of spending. Specifically, spending should not rise after the arrival of a regular paycheck or Social Security payment. We estimate the excess sensitivity of total, non-recurring, and coffee shop and fast food spending to the arrival of regular paychecks or Social Security payments. We thus evaluate the possibility that the benchmark theory describes behavior well, and that excess sensitivity reflects either the convenience of coordinating recurring expenses with the arrival of regular income, or the intrinsic difficulty of smoothing some forms of spending. We also estimate the excess sensitivity of spending separately for users with different levels of liquidity and different levels of available credit. We thus evaluate the possibility that, as standard enhancements to the benchmark theory indicate, excess sensitivity is a phenomenon of those with inadequate liquidity or credit.

We restrict attention to approximately 23,000 users observed to receive paychecks or Social Security payments at a regular frequency and in regular amounts.[2] Our main econometric specification is:

$$x_{ict} = \sum_{j=Mon.}^{Sun.} \delta_{jc} + \sum_{k=-7}^{6} \beta_{kc} I_i \left(Paid_{t-k}\right) + \varepsilon_{ict}, \tag{1}$$

where x_{ict} is the ratio of spending of individual i to i's average daily spending in category c, at date t, δ_{jc} is a day-of-week fixed-effect, and $I_i \left(Paid_{t-k}\right)$ is an indicator equal to one if i received a payment at time $t-k$, and equal to zero otherwise. The β_{kc} coefficients thus measure the fraction by which individual spending in category c deviates from average daily spending in

[2] A payment is classified as regular in frequency if the median number of days between its arrival is from 13 to 15 or from 26 to 34 and if its coefficient of variation is less than 0.5. The demographic characteristics of users who receive either regular paychecks or regular Social Security payments is remarkably similar to those of the entire sample, as are the distributions of their income, spending, and balances.

the days surrounding the arrival of a payment. The day-of-week dummies capture within week patterns of both income and spending.

Fig. 2 shows estimates of β_{kc} for the following categories of spending: (a) total, (b) non-recurring, and (c) coffee shop and fast food spending. The dashed lines are the bounds of the 95% confidence intervals of these estimates based on heteroskedasticity-robust standard errors with clustering at the individual level. Fig. 2(a) shows that, on average, a user's total spending rises about 70% above its daily average the day a regular paycheck or Social Security payment arrives, and remains high for at least the next 4 days.

Total spending includes, however, expenditures such as rent, cable bills, or tuition that are recurring and predictable and whose timing can be adjusted to match the arrival of regular income. Fig. 2(b) shows the excess sensitivity of only non-recurring spending, confirming that a substantial part (40%) of the excess sensitivity of total spending can be attributed to the convenience of paying major bills automatically and avoiding the bad consequences of temporary illiquidity. Given that we defined recurring spending conservatively (i.e., required that it be the same amount to the cent), this estimate is likely a lower bound on how much accounting for it reduces excess sensitivity.

Fig. 2(c) provides still more evidence that the benchmark theory is a better description of behavior than the total spending estimates would suggest. For this imminently divisible and easily smoothed discretionary spending, we observe very modest excess sensitivity to the arrival of predictable income.

Heterogeneity of Excess Sensitivity

We find evidence of individual heterogeneity of excess sensitivity that is consistent with the theory that predicts such behavior among those with insufficient liquidity or available credit, perhaps due to imperfections in credit markets. Fig. 3 plots estimates of β_{kc} for non-recurring spending by terciles of liquidity. We define liquidity for each user as the average daily balance of checking and savings accounts over the entire sample period, normalized by the user's average daily spending. The average user in the lowest tercile has 5 days of spending in cash on hand; the average user in the highest tercile has 159 days. The estimates show that excess sensitivity is significantly more pronounced among those in the lowest tercile of the liquidity distribution.

Fig. S2 in the SOM plots estimates of excess sensitivity by terciles of the available credit utilization distribution. Excess sensitivity is concentrated among users near the limit of their ability to borrow with credit cards. Those who have little liquidity or take their debt levels very close to their limits may be poor at planning or optimizing. The evidence indicates that differences in liquidity and constraints drive heterogeneity of excess sensitivity among check users.

Comparison with Prior Studies

Many prior studies of spending responses to income have used the CEX quarterly retrospective survey, which records self-reports of income, but does not measure its timing precisely. Souleles, for example, uses it to estimate the spending response to the arrival of income tax refunds and overcomes the lack of timing information by calculating from aggregate statistics the likelihood of receiving a refund at various dates (*1*). Parker takes a similar approach and exploits anticipated changes in take-home pay when workers hit the annual cap on the Social Security payroll tax (*2*). Johnson et al., and Parker et al. measure the timing of some income more precisely by adding special questions to the CEX about tax rebates (*3, 4*).

Some studies use higher frequency data to estimate spending responses to income. The CEX diary survey records spending daily for two weeks, but does not collect high frequency income data. Stephens overcomes this limitation by studying the spending response to the receipt of Social Security benefits, which used to arrive on the same day of each month (*5*). The UK's

Family Expenditure Survey collects the most recent paystub of respondents and asks them to track spending for two weeks. Stephens uses the paystub to infer the amount and timing of paychecks and estimates the spending response to them (*6*).

These prior studies use a variety of methods, but share an interest in estimating either an elasticity, defined as $\frac{\partial \log(spending)}{\partial \log(income)}$, or a marginal propensity to consume (MPC), defined as $\frac{\partial(spending)}{\partial(income)}$. Table S3 in the SOM summarizes the key features of these prior estimates and compares them to analogous aspects of our study.

The studies differ in the timeframe over which they measure spending changes in response to a change in income. This makes the levels of their estimated elasticities or MPCs difficult to compare. For our study, we present the point estimate of effects on the first day after the income arrives; that is β_{1c} from equation (1) for the elasticity of spending in category c. For the MPC we present the γ_{1c} from the equation

$$x_{ict} = \alpha_{ic} + \sum_{j=Mon.}^{Sun.} \delta_{jc} + \sum_{k=-7}^{6} \gamma_{kc} Payment_{ic,t-k} + \varepsilon_{ict}, \quad (2)$$

where x_{ict} is the ratio of spending of individual i to by i's average daily spending in category c, at date t, δ_{jc} is a day-of-week fixed-effect, α_{ic} is a user fixed effect, and $Payment_{ic,t-k}$ is the ratio of the amount of the payment received by individual i divided by i's average daily spending in category c, at date $t-k$. Analogously, Table S3 presents only the shortest-run effects reported in all the other studies. While our and other studies estimate larger impacts at longer horizons, the central conclusion of Table S3 about the relative precision of the estimates is not affected by choice of horizon.

The prior estimates are important and influential but, as Table S3 shows, they often lack precision. Among studies of the quarterly CEX data, Hsieh is unusual in its precision (*7*). The last 4 rows of Table S3 include the confidence intervals for our estimates of both the elasticity and the MPC. These intervals are small, both economically and relative to other studies. Only

Broda and Parker provides estimates that are as precise as those from the Check data (*8*). That paper uses Homescan data and estimates precisely a MPC out of tax rebates near zero. These estimates rely on surveys to determine receipt of the rebate, however, and would be attenuated if those reports are subject to error. The Homescan spending data are also limited in scope, largely capturing only goods with Universal Product Codes.[3] Moreover, the Check data allow estimates of the response to routine payments such as paychecks and Social Security payments not just particular payments such as tax rebates.

In policy discussion prior to the 2008 tax rebates, the Congressional Budget Office and others cited the point estimates of the effect of the 2001 rebate from Parker, Johnson, and Souleles, but not the substantial uncertainty about that estimate documented in that paper and in Table S3 (*3, 12*). More generally, estimates of spending rates from different changes to income play a key role in the evaluation of the American Recovery and Reinvestment Act (*13*), making the stakes in getting credible and precise estimates of these parameters very high. This paper shows how economic theory and policy can benefit from analysis made possible with naturally-occurring data like those provided by Check.

References and Notes

1. N. S. Souleles, *The American Economic Review* **89**, 947 (1999).

2. J. A. Parker, *American Economic Review* **89**, 959 (1999).

3. D. S. Johnson, J. A. Parker, N. S. Souleles, *American Economic Review* **96**, 1589 (2006).

[3]Related studies of administrative data also provide accurate measures of spending but do not cover it comprehensively. For example, Agarwal et al. use data from a single credit card company to study the spending response to tax rebates; they can thus track the effects of the rebate on a single account, but not on overall spending (*9*). Kuchler makes use of more comprehensive administrative data collected from a debt management website, but the number of users (556) is relatively small (*10*) . The financial application Mint (www.mint.com) has a complementary data infrastructure that it is using to construct monthly time series of spending by types of good (*11*). It has not been used for research along the lines of the estimates in this paper.

4. J. A. Parker, N. S. Souleles, D. S. Johnson, R. McClelland, *American Economic Review* **103**, 2530 (2013).

5. M. Stephens, *The American Economic Review* **93**, 406 (2003).

6. M. Stephens, *The Economic Journal* **116**, 680 (2006).

7. C.-T. Hsieh, *American Economic Review* **93**, 397 (2003).

8. C. Broda, J. A. Parker, The economic stimulus payments of 2008 and the aggregate demand for consumption, *Manuscript*, Kellogg School of Management (2012).

9. S. Agarwal, C. Liu, N. S. Souleles, *Journal of Political Economy* **115**, 986 (2007).

10. T. Kuchler, Sticking to your plan: Hyperbolic discounting and credit card debt paydown, *Manuscript*, Stern School of Business (2013).

11. Mint, Intuit consumer spending index, *Tech. rep.*, Intuit Inc. (2013).

12. Congressional Budget Office, Options for responding to short-term economic weakness (2008). Washington, D.C.

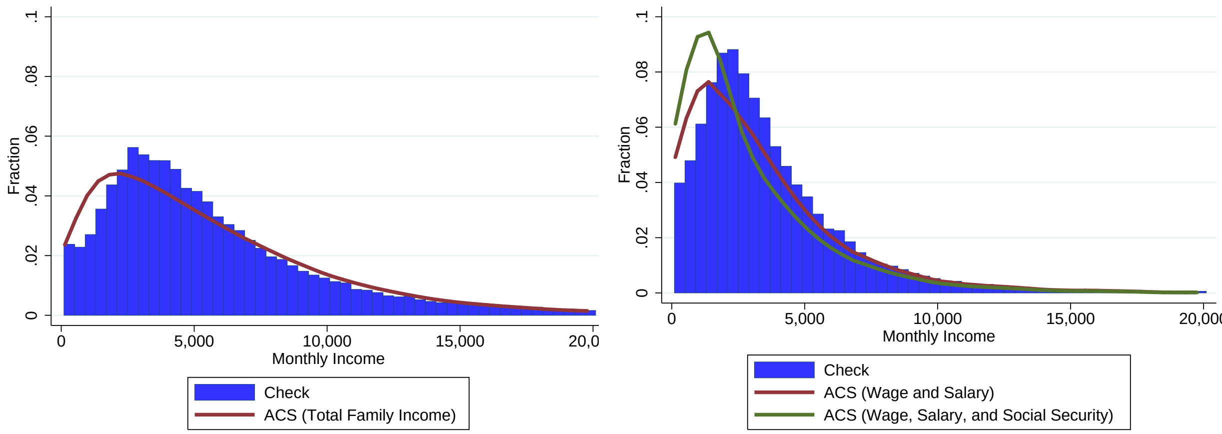
13. Congressional Budget Office, Estimates impact of the american recovery and reinvestment act on employment and economic output from july 2010 through september 2010 (2010). Washington, D.C.

14. This research was supported by a grant from the Alfred P. Sloan Foundation. Shapiro acknowledges additional support through the Michigan node of the NSF-Census Research Network (NSF SES 1131500). This paper has benefited from suggestions by the participants of the NBER Summer Institute, the Conference on Economic Decisionmaking (Aspen, Colorado), and several seminars.

Figure 1: Distribution of Monthly Income

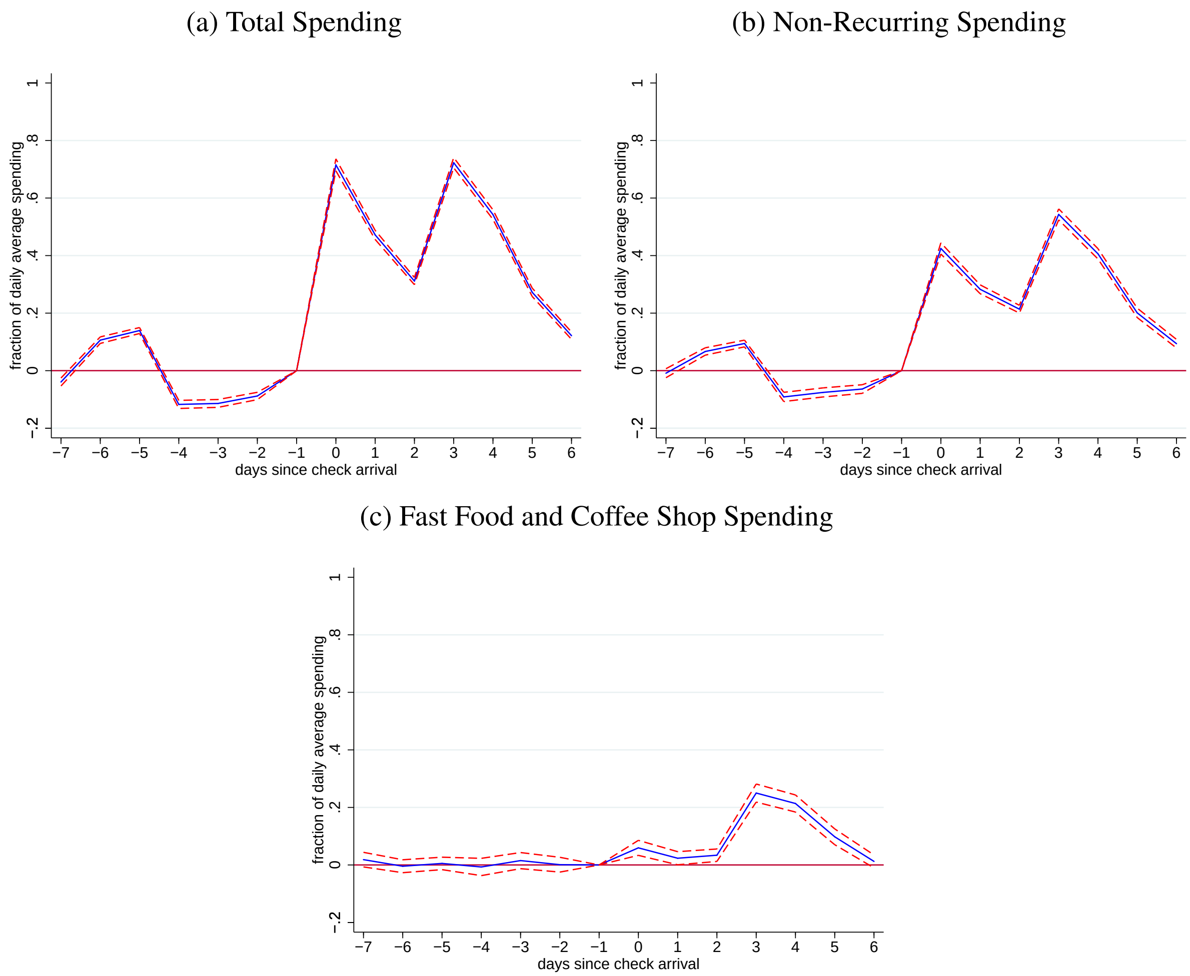
(a) Total Income

(b) Paycheck and Social Security Income



Notes: Figure shows average monthly income across users. Any month where the user has fewer than 20 days of data is dropped from computing the average. In Panel (a), the Check distribution represents 61,184 users who have at least one checking or saving account. In Panel (b), both Check and ACS distributions are conditional on having paycheck and Social Security income (47,050 users).

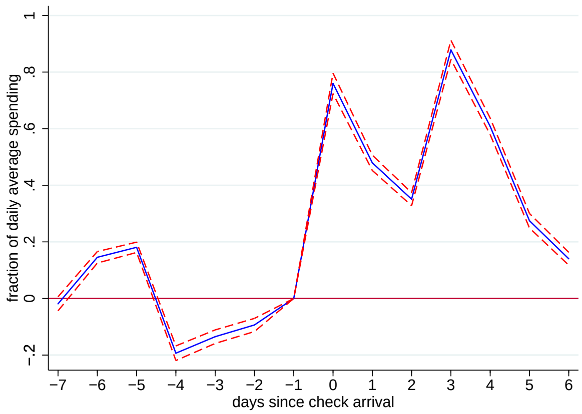
Figure 2: Response of Spending to Income: Alternative Components of Spending



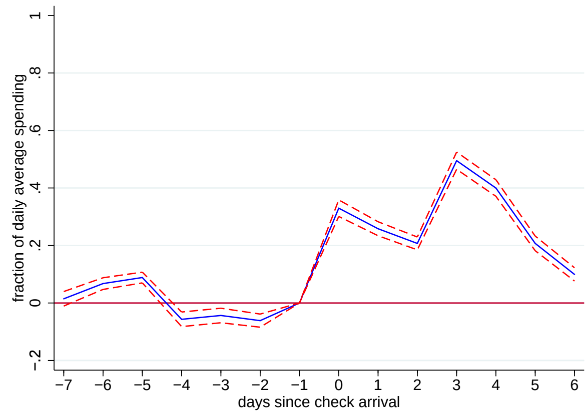
Notes: Solid line represents regression coefficients from Equation (1). Dashed lines are 95% confidence intervals. Estimates based on 5,371,244, 5,371,244, and 5,173,594 total observations from 23,985, 23,985, and 23,021 users for panels (a),(b), and (c) respectively.

Figure 3: Response of Non-Recurring Spending to Income: Liquidity Ratio

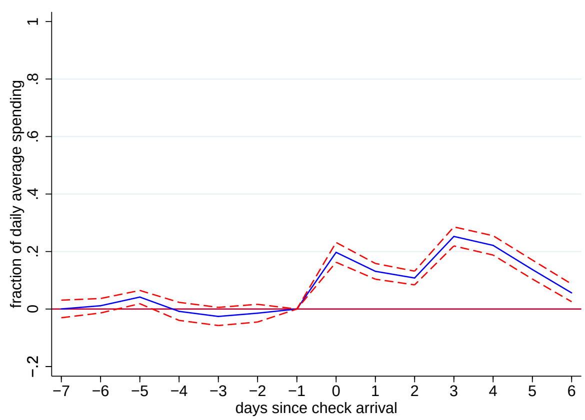
(a) Low Liquidity



(b) Medium Liquidity

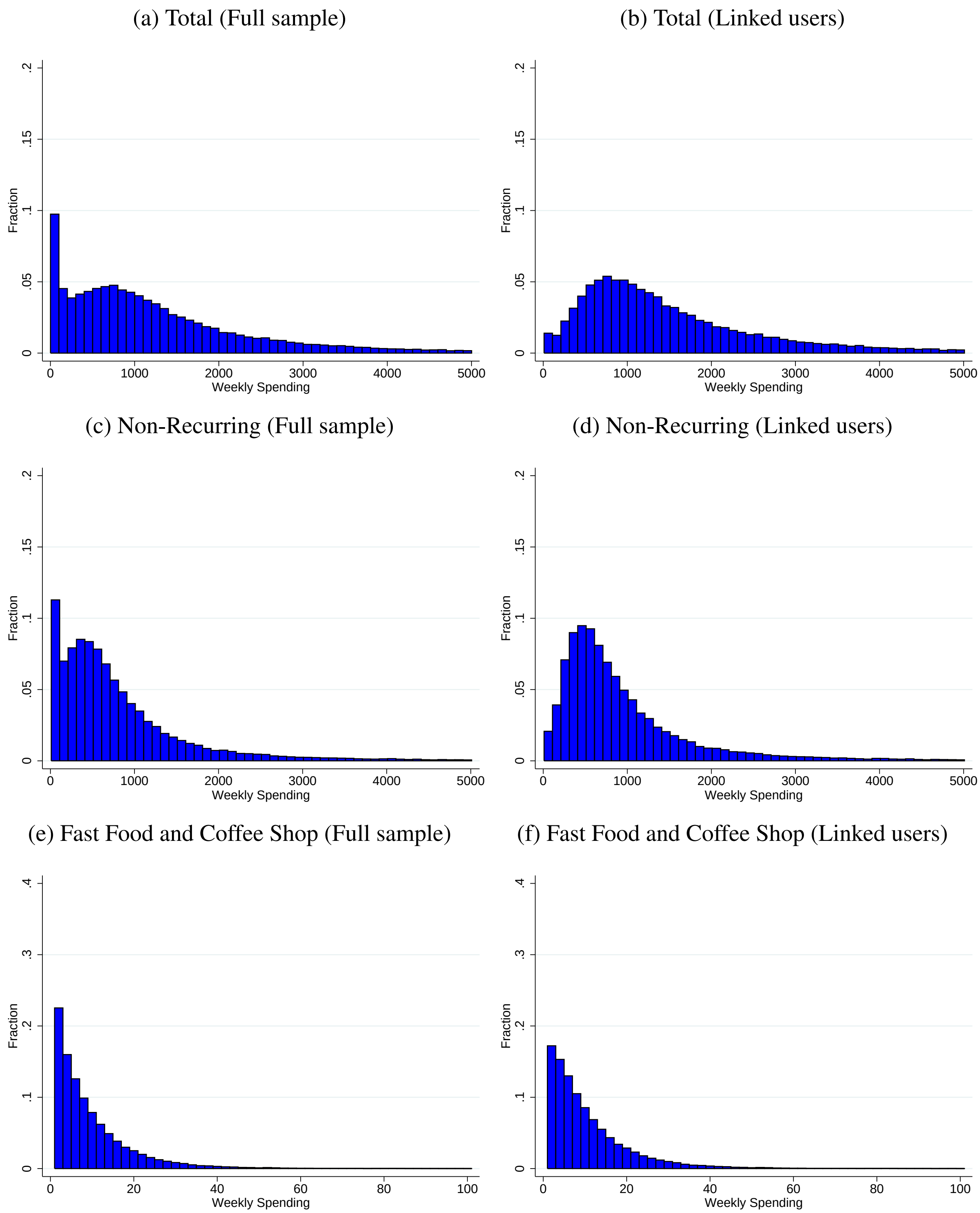


(c) High Liquidity



Notes: Solid line represents regression coefficients from Equation (1). Dashed lines are 95% confidence intervals. Estimates based on 1,784,460, 1,809,839, and 1,769,968 total observations from 7,956, 7,956, and 7,955 users for panels (a),(b), and (c) respectively. The liquidity ratio is defined as the average daily balance of checking and savings accounts normalized by daily average spending.

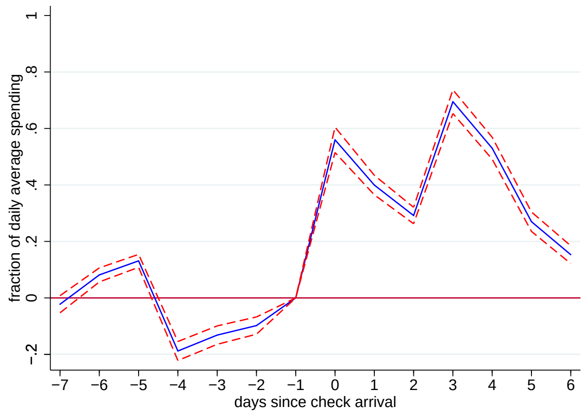
Figure S1: Distribution of Weekly Spending



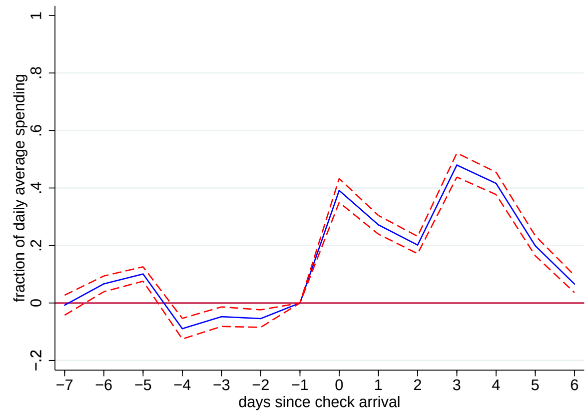
Notes: The figure contains data from 72,182, 50,510, 72,101, 50,462, 59,261, and 46,004 users out of a total 72,902 for panels (a), (b), (c), (d), (e), and (f) respectively.

Figure S2: Response of Non-Recurring Spending to Income: Credit Utilization Rate

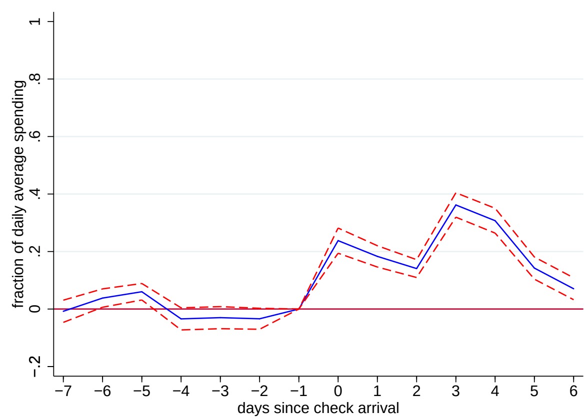
(a) High Utilization



(b) Medium Utilization



(c) Low Utilization



Notes: The utilization ratio is calculated conditional on carrying a debt balance. Solid line represents regression coefficients from Equation (1). Dashed lines are 95% confidence intervals. Estimates based on 852,709, 911,118, and 922,390 total observations from 3,832, 3,993, and 4,012 users for panels (a),(b), and (c) respectively.

Supplementary Online Materials

Transaction and Balance Data

The sample we use is based on de-identified data that we obtain from Check and that was aggregated as described below. The sample contains a panel of approximately 75,000 Check users, selected at random from the pool of U.S.-based users who had at least one bank or credit card account, and covers 300 consecutive days spanning 2012 and 2013. The sample of users is not refreshed. The average monthly rate of attrition is approximately 1.0%. Check refreshes its data daily and records every financial transaction that was posted for each user in the sample. The analysis we perform does not use the raw data collected by Check but instead uses the de-identified data to create individual-level measures that are normalized using the individual's aggregated transactions across time. We describe this in more detail below. The information about transactions includes the type of account to which they were posted, the dates of the transactions, the amount of the transactions, an indicator for whether they were a credit or debit to the accounts, and a description. The data also include daily account balances for each account of each user. Table S1 and S2 provide summary statistics of the transaction and balance data where P_i denotes the i-th percentile.

The quantity of information is notable. In only ten months, the sample of 75,000 users recorded more than 57 million transactions using an average of about six different accounts. The data show substantial variation across users. Saving accounts have an average balance of $6,476, a median of $400 and an interquartile range of $0 to $2,500. Debit and credit utilization levels are similarly heterogeneous, and credit cards are used, on average, about half as often as checking accounts.

Table S1: Transactions and Accounts

	Mean	P_5	P_{25}	P_{50}	P_{75}	P_{95}
Daily transactions	4.54	1	2	3	6	13
Credit card	1.23	0	0	1	2	5
Checking account	3.03	0	0	2	4	11
Saving account	0.22	0	0	0	0	1
Accounts	5.84	2	3	5	8	12
Credit card	3.58	1	2	3	5	9
Checking account	1.35	0	1	1	2	3
Saving account	0.79	0	0	1	1	2

Notes: In total, the 57,731,354 transactions are generated from 72,902 unique users over the study period.

Table S2: Account Balances

Panel (a): Bank	Mean	P_5	P_{25}	P_{50}	P_{75}	P_{95}
All	\$14,415	\$100	\$700	\$2,200	\$7,900	\$55,400
Checking	\$6,969	\$100	\$500	\$1,400	\$3,800	\$23,100
Saving	\$6,476	\$0	\$0	\$400	\$2,500	\$25,200
Money Market	\$12,076	\$0	\$100	\$900	\$7,700	\$57,400
C.D.	\$12,734	\$0	\$0	\$500	\$4,000	\$39,200

Panel (b): Credit Card	Mean	P_5	P_{25}	P_{50}	P_{75}	P_{95}
Balance	\$7,228	\$200	\$1,400	\$3,600	\$8,500	\$26,100
Credit Limit	\$23,019	\$800	\$4,200	\$11,900	\$29,500	\$81,800
Utilization Ratio	0.48	0.02	0.15	0.45	0.78	1.00
Revolving Debt	\$5,828	\$1,200	\$2,100	\$3,500	\$6,700	\$18,000
APR	18.46%	10%	15%	18%	23%	27%

Notes: All figures are aggregated to the user level. Panel (a) and the first three lines of Panel (b) reflect average daily balances over the last seven months of the study period. The last three lines in Panel (b) reflect average daily balances over the study period. We drop extreme values over \$10m as well as business accounts. The amounts are conditional on a user having that particular type of account. There are 64,136, 47,798, 10,530, 4,100, and 69,834 users who have Checking, Saving, Money Market, C.D., and Credit Card accounts respectively. The last three rows of Panel (b) are conditional on having revolving debt. 35,922 users have revolving debt at some point during the sample period.

Income, Cash, and Transfer Measures

We derive two measures of income: The first sums all transactions that represent credits to a user's non-credit-card accounts, excluding transfers from one account to another. The second isolates only those transactions that credit paychecks and Social Security payments, using a list of keywords commonly found in the description field. See Fig. 1 in the paper. Paychecks are identified using the following keywords: "direct", "dir dep", "dirdep", "salary", "treas 310 fed", "fed sal", "payroll", "ayroll", "payrll", "payrl", "payroll", "pr payment", "adp", "dfas-cleveland", "dfas-in."[4] Social Security payments are identified using the key word "soc sec." A paycheck or Social Security payment is classified as regular in frequency if the median number of days between its arrival is from 13 to 15 or from 26 to 34 and if its coefficient of variation is less than 0.5.

To identify ATM or cash withdrawals, we use keywords "atm" or "cash" and "withdrawal." We exclude ATM withdrawals that have "purchase" in them. Those transactions are purchases with ATM cards and are therefore classified as spending.

To identify transfers, we use keywords "transfer", "xfer", "tfr", "xfr", and "trnsfr."

Spending Measures

Fig. S1 shows the distribution of three spending measures, both for the entire sample and for those users whose accounts appear especially well-linked to the application.[5] Fig. S1 (a) and (b) show total spending. Fig. S1 (c) and (d) show *non-recurring* spending. Finally, Fig. S1 (e) and

[4]ING direct transactions and direct deposit advances identified by the keywords "ing direct", "direct deposit advance", and "dir dep adv" were not included in the data because they appear not to be payroll-related in general.

[5]Well-linked users are defined as those for whom we have at least 80% coverage of their credit card balance payments both as credits to the card, and as debits to a checking account. The denominator in this coverage is the number of credit card balance payments that are recorded on a user's credit card. The numerator is the number of payments of the same amount that show up in the user's checking account. This ratio proxies for whether users are linking their major accounts on Check.

(f) isolate spending on certain fast food and coffee shops.[6] The main text provides descriptions of how these three measures are derived from the raw transaction data.

The combined content and frequency of the spending measures shown in Fig. S1 has no analogue in other data sources. In particular, unlike other sources, the Check data allow us to isolate irregular spending using the actual amount and pattern of spending, rather than an a priori categorization based on goods and services.[7] While these measures cannot be directly compared with those from other data sources, we can make qualitative assessments of their features. We note first that well-linked users exhibit fewer weeks with very low expenditure because they are less likely to spend from accounts that go unobserved by the app. Second, the approximate lognormal shape of the distributions of aggregate spending among linked users in Fig. S1 (b) and (d) is typical, and the spike at zero for fast food and coffee shops spending in Fig. S1 (e) and (f) is to be expected.

Heterogeneity of Excess Sensitivity

In the main text we documented heterogeneity of excess sensitivity by levels of liquidity. Here we see evidence of similar heterogeneity by level of credit utilization. Fig. S2 plots estimates of β_{kc} for non-recurring spending, by terciles of credit utilization. We define credit utilization as the ratio of the average daily balance on all credit cards to the average daily credit limit

[6] The non-case sensitive keywords are starbucks, dunkin, coffee, mcdonalds, kfc, taco bell, burger king, wendy's, subway, jack in the box, and panera.

[7] The CEX, which is the leading survey of expenditure in the U.S., conducts two separate surveys, a quarterly retrospective survey and a daily diary collected over two-week intervals. The measures of spending in the two surveys do not completely overlap; the daily diary emphasizes small, and frequently purchased items that would be hard to recall in the retrospective. The CEX, thus, cannot be used to calculate total or non-recurring expenditure at a weekly frequency. The American Life Panel (ALP) is an Internet survey that collects monthly self-reports of expenditure, in 25 categories, from approximately 2,000 respondents. These 25 categories capture high to medium frequency purchases. The Nielsen Panel of Consumers (Homescan) provides weekly expenditure data primarily for products with a Universal Product Code, or barcode. These expenditures, concentrated in the grocery and mass merchandise sectors, represent merely 40% of all expenditure on goods included in the government's measure of inflation. Thus, neither the ALP or the Homescan data can be used to measure total expenditure or non-recurring expenditure at daily and weekly frequency, nor do they have income closely aligned with the spending data.

summed across those cards. The average utilization ratio in the first, second, and third terciles of this measure is 0.22, 0.58 and 0.90, respectively. The results show that excess sensitivity is by-and-large a phenomenon of those near the limit of their ability to borrow with credit cards. While almost all Check users have access to credit, many are close to their borrowing limits. Among those close to the limit, excess sensitivity is much greater.

Summary of Comparison with Other Studies

Table S3 compares estimates of the marginal propensity to consume (MPC) or elasticity across studies.

Table S3: Comparison of results with previous studies

Study	Spending	Data	Frequency	Observations	Estimate	Confidence Intervals	
						Range	Length
Parker (1999)	Non-Durable	CEX	Quarterly	133,820	Elasticity	[0.22 – 1.01]	0.79
Souleles (1999)	Total	CEX	Quarterly	4,525	MPC	[0.21 – 1.08]	0.87
Hseih (2003)	Non-Durable	CEX	Quarterly	806	Elasticity	[-0.06 – 0.07]	0.13
Stephens (2003)	Total	CEX-D[1]	Weekly	56,649	Elasticity	[-0.16 – 0.67]	0.84
JPS (2006)	Non-Durable	CEX	Quarterly	4,739	MPC	[-0.33 – 0.71]	1.04
Broda, Parker (2012)	Total	Neilsen	Weekly	28,937	MPC	[0.01 – 0.02]	0.01
PSJM (2013)	Non-Durable	CEX	Quarterly	10,362	MPC	[-0.03 – 0.55]	0.58
GKSST (2013)	Total	Check	Daily	5,371,244	Elasticity	[0.69 – 0.74]	0.05
GKSST (2013)	Non-Recurring	Check	Daily	5,371,244	Elasticity	[0.41 – 0.44]	0.03
GKSST (2013)	Total	Check	Daily	5,371,244	MPC	[0.06 – 0.08]	0.02
GKSST (2013)	Non-Recurring	Check	Daily	5,371,244	MPC	[0.00 – 0.01]	0.01

Notes: Confidence intervals are for estimates of the percent change in consumption with respect to a percent change in income ("elasticity") and the change in consumption with respect to a change in income ("MPC") for various studies. The last two columns present the range and length of the 95% confidence interval for these estimates.

[1]Stephens (2003) utilizes the diary portion of the CEX (CEX-D). The diary is collected at a daily frequency and then aggregated to weekly for the cited analysis.