Crowdsourcing Peer Information to Change Spending Behavior

Francesco D'Acunto[*]
Georgetown University

Alberto G. Rossi[†]
Georgetown University

Michael Weber[‡]
University of Chicago & NBER

This version: June 2023

Abstract

We isolate the information channel of peer effects in consumption in a setting that excludes a role for common shocks or social pressure—a spending panel paired with crowdsourced information about anonymous "peers" elicited at different times. Consumers converge to peers' spending, and more so when peer signals are more informative. Convergence is asymmetric: within 12 months of information provision, overspenders close 17% and underspenders 5% of their gap relative to peers. We exploit the quasi-random assignment to peer groups in an instrumental-variable strategy and implement an experiment for external validity. Our results are consistent with information-based theories of overconsumption.

JEL: D12, D14, D91, E22, G41.

Keywords: Social Finance, Visibility Bias, Conspicuous Consumption, Robo-advising, FinTech, Information Economics, Household Finance.

This research was conducted with restricted access to data from *Status Money Inc.* The views expressed here are those of the authors and do not necessarily reflect the views of *Status Money Inc.*, which had no right to review the content of this paper. We are indebted to the founders, Majd Maksad and Korash Hernandez, for providing us with invaluable feedback. For very helpful comments and discussions, we thank Sumit Agarwal, Marieke Bos, Stephen Dimmock, David Hirshleifer, Yi Huang, Byoung-Hyoun Hwang, Samuli Knüpfer, Tao Li, Chen Lin, Christian Lundblad, Michaela Pagel, Wenlan Qian, Tianyue Ruan, Patricio Toro, Michael Ungeheuer, Qi (Jacky) Zhang, and conference and seminar participants at the 2023 ASSA Annual Meetings, 2022 AI & Big Data in Finance Research (ABFR) Forum, 2022 ABFER Webinar Series, 2020 AFA, 2020 GSU-RFS Fintech Conference, 2019 TAU Finance Conference, 2019 Santiago Finance Workshop, 2019 Miami Behavioral Finance Conference, the 2019 EFA, 2019 Helsinki Finance Summit, 2019 Cherry Blossom Conference, 2019 Olin Household Finance Conference, 2019 Tsinghua SEM Finance Workshop, and 2019 Guanghua FinTech Conference, as well as at CMU, IMF, NSU, SMU, NTU, HKU, CUHK, HKUST, the Investor and Financial Education Council of Hong Kong, ESSEC, UIBE, IFEC, Tinbergen Institute, University of St. Gallen, and University of Bergen. D'Acunto acknowledges financial support from the Ewing Marion Kauffman Foundation and Weber from the Fama Research Fund at the University of Chicago Booth School of Business. All errors are our own.

[*] McDonough School of Business, Georgetown University, Washington, DC, USA. e-Mail: francesco.dacunto@georgetown.edu

[†] McDonough School of Business, Georgetown University, Washington, DC, USA. Fellow of the Luohan Academy. e-Mail: agr60@georgetown.edu.

[‡] Booth School of Business, University of Chicago, Chicago, IL, USA and NBER. e-Mail: michael.weber@chicagobooth.edu.

1 Introduction

Low savings limit the wealth accumulation of households, many of whom reach the time of retirement holding insufficient resources to maintain their pre-retirement lifestyle (e.g., see Banks et al., 1998; Bernheim et al., 2001; Lusardi and Mitchell, 2007). Beyond financial constraints, consumers often lack the information and sophistication needed to form beliefs about optimal spending (D'Acunto et al., 2019). Consumers might thus rely on rules of thumb such as conforming to the choices of people whose demographic characteristics are similar to theirs as if such choices provided valuable information about optimal spending (Bikhchandani et al., 1998; Kaustia and Knüpfer, 2012; Agarwal and Qian, 2014; Bailey et al., 2018; Maturana and Nickerson (2020).

In this paper, we empirically isolate this information channel of peer effects in a unique field setting that, by construction, rules out any role for common shocks across peers or peer pressure and in which connections among peers are not formed endogenously. We exploit a transaction-level panel dataset of spending choices paired with a FinTech tool (*Status Money*) that provides users with information about the overall spending of anonymous demographically similar consumers ("peers"). The information is crowdsourced from the account-level data of a large representative US population outside the *Status Money* platform. In this setting, users do not know peers' identities and do not interact with peers socially, and hence socialization or peer pressure due to direct observation of peers' choices have no scope.[1] Moreover, the information on peers' spending is not measured at the same time at which users make their own choices, which dismisses a role for common shocks that both users and peers might face as a potential explanation for correlated choices between the two groups. Because the extent to which peer groups' characteristics are similar to users' characteristics varies across users, the informativeness of peer signals varies in our setting—a feature that allows testing whether users' reaction to peer signals depends on their informativeness.

Upon subscribing to the platform, users provide a set of demographic characteristics that

[1] As we discuss in more detail below, users who learn they spend more than their peers might be subject to "peer pressure" in the sense that they feel pressured to conform to the social norm of thrift that they learn exists in a population demographically similar to them.

include their annual income, age, homeownership status, location, and type of residence. The platform obtains users' credit scores via credit reports as well as the pre-signup transaction-level spending history directly from the financial accounts users link. We can thus compute directly users' pre- and post-signup monthly spending. The platform matches each user automatically with one of a set of pre-defined peer groups constructed based on US representative transaction-level data outside the platform. Users are not aware of the rules used to match them to a peer group.[2]

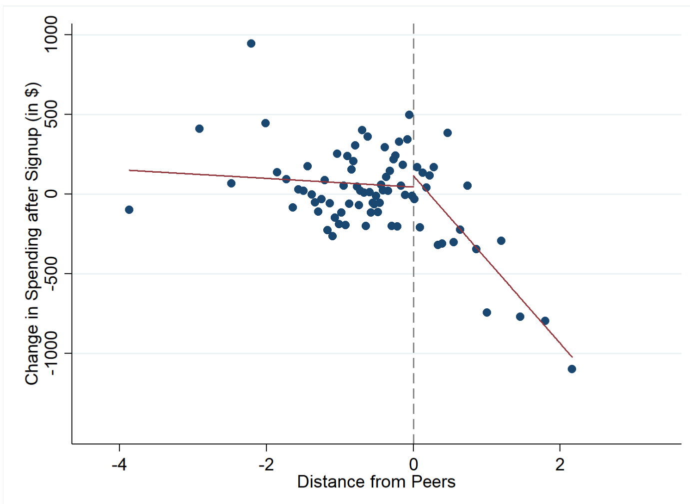


Figure 1. Distance from Peers' Spending and Change in Spending: Raw Data

Figure 1 is a binned scatterplot of dollar changes in spending (y-axis) for users who sign up to *Status Money* three months after signup relative to three months before. Dollar changes are the residuals from regressing actual changes on month-of-year signup dummies to account for seasonality in spending. The x-axis sorts users based on the difference in spending with respect to peers over the 30 days before signup, which we standardize to have a unit standard deviation. The binned scatterplot divides the 20,679 users into 80 groups. The solid lines plot the fitted values of a threshold regression that estimates different linear regression coefficients below and above the level of zero distance from the peers.

Figure 1 is a raw-data plot that motivates our analysis. In this binned scatterplot, each point represents approximately 250 users. On the y-axis, we plot the average change in monthly dollar spending for users in the three months after signup relative to the three months before signup. On the x-axis, we sort users based on the standardized distance of

[2]After signup, users can set up *additional* peer groups. Our analysis does not use such self-designed groups because less than 2% of users in our sample set them up. Even for the latter users, information from the pre-defined groups is always visible.

their spending relative to the peer group's average spending over the 30 days before signup.[3] The distance from peers is positive if users spent more relative to the average peer spending before signup and negative otherwise. The raw data reveal three patterns. First, users adjust their spending toward the level of their peers' spending after signup, irrespective of whether they spent more or less than their peers. Second, the size of the change is monotonically related to the distance from peers in both directions. Third, the reaction is asymmetric—users above their peers change their spending by more than users at a similar distance below their peers.[4]

We show that the static asymmetric effect in Figure 1 holds dynamically: both overspenders and underspenders keep adjusting their spending toward the levels of their peers over the 12 months after signup. Overspenders react more, though: they close their gap by 17%, whereas under-spenders only by 5%.

A potential mechanical relationship between the observed change in spending and users' distance from peers' spending before signup is an important concern with the baseline facts: users who face especially large or small expenses in the pre-signup period might merely revert to their usual levels of spending resembling a reaction to peer information. Moreover, users might be reacting to other pieces of information or other features of the platform rather than to the information about peers' spending. The raw-data patterns also face a set of endogeneity concerns. For instance, users might first decide to change their spending and only sign up to the platform to use the income aggregation feature and monitor their transactions (e.g., see D'Acunto et al., 2019; Rossi and Utkus, 2020).

We tackle these concerns in the second part of the paper. We first show directly that mean reversion cannot drive our results. The facts are robust to controlling directly for users' pre-signup spending as well as to using their spending levels two months or three months before signup to compute the distance from peers, thus dismissing unusual spending in the

[3] Several platform features have changed over time and especially after the sample period for which we have information, which ends in January 2019. In section 2, we discuss in detail the features and timing of these changes based on the timeline the platform founders shared with us and official press releases. All the features we discuss in the paper only refer to what users saw during our sample period.

[4] The raw-data results assume that spending adjustments differ systematically above and below the level of peer spending rather than based on other values. We verify below that the cross-sectional spending changes have a kink at the peer level by estimating endogenous threshold regressions.

weeks before signup as a potential explanation for our results.[5] The dynamics of the effects over time are also inconsistent with mean reversion because the reaction builds up over time for both overspenders and underspenders.

We then establish that users react to the information about peer spending rather than to other information they see on the platform, such as the average spending of all US consumers and information about users' own average income. We do so by considering cases for whom peer spending and other pieces of information predict reactions of opposite signs. This analysis also helps reconcile our results with earlier research that did not detect substantial effects of peer information on households' financial choices (e.g., see Beshears et al., 2015): even in our setting, similar to earlier work, providing information about broad groups with varied demographics, such as the average of all US consumers, is uninformative and has no effect on consumers' choices.

To tackle the broader endogeneity concerns directly, we move on to propose an instrumental-variable (IV) analysis that exploits a design feature of the platform—the quasi-random assignment of users to peer groups based on threshold rules that users ignore. Because users do not know the assignment rules, they cannot strategically manipulate their reported demographics, for instance if they wanted to avoid negative news about their spending relative to peers. And indeed, we show that the pre-signup spending levels of similar users who end up being assigned to different peer groups are indistinguishable. The IV strategy confirms our baseline results. Users who look similar along all dimensions we observe, including pre-signup spending, converge to the different average spending of the peer groups to which they are assigned. We corroborate the IV results through a placebo IV analysis and a set of falsification tests.

We also exploit this identification strategy to assess the role of the informativeness of peer signals by designing heterogeneity tests that compare the reactions of peers who observe more or less informative peer signals. In these tests, channels other than peer information, such as design features of the platform, are kept constant because they are equally active for all

[5] We do not argue that mean reversion in spending does not exist. In fact, we do find evidence of mean reversion in spending over time. What we argue is mean reversion is inconsequential to our results and that users do react to the information about peers above and beyond other changes in their spending profile, including mean reversion.

users. Instead, we exploit design features of the platform that allow for variation of peer groups' informativeness, such as the precision with which users' characteristics are matched to those of the peer group and peer group sizes. Across all these dimensions, we find that users who are matched with more informative peer groups react substantially more after signup, whereas users who are matched with less informative groups barely react. Users thus do not just anchor their monthly spending to a number they read on the platform, but they interpret the number as material to their choices only when the group of consumers based on which such number is computed is more precisely tailored to their own characteristics.

We then move on to tackle another issue that research based on spending aggregators cannot address directly—external validity. Because *Status Money* is marketed as a tool that provides comparisons with peers, the population that selects into this service might be more sensitive than the average US consumer to the differences between their spending and peers' spending. This issue hinders us from drawing inference about the effects of providing peer spending information to the broader population based on our baseline results.

To tackle this external-validity concern, we implemented a randomized control trial (RCT), which we ran on a representative US population that we recruited online without any mention of peers, peer information, or household finance. Similar to *Status Money*, we gave respondents truthful information regarding the spending choices of demographically similar consumers and tested how such information affected respondents' spending plans in a within-subject experimental design. Following the literature on consumption and saving decisions (Parker and Souleles, 2019), we first elicited respondents' marginal propensities to consume (MPC) out of an unexpected reimbursement, which scales directly respondents' spending by their (unobserved) income.[6] We then provided information on the MPC of consumers in the same income bracket as the respondent (income peers) and elicited the respondent's MPC again. This non-selected population responds to peer spending information in a strikingly similar way as *Status Money* users. In particular, all respondents react to peer information, but the reaction is asymmetric and stronger for overspenders than underspenders.[7]

[6] We followed the literature and elicited MPCs rather than asking for the full set of individual transactions across spending categories, which would have been too burdensome for respondents and would have likely generated substantial measurement error.

[7] The RCT also made it possible to elicit demographic information and economic preferences and beliefs, which were

Stressing what we do not claim in this paper is also important. We do not claim the information users obtain captures precisely the actual spending of their peers or that the rules used to create peer groups are optimal. In fact, our setting includes cases in which peer groups' characteristics are more or less similar to users' characteristics, which is a crucial feature to assess the role of peer signals' informativeness.

Moreover, we do not claim that reacting to peer information is optimal in our or other contexts. Related, our paper has no normative implications in terms of whether reacting to peers' spending increases or decreases users' welfare. The asymmetric reactions we uncover, though, suggest that providing consumers with information about the spending of a targeted population that shares the same demographic characteristics might moderate excessive spending in the aggregate. Targeted peer information might thus be used as a low-cost communication-based policy to manage aggregate consumption and savings (D'Acunto, Hoang, and Weber, 2021).

1.1 Contributions to the Literature

Our paper contributes to several strands of literature in economics and finance. First and foremost, we contribute to research that studies the role of peer effects on economic decision-making, that is, the extent to which consumers' choices conform to those of peers. In particular, our paper aims to assess the existence and estimate the economic magnitude of the information channel of peer effects, whereby consumers conform their choices to those of peers because they think they learn about optimal decision-making based on peers' choices. Isolating this information channel of peer effects is notoriously difficult for at least two reasons. First, the literature finds that socialization and common shocks might explain correlated choices across peers even if agents do not react actively to peers' choices (Manski, 1993). Whereas this concern can be tackled by tailoring convincing causal identification designs, a second important concern is that peers' choices might generate peer pressure above and be-

not available in our baseline observational data (D'Acunto, Rauter, Scheuch, and Weber, 2020; D'Acunto, Malmendier, and Weber, 2020; Coibion et al., 2019; Coibion et al., 2020). We find that demographic dimensions typically related to spending attitudes, such as gender, marital status, number of children, as well as risk aversion, and patience, do not relate to systematically different reactions to peer information, which further reduces the concern that our baseline results are not externally valid.

yond providing informative signals (e.g., see Mas and Moretti, 2009) and hence even carefully identified reactions to observing the choices of peers might confound these two effects.

For instance, Allcott (2011) and Allcott and Rogers (2014) study the Opower experiment, whereby households were provided information about the energy consumption of their neighbors and their energy consumption was subsequently monitored. They find that high-consumption households cut their energy consumption when they learn that neighbors consume less, an effect that might be due to learning about optimal energy consumption from peers' signals (information channel), due to avoiding being perceived as overconsumers by peers the next time information is distributed (social pressure channel), or to both channels.

The literature has proposed settings to isolate the information channel of peer effects. Bursztyn et al. (2014) design a RCT in which financial advisors provide retail investors with information about financial investments made by their relatives. They argue that the peer effects they estimate are likely driven by an information channel rather than peer pressure because financially illiterate investors, such as medical doctors, react more to information about the choices of relatives who have a background in finance. Maturana and Nickerson (2019) exploit information on teachers' randomized schedules that create groups of teachers who have breaks from classes at the same time and hence presumably discuss and spend time with each other during the breaks. They find that teachers who share the same break time tend to refinance their mortgages after one among them does so arguably due to the exchange of information between them. This effect is not present in teachers who don't share the same break time. Ouimet and Tate (2020) analyze the employee stock purchase plans (ESPPs) choices of coworkers who have the same eligibility window and find that the choices of highly-informed workers have a stronger influence on the choices of coworkers than the choices of others.

Our paper's contribution in this area is the study of a setting that, by construction, cannot generate peer effects in the form of common shocks or social pressure. On the one hand, the peers in our setting are anonymous; that is, users observe information about the spending of a representative set of individuals with similar demographic characteristics without knowing their identities or being observed by them. Therefore, whether users update their spending behavior or not, none of the individuals whose information was part of the signal will ever observe the change in spending behavior. Thus no scope exists in this setting for reacting due to social pressure. On the other hand, the app's information about peer spending is static and was obtained for a representative US consumer cohort in July 2017. In contrast, users' spending behavior was observed from August 2017 until the end of our sample period in January 2019. By construction, our users and peers do not make contemporaneous spending decisions, and hence potential common shocks faced by all peers at the same time cannot explain our results.

The paper also fits into a growing literature on the role of FinTech applications in consumption and household finance choices (see D'Acunto and Rossi (2023) for a recent review of this literature). FinTech applications have been shown to modify the consumption, saving, and investment choices of households in the direction of those of rational economic agents through at least three channels. First, they provide consumers with information about their own characteristics through income aggregation features (e.g., see Olafsson and Pagel (2017) and Lee (2019)), as well as information about other economic agents, like in this paper, and about macroeconomic variables. Second, they reduce distortions in beliefs about own characteristics, others' characteristics, and the aggregate economy (e.g., see Gargano and Rossi (2020)). Third, they reduce the transaction costs that have been shown to lead to inertia in several realms of household decision-making, which can take the form of monetary costs (Reher and Sokolinski (2021)) as well as cognitive and psychological costs (e.g., see Chak et al. (2022)).

The empirical results in our paper also speak to the theoretical literature on peer effects on consumption and saving choices, especially theories of overconsumption and undersaving behavior. Overall, we note that the implications of visibility bias for consumption can rationalize the full set of results we document empirically (Han, Hirshleifer, and Walden (2019)).[8] We discuss the implications of our results for extant models in Section 9.

[8]We would like to thank David Hirshleifer for helping us draft this section based on a very detailed conference discussion.

2 Institutional Setting

In this section, we discuss the characteristics of *Status Money*, the signup procedure, and the information users observe after signup. We also discuss how the platform (which we label "app" thereafter) has evolved over time, including features that a contemporary reader might observe but did not exist during our sample period.

2.1 Purpose of the App and Signup Process

Similar to other FinTech apps in the consumer-finance space in the US and abroad, *Status Money* provides an income-aggregation feature, which allows users to visualize their full balance sheet in a simple and dynamic way and to access a set of metrics about their finances, which would be hard to compute for a household without a holistic view of all their debt, credit, and investment accounts (D'Acunto and Rossi, 2021).

Status Money also shows users how consumers similar to them in terms of observable characteristics (*peers*) manage their finances. The information about peers is crowdsourced from proprietary transaction-level data for a large sample of individuals outside the app and representative of the US population. The app thus enables users to obtain complex information that they could only acquire and process on their own if they had access to large-scale proprietary transaction-level data and if they were able to manage and analyze such big data.

Importantly, peers in this setting are not individuals that interact socially with users and users do not know their peers' identities. Users know that peer groups are defined based on an anonymous representative population outside the app. This feature departs from most of the research studying peer effects and is crucial to argue that our analysis isolates the information channel of peer effects relative to other channels that transmit peer effects, such as common shocks faced by peer groups, socialization, or peer pressure stemming from direct interaction.

When signing up, users provide their date of birth, their annual income, and their housing type—whether they own or rent. Users are then prompted to insert their address and the last four digits of their Social Security number. This information allows the app to connect to a credit bureau that returns all of the users' credit-score-related information.[9] Finally, the app asks users to link their checking and savings accounts, their credit card and other debt accounts, and their taxable and non-taxable investment accounts.

Users' incentive to link their financial accounts consists of accessing the aggregator features of the app, which are only meaningful if users link all their accounts. Consistent with the relevance of this incentive, the users in our sample link, on average 5 accounts, and the median user links 4 accounts.

2.2 Peer Groups Construction

For each user, the app matches a peer group from a large set of pre-designed peer groups. Peer-group matching depends on the user's age, income, location and type, credit score, and housing type, and each group must include a minimum of 5,000 underlying observations. In Figure 2, we provide an example of the screenshot that *Status Money* users observe about their own characteristics (Panel (a)) and the characteristics based on which the peer group is defined (Panel (b)). In this fictitious example, the user is 42 years old, has an annual income of $140K, lives in New York, has a credit score of 769, and is a renter. The peer group assigned to this user contains individuals whose age ranges between 40 and 49, whose income ranges from $100K to $150K, who live in New York, pay rent, and have a credit score that ranges between 720 and 779.

The constraint of at least 5,000 observations underlying each peer group is important for our analysis because it varies the precision with which peer groups' average characteristics match users' characteristics. Varying precision produces variation in the informativeness of peers' spending information across users. For instance, Figure 2 shows that, in order to cross the threshold of 5,000 peer observations, the fictitious user needs to be matched with peers living anywhere in New York and not just the urban areas (location type: "All"). Another New York urban user whose demographic characteristics allow finding 5,000 peers living in the urban area would be matched to a more precise peer group. The width of the ranges of quantitative demographic characteristics, such as income, also varies across users, and

[9] We as researchers do not observe any individually-identifiable information about *Status Money*'s users.

broader ranges imply less precise peer group comparisons. We will exploit these features in Section 7 to test for the heterogeneity of users' reactions to peer signals based on peer groups' informativeness.

2.3 Main Characteristics of the App During Our Sample Period

Once the user enrolls, the app automatically retrieves information from her financial accounts. The app stores all transactions and computes the user's net worth as the difference between assets and liabilities. To give the reader a sense of the information users observe, we describe the content of the home page below based on the app's appearance during our sample period.

The main feature of the home page is the comparison of the user's spending with her peers' spending. Figure 3 displays the salient graphic that compares the user's own daily spending based on daily transactions with the projected average daily spending of the peer group and the US national average. The screenshot was taken on October 30, 2018. We argue that this picture captures the main feature of the home page because during our sample period and in both smartphone and desktop versions of the app it was the first item users would see when logging into the app, it was placed at the center of the screen, and occupied most of the width of the screen.

The top part of Figure 3 shows the user's total spending next to the average peer spending and the national average. The blue line presents the user's cumulative spending over the course of the month until October 30. It also presents a forecast of total spending until the end of the month. On the same graph, the light- and dark-red lines present the peer and national average cumulative spending over the month. The app also displays as a grey-dotted line the user's average monthly income flows. This feature is unlikely to provide new information to most users and especially to employees with a fixed salary.

Note that users' spending is based on their own daily transactions, whereas peers' and US national average information are computed using a proprietary algorithm that aggregates spending information for a large sample of US consumers whose transactions *Status Money* observes. These transactions are aggregated to the monthly level and projected linearly for each day of the month, and are not computed on a daily basis. This difference in the frequency with which users' and peers' data are processed is not relevant to the scope of our tests because it does not vary by over- or underspenders, users above or below the income thresholds, which we exploit for identification, or by levels of informativeness, which we use in our heterogeneity tests.

The bottom of the home page provides links to more comprehensive statistics regarding users' debts, assets, net worth, and credit score (see Figure A.1 in the Online Appendix). Below, we show that whereas users react to the salient information about peers' spending, they do not appear to react to information about peers' debts, assets, or net worth.[10]

In pages different from the home page, which can be accessed through hyperlinks, users can obtain more detailed information about the comparison of their monthly spending with the average peer spending not only for the month in which users sign up but also for each of the 6 months before sign up. We do not know whether users click on such hyperlinks and hence whether they observe the longer comparison of monthly spending relative to the information on the home page, which we use in our baseline analysis, but in the data we find a large degree of persistence in whether users are classified as overspenders based on their most-recent monthly spending or any of the previous months. In fact, the within-individual correlation of the overspending assignment variable across the three months before signup is 73.3%. Because comparisons for previous months are typically similar to the most recent comparisons users see on the homepage, observing the full history should if anything reinforce the signal about peers' spending levels. Consistently, as we discuss below, when we assign users to overspender or underspender status based on their spending in any of the available months before signup we find results that are barely distinguishable from our baseline analysis.

As is the case for most FinTech apps, the graphics and features of *Status Money* have evolved over time. In Figure A.4 of the Online Appendix, we report the timeline of the changes that might be relevant to our analysis. In essence, any relevant change, such as the introduction of alerts, cash incentives, and social feeds only happened after the end of our sample period—which starts in July 2017 and ends in January 2019—and therefore cannot

[10]This non-reaction could be due to the facts that users do not click on the links at the bottom of the page and hence do not see this information, or that adjusting assets and liabilities is more complicated than adjusting spending, among other possible explanations, which we cannot disentangle empirically.

affect our analysis. For completeness, in Online Appendix A.1, we provide a more detailed discussion about the features that were introduced after our data sample ends and before the paper's publication, which a reader might see when accessing the app even if they were not present during our sample period.

3 Data

The first group of variables we observe is the set of users' self-reported demographic characteristics at the time of signup, including users' age, income, whether the user owns or rents the house in which she lives, the city in which the user lives, and her credit score.

The second set of characteristics we observe relates to peers. For each user, we observe the average characteristics of the peer group computed by *Status Money*. As we discussed above, the app does not use the characteristics of other app users to construct peer groups, but uses external proprietary data of a representative set of US consumers. This procedure rules out that any selection in the types of consumers who sign up are reflected in the average demographics of the peer groups. Note also that the information on peer spending was static during our sample period, that is, peer-group spending values were collected for a specific historical month (July 2017) and were not dynamically updated. Because spending is seasonal throughout the year, and July is one of the highest expense months in the year, this procedure contributes to explaining why the majority of users in our sample underspend relative to their peers. Hence, we have to account for seasonality in spending in our empirical analyses below. The demographics we observe for peers include their average credit score, average level of debt, average value of assets, average net worth, and average income. Finally, we observe the number of individuals populating each group.

We also observe information about the usage of *Status Money* accounts. These variables include the signup date, the number of monthly logins, and the number of financial accounts users link to the app.

Finally, we observe data on users' and peer groups' spending amounts, which represent the main variables in our analyses. We observe users' spending for up to 3 months before signup and up to 12 months after signup.

3.1 Sample Selection

To ensure our working sample includes individuals for which we meaningfully observe inflows and outflows before and after signup, we only consider users who have linked at least one spending account because, by construction, we cannot observe the spending behavior of those who linked no accounts.

Moreover, we select the raw sample according to a set of steps that are inspired by the sample selection used in earlier app-based research in household finance (e.g., see Olafsson and Pagel (2018) and Ganong and Noel (2019)).[11] We verify that none of our results depend on such sample selection steps in Section 5.1. Despite the similarity of the results for the full sample and across each selection step, we proceed with the sample selection steps to guarantee consistency with the empirical analysis earlier research has performed using other income-aggregating apps.

First, we only include users whose number of linked accounts does not change through the 90 days after signup. This step is important for two reasons. On the one hand, users might link one account at signup and start linking other accounts over time, as they build trust in the app. In this case, we might miss the reaction of users in accounts that were not linked at the time of signup or we might categorize as an increase in spending the mere fact that we observe more spending sources for the user over time. On the other hand, a concern with the FinTech apps in the literature is that users' accounts might be de-linked from the app due to inactivity, changing passwords, or other technical issues related to their account settings. Although these types of cases are dropped from the app, and hence we should not observe them, to avoid any risk that we falsely categorize as a drop in users' spending a mere delinking of an account, we drop users whose number of accounts linked changes over time after signup. By limiting the sample to users whose number of accounts does not change, we eliminate both possibilities from the analysis.

Second, we only include users with at least one monthly login to the app after signup. This step further alleviates the concern that, because of users' inaction, the app might not observe users' activity and categorize their spending as declining. *Status Money* has not detected this

[11] We thank Michaela Pagel and Byoung Hyoun-Hwang for raising points that guided our sample selection steps.

type of issue on their app, but because this issue was documented in other income aggregators, we also implement this selection step in our main working sample.

Third, we only include users who spend at least $100 per month in food-related transactions during the sample period. This step ensures that we do not keep individuals who are not actively using the accounts they link for spending purposes. We consider food-related expenses because after the third selection step all users have at least one transaction categorized as food expense each month after signup.

We also verify that no users are duplicates, that is, individuals who might have opened different accounts at different points in time, for instance, because they forgot their login credentials. Repeated accounts would overestimate the statistical significance of our results. To identify duplicates, we search for individual accounts with the same balances at the end of each month and who connect from the same IP address and have the same demographic characteristics.

3.2 Summary Statistics

The users in our sample are spread across the whole US (please see Figure A.5 in the Online Appendix for a map of users' location), which is important to ensure that individuals with varied demographic and social backgrounds, such as different attitudes towards spending and saving, enter the sample (see D'Acunto, 2018 and D'Acunto, 2019).

Panel A of Table 1 reports the user characteristics in the main sample. For each variable, we report the number of observations, averages, standard deviations, and a set of percentiles. The average user is 32 years old, with a standard deviation of about eight years. This figure indicates that *Status Money* users are younger than the average US population, which is common for app-based samples. The average credit score is 736, which is higher than the average credit score for the US population.[12] Thirty-nine percent of users are homeowners, which is below the US average, in line with the fact that *Status Money* users are, on average, younger than the average US consumer.

[12] Although we do not have access to administrative data on credit scores for a representative US sample, commercial reports indicate lower average credit scores for representative US samples. For instance, VantageScore computes that, in February 2021, a month that is close in time to the sample period we use in this paper, the average credit score was 698.

In terms of annual income, the average user earns a little more than $92,000 per year, but this average figure masks substantial heterogeneity as the large standard deviation suggests ($62,838). The identification sample in Panel B of Table 1, which excludes high-income users by design as we explain below, reveals that the baseline sample in Panel A has a fat right tail in terms of income. In that sample, the average income is $71,917, which is substantially lower than the average in the main sample and close to the US average. Panel A of Table 1 further shows that in the baseline sample, monthly spending in the 30 days before signup equals $4,963, with a standard deviation of $4,007, and monthly spending is similar in the second or third month before signup.

Panel B of Table 1 reports the summary statistics for the identification sample we use in section 6. This sample contains individuals whose income ranges between $6,000 below and $2,000 above each income threshold used to compute peer groups: $35K, $50K, $75K, $100K, and $150K. The summary statistics of this identification sample are in line with the ones of the baseline sample, with the exception of annual income and monthly spending, which equal $71,917 and $4,375, respectively. Panel B also describes four dummy variables that we use in the analysis of the causal effects of peer spending on users' spending based on the informativeness of the signals users observe. We define and discuss these variables, some of which are defined only for a subset of our sample and their interpretation in section 7.

Panel C of Table 1 describes our RCT sample, which we introduce and discuss in detail in section 8. The RCT allows us to elicit a set of demographic characteristics that are not available in observational data but can be important determinants of spending behavior according to earlier research, such as risk preferences, beliefs, gender, education levels, marital status and family size, political views, and financial literacy. At the same time, in the RCT sample we do not observe actual spending from transaction-level observational data, but we need to rely on subjects' reported levels of spending and especially on their marginal propensities to consume out of hypothetical windfalls, whose elicitation we discuss in section 8.

4 Peer Information and Spending: Raw-Data Evidence

We start our empirical analysis by describing the changes in spending after users observe their peers' spending relative to before signup. We then estimate the dynamics of the change in spending over time.

4.1 Reaction to Peer Information at Signup

Our baseline analysis tests whether two pieces of information users receive at signup—whether they spend more or less than their peers, and to what extent their spending differs from their peers' average spending—relate to users' subsequent spending behavior. We first compute the average monthly spending of each user for the 90 days before signup and the 90 days after signup and measure the change in aggregate spending across the two periods.

Because spending is seasonal and varies systematically across time within a year, we regress user-level spending on a full set of month-of-year dummy variables for the month in which the user signed up and use the residuals from this regression as the seasonally-adjusted level of spending at the user level.

Figure 1 in the introduction showed a graphical depiction of the change in users' spending after and before signup as a function of users' distance from the level of peer spending. We define distance from peers as the difference between user's spending in the month before signup and the spending of the assigned peer group. We classify users that spent more than their peers as overspenders, even though this terminology does not imply that users are deviating from an optimal spending path or that they are making a mistake. Figure 1 is a binned scatterplot that divides our 20,679 users into 80 groups. To ease the interpretation, we standardize the "Distance from Peers" x-axis variable to have a unit standard deviation. Figure 1 documents three features of the raw data. First, both underspenders and overspenders converge to the value of peer spending in the 90 days after signup, relative to before signup.

Second, the sensitivity of users' change in spending based on whether users spend more or less than their peers is asymmetric and larger in absolute value for overspenders than for underspenders. Overspenders reduce their spending on average by $231 per month and underspenders increase their spending on average by $71 per month. Third, the distance of users from their peers' spending is monotonically related to users' change in spending—the further an individual is from the peers' spending level, the higher the change in her spending, irrespective of the sign. The slope coefficients associated with the regression lines in Figure 1 are -27 for underspenders and -526 for overspenders. These coefficients imply that a standard-deviation increase in the distance from peers is associated with $27 higher monthly spending for underspenders. On the other hand, a standard-deviation increase in the distance from peers is associated with $526 lower monthly spending for overspenders.

Note that the average subscriber underspends relative to her peers, perhaps because app users are more attentive to their own finances relative to the broader population, as their decision to sign up to *Status Money* reveals. We tackle this important external-validity issue in the RCT we present in Section 8.

The results reported in Figure 1 consider the change in overall spending. Intuitively, we would expect that users can adjust their discretionary spending in the short run but not their non-discretionary spending. The app categorizes expense types (see Figure A.3 for a few examples). We report the changes in spending separately for discretionary and non-discretionary categories in Figure A.6,[13] which shows discretionary spending indeed fully drives the baseline results.

To exclude the possibility that less active users might drive our results above and beyond the sample selection steps we discussed above, in Figure A.7 of the Online Appendix we repeat the raw-data analysis on various sub-populations that are more and less likely to be affected by measurement error.[14] Specifically, in panels (a) to (c), we replicate our raw-data facts in the subset of users who link at least two accounts, users who are 35 years old or younger—and are less likely to hold many spending and investments accounts—and users whose income is below $200K, who again might be less likely to hold several accounts. All our raw data facts are similar within each sub-population.

Finally, in panel (d) of Figure A.7, we re-estimate our results excluding users whose spend-

[13]Spending is broken down into categories based on classifying the vendors of each transaction. Discretionary spending includes checking-account withdrawals, entertainment, restaurants, shopping, travel, and fees. Non-discretionary spending includes groceries, utilities and bills, health and medical, auto and gas, and education.

[14]We thank Byoung-Hyuon Hwang for proposing these tests.

ing is above or below the extreme 2.5% of the distribution of spending in both directions. The patterns we uncover are similar to the ones we obtain in the full sample, which shows our findings are not driven by the change in behavior of severe outlier/unusual users.

4.2 Endogenous Threshold Estimates

Figure 3 indicates that the app shows users information over and above their peer group spending, such as the average spending of US consumers or users' own average monthly inflows. The app also provides other pieces of information in locations other than the main home-page picture. One might be concerned that users reacting to information different from peer spending levels drive the cross-sectional distribution of spending changes around peers' spending levels.

To tackle this concern, first we test formally whether peer groups' spending does represent a threshold around which users react differently with two complementary methods, which are based on different assumptions. The first method builds on Hansen (1996, 2000). It estimates a threshold model with unknown threshold. To build intuition, consider the case of one regressor. The threshold regression estimates the optimal threshold for a linear model that has different intercept and slope estimates below and above the threshold. Hansen (1996) also proposes a test for whether the coefficient estimates below and above the threshold are statistically different from each other.

Based on this method, we estimate that the location of the threshold is -0.2 standard deviations away from peers' spending—a value that is economically close to zero and for which we cannot reject the null hypothesis that the threshold is at 0 at any plausible level of significance. In addition, the regression coefficient estimates below and above the threshold are statistically different from each other at the 1% level. Panel A of Figure 4 presents a binned scatterplot of our raw data in which we set the threshold to the non-parametric estimate we obtain when implementing this estimation method. None of the baseline facts we document in the raw data changes when considering this estimated threshold. These results suggest that a threshold in users' response to peers' spending information exists and we cannot reject the null that this threshold lies exactly at the value of peer spending either economically or statistically.

For the second method, we follow Hansen (2017) and estimate a regression kink model with unknown threshold. Contrary to the first method, this model does not allow for different intercepts above and below the kink. For this reason, this approach is similar to estimating a linear spline model that has a single endogenously determined node. Panel B of Figure 4 presents graphically the results when using this second method. In this case, the non-parametric estimate for the kink is 0.22 standard deviations away from peers' spending. Again, the estimated kink is close to zero economically and is not different from zero statistically. The test also reveals that the slope estimates below and above the kink are statistically different from each other at the 1% level.

Overall, both methods estimate the value of the threshold to be close to the level of peer spending, which users see on the app. We interpret the small absolute values of the two estimates around zero, their lack of statistical difference from zero, and the opposite signs when using different assumptions as evidence supporting the conjecture that users do indeed react to information about the level of peers' spending rather than to other information. Below, we implement additional tests focusing on users for whom different pieces of information predict reactions of opposite directions, which all suggest that users react to the information about peer spending rather than to other pieces of information they observe on the app.

4.3 The Dynamic Effect of Peer Information Over Time

The results so far compare users' average monthly spending three months after sign up to three months before sign up based on the information users observe about their peers' spending. This analysis cannot assess the dynamics of the effects and in particular cannot tell us whether the change in spending is temporary or permanent. Addressing this question is important to understand the mechanisms that drive our results as well as for the policy implications. If a temporary drop in spending by overspenders mainly drove the effects we detect and users reverted back to pre-signup spending in the medium run, the effectiveness of providing information about peers on adjusting consumption would be short lived. If, instead, users would consistently reduce their spending over time to reach the level of spending of their peers, providing information about peers could be a powerful tool to moderate overspenders' behavior.

To tackle this question, we propose a dynamic version of our baseline regressions in which we use data on monthly spending up to twelve months after signup. We compute the monthly average change in spending using the spending up to the n-th month after signup. The pre-signup spending is instead computed using the average monthly spending over the three months before signup.

Panel A of Figure 5, which refers to overspenders, reveals a gradual cut in spending that builds up over the 12 months after signup: after 12 months, overspenders have cut their spending by \$685 per month on average relative to their spending before signing up. Note that the average overspender spends \$4,000 more than her peers at signup. Hence, over twelve months, overspenders close this spending gap on average by \$685/ \$4,000=17.1%. The dynamic pattern for overspenders is consistent with the possibility that these users strive to reduce their monthly spending gradually and target the level of peers' monthly spending, which they learn on the platform.

Panel B of Figure 5 focuses on underspenders. First of all, note that because the number of underspenders in our sample is larger than the number of overspenders the confidence intervals around the estimated averages are tighter relative to those in Panel A. Moreover, we find that underspenders also increase their spending gradually over time, but this change is relatively smaller than the change for overspenders. After 12 months, underspenders have increased their spending on average by \$198, while their average underspending at signup equals \$4,400, so their spending gap closes only by \$198/\$4,400=4.5% on average.

Note that the results in Figure 5 include different sample sizes for each month after signup based on the number of users that survives on the app over time. For example, if we focus on under-spenders and 6-, 9-, and 12-month horizons, we use, respectively, 82%, 68%, and 57% of the data we had in our baseline estimates for each of those longer horizons, which correspond to the accounts that are not canceled over time. The corresponding percentages for over-spenders are 84%, 68%, and 58%. Censoring accounts for 16% of the attrition we observe. All our results are very similar if we repeat the analysis described in Figure 5 only within the sample of users who survive for at least 12 months after signup. Moreover, we detect no difference in the reaction to peer information by users that survive on the app at different horizons in the months in which they had an account on the app.

These patterns corroborate the static facts we documented above and in particular that all users converge to peers' spending level after learning about it, but overspenders display a larger convergence than underspenders at a similar distance from peers.

5 Multivariate Analysis and Alternative Explanations

In this section, we assess the robustness of the raw-data facts to alternative specifications and definitions of the change in spending. We also tackle a set of potential explanations alternative to users' reaction to peer information.

5.1 Variation in Spending Levels

First, the raw-data results do not account for the fact that a reduction in spending of $100 has different implications for users who have different levels of spending before signup. We thus harmonize the spending response and make it comparable across users with different levels of spending pre-signup by considering the ratio of spending over the 3 months after signup divided by the spending over the 3 months before signup. To ease the interpretation, we standardize this measure so that it has a mean equal to 0 and a unit standard deviation.

Table 2 reports the results when using this spending ratio as outcome variable that controls directly for pre-signup spending levels. The first two columns compute the average change in spending for those above peer consumption at signup—the overspenders—and those below peer consumption at signup, the underspenders, respectively. Confirming the robustness of the results in section 4.1 to normalizing the outcome variable by pre-signup spending levels, we find that overspenders significantly reduce spending, whereas underspenders significantly increase it after signing up.

The remaining columns of Table 2 estimate variations of the following specification:

$$\frac{Spending_{i,post}}{Spending_{i,pre}} = \alpha + \gamma \ Distance\ Peers_i + \delta \ \boldsymbol{x}_i + \epsilon_i, \tag{1}$$

where $\frac{Spending_{i,post}}{Spending_{i,pre}}$ is the standardized change in spending, $Distance\ Peers_i$ is the difference in spending in the one month before signup between the user and the peer group, $\boldsymbol{x}_i$ is a vector of regressors containing the following covariates: a dummy for homeownership, the log of the user's credit score, the log of the user's age, and the log of the user's asset and debt balances. Unfortunately, we do not observe all these covariates jointly for the full sample and hence the sample size is smaller for the multivariate specifications.

In columns (3) and (4) of Table 2, we find a significant relationship between distance from peers and the change in spending for both underspenders and overspenders, with a stronger effect for overspenders. Columns (5) and (6) replicate the analysis after adding the set of demographic characteristics we observe in the app. For both underspenders and overspenders, controlling for observable determinants of spending behavior barely changes the estimates of the regressor of interest either economically or statistically.

In Table A.1 we show that the results are similar across each of the sample selection steps we describe in section 3.1, including the full unselected sample.

5.2 Robustness and Alternative Specifications

Based on the day of the month in which users sign up, the graphics of their comparison with peers differ. In particular, Figure 3 shows that the number of days for which users' daily transactions are compared with peers' spending varies. Because we categorize all users as overspenders or underspenders based on their spending in the 30 days before signup, for users who sign up early in the month our categorization might in principle not align with what the users observe on the app. For instance, a user who faces large expenses early in the month and signs up at that time might see that she overspends relative to the peers at signup, even though the cumulative spending over the previous 30 days might be lower than the average monthly peer spending. In Table A.2 of the Online Appendix, we show that our results are similar for users who sign up in the first half or the second half of the month, which alleviates this concern.

In Table A.3 of the Online Appendix, we further assess the robustness of the results of Table 2 when estimating alternative specifications. We interact the distance from peers separately for users who lie above or below their peers in terms of pre-signup spending. This table confirms our baseline results both in terms of the significance of the spending changes and the larger effect for overspenders.

Moreover, in Table A.4 of the Online Appendix, we assess the robustness of our baseline results to alternative assumptions about statistical inference. In our baseline results, we estimate Huber-White heteroskedasticity-robust standard errors, because the outcome variables in our specifications are changes in spending for users who sign up at different points in time and the signup dates are unlikely to coincide with common shocks across users. For this reason, we do not expect that cross-sectional dimensions might correlate across users and hence that our estimated standard errors might be downward biased and the t-statistics upward biased. To assess this argument, in Table A.4, we cluster the standard errors of our baseline specification across each of three dimensions that might be relevant for common shocks in our setting—belonging to the same peer group, living in the same location, and signing up on the same date. We find that the estimated t-statistics are similar to those implied by Huber-White standard errors. Statistical inference is also similar when we estimate standard errors clustered three ways across all of these dimensions.

5.3 Reaction to Information or Mean Reversion in Spending?

Mean-reversion in spending could in principle account for some of the effects we document in section 4.1 and especially the larger effect of peer information on overspenders. Intuitively, a user who buys an iPhone in the month before signup is likely to spend less in the following month simply because the extraordinary iPhone expense is not recurring. To verify that mean reversion does not drive our results, we add the covariate $Spend\ Before_i$ to equation 1, which controls directly for users' spending in the month before signup.

Column (1) of Table 3 reports the results of column (3) of Table 2 as a benchmark. Column

(2) adds past spending as a control and column (3) includes past spending and additional user characteristics. The coefficients on past spending are negative, indicating that spending does mean revert. However, the coefficient on $Distance\ Peers_i$ is still highly economically and statistically significant. Mean reversion thus cannot drive our results because, once we control for previous spending directly, the results are qualitatively unaffected.

Although the coefficients on $Distance\ Peers_i$ remain significant, the inclusion of pre-signup spending as a control has an impact on the magnitude of the estimates. To ensure that controlling for mean reversion in alternative ways does not alter our findings, we propose additional tests in which we use alternative definitions of the distance from peer spending. In columns (4)-(6) we repeat the estimations of columns (1)-(3) but we compute the distance with respect to peers using spending in the second month before signup rather than the month before signup; columns (7)-(9) repeat the exercise using the third month before signup. These specifications also rule out directly that exceptionally large expenses in the month before signup drive our assignment of users to overspenders and hence that we misattribute changes in spending over time due to mean reversion to a reaction to peer information. These alternative ways of controlling for mean reversion do not alter our results.[15]

5.4 Information about Peers' Spending or Other Information?

As we discussed in Section 3, the app provides additional information to its users beyond the comparison with peers. The comparison between users' spending and peer spending is the first, most prominent, and most salient piece of information users see upon accessing the app during our sample period. The same picture, however, provides information about the average US consumers' spending, which is computed as the average spending of the observations used to create all peer groups, as well as information about the user's own average monthly inflows. Moreover, other pages of the app include comparisons between users' and peers' debt, assets, and net worth.

To further corroborate the role of peer spending information, in addition to the results for the endogenous threshold models we discussed above, we focus on users who lie above their

[15] Figure A.8 of the Online Appendix shows that also the raw-data facts discussed above are similar if we compare users with peers based on their spending well before signup.

peers in terms of spending but below their own average monthly inflows. If these users reacted to peers' spending information, we should observe a decline in spending that is proportional to their distance from peers. To the contrary, if these users reacted to learning that they spent less than their monthly inflows they would increase their spending or keep it unchanged. Columns (1)-(2) of Table A.5 in the Online Appendix show that users cut their spending if they are overspenders relative to their peers even if they spent less than their monthly inflows.

We then focus on overspenders relative to peers who learn that they spent less than the average US population. In this case, if users were reacting to information about the average US consumer, their spending should increase or stay the same, rather than decrease, after signup. Columns (3)-(4) of Table A.5 show that such users decrease their spending.

In the remaining columns of Table A.5, we assess whether other comparisons users can observe on the app—the level of their debt and net worth relative to peers—drive their change in spending. We define distance from peers debt and net worth in an equivalent way to how we measure distance from peers' spending. In both cases, we find that, once we control for users' distance from peers' spending, their distance from peers' debt or peers' net worth has no explanatory power for users' change in spending.

6 Assessing Causality: Quasi-Exogenous Matching to Peer Groups

The results so far cannot rule out a few endogeneity concerns. For instance, the endogenous timing of signup might be correlated with users' intentions to change their spending behavior irrespective of peer information. Signing up might be a way to track expenses more easily. In this case, overspenders might have cut their spending irrespective of whether the app compared them to their peers. Another potential concern, which we have already discussed, is that mean reversion in spending drives our results. Although we proposed direct tests to tackle the issue, one might still be concerned.

To address these concerns, ideally we would compare the reaction of users who have similar characteristics and spend similar amounts before signup, but observe different information about peers at signup. Our setting allows for this type of test. We exploit the fact that the app engages in bucketing to assign users to peer groups. In this way, users that are observationally similar but happen to lie just within or just outside a peer-group bucket will be matched to different peer groups and hence observe different information about peers' spending.

In particular, we consider income bucketing, because we observe income as a continuous variable and income is monotonically related to spending, whereas the association of other continuous dimensions such as age with spending is non-monotonic. *Status Money* matches users with peer groups based on whether users fall above or below a set of income thresholds of which users are not aware of at signup. The income thresholds are \$35K, \$50K, \$75K, \$100K, and \$150K. Because income is a continuous variable, small differences in income relate to otherwise similar consumers, as we document below. And, yet, the spending of similar users who are matched to different peer groups reacts differently after sign up relative to before.

For example, consider a user who reports an annual income of \$99K and one who reports an annual income of \$101K. We can test directly that these users are not distinguishable along the set of characteristics we observe, which include demographics and the spending behavior in the months before signup.

Whereas users around the income values are similar, the peer groups to which they are matched, and hence the information about peers' spending they observe, can be substantially different. Due to bucketing, in our example a user who reports an income of \$99K receives information about the average spending of peers whose income is between \$75K and \$99K, whereas a similar user who reports an income of \$101K receives information about the average spending of peers whose income is between \$100K and \$149K. The user matched to the lower peer group will thus be more likely to be classified as an overspender relative to her peers, whereas the similar user matched to the higher peer group will be more likely to be classified as an underspender.

To implement our identification test, we need to restrict the sample to users who are close enough to each income threshold so that they do not differ based on observables. Moreover, we need to have a large enough number of users both below and above each threshold. Because some users tend to report rounded values of annual income, the mass of users just above each income value (which includes the value) is higher than the mass of users just below. In our identification sample, we thus include users with an income up to \$6K below each income value and users with an income up to \$2K above the value to guarantee samples of similar sizes above and below each value. We show that our results do not change if we use symmetric bandwidths around income values, which though imply samples of different sizes for underspenders and overspenders.

We exploit the quasi-random assignment of users to peer groups based on these pre-defined income bucketing rules for an instrumental-variable identification strategy. Note that we cannot implement a regression discontinuity design in this setting because peer groups are determined based on six different observable characteristics at the individual level and not just based on income. For dimensions other than income, we do not have a clearcut prediction about whether average peer spending should be higher or lower at either side of the bucketing threshold.

6.1 Balancing of Observables across Income Bucketing Levels

Before proceeding with the analysis, we verify that we cannot reject the null hypothesis that users who end up just below and just above the income bucketing values differ on observables. We start by providing graphical evidence in Figure 6. For each income value, we plot the estimated average monthly spending and confidence intervals for the three months before signup (in thousands \$) for users around the income values at intervals of \$500 of annual income.[16] For each income level, we cannot reject the null that users above and below spent the same amounts before signup. The highest threshold (\$150K) is the one for which average pre-spending is more varied across income levels, largely because of the smaller sample, but even for this threshold we do not detect systematic economic or statistical patterns.

We also assess the balancing of the observable characteristics that entered our baseline analysis in Table 4. For each variable and each income bucketing level, we report the results of regressing the variable on a dummy that equals 1 if the user is above the threshold and

[16] Confidence intervals are based on one third of a standard deviation above and below the average.

0 otherwise. Each panel of Table 4 refers to one of the income buckets. Overall, we fail to detect any economically or statistically significant patterns. Although for two individual coefficients we reject the null, the vast majority of estimated coefficients is not different from zero economically or statistically. Crucially, even for the coefficients that appear significant, we detect no systematic patterns in terms of signs. Overall, the observables appear balanced above and below each income level.

To further support the exclusion restriction we need to assume to interpret our IV results causally, in Figure 7 we show that demographic characteristics such as age, credit score, log of asset balance and log of debt balance are similar above and below the income bucketing levels.[17] This evidence dismisses the potential concern that some users manipulate their position relative to their peers and supports the assumption that the assignment to one or the other peer groups is quasi-exogenous in this identification sample.

6.2 IV Results

Armed with our identification sample, we use the dummy for whether users are above the income thresholds that place them in different peer-group buckets as an instrument for the peer spending such users see after signup. A user at or just above the income threshold is assigned to a peer group whose income is, on average, higher, because the user is at the bottom of the peer distribution based on income. The user will thus be more likely to observe peer spending that is higher than her own spending. The opposite is true for users just below the income threshold, who are assigned to a peer-group bucket for which they are at the top of the peer distribution by income. These users are more likely to be classified as overspenders with respect to their peers, relative to users above the thresholds. We can test for the relevance of our instrument in the first stage of our two-stage least-squares model:

$$Peer\ Spending_i = \alpha + \beta\ Dummy\ Above_i + \zeta\ Spending\ Before_i + \epsilon_i, \tag{2}$$

[17]Figure 7 focuses on the \$50K income value to avoid reporting many panels, but the results are similar for all the other income values.

where $Peer\ Spending_i$ is the peer spending user i sees at signup, and $Dummy\ Above_i$ is a dummy variable that equals one for users' with incomes at or just above the threshold. Based on the design, we predict that $\hat{\beta} > 0$—users at or above the threshold will be matched to a peer group whose spending is higher, on average, relative to those below the threshold.

Note that the endogenous variable we instrument in this design is peer spending rather than the dummy variable for whether users are overspenders relative to their peers, because instrumenting the dummy would create a forbidden regression problem whereby the assumption on the linearity of the conditional expectations function of the 2SLS design does not hold. This problem does not arise when instrumenting a continuous variable such as the level of peer spending.

The second stage uses the instrumented $Peer\ Spending_i$ in equation (2) as the main covariate in the following specification:

$$\frac{Spending_{i,post}}{Spending_{i,pre}} = \alpha + \gamma\ \widehat{Peer\ Spending}_i + \zeta\ Spending\ Before_i + \epsilon_i, \tag{3}$$

where $\frac{Spending_{i,post}}{Spending_{i,pre}}$ is the ratio of post- to pre-signup spending in the three months around signup.

We first verify the relevance of our instrument by assessing the results from the first-stage regressions in column (1) of Table 5. Consistent with the app's assignment rule, users just above each income level do observe a higher peer spending relative to users below, who are assigned to a different peer group—about three quarters of a standard deviation higher. The first-stage F statistics is 606, which limits concerns of a weak instrument.

The second-stage estimates are reported in column (2). Even in the second stage, our hypothesis is that $\hat{\gamma} > 0$—the higher the peers' spending, the more likely the user is an underspender and the less likely she is to cut her spending. Conversely, the lower the peers' spending, the more likely she is to be an overspender and the more likely she is to cut her spending.

In terms of economic magnitudes, note that the variable $Peer\ Spending_i$ is standardized and the sample standard deviation is \$3,100. Therefore, our estimate in column (2) implies that users who are assigned to peers whose spending is \$3,100 lower than other peer groups cut their spending by about 11% more of their pre-sign up spending relative to other users.

As discussed above, our IV analysis uses asymmetric bandwidths around the income bucketing levels to allow for a similar number of observations both above and below the levels. This choice is not relevant as long as the instrument is valid in the identification sample, but to further dismiss any concerns, in Table A.6 of the Online Appendix we show that the IV results do not change qualitatively or quantitatively when we use symmetric bandwidths to select the observations in the identification sample and hence samples that are not balanced in terms of the number of observations around the income bucketing levels.

6.3 Placebo IV Analysis

To corroborate our interpretation of the IV analysis and the exclusion restriction we assume for a causal interpretation, we propose a placebo test. We set alternative placebo income bucketing thresholds different from the true thresholds that *Status Money* uses to assign users to peer groups but including a similar-sized mass of users.

To determine the placebo levels, we use three criteria. First, we ensure that none of the users in the IV sample also enters the placebo IV sample; that is, we guarantee no overlap across the income values of the two IV analyses. This criterion ensures that none of the results might be driven by users' reaction to the actual assignment to different peer groups. Second, we ensure the placebo income levels are set at round dollar values like the true levels. This criterion allows us to rule out that our IV results might be related to systematic differences of the behavior of reported-income rounders and non-rounders. Third, we ensure that the placebo IV sample is large enough to dismiss that the lack of statistical power can be responsible for any non-results in this sample.

These three criteria lead us to select the following values as placebo income thresholds: \$45K, \$60K, \$90K, \$110K, and \$140K. We construct the placebo IV sample exactly as we did in the IV analysis, and we estimate equations (2) and (3) for this placebo sample. Of course, our predictions about the coefficients are now different: In equation (2), the first stage, we expect that $\hat{\beta}$ is not economically or statistically different from zero—being above or below the placebo value should not predict the dollar amount of peer spending, because users above and below the peer cutoff are not systematically assigned to different peer groups. In terms of second stage, estimating it when the first stage shows the placebo instrument is not relevant is meaningless. If we were still proceeding with the estimation, we would expect that $\hat{\gamma}$ "blows up;" that is, it becomes large in any direction and statistically insignificant, because the variation we use to instrument for the endogenous variable is unrelated to the actual variation in the data. Columns (3)-(4) of Table 5 show that the first stage of this placebo IV analysis does not go through and consistently the second stage blows up.

The placebo IV analysis also dismisses that users who bunch at rounded income values might behave differently from others. This phenomenon would be worrisome if the unobservables that make certain users provide rounded values of annual income also predict systematically different reactions to peers' information. Figure 8 shows that bunching at rounded values holds for both the actual bucketing income thresholds and the placebo thresholds. If anything, the largest mass of respondents bunches at one of the placebo thresholds we use—$60K. Despite the commonality of bunching across all rounded values, our placebo IV analysis finds no effect of being at the (rounded) values relative to below, which dismisses the concern.

7 Informativeness and Spending Behavior

As we discussed in section 2, another unique feature of our setting is the heterogeneity in the informativeness of peer groups, whose average characteristics are more or less tightly defined around users' own characteristics. We can test whether users who are assigned to more similar, and hence informative, peer groups react more to peer information relative to others.

On top of providing another test for the information channel, this analysis allows us to separate the role of information about peer spending from the role of any other design features overspenders might see on the app, such as colors, bars, or text (e.g. see (Bazley et al. (2020))), because these features are the same for all overspenders irrespective of the informativeness of their peer groups.

7.1 Tightness of Peer-Group Characteristics

Our first proxy for informativeness exploits the fact that *Status Money* imposes a minimum number of 5,000 underlying observations when constructing each peer group to make peer-group averages meaningful. Based on this rule, the ranges of peers' characteristics are tighter or wider based on how many observations similar to the user exist in the external representative data used to construct peer groups.

For instance, suppose two users have the same characteristics under all dimensions except their location, which might be Manhattan, NY, for user A and Helena, MT, for user B. Whereas the app might easily find 5,000 observations in Manhattan with the characteristics of user A, it is likely to miss the same number of observations for Helena, MT. For user B, then, the app would enlarge the geographic area of peer comparison to the overall state of Montana, and if enough observations were still not available, to the overall US. Users know the characteristics of the peer group, and hence user A knows she is compared with similar peers living in her same location, whereas user B knows she is compared with peers who are similar along all the characteristics except for their location. Intuitively, then, if users attach any information value to the peer signal, user A should believe that the signal she gets is informative, whereas user B might not.

Based on this intuition, we add an interaction term between peer-group spending and a dummy that equals one for users whose characteristics are all matched to the closest possible range and zero for users for whom at least one characteristic is matched to a broader range (less informative). We then estimate the two-stage least-squares specification in equations (2) and (3) and report the second-stage estimates in column (1) of Table 6. In line with our conjecture, users for whom the peer spending signal is more precise drive the spending reaction. The difference in the change in spending after signup is about three times larger for users who observe precise peer groups, and the difference is not statistically different from zero for other users.

In column (2) of Table 6, as in all the even columns of the table, we repeat the analysis after controlling directly for users' income and the interaction of users' income with the informativeness of peer groups. This analysis aims to address the concern that the variables we use to proxy for the informativeness of peer groups might at the same time capture systematic differences in users' income. This concern is compelling, because users whose income is in the right tail of the distribution are by construction more likely to be matched with an imprecise peer group relative to users with lower levels of income. At the same time, we find that high-income users react less to peer information relative to low-income users, as we show in panel (a) of Figure A.9 in the Online Appendix.[18] Controlling for income and its interaction with the peer-group-similarity dummy does not change our results in terms of the effect of informativeness on spending reaction.

7.2 Size of Peer Groups

Our second proxy for informativeness is the size of peer groups. The number of peers in the group to which users are assigned is displayed in a salient fashion to users in the top-right corner of the peer group figure, as we show in Figure 2. The larger is the sample size of the peer group whose average spending is computed, the less noisy and more informative is the estimated sample mean. One might worry that, if users lack statistical literacy, they might ignore that a sample average estimated on a larger sample is likely to be closer to the true population parameter than a sample average estimated on a smaller sample. But due to the principle of the "wisdom of the crowd," at a very minimum users should think that a behavior in which a large number of agents engages in is more likely to be common than a behavior in which only a small number of agents engages in.

In columns (3) and (4) of Table 6, we interact peer spending with a dummy variable that equals one if users are matched to peer groups in the top quarter of the distribution by size (larger or equal to 21K peers), and zero if they are matched to peer groups in the bottom quarter (lower or equal to 6K, with 5K being the minimum size of peer groups). Consistent with a stronger reaction for more informative peer groups, the effect is about twice as large for users assigned to a larger peer group. We also reject the null that the effect is zero in the specification that includes controls for income in column (4).

[18] In Figure A.9, we show that we do not detect systematic differences in reactions across other demographic characteristics of users, such as age or credit scores.

7.3 Peer-Group Income-range Width

Third, we consider the fact that the peer categories users observe have different ranges of values. For example, consider income. Because the income distribution in the population is far from uniform—the mass of individuals with middle income levels is substantially larger than the mass of individuals with high levels of income—in a representative sample, obtaining a large mass of peers for tight ranges of income values among mid-income earners is relatively easy, whereas ranges need to be wider at higher income levels to ensure a similarly large number of peers.

For this reason, the thresholds of income we use in the IV analysis imply tighter and wider peer groups. For instance, a user who earns \$42K a year will be assigned to a peer groups whose range of income is between \$35K and \$49K, which is relatively tight. Instead, a user who earns about \$300K will be assigned to a peer group including individuals who earn more than \$150K, which is a substantially wider group. For such a broad group, average peer spending might not be informative.

In columns (5) and (6) of Table 6, we interact the peer group spending with a dummy variable that equals 1 for users who are matched to a peer group whose income width is \$50K or above, and zero if the range is \$15K or below. Users for whom the peer group is more narrowly defined display an economically large reaction. By contrast, the estimated effect for users who are assigned to wider peer groups is close to zero economically and statistically. Note that these tests keep constant users' income and hence we only capture the effect of the width of peer groups.

7.4 Direct Access to Peer Information before Signup

Another dimension along which the signal about peer spending might be more or less informative is the extent to which users could have been exposed to information about their peers' actual overall spending before signing up. To obtain variation, we use users' location as a proxy for the density of information. Specifically, we conjecture that users who live in less crowded and dense areas might have fewer peers, that is, people who look similar to them among most characteristics, while both types of users might access similarly biased views on spending through social media. Instead, users who reside in highly dense urban locations are exposed to more peers in their daily lives and hence to a higher density of information about peers before signup. Table A.7 in the Online Appendix reports examples of locations that are categorized as urban or rural. Urban locations include the largest metropolitan conglomerates in the US. The rural group includes smaller cities and towns, such as Tucson (AZ), Tallahassee (FL), and Chapel Hill (NC).

Based on our conjecture, users from less dense towns should react more than urban users, which is what we document in columns (7) and (8) of Table 6.[19] We find large and significant effects of peer spending on users in rural areas, whereas the effect for users in urban areas is close to zero.

Although the test based on urban-rural locations can be interpreted as our conjecture suggests, we acknowledge that living in urban or rural environments might be correlated with many unobserved characteristics of users and hence this fourth test is a less precise test for the role of peer signal informativeness compared to the others we discussed above.

8 External Validity: Randomized Control Trial

Due to the setting, our results so far cannot speak to two important issues. The first is external validity. All our results are based on a sample of users who self-select into the app. Even though our IV design supports the causality of peer-group information on spending changes, this effect might be peculiar to users of the app and might not hold in the general population, because app users in our setting cared enough for obtaining information about their peers that they decided to sign up. Other parts of the general population might not be interested in peer information, and providing them with peer information might not change their spending behavior. Moreover, like in all other income aggregators studied in the literature,[20] the app

[19]Note the urban and rural subsamples do not sum up to the full identification sample. This discrepancy arises because the app allows for a third location category labeled "suburban." This category includes locations that are at times highly urbanized, such as the suburbs of large metropolises, or quite rural. For this reason, we exclude this group from the split analysis. When we look at this group separately, we find an effect similar in size to the effect for the urban subsample.

[20]For instance, see Gelman, Kariv, Shapiro, Silverman, and Tadelis (2014), Gelman, Kariv, Shapiro, Silverman, and Tadelis (2018), Olafsson and Pagel (2018), Hau, Huang, Shan, and Sheng (2019).

is only meaningful to users if they provide sensitive information. One might wonder whether individuals who are willing to provide such sensitive information might also be more likely to react to peer information than the general population, which would cast doubt on the external validity of our results.

The second issue we face is that *Status Money* does not elicit some demographics that earlier research shows are important to determine spending behavior (D'Acunto et al., 2016, 2021). For instance, we do not observe users' gender, their education levels, their marital status, or whether they have children. We have no compelling reasons to think that any of these unobservables might vary systematically around the thresholds of our IV analysis, even though we cannot test this claim. The main reason why we are concerned about unobserved characteristics is not that the baseline results could be explained by them, but that the baseline effects might be higher or lower across different demographics—unobserved characteristics might interact with peer information to determine heterogeneous effects.

The ideal test to address these issues would be an RCT in which we randomly provide a representative sample of the general population with the app, and we test the effect of providing peer information on the change in the spending of these agents. Unfortunately, we cannot run this ideal test, because, by construction, we need users to give us their account information and we cannot force members of the general population to do so.

To get as close as possible to the ideal test, we designed the following RCT. We recruited a representative US population through an online platform, Amazon Mechanical Turk (mTurk), which is increasingly used for experimental research in the US and abroad.[21] Figure A.10 of the Online Appendix plots the geographic distribution of the respondents, who are distributed across the whole US. We invited respondents to answer a survey without any references to peer information, peers, or household finances. In this way, none of the respondents we recruited could have decided to participate because they were interested in information about their peers or in their own finances, which is the main concern of our baseline analysis in terms of external validity.

[21]For instance, see Kuziemko et al. (2015), DellaVigna and Pope (2017), Bazley, Cronqvist, and Mormann (2020), Lian, Ma, and Wang (2019), D'Acunto, Hoang, Paloviita, and Weber (2020), D'Acunto (2018), D'Acunto (2019), and D'Acunto et al. (2021).

We report the survey questions in the Online Appendix. We followed the work of Attanasio and Weber (1995) and subsequent literature (Jappelli and Pistaferri (2010)) to elicit respondents' marginal propensity to consume (MPC) before and after receiving truthful information about the MPC of representative individuals with similar income levels (income peers).

After eliciting users' MPC, we provide them with truthful information about the MPC of US consumers that are similar to them in terms of income (income peers). Theoretically and empirically, income is a strong predictor of MPCs (Coibion et al. (2020)). We call participants with a MPC higher than their income peers overspenders and others underspenders. Finally, we elicit survey participants' MPCs again after observing their income peers' MPCs.

The RCT mimics our baseline field setting by providing peer spending information, comparing it to respondents' own MPC, and assessing any changes in respondents' reported spending behavior. The RCT has the advantage of keeping constant dimensions that we cannot observe in the baseline field analysis.

The RCT faces the potential issue of demand effects: respondents might infer the hypotheses we want to test and, in a setting in which stakes are no as high as in the field, might conform to our hypothesis. De Quidt et al. (2018) show that the scope for demand effects is minimal in settings like ours. Moreover, we document below asymmetric responses in the survey as we see in the field setting, which directly rules out demand effects or anchoring as drivers of our RCT findings. To further reduce the concern of demand effects, we told respondents explicitly on two occasions that there were no right or wrong answers and that all answers would be of extreme interest to us as long as they reflected the respondents' true opinions.

The survey consisted of three parts. In the first part, we elicited respondents' age and income groups. We proposed three income buckets based on the values of income groups for which Parker and Souleles (2019) report the marginal propensity to spend an unexpected reimbursement.

In the second part of the survey, subjects read information about their income peers' average MPC. We then elicited subjects' MPC following Jappelli and Pistaferri (2014) both before and after delivering peer information. Moreover, because the survey was run during the

COVID-19 pandemic, we asked respondents to provide their MPC based on their situation at the time of the survey as well as before the COVID-19 pandemic started. The treatment's size depends on subject's income group as well as on their reported MPC before the information provision.

In the third part of the survey, we elicited a set of demographics such as gender, marital status, number of children, and political leanings, as well as economic preferences and beliefs—risk aversion, patience, financial literacy, and generalized trust.

8.1 Effect of Peer Information on a Non-selected Population

Armed with this RCT, we can first assess if this non-selected draw of the US population responds to the provision of peer information about spending in a similar manner as *Status Money* users. In Figure 9, we plot the changes in reported MPCs for the whole sample after sorting respondents based on the distance of their MPC from income peers' MPC. In Table 7, we report the average change in reported MPC before and after the provision of income peers' information, for both MPCs in normal times and MPCs during the COVID-19 pandemic. In both cases, respondents converge to income peers' behavior and the effect is stronger for respondents whose MPC was higher than their income peers' and for respondents who were further away from their income peers before the information provision.

We then ask if the spending response to peer information varies by demographic groups, some of which might be overrepresented on *Status Money*. Reassuringly, in Table 8 we find that the reactions are economically and statistically significant for each subsample of RCT respondents split across demographics that are important determinants of spending behavior, such as gender, marital status, number of children, and individuals who are more or less risk averse.

Finally, we can use our RCT to assess the average overall effect of the intervention—providing consumers with information about their income peers' MPC—on the population. Because consumers who are especially concerned about their finances or especially sensitive to peers' choices might select into *Status Money*, we cannot extrapolate the baseline analysis results to make claims about the general population.

The drawback from inferring population-level effects from our RCT is that the MPC is self-reported and not based on actual spending data. We assess spending plans rather than actual amounts spent. Note, though, that economic research based on which we design our MPC elicitation questions finds that elicited MPCs in surveys map closely into average MPCs observed in actual consumption data (e.g., see Parker and Souleles (2019); Jappelli and Pistaferri (2014); D'Acunto et al. (2016)).

The average reported MPC in normal times before any information about peers in our survey sample is 45.6%, and declines to 44.7% after respondents observe peer information. Overall, providing the whole population with information about peers reduces the average MPC by 0.9 percentage points, which is about 2% of the pre-information average MPC. This effect is, on average, negative, because we detect an asymmetric effect of peer information based on whether respondents are above or below the peers: Those who have higher MPCs than their income peers react disproportionally more to the information than those who have lower MPCs than their income peers. This asymmetry suggests that, when peer information is provided to a population who did not select into observing such information, this information has a moderating effect, on average, in terms of spending plans. A moderating effect of peer information has also been detected for expert decision-makers such as the CEOs of listed companies (e.g., see D'Acunto, Weber, and Xie (2019)).

For respondents who accessed the RCT with a higher MPC than their peers, the average MPC was 66.6% and declined to 62.9% after observing peer information, which is a drop of 3.7 percentage points, or 5.6% of the pre-information average MPC.

Interestingly, respondents' MPCs are *lower* in times of crisis than in normal times, which seems at odds with the standard life-cycle consumption framework but in line with evidence from recent episodes of spending behavior in response to government transfers during the COVID-19 crisis (Coibion et al. (2020)). This result is consistent with recent evidence in macroeconomics about heightened precautionary savings motives in times of economic crises (e.g., see Coibion et al. (2021) and D'Acunto et al. (2020))

9 Implications for Competing Theories of Overconsumption

The main contribution of our paper is to isolate and quantify the information channel of peer effects in a setting that rules out a role for common shocks across peers or peer pressure and in which connections among peers are not formed endogenously. However, an additional key value added of our setting and identification strategy is that we can provide insights on the extent to which several mainstream theories of overconsumption/ undersaving behavior are more or less consistent with empirical evidence.[22]

In what follows, we consider both preference-based theories (Present Bias and Keeping Up with the Joneses) and information-based theories (Wealth Signaling, Speculative Disagreement, and Visibility Bias).

Present Bias. In Laibson (1997), present-biased agents overconsume because they discount the future by more than a standard neoclassical agent. Specifically, the model posits a functional form for the discount rate whereby the present-biased agent's preferences converge to those of a neoclassical agent in the long run but differ in the short run ("$\beta - \delta$ preferences"). When optimizing intertemporal consumption based on present-biased preferences, this feature of discounting leads agents to systematically frontload consumption in each period and discount the utility from subsequent-periods' consumption by more than a standard agent, and hence ultimately to undersave relative to standard agents. Present bias theory is, therefore, an asocial preference-based theory in the sense that others' behavior does not affect agents' consumption choices. By construction, a standard form of present bias theory cannot explain the key mechanism we uncover in the paper, whereby agents change their consumption choices after observing information about peers' spending. To explain this finding, one would need to incorporate a form of learning about one's own discount rate based on signals about others' discount rates in the baseline present bias theory. In this case, agents would update their discount rates based on signals about others' consumption. At the same time, such adapta-

[22]We are indebted to David Hirshleifer for inspiring this section based on his discussion of our work at the AI & Big Data in Finance Research (ABFR) Forum.

tion would still not produce the asymmetric effects of peers' spending information on agents' own choices based on whether they spend more or less than their peers. More generally, such adaptation would change the nature of the present bias theory from a preference-based model of overconsumption into an information-based model that includes a role for socialization.

Keeping up with the Joneses. Dupor and Liu (2003) proposed another successful preference-based theory of overconsumption in which agents consume in coordination with the consumption of others. This preference-based model thus includes a role for socialization. At the same time, in its standard form, the model is deterministic, and agents are assumed to know the consumption levels of their neighbors (”Joneses”)—they do not need to learn it. In this setting, disclosing information about peers' expenditures should not affect agents' consumption choices because agents learn nothing from the information that is provided to them, which they already know. The original model could be extended by introducing uncertainty about peers' consumption levels, in which case a reaction in terms of consumption choices when agents learn information about their peers' spending would arise, and this reaction would be stronger the more informative are peers' signals. This adjusted model, though, could barely explain agents' asymmetric reactions based on their position relative to peers. If anything, the model would predict a stronger reaction and convergence for underspenders, who want to keep up with peers' consumption levels, relative to overspenders, who are already "ahead of the Joneses." These patterns are the opposite of what we find non only for the users on the app (who might care about savings more than others) but also in the general population through our RCT.

Wealth Signaling. Bagwell and Bernheim (1996) and Charles et al. (2009) propose an information-based theory of overconsumption whereby agents want to be perceived as wealthy and learn about others' wealth by observing others' consumption patterns. Agents do not, however, learn about how much they should consume by observing others' consumption patterns, at least partly because agents do not know each individual's wealth level but just the overall wealth distribution in the economy. For this reason, in its original form, this theory does not accommodate all of our main findings. One could envision an extension of the original model whereby individuals do not know how wealth is distributed in society. In this case, disclosing information about peers' spending, like in the app we study, might help agents learn about such distribution. This extension would be consistent with the consumption changes by the users in our setting after observing the information displayed by the app. However, even in this case, reactions would not necessarily be asymmetric based on agents' position relative to their peers. To explain all our facts, additional modifications to the original model should be introduced.

Speculative Disagreement. In this information-based theory, Heyerdahl-Larsen and Walden (2022) posit that agents with strong priors disagree about future economic conditions and bet against each other in asset markets. Each agent believes she will become wealthy by trading against those with starkly different priors about future economic conditions. For this reason, all agents (irrespective of the sign of their priors) have high expectations about future wealth levels, which might lead them to overconsume. Based on the original formulation of the model, disclosing information through the app should not affect agents' consumption choices because agents have strong priors and do not update their beliefs over time. One could envision an extension of the original model whereby agents update their beliefs after observing others' consumption choices. In this case, agents should change their consumption behavior after observing the information displayed by the app. However, and similar to the other theories discussed so far, the asymmetric reaction to peer information provision across overspenders and underspenders would remain unexplained.

Visibility Bias. Han et al. (2019) propose another information-based theory of overconsumption, according to which agents do not know their optimal savings rate and rely on others' consumption as a signal to drive their choices. Agents do not observe the full set of others' consumption transactions but only observe the most visible portion of others' consumption, which tends to be the most conspicuous. As a result, agents tend to overestimate other people's consumption, which leads them to overspend/undersave. In this model, the disclosure of information about peers' full consumption expenditures should cause them to change their consumption in order to converge to their peers' consumption, which is in line with our main findings. In addition, the effect should be stronger for more informative disclosures—another fact we find in our empirical analysis. In addition, some model parameterizations in Han et al. (2019) can also generate the asymmetric effect we document, whereby overspenders decrease consumption by more than underspenders increase consumption after everybody learns information about peers' overall spending.

To sum up, our empirical results have implications for theories of overconsumption: They provide novel empirical support for information-based theories of overconsumption given that the agents in both our settings—app users and a non-selected RCT population—react to information about peers' overall spending and MPCs. Within the set of information-based models, the Visibility Bias Theory of Han et al. (2019) can, even in its original form, explain the broader set of empirical facts we document in the paper. Non-information theories, on the other hand, are unable to explain our findings because they would predict no reaction on the part of app users upon receiving information about their peers' consumption unless a role for learning about uncertain model parameters was introduced.

10 Conclusions

We study the effects of providing consumers with crowdsourced information about unknown demographic peers' spending through a FinTech app. We find that all users, on average, converge to their peers' spending, but the effect is stronger for users who learn that they overspend relative to their peers. Moreover, users who are exposed to a more informative signal react more to peer spending relative to other users and the effects build up persistently over time. We replicate these results in an RCT on a population that is not selected based on their interest in observing peer information, which supports the external validity of our findings.

Our findings speak to the growing use of algorithmic advice in economic agents' decision making. To the best of our knowledge, this is among the very first studies of the effects of robo-advising tools that target consumption/saving choices (D'Acunto and Rossi (2020)), which is perhaps the most important economic decision consumers make through their life.

Future research should investigate the optimal design of such tools as well as the economic channels through which their effects are transmitted to decision-makers.

Moreover, whether reacting to peer information is optimal and whether peer spending contains any relevant information for consumers is an open question that should be addressed in future research. If increasing the average saving rate in the economy was the policy objective, peer information might be a successful tool due to the larger reaction of users who lie above their peers. Our results, though, indicate the scope for heterogeneous and redistributional effects that future research should isolate and study.

References

Agarwal, S. and W. Qian (2014). Consumption and debt response to unanticipated income shocks: Evidence from a natural experiment in singapore. *American Economic Review 104*(12), 4205–30.

Allcott, H. (2011). Social norms and energy conservation. *Journal of public Economics 95*(9-10), 1082–1095.

Allcott, H. and T. Rogers (2014). The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. *American Economic Review 104*(10), 3003–37.

Attanasio, O. P. and G. Weber (1995). Is consumption growth consistent with intertemporal optimization? evidence from the consumer expenditure survey. *Journal of Political Economy 103*(6), 1121–1157.

Bagwell, L. S. and B. D. Bernheim (1996). Veblen effects in a theory of conspicuous consumption. *The American economic review*, 349–373.

Bailey, M., R. Cao, T. Kuchler, and J. Stroebel (2018). The economic effects of social networks: Evidence from the housing market. *Journal of Political Economy 126*(6), 2224–2276.

Banks, J., R. Blundell, and S. Tanner (1998). Is there a retirement-savings puzzle? *American Economic Review*, 769–788.

Bazley, W., H. Cronqvist, and M. Mormann (2020). Visual finance: The pervasive effect of red on investor behavior. *Management Science, forthcoming*.

Bernheim, B. D., J. Skinner, and S. Weinberg (2001). What accounts for the variation in retirement wealth among us households? *American Economic Review 91*(4), 832–857.

Beshears, J., J. J. Choi, D. Laibson, B. C. Madrian, and K. L. Milkman (2015). The effect of providing peer information on retirement savings decisions. *The Journal of Finance 70*(3), 1161–1201.

Bikhchandani, S., D. Hirshleifer, and I. Welch (1998). Learning from the behavior of others: Conformity, fads, and informational cascades. *Journal of economic perspectives 12*(3), 151–170.

Bursztyn, L., F. Ederer, B. Ferman, and N. Yuchtman (2014). Understanding mechanisms underlying peer effects: Evidence from a field experiment on financial decisions. *Econometrica 82*(4), 1273–1301.

Chak, I., K. Croxson, F. D'Acunto, J. Reuter, A. Rossi, and J. Shaw (2022). Robo-advice for borrower repayment decisions. *FCA Occasional. paper* (61).

Charles, K. K., E. Hurst, and N. Roussanov (2009). Conspicuous consumption and race. *The Quarterly Journal of Economics 124*(2), 425–467.

Coibion, O., D. Georgarakos, Y. Gorodnichenko, G. Kenny, and M. Weber (2021). The effect of macroeconomic uncertainty on household spending. Technical report.

Coibion, O., D. Georgarakos, Y. Gorodnichenko, and M. Weber (2020). Forward guidance and household expectations. Technical report, National Bureau of Economic Research.

Coibion, O., Y. Gorodnichenko, and M. Weber (2019). Monetary policy communications and their effects on household inflation expectations. Technical report, National Bureau of Economic Research.

Coibion, O., Y. Gorodnichenko, and M. Weber (2020). How did us consumers use their stimulus payments? Technical report, National Bureau of Economic Research.

D'Acunto, F. (2018). Identity and choice under risk. *Working Paper*.

D'Acunto, F. (2019). Tear down this wall street: Anti-finance rhetoric, subjective beliefs, and investment. *Working Paper*.

D'Acunto, F., A. Fuster, and M. Weber (2021). Diverse policy committees are more effective. *Working Paper*.

D'Acunto, F., D. Hoang, M. Paloviita, and M. Weber (2019). Human frictions in the transmission of economic policy. *Working Paper*.

D'Acunto, F., D. Hoang, M. Paloviita, and M. Weber (2020). Iq, expectations, and choice. *Working Paper*.

D'Acunto, F., D. Hoang, and M. Weber (2016). The effect of unconventional fiscal policy on consumption expenditure. Technical report, National Bureau of Economic Research.

D'Acunto, F., D. Hoang, and M. Weber (2021). Managing households' expectations with unconventional policies. *Review of Financial Studies (forthcoming)* (19-16).

D'Acunto, F., U. Malmendier, and M. Weber (2020). Gender roles and the gender expectations gap. Technical report, National Bureau of Economic Research.

D'Acunto, F., N. Prabhala, and A. G. Rossi (2019). The promises and pitfalls of robo-advising. *The Review of Financial Studies*.

D'Acunto, F., T. Rauter, C. K. Scheuch, and M. Weber (2020). Perceived precautionary savings motives: Evidence from fintech. Technical report, National Bureau of Economic Research.

D'Acunto, F. and A. Rossi (2023). Robo-advice: Transforming households into rational economic agents. *Annual Review of Financial Economics*.

D'Acunto, F. and A. G. Rossi (2020). Robo-advising. *Handbook of Technological Finance*.

D'Acunto, F. and A. G. Rossi (2021). New frontiers of robo-advising: Consumption, saving, debt management, and taxes. *Handbook of Machine Learning for Financial Markets*.

D'Acunto, F., M. Weber, and J. Xie (2019). Punish one, teach a hundred: The sobering effect of punishment on the unpunished. *Fama-Miller Working Paper*, 19–06.

De Quidt, J., J. Haushofer, and C. Roth (2018). Measuring and bounding experimenter demand. *American Economic Review 108*(11), 3266–3302.

DellaVigna, S. and D. Pope (2017). Predicting experimental results: Who knows what? *Journal of Political Economy*.

Dupor, B. and W.-F. Liu (2003). Jealousy and equilibrium overconsumption. *American economic review 93*(1), 423–428.

Ganong, P. and P. Noel (2019). Consumer spending during unemployment: Positive and normative implications. *American economic review 109*(7), 2383–2424.

Gargano, A. and A. G. Rossi (2020). There's an app for that: Goal-setting and saving in the fintech era. *Available at SSRN*.

Gelman, M., S. Kariv, M. D. Shapiro, D. Silverman, and S. Tadelis (2014). Harnessing naturally occurring data to measure the response of spending to income. *Science 345*(6193), 212–215.

Gelman, M., S. Kariv, M. D. Shapiro, D. Silverman, and S. Tadelis (2018). How individuals respond to a liquidity shock: Evidence from the 2013 government shutdown. *Journal of Public Economics*.

Han, B., D. Hirshleifer, and J. Walden (2019). Visibility bias in the transmission of consumption beliefs and undersaving. Technical report, National Bureau of Economic Research.

Hansen, B. E. (1996). Inference when a nuisance parameter is not identified under the null hypothesis. *Econometrica*, 413–430.

Hansen, B. E. (2000). Sample splitting and threshold estimation. *Econometrica 68*(3), 575–603.

Hansen, B. E. (2017). Regression kink with an unknown threshold. *Journal of Business & Economic Statistics 35*(2), 228–240.

Hau, H., Y. Huang, H. Shan, and Z. Sheng (2019). How fintech enters china's credit market. In *AEA Papers and Proceedings*, Volume 109, pp. 60–64.

Heyerdahl-Larsen, C. and J. Walden (2022). Distortions and efficiency in production economies with heterogeneous beliefs. *The Review of Financial Studies 35*(4), 1775–1812.

Jappelli, T. and L. Pistaferri (2010). The consumption response to income changes. *Annu. Rev. Econ. 2*(1), 479–506.

Jappelli, T. and L. Pistaferri (2014). Fiscal policy and mpc heterogeneity. *American Economic Journal: Macroeconomics 6*(4), 107–36.

Kaustia, M. and S. Knüpfer (2012). Peer performance and stock market entry. *Journal of Financial Economics 104*(2), 321–338.

Kuziemko, I., M. Norton, E. Saez, and S. Stantcheva (2015). How elastic are preferences for redistribution? evidence from randomized survey experiments. *American Economic Review 105*(4), 1478–1508.

Laibson, D. (1997). Golden eggs and hyperbolic discounting*. *The Quarterly Journal of Economics 112*(2), 443–478.

Lee, S. K. (2019). Fintech nudges: Overspending messages and personal finance management. *NYU Stern School of Business*.

Lian, C., Y. Ma, and C. Wang (2019). Low interest rates and risk-taking: Evidence from individual investment decisions. *The Review of Financial Studies 32*(6), 2107–2148.

Lusardi, A. and O. S. Mitchell (2007). Baby boomer retirement security: The roles of planning, financial literacy, and housing wealth. *Journal of Monetary Economics 54*(1), 205–224.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The review of economic studies 60*(3), 531–542.

Mas, A. and E. Moretti (2009). Peers at work. *American Economic Review 99*(1), 112–45.

Maturana, G. and J. Nickerson (2019). Teachers teaching teachers: The role of workplace peer effects in financial decisions. *The Review of Financial Studies 32*(10), 3920–3957.

Maturana, G. and J. Nickerson (2020). Real effects of workers' financial distress: Evidence from teacher spillovers. *Journal of Financial Economics 136*(1), 137–151.

Olafsson, A. and M. Pagel (2017). The ostrich in us: Selective attention to financial accounts, income, spending, and liquidity. Technical report, National Bureau of Economic Research.

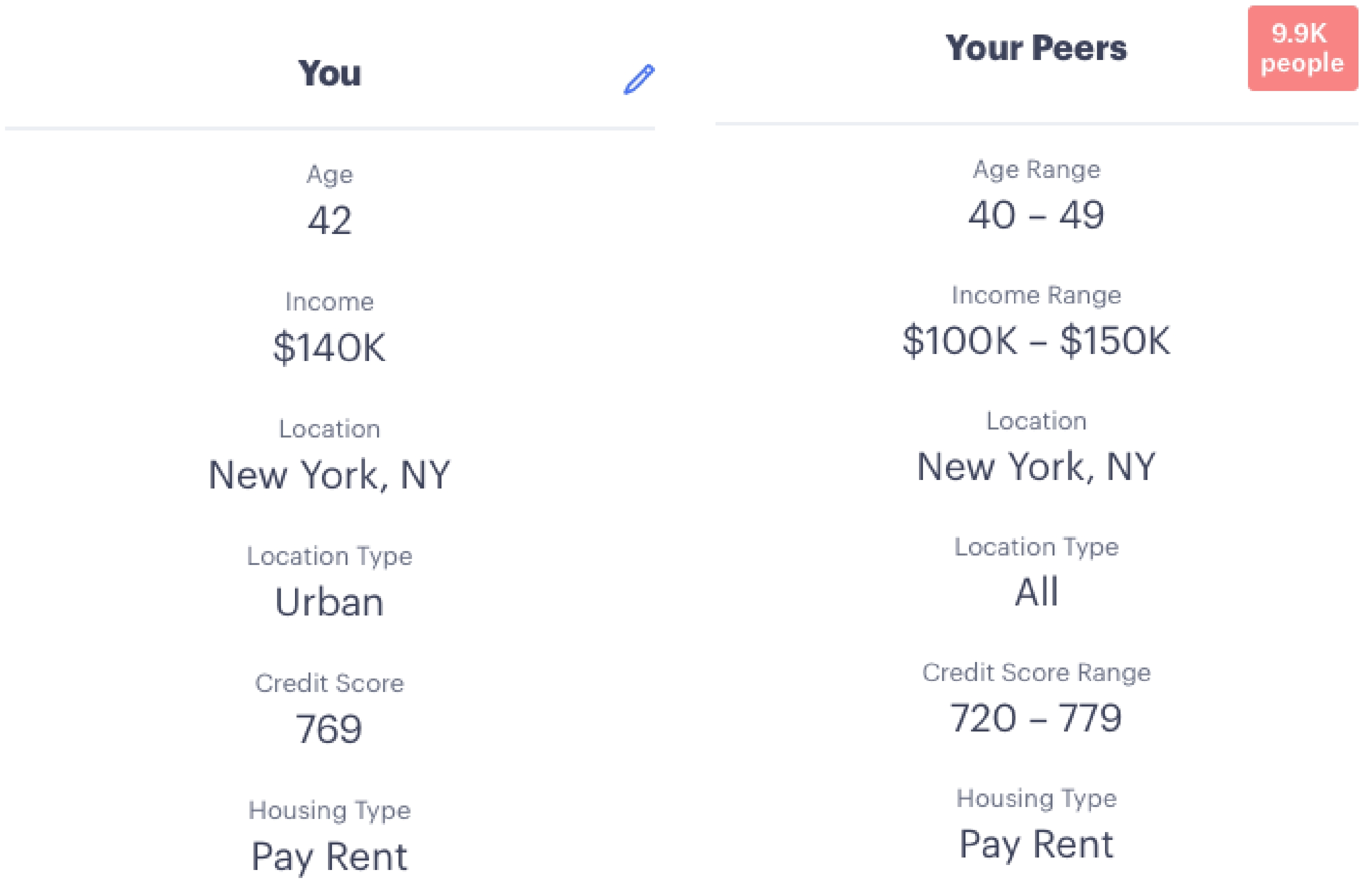
Olafsson, A. and M. Pagel (2018). The liquid hand-to-mouth: Evidence from personal finance management software. *The Review of Financial Studies 31*(11), 4398–4446.

Ouimet, P. and G. Tate (2020). Learning from coworkers: Peer effects on individual investment decisions. *The Journal of Finance 75*(1), 133–172.

Parker, J. A. and N. S. Souleles (2019). Reported effects versus revealed-preference estimates: Evidence from the propensity to spend tax rebates. *American Economic Review: Insights 1*(3), 273–90.

Reher, M. and S. Sokolinski (2021). Automation and inequality in wealth management. *Available at SSRN 3515707*.

Rossi, A. G. and S. P. Utkus (2020). Who benefits from robo-advising? evidence from machine learning. *Working Paper*.



(a) User Profile

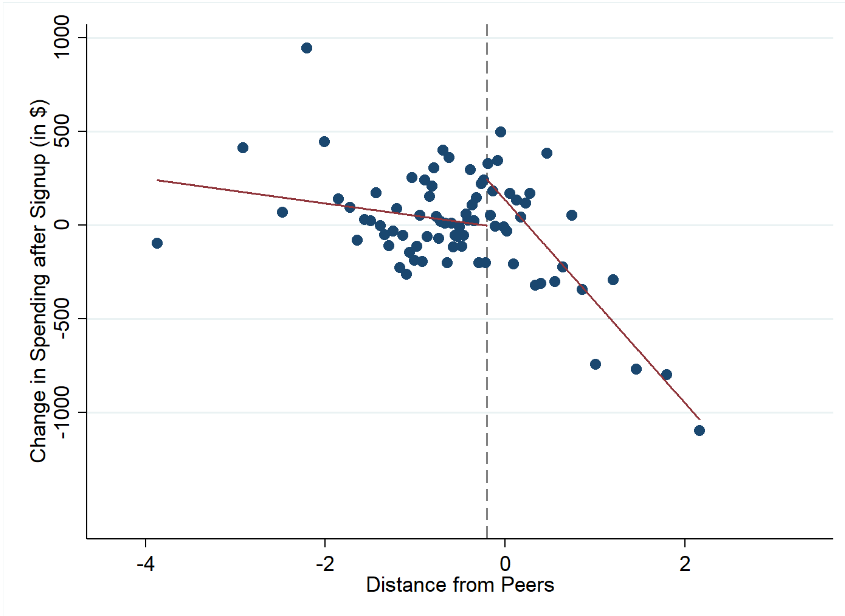
(b) Peer Group Information

Figure 2
Sample Peer Group Assignment for a Demo Account

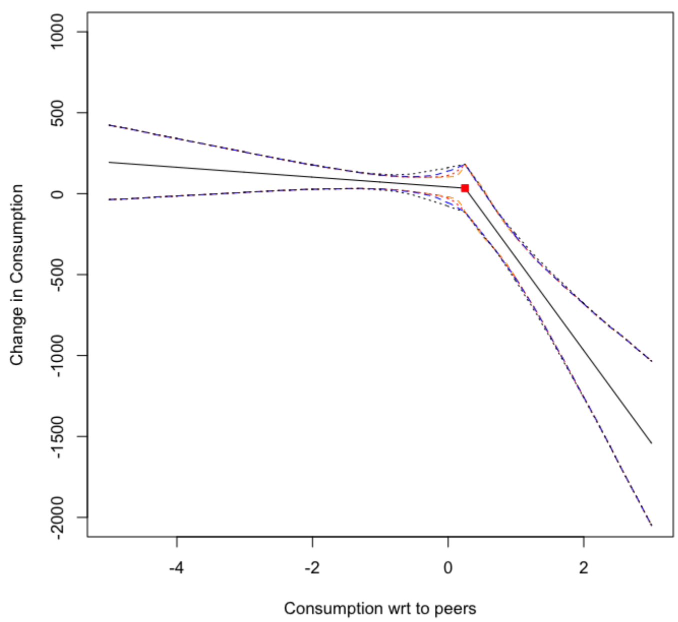
This figure shows the graphic users observed during our sample period after signing up to the app, providing their individual characteristics, and being assigned to a peer group. Note that the users are required to report exact values for continuous variables such as their age, income, and credit score. After being assigned to a peer group, users are shown the characteristics of the peer group to which they are assigned as in the reported graphic.

Figure 3
Top of *Status Money* Home Page During Our Sample Period

This figure shows the top part of the homepage of *Status Money*—the first image users see when signing in to the app during our sample period. The blue line reports the user's spending aggregated across his/her account based on transaction-level data and varies at the daily level throughout the month (x-axis). The orange line represent a daily linear projection of the average monthly spending of the peer groups assigned to the user, which is based on monthly transaction-level data for a representative US population. The red line represents a daily linear projection for the average monthly spending of all observations in the same representative US population. The horizontal dashed line represents user's average monthly income inflows. On top of the graph, users read their daily cumulative amount of spending within the month at the time they login to the app as well as the projected cumulative amount of spending for peers and the overall representative population. Note that some of this figure's features, including its position and size, have changed after our sample period, as we show in the Online Appendix.



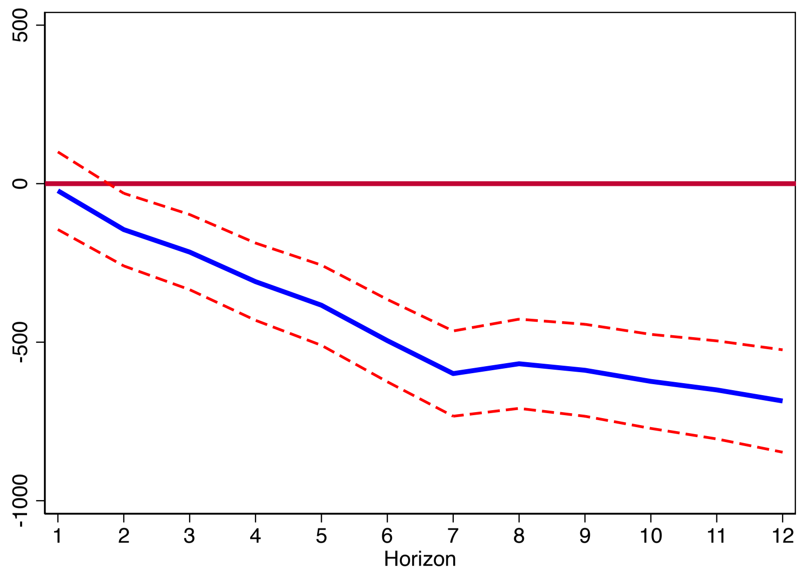
(a) Threshold Regression with Unknown Threshold



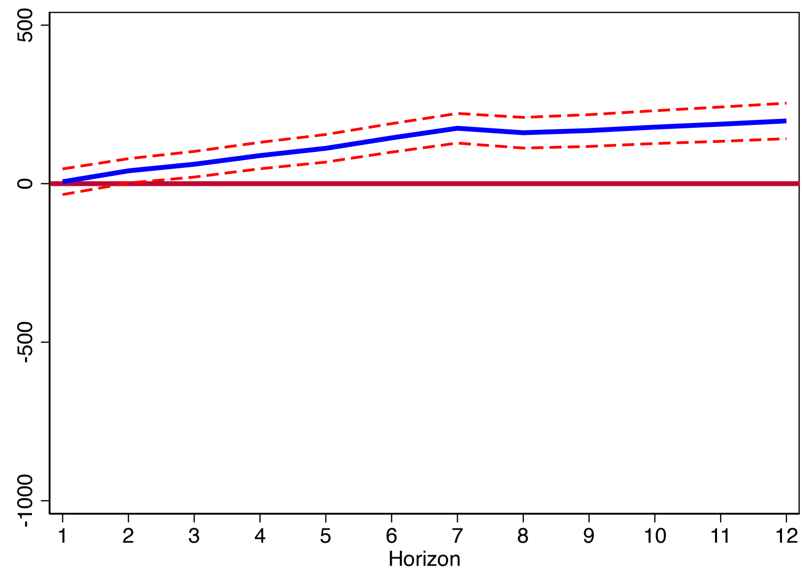
(b) Kink Regression with Unknown Threshold

Figure 4
Distance from Peers' Spending and Changes in Spending after Signup—Endogenous Threshold Models

This figure reports the fitted values of a threshold regression model, with the optimal threshold estimated using the procedure in Hansen (2000) in subfigure (a). Subfigure (b) reports the fitted values of a kink regression model with the optimal threshold estimated using the procedure in Hansen (2015). In addition to the fitted values, subfigure (b) reports 90% confidence intervals.



(a) Dynamic Change in Spending for Overspenders

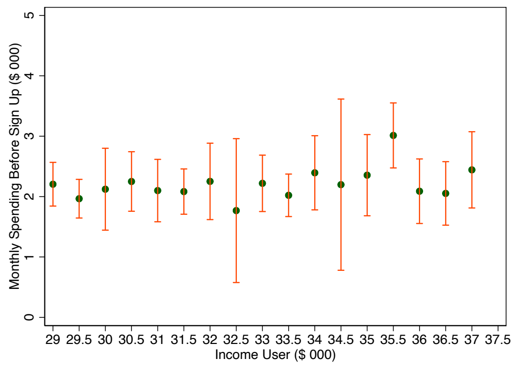


(b) Dynamic Change in Spending for Underspenders

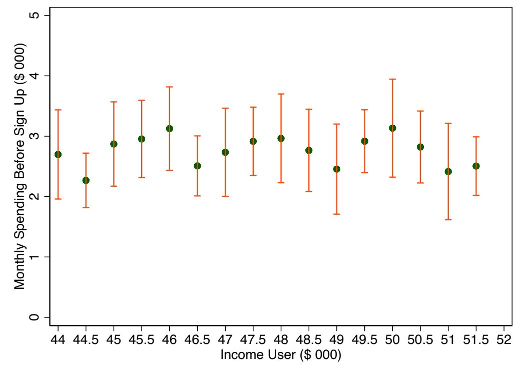
Figure 5

Dynamic Effect of Peer Information on Spending After Sign-up

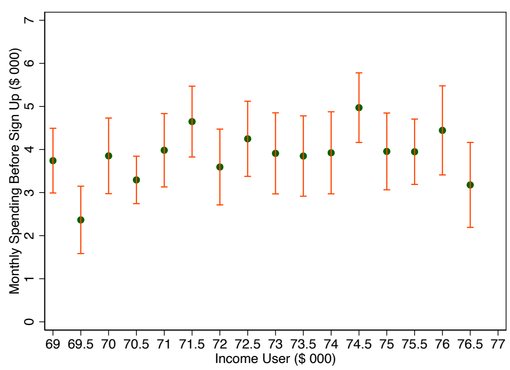
This figure plots average monthly changes in spending (and associated 95% confidence intervals) from the first month after sign up until the twelfth month after sign up with respect to the average monthly spending in the three months before signup. The left panel includes users who spent more than peers before signup, whereas the right panel includes users who spent less than peers before signup. Note that each monthly average is estimated separately for the subsamples of users who survive each month after signup and the connected lines do <u>not</u> imply that the estimated associations are continuous over time. We report these results in the form of connected lines rather than discontinuous point estimates for graphical purposes.



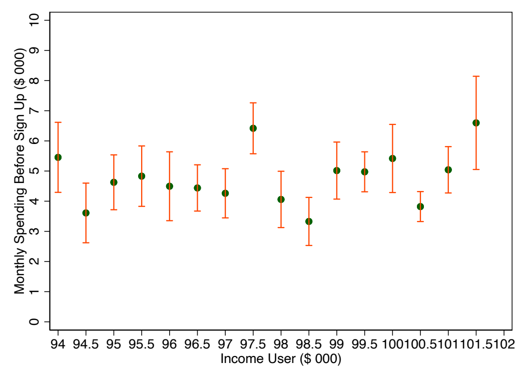
(a) Income Threshold $35,000



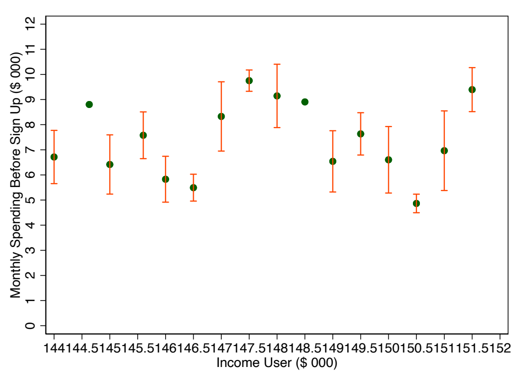
(b) Income Threshold $50,000



(c) Income Threshold $75,000



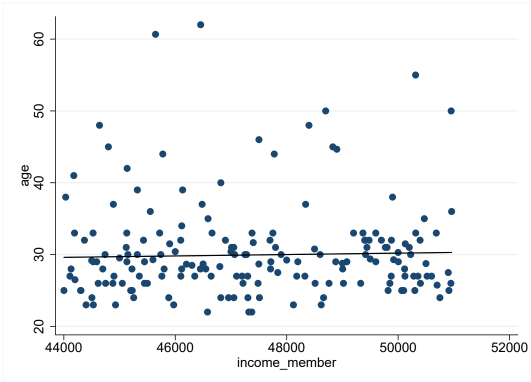
(d) Income Threshold $100,000



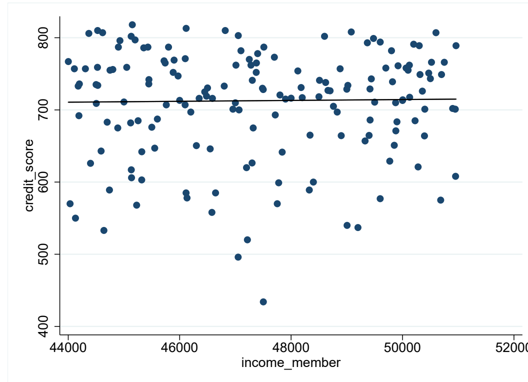
(e) Income Threshold $150,000

Figure 6
Average Monthly Spending over the Three Months before Sign-up around Income Thresholds

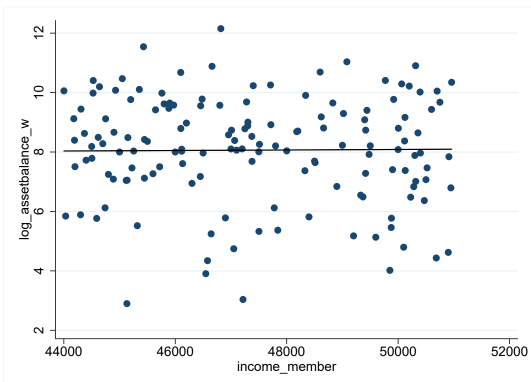
This figure reports the average users' monthly spending before signup to *Status Money* when sorting users by their reported income levels (x-axis). In each panel, income levels are reported at incremental intervals of $500. Each panel reports average spending pre-signup for users around the thresholds in our instrumental-variable analysis. Our strategy assumes that the levels of pre-spending are similar for users above and below each threshold, and hence we should fail to reject the null hypothesis of no differences in average spending across contiguous groups of income. To be conservative, we report tight confidence bandwidths of one third of a standard deviation above and below the point estimate.



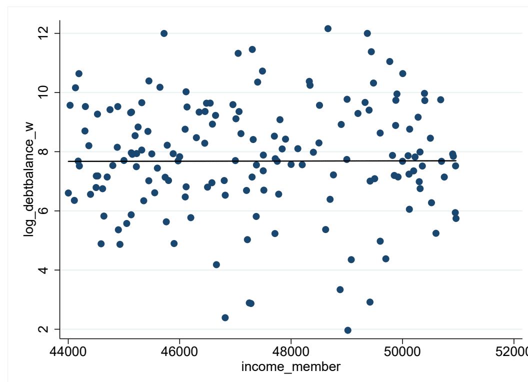
(a) Age



(b) Credit Score



(c) Log Asset Balance



(d) Log Debt Balance

Figure 7
Observable Characteristics of Users around the $50K Income Threshold

This figure reports the distribution of age, credit score, the log of asset balance and the log of debt balance around the $50,000 income threshold.

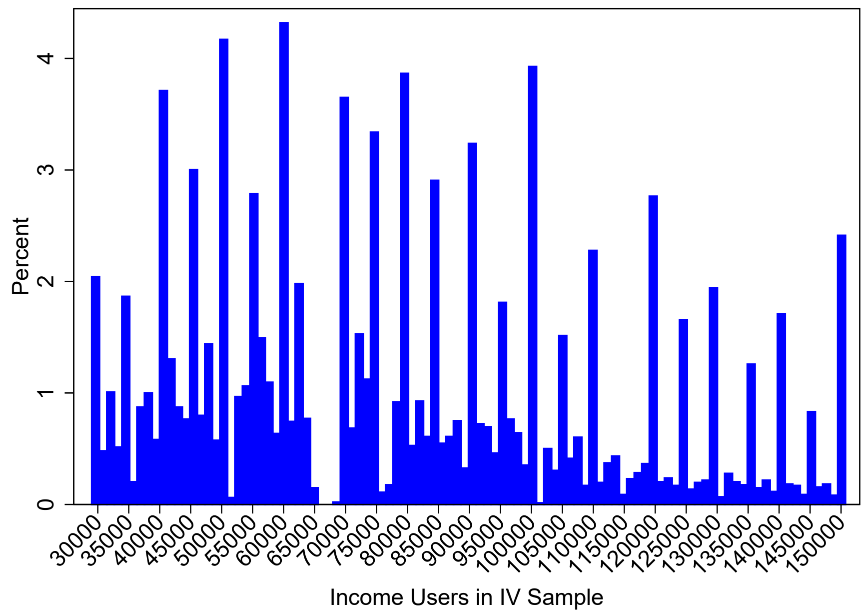
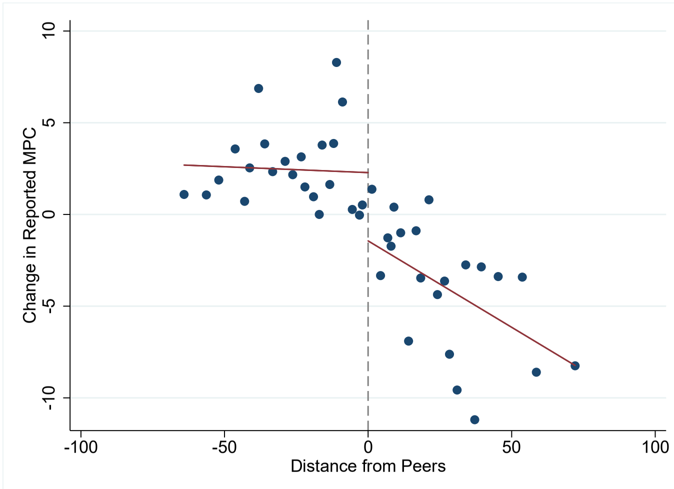
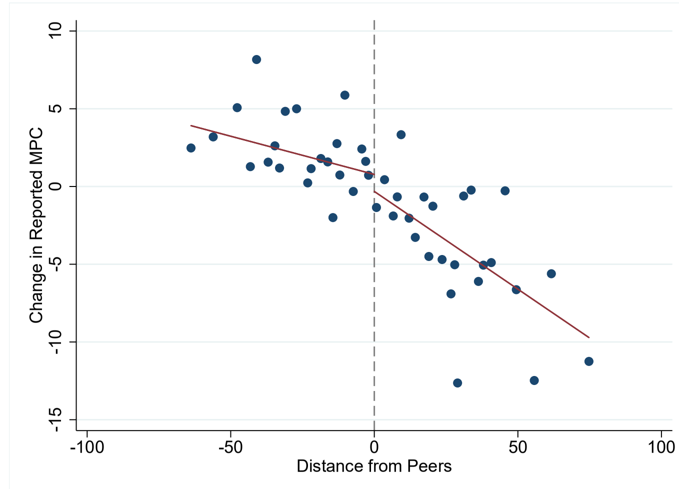


Figure 8
Bunching of Users' Reported Income at the IV and Placebo IV Thresholds

This figure plots the distribution of income levels reported by users in the IV sample to observe the bunching patterns at round values and especially at the IV and placebo IV thresholds. The IV thresholds are \$35K, \$50K, \$75K, \$100K, and \$150K. The placebo IV thresholds are \$45K, \$60K, \$90K, \$110K, and \$140K.



(a) Current Times (COVID-19 Crisis)



(b) Normal Times

Figure 9
Change in Reported MPC After Randomized Peer Information

This figure shows binned scatterplots of changes in reported marginal propensity to consume (MPC) elicited before and after provision of the MPC of income peers for both the time at which the randomized control trial (RCT) was run (during the COVID-19 crisis) and in normal times. Information about income peers' MPC is based on a representative US sample. In both subfigures, the x-axis measures the difference in MPC with respect to peers. The y-axis reports results for changes in reported MPC. Each binned scatterplot divides the 1,015 respondents into 50 groups. In addition to the scatterplot, we report in red the fitted values of a threshold regression that estimates different linear regression coefficients below and above the zero threshold.

Table 1. Summary Statistics

Panel A. Main sample

	Obs	Mean	St. Dev.	p10	p25	p50	p75	p90
Age	20,679	32	8	24	27	30	35	42
Credit Score	19,051	736	74	630	704	756	790	810
Home Ownership	20,679	0.39	0.49	0	0	0	1	1
Annual Income ($)	20,594	92,633	62,838	34,000	50,000	75,000	115,000	170,500
Distance Peers	20,679	-0.53	0.97	-1.59	-1.04	-0.53	-0.02	0.61
Assets ($)	17,190	41,101	63,870	511	2,208	10,699	47,365	142,985
Debts ($)	19,392	20,466	45,545	218	875	3,335	11,320	56,145
Spending (30 Days, $)	20,679	4,963	4,007	840	2,039	3,877	6,695	11,107
Spending (60 Days, $)	20,679	4,886	4,040	820	1,971	3,736	6,521	11,106
Spending (90 Days, $)	20,679	4,670	3,894	698	1,805	3,589	6,351	10,680
Urban	13,509	0.40	0.49	0	0	0	1	1
Suburban	13,509	0.34	0.47	0	0	0	1	1
Rural	13,509	0.26	0.44	0	0	0	1	1

Panel B. Identification sample

	Obs	Mean	St. Dev.	p10	p25	p50	p75	p90
Age	5,629	31	8	24	26	30	34	40
Credit Score	5,236	730	76	624	693	751	784	807
Home Ownership	5,629	0.36	0.48	0	0	0	1	1
Annual Income ($)	5,629	71,917	33,399	33,000	46,644	70,000	95,000	100,000
Distance Peers	5,629	-0.47	0.91	-1.47	-0.96	-0.51	-0.00	0.59
Assets ($)	4,616	26,878	39,880	381	1,720	8,015	32,542	92,671
Debts ($)	5,290	13,680	28,721	211	789	3,012	9,928	31,862
Spending (30 Days, $)	5,629	4,375	3,505	783	1,931	3,530	5,777	8,949
Spending (60 Days, $)	5,629	4,332	3,555	798	1,868	3,427	5,645	9,038
Spending (90 Days, $)	5,629	4,060	3,340	657	1,694	3,241	5,410	8,393
Low Similarity to Peers	5,629	0.38	0.48	0	0	0	1	1
High Number of Peers	2,600	0.62	0.48	0	0	1	1	1
High Peer Group Income Width	4,552	0.47	0.50	0	0	0	1	1
Had More Peer Info Before	3,613	0.72	0.45	0	0	1	1	1

Panel C. Randomized-control-trial sample

	Obs	Mean	St. Dev.	p10	p25	p50	p75	p90
Age	1,014	38.03	10.70	24.00	24.00	40.50	40.50	55.00
Annual Income ($)	1,015	67,002	40,797	27,499	27,499	54,999	54,999	137,500
Distance Peers (pre-COVID)	1,015	1.33	31.68	-40.00	-23.00	2.00	27.00	42.00
Distance Peers (during COVID)	1,015	-3.57	30.93	-43.00	-24.00	-4.00	18.00	38.00
Male	1,004	0.57	0.50	0.00	0.00	1.00	1.00	1.00
College	1,014	0.73	0.44	0.00	0.00	1.00	1.00	1.00
Partnered	1,010	0.66	0.47	0.00	0.00	1.00	1.00	1.00
Children	1,008	0.35	0.48	0.00	0.00	0.00	1.00	1.00
Liberal	1,015	0.51	0.50	0.00	0.00	1.00	1.00	1.00
Financially Literate	1,015	0.46	0.50	0.00	0.00	0.00	1.00	1.00
Change MPC (pre-COVID)	1,015	-0.89	11.14	-11.00	-1.00	0.00	0.00	9.00
Change MPC (during COVID)	1,015	-0.27	10.51	-10.00	0.00	0.00	1.00	10.00

This table reports summary statistics of the variables that enter our analysis across the three samples we use—the main sample of *Status Money* users for the baseline analysis (Panel A), the identification sample for the instrumental-variable analysis (Panel B), and the randomized-control-trial sample for the external-validity analysis (Panel C). For each variable, we report the number of observations, the sample average, the sample standard deviation, and various percentiles of the distribution.

Table 2. Change in Spending After Peer Spending Information

	Above	Below	Above	Below	Above	Below
Average Change	-0.233***	0.074***				
	(-42.00)	(8.34)				
Distance Peers			-0.103***	-0.086***	-0.104***	-0.096***
			(-11.31)	(-7.03)	(-10.47)	(-7.43)
Homeownership					0.002	0.062***
					(0.17)	(2.82)
log of Credit Score					-0.040	-0.293***
					(-1.46)	(-3.82)
log of Age					0.014	-0.149***
					(0.46)	(-2.99)
log of Asset Balance					-0.001	-0.010**
					(-0.40)	(-2.17)
log of Debt Balance					0.002	-0.011**
					(0.55)	(-2.48)
Constant			-0.166***	-0.002	0.048	2.543***
			(-20.14)	(-0.11)	(0.23)	(4.88)
Observations	5,012	15,667	5,012	15,667	4,179	10,688

Columns (1)-(2) of this table report the average change in spending of users three months after signup relative to three months before based on whether users spent more or less than the average spending of their assigned peer group in the 30 days before signup.
Columns (3)-(6) of this table report the results for estimating the following OLS specification:

$$\frac{Spending_{i,post}}{Spending_{i,pre}} = \alpha + \gamma \ Distance \ Peers_i + \delta \ \boldsymbol{x}_i + \epsilon_i,$$

where $\frac{Spending_{i,post}}{Spending_{i,pre}}$ is the ratio of user's i spending three months after signup relative to three months before. $Distance \ Peers_i$ is the difference between user's i spending and the average spending of his/her peer group in the 30 days before signup. This difference is standardized to have a unit standard deviation. The vector of individual controls x_i are the individual-level observables we have available from the app, which include a homeownership dummy, logarithm of credit score, logarithm of age, logarithm of asset balance, and logarithm of debt balance. Regression estimates are computed for users with above-peer spending in columns 1, 3, and 5 and for users with below-peer spending in columns 2, 4, and 6. Numbers in parentheses are t-statistics based on Huber-White standard errors.

Table 3. Mean Reversion? Assigning Users to Peer Groups Based on Spending Well Before Signup

	30 Days before Signup			60 Days before Signup			90 Days before Signup		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Distance Peers	-0.103***	-0.044***	-0.039***	-0.110***	-0.082***	-0.083***	-0.099***	-0.077***	-0.075***
	(-11.31)	(-4.25)	(-3.54)	(-13.83)	(-8.99)	(-8.61)	(-11.75)	(-8.05)	(-7.38)
Spend Before		-0.070***	-0.096***		-0.035***	-0.062***		-0.029***	-0.058***
		(-11.72)	(-13.11)		(-6.09)	(-8.91)		(-5.08)	(-8.25)
Homeownership			0.037***			0.035***			0.034***
			(2.76)			(2.80)			(2.68)
log of Credit Score			-0.036			-0.042			-0.078**
			(-1.35)			(-1.59)			(-2.18)
log of Age			0.081***			0.075***			0.090***
			(2.62)			(2.63)			(3.01)
log of Asset Balance			0.009***			0.012***			0.012***
			(3.05)			(3.97)			(3.92)
log of Debt Balance			0.007**			0.008***			0.007***
			(2.36)			(2.92)			(2.68)
Constant	-0.165***	-0.144***	-0.328	-0.187***	-0.175***	-0.319	-0.209***	-0.198***	-0.161
	(-20.21)	(-17.51)	(-1.62)	(-24.95)	(-22.66)	(-1.62)	(-27.18)	(-24.88)	(-0.64)
Observations	5,012	5,012	4,179	4,791	4,791	3,970	4,473	4,473	3,697

This table reports results for estimating the following OLS specification:

$$\frac{Spending_{i,post}}{Spending_{i,pre}} = \alpha + \gamma\ Distance\ Peers_i + \zeta\ Spending_{i,pre} + \delta\ \boldsymbol{x}_i + \epsilon_i,$$

where $\frac{Spending_{i,post}}{Spending_{i,pre}}$ is the ratio of user's i spending three months after signup relative to three months before. $Distance\ Peers_i$ is the difference between user's i spending and the average spending of his/her peer group. Listed on top of each column is the period used to compute user's distance from peers' spending and hence whether they are categorized as overspenders or underspenders. We alternatively use the month before signup (as in the baseline results in Table 2), two months before signup, or three months before signup. In all columns, this difference is standardized to have a unit standard deviation. The vector of individual controls x_i are the individual-level observables we have available from the app, which include a homeownership dummy, logarithm of credit score, logarithm of age, logarithm of asset balance, and logarithm of debt balance. Numbers in parentheses are t-statistics based on Huber-White standard errors.

Table 4. IV Sample: Balancing of Observables around Peer Assignment Income Rules

	Home ownership	log of Credit Score	log of Age	log of Asset Balance	log of Debt Balance
	Panel A: Income Threshold: $35,000				
Above Dummy	0.031	-0.009	0.018	-0.160	0.324**
	(1.06)	(-0.95)	(1.02)	(-0.85)	(2.10)
Observations	896	834	896	675	837
	Panel B: Income Threshold: $50,000				
Above Dummy	0.038	-0.001	0.014	0.021	0.009
	(1.63)	(-0.09)	(1.31)	(0.17)	(0.08)
Observations	1,516	1,410	1,516	1,227	1,415
	Panel C: Income Threshold: $75,000				
Above Dummy	0.013	0.002	0.012	0.017	0.027
	(0.49)	(0.25)	(1.09)	(0.14)	(0.23)
Observations	1,546	1,435	1,546	1,278	1,457
	Panel D: Income Threshold: $100,000				
Above Dummy	0.004	0.019	0.024**	0.199	-0.163
	(0.14)	(1.24)	(2.09)	(1.62)	(-1.21)
Observations	1,128	1,047	1,128	954	1,065
	Panel E: Income Threshold: $150,000				
Above Dummy	-0.015	0.002	-0.000	-0.074	-0.322
	(-0.35)	(0.24)	(-0.00)	(-0.44)	(-1.54)
Observations	543	510	543	482	516

This table reports results of regressing all demographics we use as controls in our baseline analysis on a dummy that equals one for users just below and users at or above the income values that *Status Money* uses to define peer groups—$35K, $50K, $65K, $75K, $100K, and $150K. Numbers in parentheses are t-statistics based on Huber-White standard errors.

Table 5. Instrumental-Variable (IV) Analysis:
Peer Spending Information and Change in Spending

			Placebo IV	
	First Stage	Second Stage	First Stage	Second Stage
Above Dummy	0.743***		0.078	
	(24.62)		(0.795)	
Peer Spending		0.111***		0.942
		(3.08)		(0.432)
Spending Before	0.344***	-0.305***	0.120***	-0.566***
	(23.33)	(-15.63)	(3.46)	(-2.02)
First stage F-stat	606.1			
Observations	5,629	5,629	678	678

Columns (1)-(2) of this table report the results for implementing our instrumental-variable strategy that compares users just below and users at or above the income values that *Status Money* uses to define peer groups—\$35K, \$50K, \$65K, \$75K, \$100K, and \$150K. Columns (3)-(4) refer to our placebo analysis in which we compare users just below and users at or above a set of income values that *Status Money* does *not* use to define peer groups—\$45K, \$60K, \$90K, \$110K, and \$140K. For each income value, the sample includes users whose income is at most \$6K below the value and at most \$2K above the value.
In all columns, we estimate a set of two-stage least-square specifications in which the uncertainty in the estimate of first-stage coefficients is taken into account directly. The first stage consists of the following specification:

$$Peer\ Spending_i = \alpha + \beta\ Dummy\ Above_i + \zeta\ Spending\ Before_i + \epsilon_i,$$

where $Peer\ Spending_i$ is the peer-spending value for user i and $Dummy\ Above_i$ is a dummy variable for whether the income is at or above the threshold value. The second stage specification, which is estimated jointly with the first stage, uses the instrumented $Peer\ Spending_i$ of the first stage as the main covariate in the following specification:

$$\frac{Spending_{i,post}}{Spending_{i,pre}} = \alpha + \beta\ \widehat{Peer\_Spending\_i} + \zeta\ Spending\ Before_i + \epsilon_i,$$

where $\frac{Spending_{i,post}}{Spending_{i,pre}}$ is the ratio of post and pre consumption in the three months around signup. All other variables are the same as in the first stage. Numbers in parentheses are t-statistics based on Huber-White standard errors.

Table 6. Effects by Informativeness of Peer Groups

	Similarity to Peers		Number of Peers		Peer Group Income Width		Had Peer Info Before	
Peer Spending×High	0.123**	0.181*	0.143	0.378**	-0.371***	-0.460***	-0.268**	-0.537***
	(1.97)	(1.71)	(1.56)	(2.03)	(-2.36)	(-2.59)	(-2.54)	(-2.84)
Peer Spending	0.060	0.051	0.040	-0.035	0.386**	0.492**	0.360***	0.615***
	(1.37)	(0.66)	(0.64)	(-0.33)	(2.56)	(2.83)	(3.56)	(3.51)
log of Income ×High		-0.296		-0.712**		0.422**		0.808***
		(−1.57)		(2.34)		(2.07)		(2.87)
log of Income		0.001		0.257		-0.195		-0.796***
		(0.01)		(1.50)		(-1.58)		(-3.22)
High	0.035	3.332	0.051	7.920**	-0.191*	-4.957**	-0.292**	-9.272***
	(1.30)	(1.59)	(1.38)	(2.34)	(-1.76)	(-2.15)	(-2.54)	(-2.94)
Spending Before	-0.307***	-0.286***	-0.308***	-0.290***	-0.282***	-0.296***	-0.290**	-0.272***
	(-15.71)	(-18.88)	(-11.22)	(-13.74)		(-16.50)	(-15.16)	(-14.74)
First stage F-stat	264.3	894.6	111.1	339.2	383.9	960.3	201.1	481.8
Observations	5,629	5,629	2,600	2,600	4,552	4,552	3,613	3,613

This table reports the results for a set of two-stage-least-squares specifications that compares users just below and users at or above the income values that *Status Money* uses to define peer groups, that is, \$35K, \$50K, \$65K, \$75K, \$100K, and \$150K. For each value, the sample includes users whose income is at most \$6K below the value and at most \$2K above the value. The first stage consists of the following two specifications:

$$Peer\ Spending_i = \alpha + \beta\ Dummy\ Above_i + \zeta\ Spending\ Before_i + \epsilon_i; \qquad Peer\ Spending_i \times High = \alpha + \beta\ Dummy\ Above_i \times High + \gamma High + \zeta\ Spending\ Before_i + \epsilon_i,$$

where $Peer\ Spending_i$ is the peer-spending value for user i; $Dummy\ Above_i$ is a dummy variable for whether the income is at or above the income value; $High$ is a dummy variable whose definition changes across columns: in *Similarity to Peers*, High equals 1 if users are matched to the tightest possible peer group characteristics for all categories, and zero if they are matched to wider peers groups for at least one category; in *Number of Peers*, High equals one if users are matched to peer groups in the top quarter of the distribution by size (higher or equal to 21K peers), and zero if they are matched to peer groups in the bottom quarter (lower or equal to 6K, with 5K being the minimum size of peer groups); in *Peer Group Income Width*, High equals 1 if users are matched to peer groups whose range of income is \$50K or above, and zero if the range is \$15K or below; in *Had Peer Info Before*, High equals 1 if users reside in urban locations—our proxy for being exposed to dense information about similar demographics—and zero if they reside in rural locations. X_i is a set of individual-level observables that include the logarithm of users' income, the same variable interacted with *High*, and users' spending in the month before signup.
The second stage specification, which is estimated jointly with the first stage, uses the instrumented $Peer\ Spending_i$ and $Peer\ Spending_i \times High$ of the first stage as the main covariates in the following specification:

$$\frac{Spending_{i,post}}{Spending_{i,pre}} = \alpha + \beta_1\ \widehat{Peer\_Spending\_i \times High} + \beta_2\ \widehat{Peer\_Spending\_i} + \gamma\ High + \zeta\ Spending\ Before_i + \epsilon_i,$$

where $\frac{Spending_{i,post}}{Spending_{i,pre}}$ is the ratio of post and pre consumption in the three months around signup and all other variables are the same as in the first stage. We report the second-stage estimates for different sample splits. Numbers in parentheses are t-statistics based on Huber-White standard errors.

Table 7. External Validity 1: Peer MPC Information and Change in MPC

	Current Times (COVID-19)				Normal Times			
	Above	Below	Distance Above	Distance Below	Above	Below	Distance Above	Distance Below
Average Change	-3.854***	2.462***			-3.685***	2.010***		
	(-6.99)	(6.78)			(-6.64)	(5.31)		
Distance Peers			-0.094***	-0.005			-0.127***	-0.050**
			(-3.16)	(-0.24)			(-4.40)	(-2.24)
Constant			-1.448	2.331***			-0.254	0.745
			(-1.54)	(3.52)			(-0.27)	(1.10)
Observations	439	576	439	576	517	498	517	498

This table reports results for the change in the reported marginal propensity to consume (MPC) of respondents to our randomized control trial before and after being provided with information about the reported MPC of income peers based on representative US estimates. We report the average changes and the OLS coefficients for regressing changes on the respondent's MPC distance from the MPC of income peers for self-reported MPC at the time the survey was run (during the COVID-19 crisis) and in normal times. Users' distance from their peers is standardized to have a unit standard deviation. Numbers in parentheses are t-statistics based on Huber-White standard errors.

Table 8. External Validity 2: Peer MPC Information and Change in MPC by Demographics

	Gender		Education		Marital Status	
	Men	Women	College	No College	Partnered	Single
Distance Peers	-0.086**	-0.181***	-0.122***	-0.185***	-0.119***	-0.166**
	(-2.15)	(-4.29)	(-4.02)	(-2.28)	(-3.67)	(-2.53)
Observations	322	190	397	119	386	128
	Children		Political Views		Financial Literacy	
	No	Yes	Liberal	No Liberal	High	Low
Distance Peers	-0.094**	-0.170***	-0.185***	-0.072**	-0.193***	-0.084**
	(-2.46)	(-3.83)	(-3.98)	(-2.04)	(-4.33)	(-2.22)
Observations	293	221	242	275	190	327

This table reports results for changes in the reported marginal propensity to consume (MPC) of respondents to our randomized control trial before and after being provided with information about the reported MPC of income peers based on representative US estimates. We report the average changes and the OLS coefficients for regressing changes on the respondent's MPC distance from the MPC of income peers for self-reported MPC in normal times. Users' distance from their peers is standardized to have a unit standard deviation. Numbers in parentheses are t-statistics based on Huber-White standard errors. Each panel splits the sample by demographics that are unobserved in the *Status Money* sample.

Online Appendix:

Crowdsourcing Peer Information to Change Spending Behavior

Francesco D'Acunto, Alberto G. Rossi, and Michael Weber

Not for Publication

A.1 Features Introduced After Our Sample Period

The app was launched in July 2017 with the features we describe in section 2.3. The app remained largely unchanged from July 2017 until December 2018. In December 2018, the company introduced cash rewards, which allowed individuals to earn rewards by inviting other users to the app or opening accounts with partner companies such as Betterment or Airbnb. During the first quarter of 2019, the app also started to send monthly alerts about spending relative to peers and updating individuals on whether they were underspending or overspending with respect to their peers.

Finally, in July 2019, the app introduced social feed features. These features allow users associated with the same peer group to share spending and savings tips, such as what credit cards to apply for and what percentage of monthly income to spend on rent.

Between 2019 and 2021 the app's appearance has also evolved. As of 2021, for example, the app emphasizes spending categories (see Figure A.2). The version of the app during our sample period, instead, reported only total spending on the home page and relegated information about spending categories to secondary, hyperlinked pages. Articles and presentations about *Status Money* help reconstruct the various phases of the app appearance. For instance, see the description of the app reported by CreditDonkey (`https://www.creditdonkey.com/status-money-review.html`) published in September 2018, and the video posted by *Status Money* on YouTube describing the app (`https://www.youtube.com/watch?v=mWFlIYZtRD4&t=37s`) in March 2019.[1]

[1] For additional information regarding how the app looked during our sample period, the reader can refer to the following two presentations by the founders of *Status Money* at Finovate 2018 and 2019 (see, `https://www.youtube.com/watch?v=aV__wCV22ng` and `https://www.youtube.com/watch?v=gDCrmani-Os`).

A.2 Additional Figures

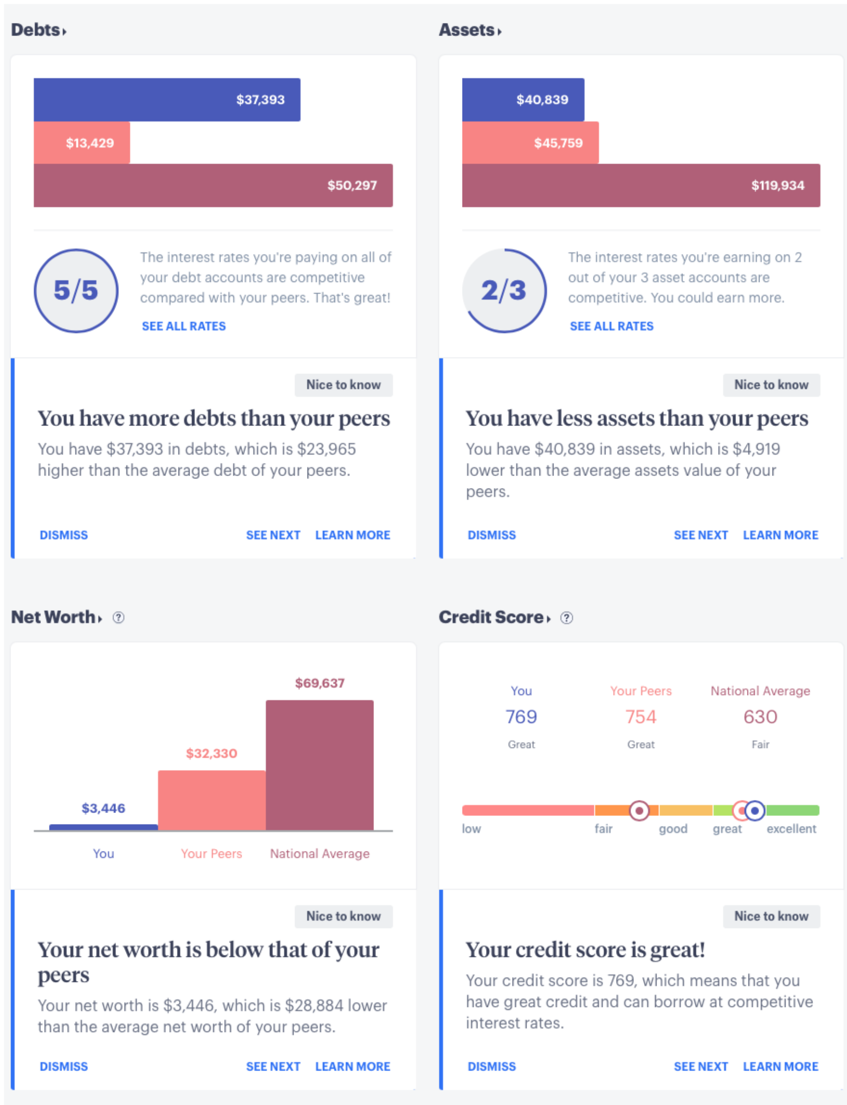


Figure A.1
Bottom Part of *Status Money* Home Page During Our Sample Period

This figure shows the bottom part of the homepage of *Status Money*, i.e. the images users see when signing in to the app during our sample period, if they are using a desktop computer. For users who sign in to the mobile-phone version of the app, these images are not visible in the first screen after signing in. If users want to access this information, they would need to click on side buttons and navigate through different pages of the mobile app.

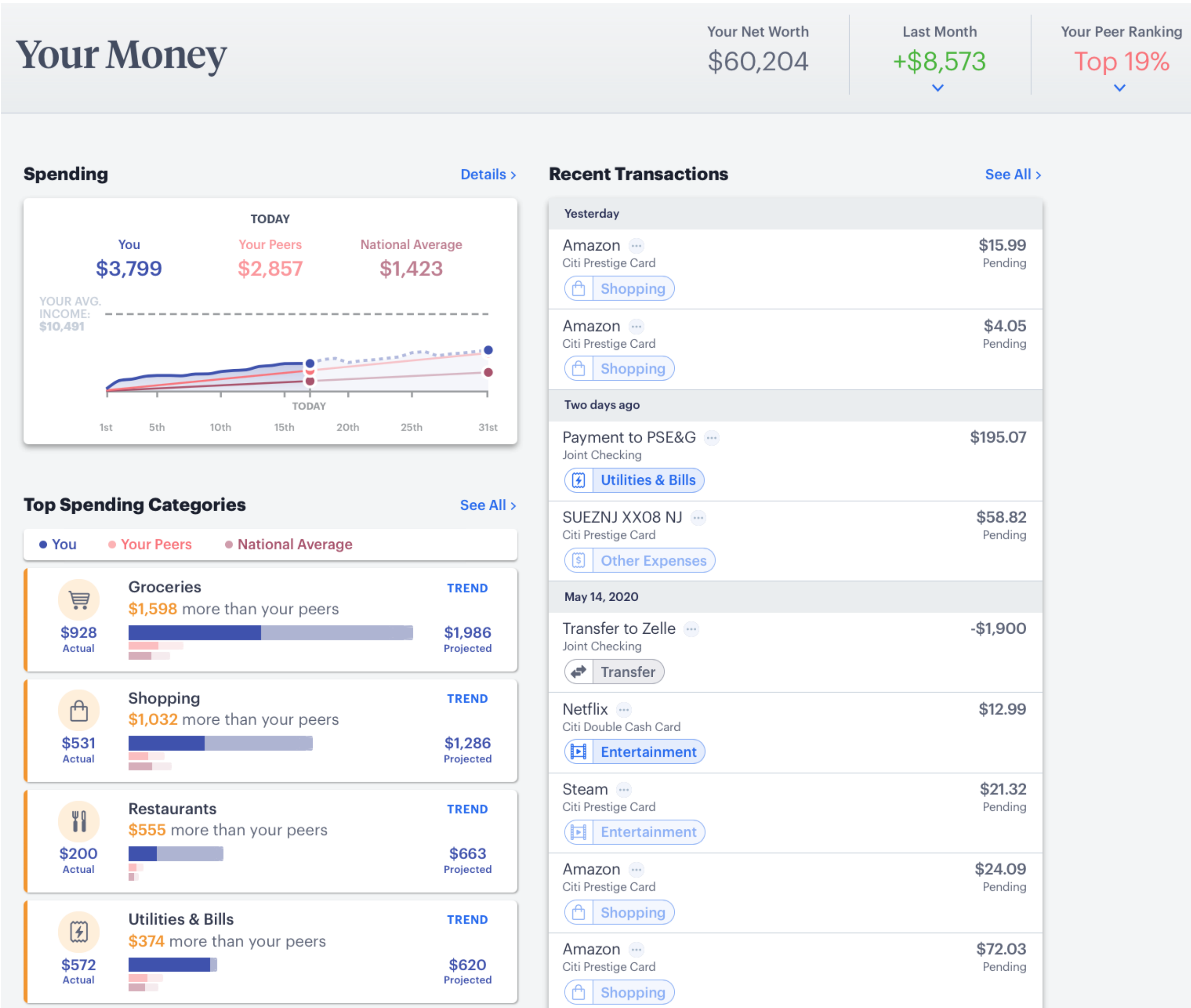


Figure A.2
Status Money **Home Page After our Sample Period**

This figure shows the homepage of *Status Money*, i.e. the main image users see when signing in to the app, after our sample period.

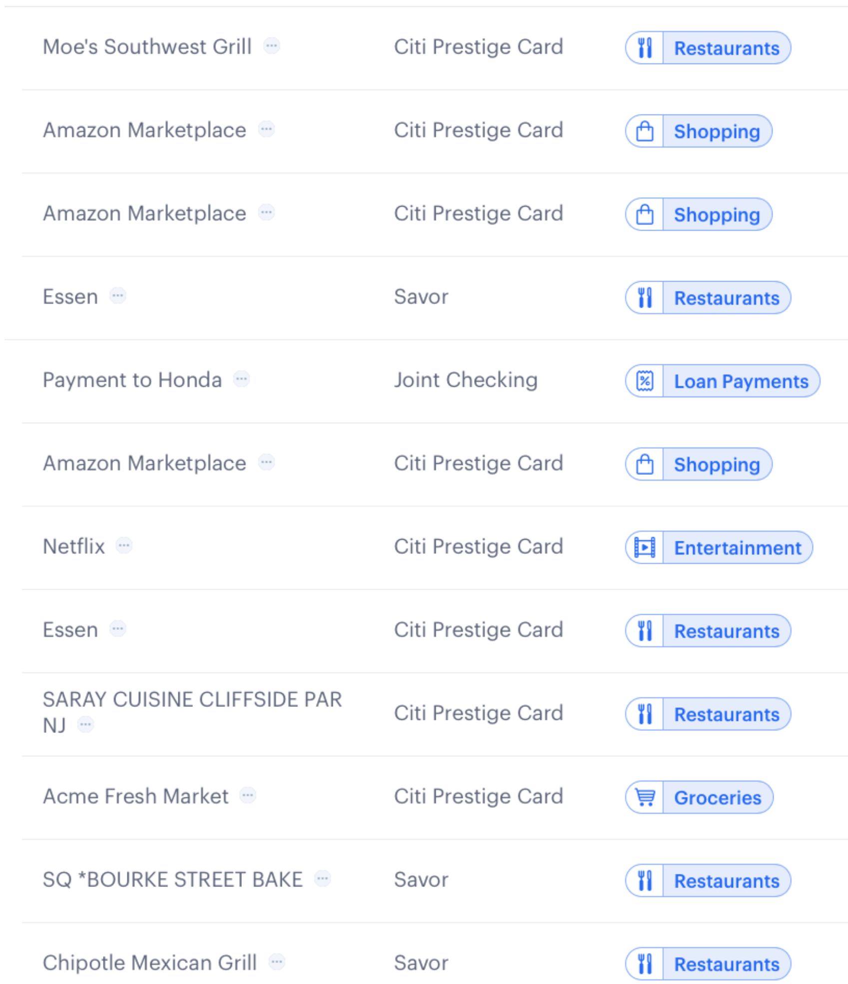


Figure A.3
Example of Transactions Categorized by Status Money

This figure shows an example of transactions categorized by *Status Money*.

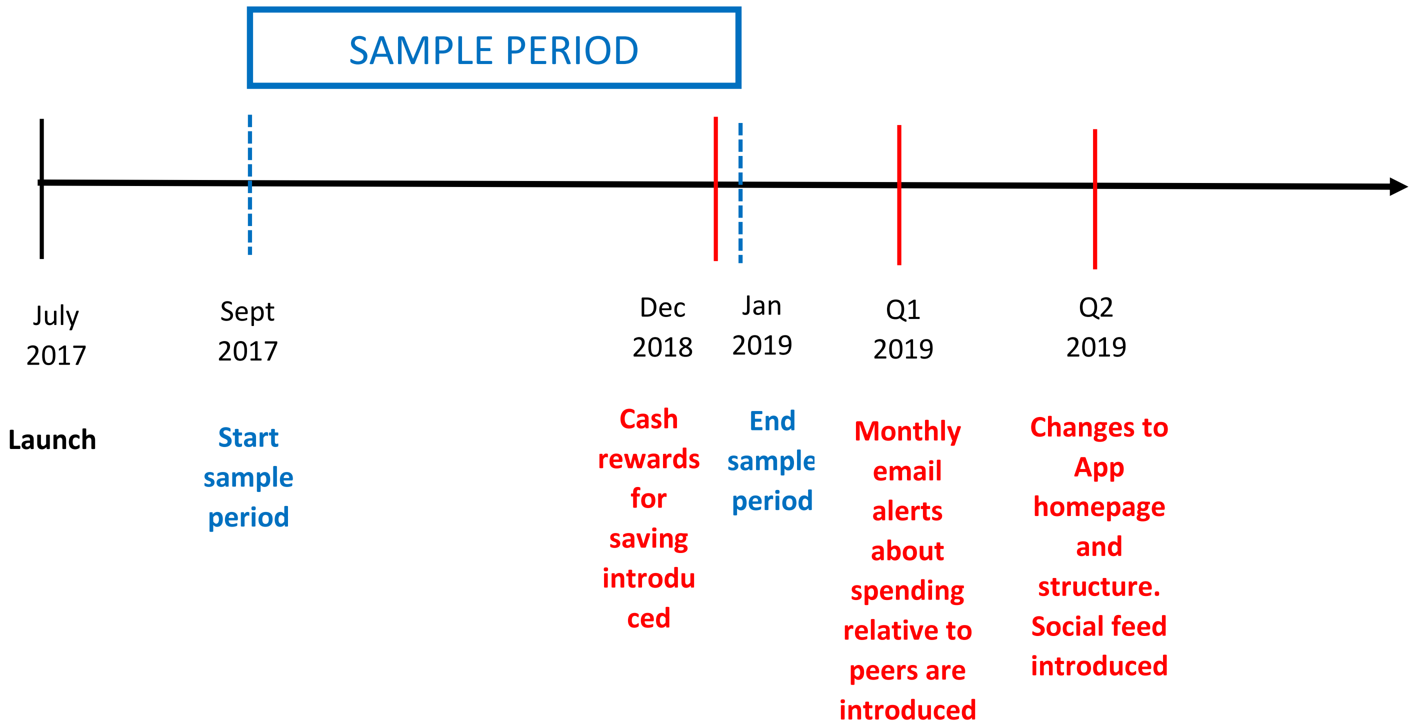


Figure A.4
Evolution of *Status Money* Features

This figure shows a timeline of the evolving features of *Status Money* since its inception and whether they happened within our sample period or outside our sample period. As the timeline shows, several features the app has adopted over time and can be observed at the time this paper is circulated were not present during our sample period, which dismisses the possibility that any of such features explain our results. Cash rewards for saving were introduced in the last month of our sample period and all our results are virtually unchanged if we exclude this months from the analysis. Note that this feature would not be able to explain the asymmetry in the spending reactions of overspenders and underspenders, because cash rewards for saving are the same for all users. Moreover, our analysis of the effects of peer-group information based on the informativeness of peer signals cannot be affected by this feature, because the rewards for saving would be the same for all users irrespective of the informativeness of their peer groups.

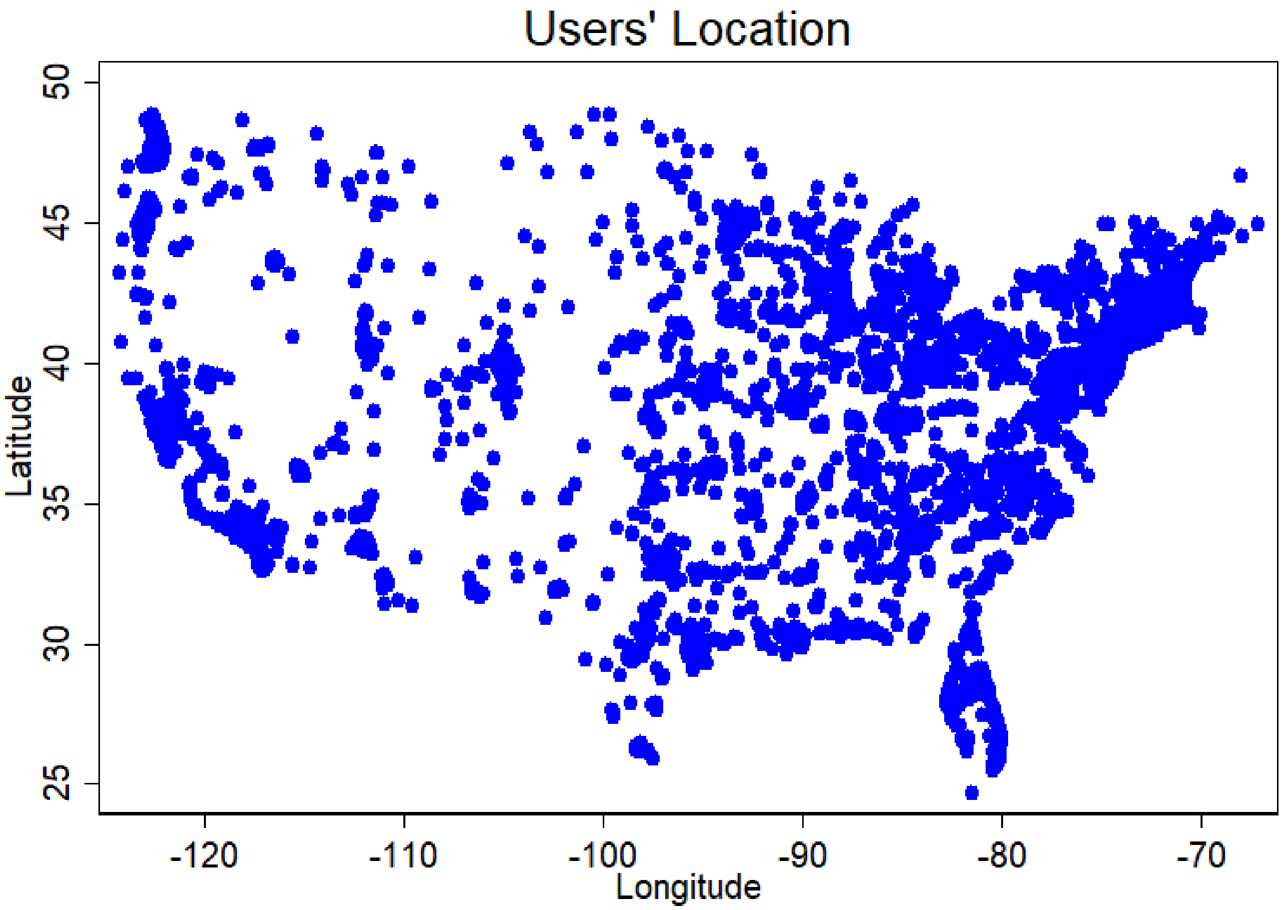
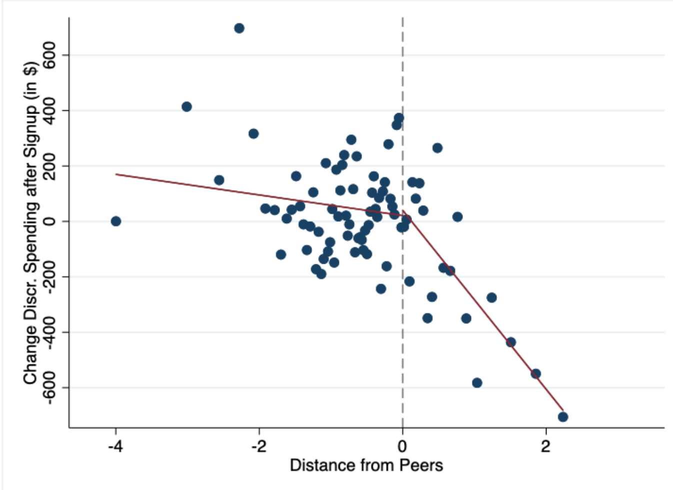
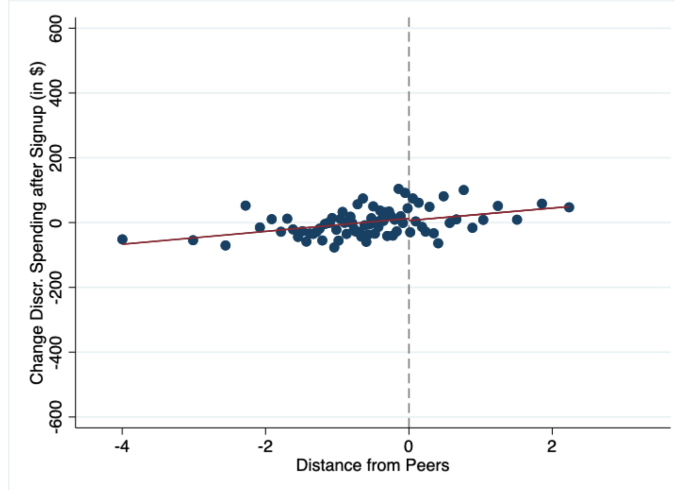


Figure A.5
Geographic Location of *Status Money* Users

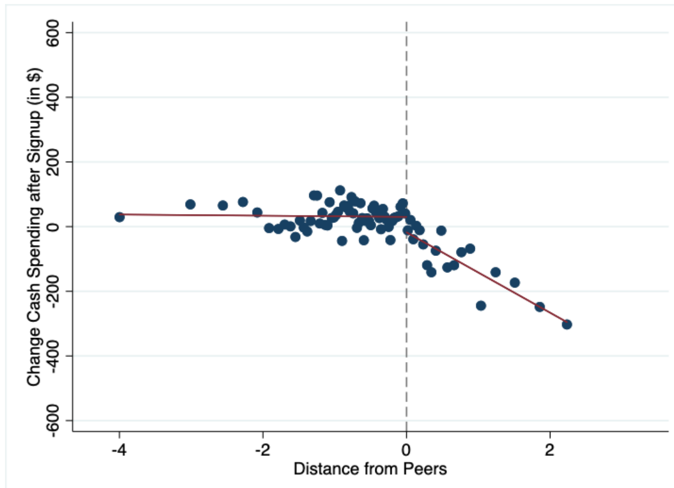
This figure plots the location of the *Status Money* users in our sample in longitude-latitude space based on their IP addresses. Each blue dot represent a user. As evidenced from the picture the dots produce, our sample of users comes from all over the United States.



(a) Discretionary Spending



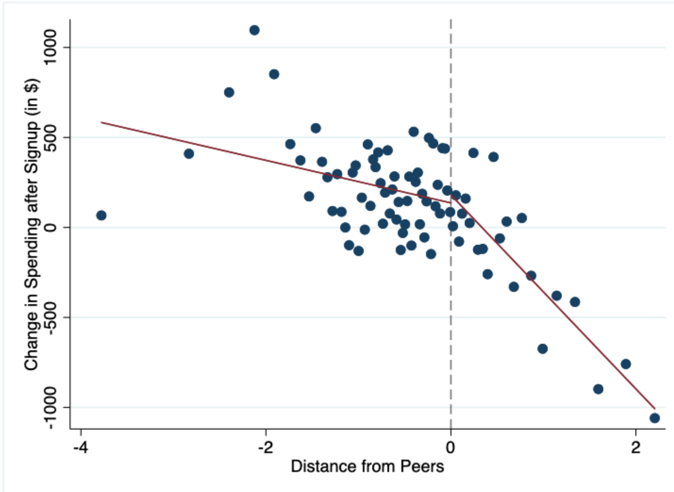
(b) Non-Discretionary Spending



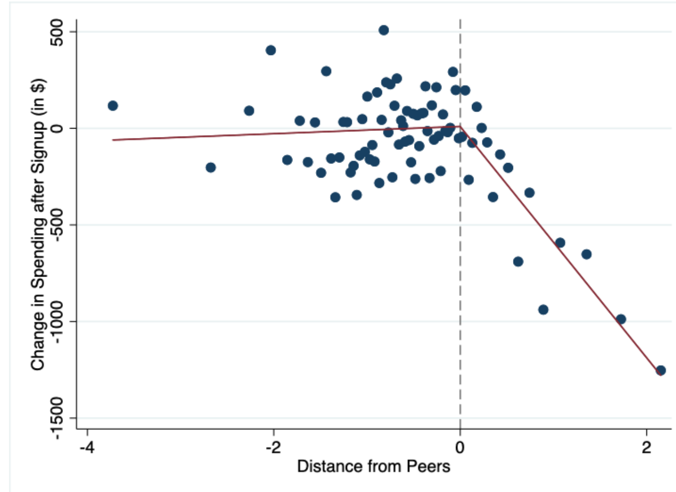
(c) Checking Account Withdrawals

Figure A.6
Quality of Transaction Categorization—Discretionary vs. Non-discretionary Spending and Cash Withdrawals

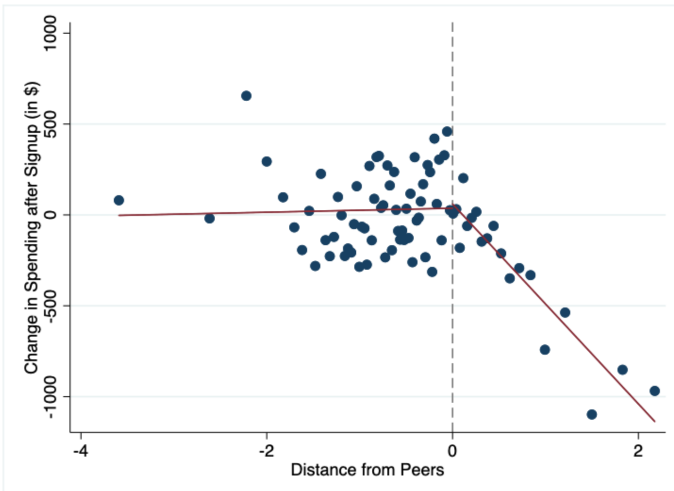
This figure shows binned scatterplots of changes in discretionary, non-discretionary consumption and cash withdrawals after signing up for *Status Money* and differences in consumption between individuals and their peer group. In all subfigures, the x-axis measures the difference in consumption with respect to peers, normalized by its standard deviation. The y-axis reports results for dollar changes in spending, computed using three months before and after signup. Subfigure (a) reports the results for discretionary consumption. Subfigure (b) reports the results for non-discretionary consumption. Subfigure (c) reports results for cash withdrawals. Each binned scatterplot divides the 21,133 users into 80 groups. In addition to the scatterplot, we report in red the fitted values of a threshold regression that estimates different linear regression coefficients below and above the zero threshold.



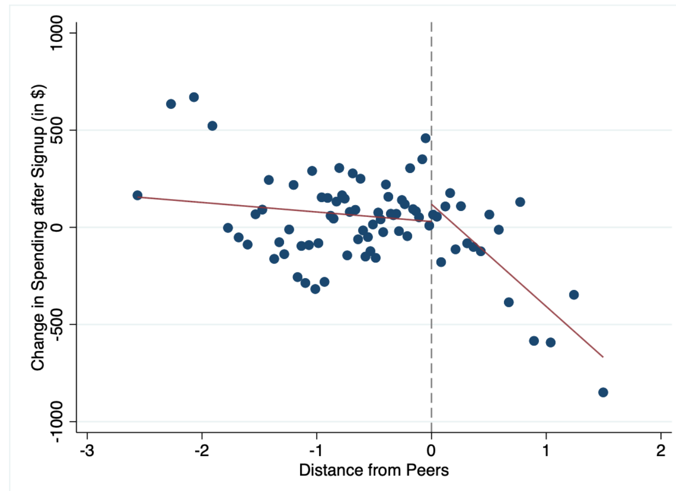
(a) At least 2 Accounts Linked



(b) Users Younger than 35



(c) Excluding Over $200K Income

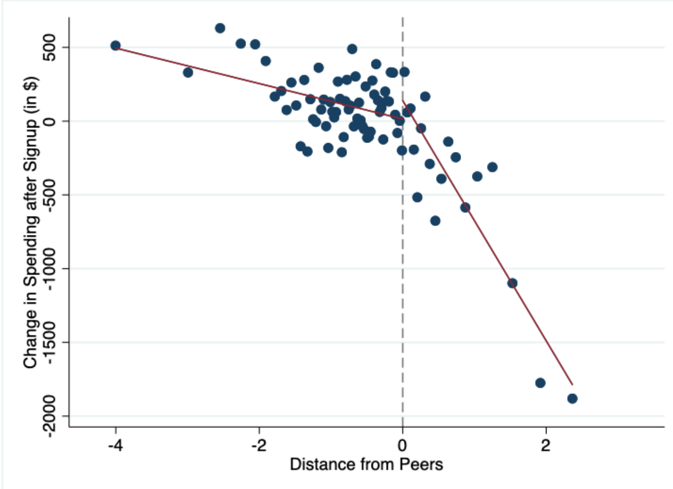


(d) Excluding Outliers

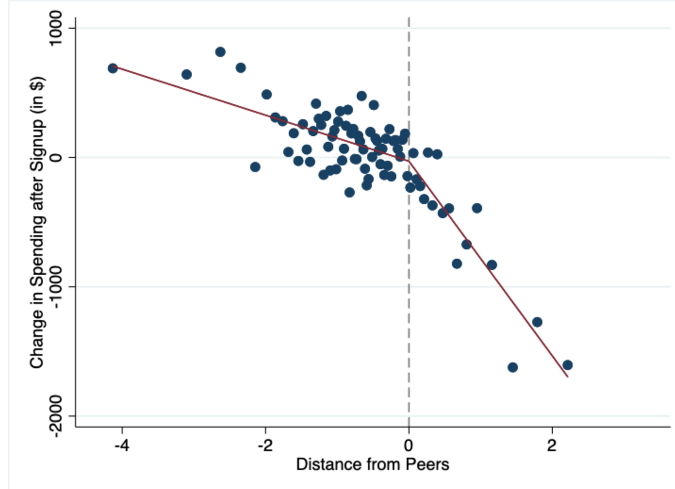
Figure A.7

Robustness: Change in Spending by Distance from Peers across Sub-groups

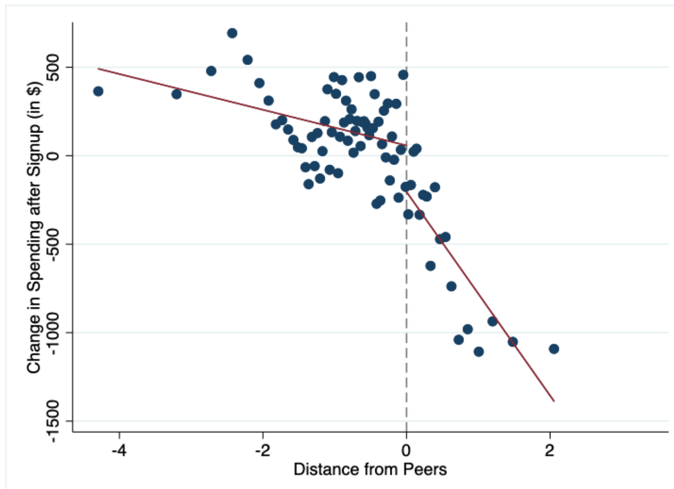
This figure shows binned scatterplots of changes in total consumption after signing up for *Status Money* and differences in consumption between individuals and their peer group. In all subfigures, the x-axis measures the difference in consumption with respect to peers, normalized by its standard deviation. The y-axis reports results for dollar changes in spending, computed using three months before and after signup. Subfigure (a) reports the results for users who have at least two accounts linked. Subfigure (b) reports the results for users younger than 35 years of age. Subfigure (c) reports results for users who earn an income of less than $200K. Each binned scatterplot divides the users into 80 groups. Subfigure (d) excludes users that lie above or below the extreme 2.5% of the distribution based on distance from peer spending. In addition to the scatterplot, we report in red the fitted values of a threshold regression that estimates different linear regression coefficients below and above the zero threshold.



(a) 2 Months Before Signup



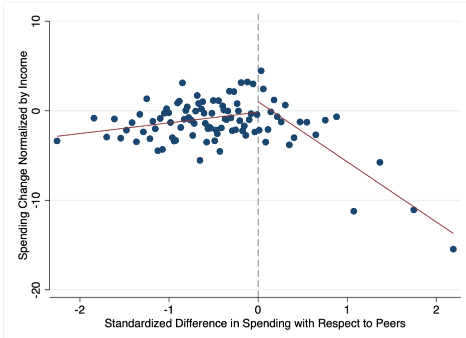
(b) 3 Months Before Signup



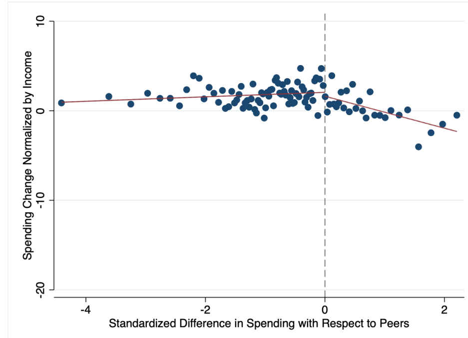
(c) Average Quarter Before Signup

Figure A.8
Mean Reversion? Assigning Over- and Underspending Based on Transactions far from Signup

This figure shows binned scatterplots of changes in total consumption after signing up for *Status Money* and differences in consumption between individuals and their peer group. In all subfigures, the x-axis measures the difference in consumption with respect to peers, normalized by its standard deviation. The y-axis reports results for dollar changes in spending, computed using three months before and after signup. Subfigure (a) computes difference in consumption with respect to peers using users' consumption two months before signup. Subfigure (b) computes difference in consumption with respect to peers using users' consumption three months before signup. Subfigure (c) computes difference in consumption with respect to peers using users' average consumption over the quarter before signup. Each binned scatterplot divides the users into 80 groups. In addition to the scatterplot, we report in red the fitted values of a threshold regression that estimates different linear regression coefficients below and above the zero threshold.

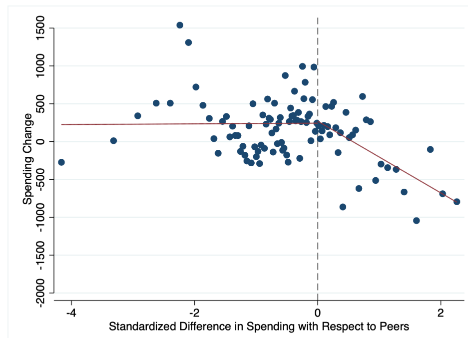


(a) Income: Below Median

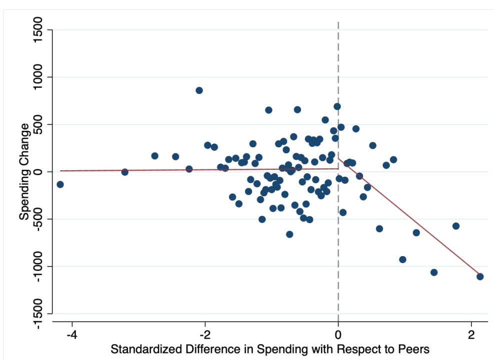


(b) Income: Above Median

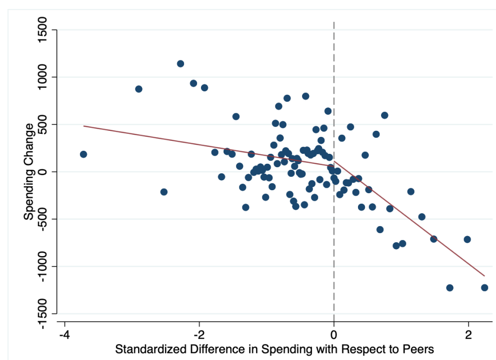
(c) Age: Below Median



(d) Age: Above Median



(e) Credit Score: Below Median



(f) Credit Score: Above Median

Figure A.9
Heterogeneous Spending Response to Peer Information by Income, Age, and Credit Score

This figure shows binned scatterplot of changes in overall consumption around signing up for *Status Money* normalized by their income and differences in consumption between individuals and their peer group. In all subfigures, the x-axis measures the difference in consumption with respect to peers, normalized by its standard deviation. The y-axis reports results for dollar changes in spending normalized by income in panels (a) and (b), which split the sample by income levels, and for dollar changes in spending in other panels. Changes are computed using three months before and after sign-up. In addition to the scatterplot, in each panel, we report in red the fitted values of a threshold regression that estimates different linear regression coefficients below and above the zero threshold.

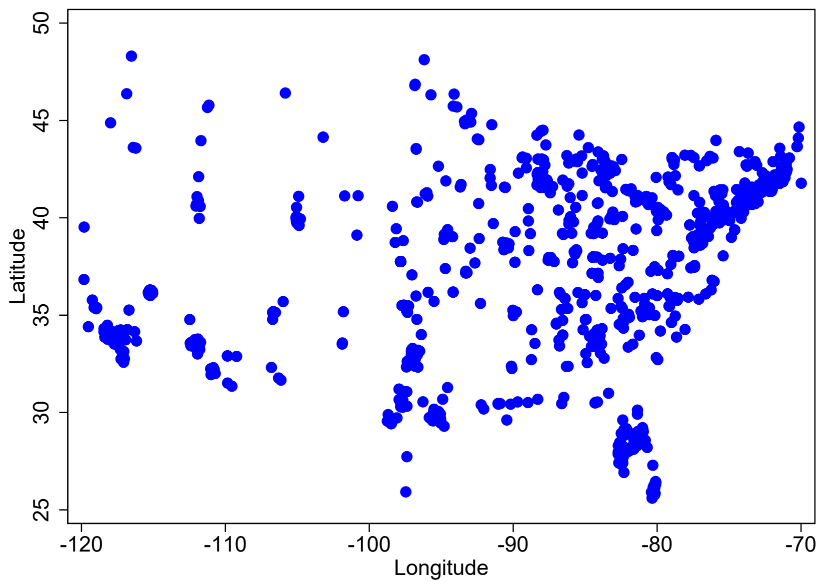


Figure A.10
Location of Respondents to the Randomized Control Trial

This figure plots the location of users based on their IP addresses in longitude-latitude space. Each blue dot represent a user in a location. As evidenced from the picture the dots produce, our sample of users comes from all over the United States and represents a varied geographic sample.

Table A.1. Robustness 1: Alternative Sample Selection Rules

	No Sample Restrictions		**Drop if Changing Num. Accounts Linked**	
	(1) Above	(2) Below	(3) Above	(4) Below
Constant	-0.094***	0.020***	-0.093***	0.022***
	(-34.97)	(3.16)	(-31.84)	(2.86)
Observations	7,560	19,061	5,522	17,339
	Drop if Food Spending <$100		**Drop if <1 Login per month**	
	(5) Above	(6) Below	(7) Above	(8) Below
Constant	-0.106***	0.033***	-0.233***	0.074***
	(-52.74)	(3.66)	(-42.00)	(8.34)
Observations	5,015	16,118	5,012	15,667

This table reports results for regressing the ratio of spending post signup to spending pre signup on a constant for users who are told they spend more than their peers (*Above*) or less than their peers (*Below*) once they sign up to *Status Money*. Spending ratios are computed using consumption three months before and after signup. Each panel refers to samples that are subject to different sample selection steps as reported on the columns' titles. Numbers in parentheses are t-statistics based on Huber-White standard errors.

Table A.2. Robustness 2: Within-month Date of Signup

Users Who Signed up in the First Half of the Month

	Above	Below	Above	Below	Above	Below
Average Change	-0.237***	0.055***				
	(-31.08)	(4.56)				
Distance Peers			-0.105***	-0.067***	-0.104***	-0.087***
			(-14.87)	(-4.04)	(-7.56)	(-4.99)
Constant			-0.167***	-0.006	0.151	1.522
			(-14.87)	(-0.31)	(0.51)	(1.48)
Individual Controls					X	X
Observations	2,695	8,398	2,695	8,398	2,272	5,720

Users Who Signed up in the Second Half of the Month

	Above	Below	Above	Below	Above	Below
Average Change	-0.228***	0.097***				
	(28.25)	(7.32)				
Distance Peers			-0.101***	-0.109***	-0.105***	-0.109***
			(-7.69)	(-5.99)	(-7.25)	(-5.59)
Constant			-0.161***	-0.002	-0.070	2.788***
			(-13.67)	(-0.10)	(-0.25)	(4.41)
Individual Controls					X	X
Observations	2,317	7,269	2,317	7,269	1,907	4,968

In this table, we report our baseline results separately for users who signed up in the first half of their signup month (top panel) and in the second half of their signup month (bottom panel). Columns (1)-(2) of this table report the average change in spending of users three months after signup relative to three months before based on whether users spent more or less than the average spending of their assigned peer group in the 30 days before signup. Columns (3)-(6) of this table report the results for estimating the following OLS specification:

$$\frac{Spending_{i,post}}{Spending_{i,pre}} = \alpha + \gamma \; Distance \; Peers_i + \delta \; \boldsymbol{x}_i + \epsilon_i,$$

where $\frac{Spending_{i,post}}{Spending_{i,pre}}$ is the ratio of user's i spending three months after signup relative to three months before. $Distance \; Peers_i$ is the difference between user's i spending and the average spending of his/her peer group in the 30 days before signup. This difference is standardized to have a unit standard deviation. The vector of individual controls x_i are the individual-level observables we have available from the app, which include a homeownership dummy, logarithm of credit score, logarithm of age, logarithm of asset balance, and logarithm of debt balance. Regression estimates are computed for users with above-peer spending in columns 1, 3, and 5 and for users with below-peer spending in columns 2, 4, and 6. Numbers in parentheses are t-statistics based on Huber-White standard errors.

Table A.3. Robustness 3: Alternative Specifications

	Above	Below	Interactions	Controls and Interactions
Average Change	-0.233***	0.074***		
	(-42.00)	(8.34)		
Distance Peers–Above			-0.200***	-0.163***
			(-11.08)	(-9.58)
Distance Peers–Below			-0.116***	-0.117***
			(-11.59)	(-11.42)
Homeownership				0.042***
				(2.64)
log of Credit Score				-0.143***
				(-3.20)
log of Age				-0.114***
				(-3.12)
log of Asset Balance				-0.009***
				(-2.84)
log of Debt Balance				-0.008**
				(-2.39)
Constant			-0.048***	1.379***
			(-4.38)	(4.42)
Observations	5,012	15,667	20,679	14,867

Columns (1)-(2) of this table report the average change in spending of users three months after signup relative to three months before based on whether users spent more or less than the average spending of their assigned peer group in the 30 days before signup.
Columns (3)-(4) of this table report the results for estimating the following OLS specification:

$$\frac{Spending_{i,post}}{Spending_{i,pre}} = \alpha + \gamma_1 \; Distance\ Peers_i \times Above_i + \gamma_2 \; Distance\ Peers_i \times Below_i + \delta \; \boldsymbol{x}_i + \epsilon_i,$$

where $\frac{Spending_{i,post}}{Spending_{i,pre}}$ is the ratio of user's i spending three months after signup relative to three months before. $Distance\ Peers_i$ is the difference between user's i spending and the average spending of his/her peer group in the 30 days before signup, which is interacted separately with a dummy for whether the user spent more than his/her peers in the month before signup (*Above*) or less than his-her peers (*Below*). The vector of individual controls x_i are the individual-level observables we have available from the app, which include a homeownership dummy, logarithm of credit score, logarithm of age, logarithm of asset balance, and logarithm of debt balance. Numbers in parentheses are t-statistics based on Huber-White standard errors.

Table A.4. Robustness 4: Statistical Inference

Panel A. Above Peers

	Coefficient	Clustering by Peer Group	Clustering by Location	Clustering by Join Date	Three-way Clustering
Distance Peers	-0.103	(-12.45)***	(-9.52)***	(-12.01)***	(-9.29)***
Observations	5,012				
Number Clusters		2,983	203	468	203

Panel B. Below Peers

	Coefficient	Clustering by Peer Group	Clustering by Location	Clustering by Join Date	Three-way Clustering
Distance Peers	-0.086	(-5.62)***	(-4.21)***	(-7.51)***	(-4.23)***
Observations	15,667				
Number Clusters		5,628	209	493	209

This table compares t-statistics estimated using alternative statistical inference methods for the results on the sensitivity of the ratio of spending post signup to spending pre signup to peer consumption. Spending ratios are computed using consumption three months before and after signup. Regression estimates correspond to the baseline specifications of our analysis, that is, the specifications in columns (3) of Table 2 for Panel A and column (4) of Table 2 for Panel B. In each Panel, the first column reports the coefficient estimates, which does not vary under alternative approaches to computing standard errors. For the other columns, we report the t-statistics estimated based on the assumptions about statistical inference indicated on top of the column. For each approach to statistical inference, we report the clustering level and the number of clusters.

Table A.5. Robustness 5: Information About Peers' Spending vs. other Information Own Income, Average US Spending, Assets, and Debts

	Below Own Income		Below Avg US Spending		Peers' Debt		Peers' Net Worth	
	Above Peers	Distance Peers	Above Peers	Distance Peers	Above Peers	Distance Peers	Above Peers	Distance Peers
Average Change	-0.202***		-0.068***		-0.031*		0.003	
	(-22.72)		(-2.65)		(-1.83)		(0.24)	
Distance Peers (Spending)		-0.122***		-0.423**		-0.157**		-0.132***
		(-4.74)		(-2.23)		(-8.91)		(-9.56)
Distance Peers (Debt)						-0.024		
						(-1.30)		
Distance Peers (Net Worth)								-0.015
								(-1.46)
Constant		-0.152***		-0.001		0.042		0.009
		(-11.07)		(-0.04)		(1.62)		(0.60)
Observations	1,885	1,885	530	530	1,933	1,933	4,824	4,824

This table reports results for the sensitivity of the ratio of spending post signup to spending pre signup to peer consumption. Spending is computed using three months before and after signup. Columns 1 and 2 condition on users that spend less than their income and columns 3 and 4 condition on users that spend less than the US average spending. Numbers in parentheses are t-statistics based on Huber-White standard errors.

Table A.6. Instrumental-Variable (IV) Analysis: Peer Spending Information and Change in Spending with Symmetric Bandwiths

	Symmetric Band: ± \$2K		Symmetric Band: ± \$3K		Symmetric Band: ± \$4K	
	First Stage	Second Stage	First Stage	Second Stage	First Stage	Second Stage
Above Dummy	0.501***		0.456***		0.435***	
	(13.28)		(15.17)		(15.95)	
Peer Spending		0.162**		0.122*		0.116*
		(2.01)		(1.74)		(1.742)
Spending Before	0.310***	-0.278***	0.318***	-0.274***	0.317***	-0.272***
	(19.45)	(-8.88)	(22.97)	(-9.80)	(24.85)	(-10.46)
First stage F-stat	176.3		230.1		254.5	
Observations	3,367	3,367	4,431	4,431	5,255	5,255

This table reports results for a two-stage-least-squares identification strategy that compares users just below and users at or above the income thresholds that *Status Money* uses to define peer groups, that is, \$35K, \$50K, \$65K, \$75K, \$100K, and \$150K. For each threshold, we only keep users whose income is at most \$2K below and above the threshold in the first two columns, \$3K below and above the threshold in the middle two columns, and \$4K below and above the threshold in the last two columns. We then estimate a set of two-stage least-square specifications in which the uncertainty in the estimate of first-stage coefficients is taken into account directly. The first stage consists of the following specification:

$$Peer\ Spending_i = \alpha + \beta\ Dummy\ Above_i + \zeta\ Spending\ Before_i + \epsilon_i,$$

where $Peer\ Spending_i$ is the peer-spending value for user i and $Dummy\ Above_i$ is a dummy variable for whether the income is at or above the threshold value. The second stage specification, which is estimated jointly with the first stage, uses the instrumented $Peer\ Spending_i$ of the first stage as the main covariate in the following specification:

$$\frac{Spending_{i,post}}{Spending_{i,pre}} = \alpha + \beta\ \widehat{Peer\_Spending\_i} + \zeta\ Spending\ Before_i + \epsilon_i,$$

where $\frac{Spending_{i,post}}{Spending_{i,pre}}$ is the ratio of post and pre consumption in the three months around signup and directly report the second stage estimate for different sample splits. Numbers in parentheses are t-statistics based on Huber-White standard errors.

Table A.7. Urban-Rural Classification

Top 20 Urban	Top 20 Rural
Chicago, IL	Las Vegas, NV
Manhattan, NY	Tucson, AZ
Brooklyn, NY	Colorado Springs, CO
Los Angeles, CA	Scottsdale, AZ
San Francisco, CA	Boise, ID
Washington, DC	Tallahassee, FL
Austin, TX	Provo, UT
Houston, TX	Fort Collins, CO
Atlanta, GA	Anchorage, AK
Denver, CO	Winter Garden, FL
Dallas, TX	Charlottesville, VA
Seattle, WA	Chapel Hill, NC
Minneapolis, MN	Lynchburg, VA
Philadelphia, PA	Morgantown, WV
Indianapolis, IN	Reno, NV
San Diego, CA	Franklin, TN
Columbus, OH	Logan, UT
Charlotte, NC	Bloomington, IN
Portland, OR	Bowling Green, OH
Arlington, VA	Grovetown, GA

This table reports the top 20 cities in the urban-rural classification.

A.3 Randomized Control Trial Instructions and Materials

In this Online Appendix, we report the crucial parts of our Randomized Control Trial.

Thank you for participating in our survey. We are interested in your views about your spending and saving plans.

No special knowledge is required. There are no right or wrong answers. Any answer is correct as long as it <u>*truly reflects your opinion*</u>*. For this reason, please do not consult any external sources of information when answering the survey.*

All responses are anonymous. If you decide to quit the survey at any stage, please let us know why, by using a special comment field available at each page. You will also be able to give us some general feedback in the end.

Thank you for your help, and welcome to the survey!

Please tell us a bit about yourself ...

Question 1 *What is your age bracket?*

- *18-30*
- *31-50*
- *50+*

Question 2 *Which category represents your total combined pre-tax income for 2019? Please include money from all jobs, income from business, farm or rent, pensions, interest on savings or bonds, dividends, social security income, unemployment benefits, Food Stamps, workers compensation or disability benefits, child support, alimony, scholarships, fellowships, grants, inheritances and gifts.*

- *Less than $35,000*
- *Between $35,000 to $74,999*

- *$75,000 or above*

Now, we would like to ask how you think you would behave under different scenarios. There is no right or wrong answer here, and it is very important for us that you just respond based on your own opinion/sentiment.

Question 3 *Imagine you unexpectedly receive a reimbursement equal to the amount you earn in a month. Please tell us your best estimate of the share you would spend over the next 30 days given your situation today:*

Slider from 0 to 100 in percent

Question 4 *Now, please, think about your situation before the COVID-19 pandemic. Imagine again that you unexpectedly receive a reimbursement equal to the amount you earn in a month. What is your best estimate of the share you would have spent over the following 30 days in that context?*

Slider from 0 to 100 in percent

Now, we will show you some information about how other people similar to you ("income peers") answered the same question.

We aim to assess whether this information affects your answers, but there is no right or wrong answer in any way. Any answers would be equally interesting to us.

Question 5 *Based on US actual spending data,* ***your income peers****—individuals similar to you in terms of income—****spend about XX%*** *of such an unexpected reimbursement.*

Now imagine again that you unexpectedly receive a reimbursement equal to the amount you earn in a month.

Please tell us your best estimate of the share you would spend over the next 30 days given your situation today.

Previously, ***you said that you would spend YY%****.*

Slider from 0 to 100 in percent

Question 6 *Now, please, think about your situation before the COVID-19 pandemic. What is your best estimate of the share you would have spent over the following 30 days in that context?*

Previously, ***you said that you would have spent ZZ%****.*

Slider from 0 to 100 in percent

Question 7 *What is your gender?*

- *Male*
- *Female*
- *Other*
- *Prefer not to answer*

Question 8 *Are you currently married/living as a partner with someone?*

- *Yes*
- *No*

Question 9 *Did your employment and/or business ownership situation change with the COVID-19 pandemic?*

- *Yes*
- *No*

Question 10 *Did your income from employment/own business change with the COVID-19 pandemic?*

- *It stayed about the same*
- *It decreased*
- *It increased*

Question 11 *What is the highest level of school you have completed, or the highest degree you have received?*

- *High school diploma (or equivalent) or less*
- *Some college but no degree (including academic, vocational, or occupational programs)*
- *College degree or equivalent (including academic, vocational, or occupational programs) or higher*

Now, we would like to ask a few more of your opinions and views

Question 12 *Let's say you have \$200 in a savings account. The account earns 10% interest per year. If you never withdraw money or interest payments, how much will you have in the account at the end of 2 years?*

- *\$200*
- *\$220*
- *\$240*
- *\$242*
- *\$280*

Question 13 *Which of the following best describes how financial decisions are made in your household?*

- *Someone else in my household makes most financial decisions*
- *I share financial decisions equally with someone else in my household*
- *I make almost all financial decisions myself*

Question 14 *On a scale from 1 to 7, how would you rate your willingness to take risks regarding financial matters, such as saving and investments?*

Slider from 1 to 7 with 1: do not trust others and 7: trust others fully

- *Yes*
- *No*

Question 15 *Imagine, you get either \$100 immediately or a higher amount of money in a month. What is the lowest amount you would be willing to wait for a month?*

- *$101*
- *$103*
- *$108*
- *$117*
- *$125*
- *$150*
- *$200*

Question 16 *Which of the following describes you best in terms of your political views?*

- *Conservative/Republican*
- *Liberal/Democrat*
- *Centrist/Independent*
- *Libertarian*

THANK YOU VERY MUCH FOR TAKING PART IN OUR SURVEY!

Question 17 *Do you have any other comments about the survey or the survey experience?*