Financial Literacy
and Personal Finance
Research Network

G53 Network Working Paper Series

Working Paper 2021-2

Crowdsourcing Peer Information to Change Spending Behavior

Francesco D'Acunto
Alberto G. Rossi
Michael Weber

DECEMBER 2021

Crowdsourcing Peer Information to Change Spending Behavior

Francesco D'Acunto[*]
Boston College

Alberto G. Rossi[†]
Georgetown University

Michael Weber[‡]
University of Chicago & NBER

This version: December 2021

Abstract

Consumers might overestimate optimal spending if forming beliefs based on others' spending, because others' conspicuous consumption is more visible than the rest of their consumption. If true, information about others' *overall* spending should change beliefs and choice. For a test, we provide crowdsourced information about anonymous "peer groups" to users of a FinTech app. Users converge to peers, especially when peer groups are more informative. For identification, we compare similar users matched to different peers based on sharp thresholds. A randomized control trial on a non-selected population supports external validity. Our results inform the design of robo-advisors for spending.

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

Keywords: FinTech, Learning, Beliefs and Expectations, Robo-advising, Visibility Bias, Social Transmission, Information Economics, Household Finance.

This research was conducted with restricted access to data from *Status* Money. The views expressed here are those of the authors and do not necessarily reflect the views of *Status* Money. We are indebted to Majd Maksad and Korash Hernandez (the founders of *Status* Money Inc.) for providing us with invaluable feedback in addition to allowing us to use the company's data. For very helpful comments and discussions, we thank Sumit Agarwal, Marieke Bos, Stephen Dimmock, Yi Huang, Byoung-Hyoun Hwang, Samuli Knüpfer, Tao Li, Chen Lin, Christian Lundblad, Michaela Pagel, Wenlan Qian, Patricio Toro, Michael Ungeheuer, Qi (Jacky) Zhang, and seminar participants at the 2020 American Finance Association, 2020 GSU-RFS Fintech Conference, 2019 TAU Finance Conference, 2019 Santiago Finance Workshop, 2019 Miami Behavioral Finance Conference, the 2019 European Finance Association, 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 Carnegie Mellon University, the International Monetary Fund, National University of Singapore, Singapore Management University, Nanyang Technological University, Hong Kong University, Chinese University of Hong Kong, Hong Kong University of Sciences and Technology, 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. Weber gratefully acknowledges financial support from the the Fama Research Fund at the University of Chicago Booth School of Business. All errors are our own.

[*]Carroll School of Management, Boston College, Chestnut Hill, MA, USA. e-Mail: dacuntof@bc.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, who often reach the time of retirement holding inadequate financial resources to maintain their pre-retirement lifestyle (e.g., see Banks, Blundell, and Tanner, 1998; Bernheim, Skinner, and Weinberg, 2001; Lusardi and Mitchell, 2007). Above and beyond financial and liquidity constraints, several consumers lack the information and ability to form beliefs about optimal spending, which would require solving complicated optimization problems (D'Acunto, Hoang, Paloviita, and Weber, 2019). For this reason, consumers might rely on the decisions of those who look similar to them as signals when forming beliefs about optimal financial decisions (e.g., see Kaustia and Knüpfer, 2012; Agarwal and Qian, 2014; Agarwal, Qian, and Zou, 2017; Maturana and Nickerson, 2019; Ouimet and Tate, 2020; Kalda, forthcoming). These signals might be informative through the *wisdom of the crowd* (Chen, De, Hu, and Hwang, 2014; Gargano and Rossi, 2018; Da and Huang, 2020).

An issue with inferring own optimal spending from what is observable about others' spending is that the overall spending and savings of others are mostly unobserved. Information about the most visible and conspicuous component of others' consumption is easier to access than information about overall spending (Charles, Hurst, and Roussanov, 2009), especially in times of widespread use of social media. By relying on what they see about others, consumers might end up with upward-biased beliefs about their own optimal spending (Han, Hirshleifer, and Walden, 2019).

In this paper, we aim to test whether this information friction—consumers' inability to correctly infer others' spending—is material to consumers' spending decisions. If this information friction were material, observing the overall spending of a large sample of individuals with similar income and demographic characteristics ("peers") might change consumers' beliefs about the appropriate spending and savings rates. This update would happen irrespective of whether peers' savings rates are optimal, as long as agents believe the signal they receive is credible and valuable (Gargano, Rossi, and Wermers, 2017). The information channel we propose does not require that individuals know their peers in real life, contrary to the more common peer effects and peer-pressure channels studied in the literature.[1]

To perform our test, we exploit a unique setting—a free-to-use FinTech application (app) called *Status Money*—in which users observe information about the overall spending of similar individuals,

[1]Note that existing research is not conclusive on whether peer effects are material—see, for instance, Duflo and Saez, 2003; Bursztyn, Ederer, Ferman, and Yuchtman, 2014; Chalmers, Johnson, and Reuter, 2014, Bailey, Cao, Kuchler, and Stroebel, 2018—or immaterial—see, for instance, Beshears, Choi, Laibson, Madrian, and Milkman, 2015; Lieber and Skimmyhorn, 2018—to households' and investors' financial decisions.

which is crowdsourced from financial-account-level data of a large representative US population. The extent to which the information that users observe relates to more or less informative peers varies substantially. This variation is important because it allows us to tease out the effect of peer-group informativeness from other potential channels, such as peer pressure, anchoring, nudges, and alerts.

Upon subscribing to the app, users provide a set of demographic characteristics that include their annual income, age, homeownership status, location of residence, and location type. *Status* also obtains users' credit scores via credit reports. For every user, *Status* constructs a peer group based on her characteristics and computes the average monthly spending among the peers. Note that peers' spending is computed using a representative sample of US consumers outside the app and that users cannot manipulate the peer groups they are assigned to.[2] Users are then prompted to link their credit, debit, and other financial accounts, thus allowing them to compare their own spending with that of their peers. Doing so also gives them access to the income aggregator features of this app, which, similar to other apps, constructs balance sheets of household finances.[3] A desirable feature of the data is that *Status* obtains the pre-signup history of transactions of its users using their financial-accounts information. Hence, we can compute directly users' pre- and post-signup average monthly spending.

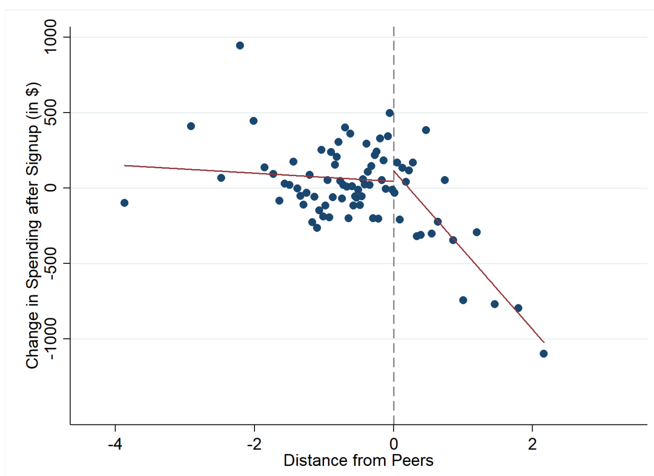


Figure 1. Distance from Peers' Spending and Change in Spending: Raw Data

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 dollar spending for

[2]After starting to use the app, users can set up *additional* peer groups, which are endogenously chosen by users. Our analysis does not use such endogenous groups, but the original groups *Status* sets up at signup, which is always visible to users. Note that less than 2% of the users in our sample set up an additional customized peer group.

[3]As we discuss in detail below, we only consider users with actively linked accounts and whose number of linked accounts does not change over time.

the users in our sample in the three months after signup, relative to the three months before signup. On the x-axis, we sort users based on the distance of their spending in the 30 days before signup relative to the peer group's average spending, which is vividly reported on the app's homepage.[4] The distance from peers is positive if users spent more relative to the average peer spending before signup, and negative otherwise. The raw data suggest two qualitative results. First, users converge to their peers' spending after signup, with an average drop in spending for users above their peers and an average increase for other users. Second, the reaction seems asymmetric—users above their peers change their spending by a larger absolute amount than users below their peers at a similar distance.

Several channels could explain this raw-data motivational evidence, potentially unrelated to a conscious reaction of users to peer information. For instance, mean reversion in spending might explain these results if users who face especially large expenses in the 30 days before signup merely reverted to their usual levels of spending after signup. In fact, similar to the decision of soliciting help when dieting, the sign-up decision might be driven by users who believe they spend excessively. Moreover, users might be reacting to other pieces of information or other features of the app, and not necessarily to the information about peers' spending. The baseline analysis also faces a set of endogeneity concerns, such as the possibility that users might have already decided to cut or increase their spending before signing up (e.g., see D'Acunto, Prabhala, and Rossi, forthcoming; Rossi and Utkus, 2020).

We dedicate most of the paper to showing that, indeed, users react to the peer information they receive. We first rule out directly the effect of mean reversion and different sources of information in multivariate analyses. To rule out mean reversion, we show that the raw-data 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 their distance from peers, thus eliminating any role of spending in the weeks just before signup.[5]

Other spending information in the app refers to the average spending of all US consumers, instead of only the user's peer group, as well as information about users' own income flows based on their transactions. We show that users react more to peer information relative to other pieces of information by considering those cases for whom peer spending and other pieces of information predict reactions

[4]Several app 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 changing features and the timing of these changes based on the timeline the app founders shared with us and official press releases. All the features we discuss in the paper refer, obviously, to what our users saw during our sample period, and not to periods outside our sample.

[5]We do not argue that mean reversion in spending does not exist. To the contrary, we do find evidence of mean reversion in spending over time. What we argue is that this 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.

of opposite directions. In these cases, users change their spending in the direction of peers and not in the alternative directions. This analysis also helps reconcile our results with earlier research that did not detect substantial effects of peer information on household finance choices (e.g., see Beshears, Choi, Laibson, Madrian, and Milkman; 2015): even in our setting, similar to earlier work, providing information measured on broad groups with varied demographics, such as the average of all US consumers, has barely any effects on choice.

To tackle the broader endogeneity concerns directly, we propose an instrumental-variable (IV) analysis that exploits a design feature of the app—the quasi-random assignment of users to peer groups based on sharp income thresholds. Two users with similar incomes (as well as similar observable demographics), but one with income slightly below a threshold and the other at or above the threshold, are matched to different peer groups and hence obtain different information about their peers' spending. For example, consider one of the yearly income thresholds the app uses—$100K. Suppose user A declares she earns $99K, whereas user B reports earning $101K. Although these reported incomes only differ by $2K, users A and B will be matched to largely different peer groups. User A will face a peer spending that is the average of the transactions of otherwise similar consumers earning between $75K and $99K, whereas user B will observe a higher peer spending—the average of the transactions of consumers earning between $100K and $149K.

We show directly that observables are not economically or statistically distinguishable across users just below or above the income thresholds. The amounts spent before signup are also economically and statistically indistinguishable around the thresholds, which rules out a potential role of mean reversion in the results. Because users do not know this assignment rule, let alone the threshold values, they cannot strategically manipulate their reported incomes to avoid receiving negative news about their spending relative to peers.

The IV strategy confirms our baseline results. Users who look similar along all dimensions we consider, including their pre-signup spending, change their spending differently by converging to the average spending of the peer groups to which they are assigned.

We also corroborate the IV results through a placebo IV analysis. We set placebo income thresholds at round dollar values, like the actual thresholds, but for values the app does not use. Users around the placebo thresholds are similar to each other in all observable dimensions, and their change in spending is also similar after signup. This test also rules out that a systematically different behavior of income rounders and non-rounders drives our results, because the placebo thresholds are set at salient round income values like the actual thresholds the app uses.

To further assess the relevance of an information channel relative to peer pressure or alerts and reminders on the app, we move on to test the heterogeneity of users' reactions based on the informativeness of the peer signals they receive. In these heterogeneity tests, channels other than peer information are kept constant, because they are equally at work for all users. We exploit a rich set of app features that allow for variation of the informativeness of peer groups, such as the precision with which users' characteristics are matched to those of the peer group to the size of the peer group. We also construct measures of the extent to which users might have accurate information about peer spending before signup. Across all these dimensions, we find that users who are matched with more informative peer groups react substantially after signup, whereas users who are matched with less informative groups barely react.

In the last part of the paper, we discuss another issue that the app-based analysis cannot tackle: the external validity of our results. Because *Status* is marketed as a tool that provides comparisons with peers, the population that selects into using this service might be more sensitive than the average US consumer to the differences between their spending and peers' spending. Users might thus react more to the peer information than the average US consumer. This feature makes *Status* an ideal laboratory to study the channels through which learning about others' spending might affect choice, but it hinders us from making claims about the predicted effects of providing peer information to the broader population.

To tackle this issue of external validity, we propose a randomized control trial (RCT), which we ran on a broader, representative US population that was selected through an online platform without any references to peers, peer information, or household finance. We only started to mention these topics after respondents had accepted to be part of the RCT. Similar to the app setting, we gave respondents information regarding the marginal propensity to consume of people in the same income bracket (income peers), and tracked their reactions in terms of reported marginal propensity to consume. First, we find this non-selected draw of the US population responds to the provision of peer spending in a strikingly similar way as *Status* users. This result casts doubts on the possibility that the results in our main analysis are only true for a selected population that cares about peers and/or managing their finances. Second, the randomized trial allows us to elicit additional demographic information as well as economic preferences and beliefs (D'Acunto, Rauter, Scheuch, and Weber, 2020; D'Acunto, Malmendier, and Weber, 2020; Coibion, Gorodnichenko, and Weber, 2019; Coibion, Georgarakos, Gorodnichenko, and Weber, 2020). We find that demographic dimensions that are unobserved in our main sample and typically relate to spending attitudes, such as gender, marital status, number of children, as well as risk aversion and patience, do not predict systematically different reactions to peer information, which further reduces the concern that our baseline results have no external validity.

Stressing what we do not claim is also important. We do not claim that the information that users obtain in our setting captures precisely the actual spending of their peers, or that the ways in which peer groups are constructed in our setting are optimal in any respect. The very notion of peers is hard to define precisely and could be based on many different dimensions relative to the ones the app and our RCT use. We find users do understand the cases in which the information they face is more or less tailored to their characteristics and react accordingly, which in fact allows us to isolate the information channel more precisely.

Moreover, we do not claim that users' reaction to peer information is optimal in any respect. Forming beliefs about optimal spending and household financial management based on easy-to-access signals and heuristics, which are often suboptimal, has been documented in the context of non-expert consumers (for instance, see Leary and Wang, 2016; Keys and Wang, 2019; and Argyle, Nadauld, and Palmer, 2020). Our paper has no normative implications in terms of whether reacting to this information makes consumers better or worse off. Having said this, in our RCT on a non-selected population, we find that the overall-sample average MPC, including both overspenders and underspenders, on average declines by 2% after exposure to peer information due to the stronger response of overspenders relative to underspenders.

Further research should be devoted to studying the optimal design of FinTech tools based on crowdsourced information to provide tailored advice to users. For instance, whether including more categories than the ones we studied (e.g., amount of mortgage debt or student debt outstanding) might have a stronger effect on behavior is an interesting avenue for future research. In fact, the provision of crowdsourced peer information through FinTech apps, which make such information readily and easily available to consumers, might be considered as a form of robo-advising for spending and saving based on the taxonomy of D'Acunto and Rossi (2020) (see also Gargano and Rossi (2020)). The randomized control trial we designed for external validity purposes could be the starting point for future research endeavors in this direction, which are also likely to deliver policy implications.

The persistence of the effects is another aspect further research should assess. Within the time frame we observe, which includes about three months around the signup to *Status*, we do not detect any dissipation or reversal of the effects. Whether peer information might have long-lasting effects on savings will require the observation of longer time series than are currently available.

2 Institutional Setting

In this section, we discuss the characteristics of *Status*, the signup procedures, and the information users observe after signup. We also discuss how the app has evolved over time.

2.1 Purpose of the App and Signup Process

Status is an app designed to help individuals make more informed decisions in the personal finance space. Similar to other apps in the US and abroad, *Status* 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. The unique feature of *Status* is that it also shows users how individuals 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 representative of the US population outside the app. The app thus enables users to easily gain access to complex information they could barely acquire and process on their own.

Importantly, peers in our setting are not individuals that interact socially with users, and users do not know their identities. This feature is crucial to test the information channel, because if we were showing users information about the spending and savings of their friends and acquaintances, users might already know this information through social interactions. Users know that the peers are a representative population with certain demographic characteristics, and that peer statistics are computed using the full set of transactions made by such individuals. We do not estimate peer effects, which the literature attributes to the extent to which social interactions change one's behavior. Rather, we estimate the extent to which obtaining crowdsourced information changes users' behavior.

To sign up for *Status*, users provide their date of birth, their annual income, and their housing type—whether they own or rent the home in which they live. 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 user's credit-score-related information.[6] Finally, the app asks users to link their checking and savings accounts, their credit-card accounts, and their taxable and non-taxable accounts.

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

2.2 Peer-Group Construction

For each user, the app constructs a peer group based on the user's age, income, location, credit score, and housing type. Peer group precision varies based on the constraint that each group should include at least 5,000 underlying observations. In Figure 2, we provide an example of the screenshot that *Status* 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 constructed for this user contains individuals whose age ranges between 40 and 49, whose income ranges from $100K and $150K, and who live in New York City, pay rent, and have a credit score that ranges between 720 and 779. Because this user belongs to the right tail of the income distribution, the range or peers' income is wide. Users with lower levels of income are assigned to narrower income ranges due to the larger mass of peers with mid-income levels.

2.3 Variation in Informativeness of Peer Groups

Our setting is unique, because the various peer groups constructed by the app are more or less informative, depending on the characteristics of the individual. This unique feature allows us to test whether users' reaction is related to the informativeness of the peer group to which they are assigned. Take the number of characteristics used to compute peer groups. For certain users, the algorithm is able to construct peer groups based on the full set of characteristics, that is, age, income, location, credit score, and housing type. For others, the algorithm does not have sufficient information to construct the peer group using all the characteristics and therefore uses broader categories. For example, it may not distinguish between renters and homeowners in smaller cities. Users whose peer-group information is based on all characteristics are likely to perceive the signal as more informative.

A second dimension is the number of people in the peer group. The app tells the user how many peers are in her peer group, and the higher the number of peers on which the group average is based, the more informative is the signal. For example, the peer group constructed for the fictitious account reported in Figure 2 is based on 9,900 individuals, as can be seen in the top-right corner of the figure.

Third, the most important category to construct peer groups is income, because peer spending varies most substantially by income. The main thresholds used to compute income groups are $35K, $50K, $75K, $100K, and $150K. The cut-offs are constructed in such a way to maintain approximately an equal number of individuals across all groups. As a result, however, individuals with incomes of

$100K dollars or more have peer groups that are less representative than individuals with incomes below $100K. For example, a finance faculty earning $240K+2/9 will be in the same peer group as an investment-banking managing director earning in excess of $1M and a white-collar employee earning $150K. Even though they are in the same peer group, these individuals are likely to have very different spending patterns and might find the signal uninformative. On the other hand, a new college graduate earning $40K a year would be compared with peers earning between $35K and $50K, making the peer comparison informative.

Fourth, the app categorizes users as *Urban* or *Rural*. *Urban* areas comprise densely populated areas, such as Chicago, Manhattan, and Los Angeles. We report the top 20 urban areas by users in Table A.1. *Rural* areas instead comprise smaller cities. The most common areas in this sample are cities such as Tucson, Colorado Springs, or Scottsdale. We expect the information to be more informative for rural users, compared to urban ones, because in high-density areas we expect individuals to be able to observe other people's spending more precisely even absent the app. We acknowledge that the cut by urban versus rural users is also open to alternative interpretation which we discuss in the results section and we do not intend to overemphasize this specific heterogeneity test.

Fifth, we can measure the intensity with which a user is exposed to peer information by the number of times the user logs into the app. Users who (endogenously) log in more frequently observe the signal more often and hence might recall it more than users who log in infrequently.

We exploit these features of the app to estimate whether the informativeness of the peer-group information and the intensity of information acquisition relate to stronger reactions by users in section 6.

2.4 Changes in App Features outside the Sample Period

As is the case for a large number of apps, *Status* has evolved over time. In Figure 3, we report the timeline of the changes to features that might be relevant to our analysis. In essence, any relevant changes to the features of the app, such as the display of peer information, other information, or other alerts and incentives, only happened after the end of our sample period—which starts in July 2017 and ends in January 2019—and therefore do not affect our analysis.

The app was launched in July 2017 in the form we describe below in section 2.5. 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 sending alerts about spending relative to peers, updating individuals on whether they are underspending or overspending with respect to their peers on the 15th day of each month.

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. During 2019, the app also evolved in its appearance. The current version of the app, for example, emphasizes spending categories (see Figure A.1). The version of the app during our sample period, instead, reported only total spending in the home page, relegating information about spending categories to secondary hyperlinked pages. Articles and presentations about *Status* help reconstruct the various phases of the app appearance. For instance, see the following description of the app reported by CreditDonkey (`https://www.creditdonkey.com/status-money-review.html`) published on 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. 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`).

2.5 Main Features of the App

Once the user enrolls, the app automatically retrieves information from her financial accounts. The app stores all transactions and investment returns 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 4 displays the vivid graphics that compare the users' own daily spending based on daily transactions with the projected average daily spending of the peer group and the national US average. The screenshot was taken on October 30, 2018. On the top, the plot shows the user's total spending, which turns out to be \$17,799, together with the average peer spending, \$8,651, and the national average, \$4,222. 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, \$10,204. As a final piece of information, the app explicitly tells the users how they are doing in terms of spending for the current month.

Note that users' spending is based on their own daily transactions. Recall that 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* observes. These transactions are aggregated at the monthly level and projected linearly for each day of the month, and 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, or by levels of informativeness.

The bottom of the home page reports more comprehensive statistics regarding users' debts, assets, net worth, and credit score (see Figure 5). On top of this information, the app tells the user whether the interest rate she is charged is competitive with the national average. In Figure 5, the interest rates are competitive for two of the three accounts, but not for the third. The third quadrant reports the information for net worth, which is simply computed as the difference between assets and liabilities, and the fourth quadrant reports information for the user's credit score. Below, we propose tests to show that users react to peer information rather than other pieces of information they receive from the app.

3 Data

Status collects and displays large amounts of information from and to its users.

First, we observe a set of 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 zip code and city in which the user lives, and her credit score.

The second set of characteristics we observe relates to users' peers. For each user, we observe the average characteristics of the peer group computed by *Status*. As we discussed above, *Status* does not use the characteristics of other app users to construct peer groups, but uses external proprietary data on a representative set of US consumers. This procedure rules out that any selection in the types of consumers who sign up for *Status* are reflected in the average demographics of the peer groups. This procedure might also help explain why the majority of users in our sample underspend relative to their peers.

The peer demographics we observe include the average credit score of the peer group, the average debt, the average value of assets, the average net worth, and the average income. Finally, we observe the number of individuals populating each group.

Third, we observe a set of variables that capture the usage of *Status* accounts by account holders. These variables include the signup date, the number of account logins by users during the first, second, and third month after signup, and the number of financial accounts users link to the app. Finally, we observe data on users' and peer groups' spending amounts across spending categories. Specifically, we observe users' spending over the first, second, and third month before users signed up to *Status* as well as their spending over the first, second, and third month 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 select the raw sample according to the following steps.

First, we only include users who have linked at least one spending account at signup to exclude users who might merely sign up to observe the peer information and whose spending information we would not observe.

Second, we only include users whose number of linked accounts does not change through the 90 days after signup, which is the longest horizon we use in the analysis. This step is important for two opposite reasons: First, 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. Second, users' accounts might be de-linked from the app due to inactivity, changing passwords, or other technical issues related to their account's settings.[7] In this case, we might falsely measure a drop in users' spending, which is instead merely due to our inability to track unobserved spending over time. This issue would be problematic for our analysis if account delinking were systematically more likely for users who overspend relative to their peers. By limiting the sample to users whose number of accounts do not change across our time period, we eliminate both possibilities from our analysis.

Third, we only include users with at least one monthly login to the app after signup. This step further alleviates concerns that, because of users' inaction, the app fails to download information from users' accounts, which would appear as a misleading drop in inflows and/or spending.

[7]We thank Michaela Pagel for raising this point.

Fourth, we only include users who spend at least $100 per month in food-related transactions during the sample period. This step ensures we do not keep individuals who are not actively using the accounts they link for spending purposes.

Finally, we 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

Figure 6 is a map of the US plotting users' geographic location based on their IP addresses. The figure shows our users are spread across the whole US, which is important to ensure that individuals with varied cultural backgrounds enter the sample (see D'Acunto, 2018 and D'Acunto, 2019).

Panel A of Table 1 reports the basic characteristics of the users in our baseline sample. For each variable, we report the number of observations, averages, and standard deviations. The first three variables are demographic characteristics: *Age*, *Credit Score*, and *Home Ownership*. The average client is 32 years old, with a standard deviation of eight years, indicating *Status* users are rather young. The average credit score is 734, higher than the average US credit score of 687. Thirty-nine percent of users are homeowners, which is below the US average, in line with the fact our app users are, on average, younger than the average US consumer.

The average client earns a little more than $90,000 per year, with a large standard deviation of $62,000, suggesting our sample spans individuals with varying levels of income. Finally, monthly spending equals $4,860, with a standard deviation of $4,026.

Figure 7 reports the distribution of monthly spending by income quartiles. We highlight two main facts that suggest our data align with intuition and are reliable. First, monthly spending increases with income. Across the four income groups, average spending equals $2,380, $3,550, $4,785, and $7,555. Second, the within-group standard deviation of spending increases with income. Higher-income individuals have more varied levels of spending than lower-income individuals.

Panel B of Table 1 reports the summary statistics for the identification sample we use in section 5. As we describe in detail below, 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 very much in line with the ones of the baseline sample, with the exception of annual income and monthly spending, which equal \$71,916 and \$2,806, respectively. These lower values are due to the fact that the identification sample includes only individuals with income up to \$151,000, whereas the baseline sample includes users whose income is well above \$200,000.

4 Peer Information and Spending: Baseline Results

In this section, we start by reporting baseline results that use the raw data, and estimate the changes in consumption after signing up for the app and observing peer-group behavior. We then present formal multivariate regression analysis results that control for a variety of observable characteristics. We conclude the section by showing results for different measures of over-spending and under-spending at signup as well as results showing individuals indeed react to peer information, as opposed to reacting to other pieces of information displayed by the app.

4.1 Raw-Data Evidence

Our baseline analysis tests whether the two pieces of information users receive at signup—whether they spend more or less than their peers, and how different their spending is with respect to peers' spending—relate to users' subsequent spending behavior. We first compute the overall spending for each subscriber 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 cyclical, we deduct the average monthly change in spending across all users from the change in spending of each user. We refer to this quantity as seasonally-adjusted spending in some cases and simply as spending in other cases.

Figure 1 in the introduction presents a graphical depiction of the change in users' spending after and before signup as a function of their distance from the value of peer spending that users observe at signup. The plot is a binned scatterplot that divides the 21,133 users in the baseline sample 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. Overspenders reduce their spending on average by \$229/3=\$76.3 per month, and underspenders increase their spending on average by \$107/3=\$35.6 per month.

Second, the sensitivity of users' change in spending based on whether users spend more or less than their peers is asymmetric. Third, the distance of users from their peers' spending is monotonically related to users' change in spending—the further the group is from the peers' spending level, the higher the change in their spending, irrespective of the sign. The regression slope coefficients associated with the regression lines in Figure 1 are -43 for underspenders and -536 for overspenders. These coefficients imply that a standard-deviation increase in "Distance from Peers" decreases spending by \$43/3=\$14 per months for underspenders. On the other hand, a standard-deviation increase in "Distance from Peers" decreases spending by \$536/3=\$178.6 per months for overspenders.

As discussed above, the average subscriber underspends relative to her peers, plausibly because peer information is from a representative US sample, whereas our users are likely more concerned about their own finances than the broader population, as their decision to sign up to the app reveals. We tackle this external-validity issue with our sample in the RCT we discuss in Section 7.

The results reported in Figure 1 compute the change in overall spending. Although the app categorizes the expenses in various groups (see Figure A.2 for a few examples), we cannot assess the precision of the categorization. As a reality check, we report the changes in spending separately for discretionary and non-discretionary spending in Figure 8.[8] Intuitively, we would expect individuals to change their consumption by adjusting their discretionary spending, rather than their non-discretionary spending, which is exactly what we observe in Panel (a) and Panel (b). Panel (a) is very similar to Figure 1, both qualitatively and quantitatively. Panel (b), on the other hand, shows virtually no change in spending by the users after signing up for the app. We cannot certify how well transactions are categorized in our setting and which observables determine the quality of the categorization. For instance, categorization may be easier for big retailers such as Costco or WholeFoods, but not for smaller retailers or family-owned shops. For this reason, in Panel (c), we show the behavior for cash withdrawals, that are unlikely to be mis-categorized, because they contain the strings "ATM" and/or "Withdrawal."[9] Overall, we observe a rather strong reaction post signup for cash withdrawals. Because of the miscategorization potential, all the results in the paper are based on overall spending and not on spending in specific categories.

[8]Spending is broken down into categories based on classifying the vendors of each transaction. The transactions that cannot be classified are labeled as other expenses. 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. Because we cannot ultimately classify the remaining categories into discretionary and non-discretionary, we exclude them for the sake of Figure 8.

[9]We thank Byoung-Hyoun Hwang for this suggestion.

Finally, for us to correctly estimate our results, subscribers must use the app. For example, users should link all their financial accounts to limit measurement error. Because we cannot check whether users link all their financial accounts, as we discussed in section 3.1, we limit our analysis to only users who link at least one account and do not change the number of linked accounts in the first three months after signup. To further assess the scope for measurement error due to misreported accounts, in Online Appendix A.1 we repeat the raw-data facts on various sub-populations that are more and less likely to be affected by measurement error.[10] Specifically, we replicate our raw-data facts in the subset of users who link at least two accounts to *Status*, for users who are 35 years old or younger—and are less likely to hold many spending and investments accounts—and for users whose income is below $200K, who again might be less likely to hold several accounts. Our baseline, raw-data facts are qualitatively similar within all sub-populations.

4.2 Multivariate Regression Results

The results reported in section 4.1 have two limitations. First, by computing changes in spending, they do not control for the fact that a reduction in spending of $100 is very different for a user who spends $2,000 before signup, than for someone who spends $6,000 before signup. Also, mean-reversion in spending can likely account for some of the effects we document in section 4.1. Intuitively, a user who buys an iPhone in the month before signup is likely to reduce spending the following month simply because the extraordinary monthly iPhone expense is not repeated in the following month.

To directly control for both issues, we construct the ratio of spending over the 90 days after signing up for the app divided by the spending over the 90 days before signing up for the app. To ease the interpretation and account for the fact that spending is seasonal, we standardize this measure so that it has a mean equal to 0 and a unit standard deviation. The standardization is such that the ratio we construct is numerically identical to the standardized percentage change in consumption after signing up for the app, relative to before.

Table 2 reports the results. 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 results in section 4.1, we find that overspenders significantly reduce consumption, whereas overspenders significantly increase consumption after signing up.

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

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

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

where $\frac{Spending_{i,post}}{Spending_{i,pre}}$ is the standardized change in spending, $Distance\ Peers_i$ is the difference in spending at the time of 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; the final covariate, $Spending\ Before_i$, controls for spending in the month before signing up and directly controls for mean-reversion in spending.

Starting from columns (3) and (4) in Table 2, we find a significant relationship between distance from peers and change in spending for both underspenders and overspenders, with a stronger effect for overspenders. Measuring the change in consumption in percentage changes rather than dollar changes, however, shows that users' sensitivity with respect to peer spending is close in magnitude for the two groups. Columns (5) and (6) repeat the results, but add control variables. For both underspenders and overspenders, introducing additional controls does not change the coefficient estimates for the regressor of interest, $Distance\ Peers_i$.

Finally, in columns (7) and (8), we add past spending as a control. The coefficient on past spending is negative, indicating spending indeed mean reverts. However, the coefficient on $Distance\ Peers_i$ is still highly economically and statistically significant for both underspenders and overspenders. The coefficient for underspenders is virtually unchanged, whereas the coefficient for overspenders decreases slightly. These results exclude that mean reversion, which exists for both overspenders and underspenders, drives our results because, once we control for previous spending directly, the results are qualitatively unaffected.

4.3 Reaction to Information or Mean Reversion in Spending?

Although the coefficients on $Distance\ Peers_i$ remain significant for both over- and underspenders, the inclusion of pre-signup spending as a control has an impact on the coefficient estimates. To ensure that controlling for mean reversion in alternative ways does not alter our main findings, we repeat the estimates in the last two columns of Table 2, but change the definition of consumption used to compute the distance from peer spending.

In particular, in Table 3, the first two columns compute the distance with respect to peers using spending in the second month before signup rather than the previous month; the middle columns repeat the exercise using the third month before signup; and the last two columns use average spending over the whole quarter before signup. These specifications are an alternative way to address the issue of mean reversion: They rule out directly that exceptionally large or small expenses in the month before signup mechanically drive our assignment of users to overspenders or underspenders, and hence that we misattribute changes in spending over time due to mean reversion to a reaction to peer information.

In all cases, our results are robust both economically and statistically. In Figure A.4 of the Online Appendix, we show also that the raw-data facts discussed above are similar if we compare users with peers based on their spending well before signup.

Overall, controlling directly for mean reversion in spending either by partialling out pre-signup spending or by assigning users to peers based on spending well before signup reveals that mean reversion in spending does not drive our results.

4.4 Information about Peers or Other Information?

As we discussed in Section 3, the app provides additional information to its users on the homepage and on additional pages. The comparison between users' spending and peer spending is the first, most prominent, and most vivid piece of information users see upon signing in to the app during our sample period. The same picture, however, provides information about the average US consumer's spending, which is computed as the average consumption of all the individuals used to create the peer groups, as well as information about the user's own average income, computed as the average monthly inflows into the user's accounts. One might wonder whether our baseline results are genuinely driven by users' reaction to information about their peers' spending, or whether users reacted to the other pieces of information they received at the same time. For instance, if a user is simultaneously above her peers' spending, above the average US spending, and spends more than the estimated monthly income inflows, the fact that the user reduces her spending after signup might be driven by a reaction to any of these three pieces of information, which all go in the same direction.

To confirm that users respond to peer information, we propose two tests in subsamples of our data. Both tests are based on the idea of isolating users for which the peer spending information and the other sources of information predict reactions in opposite directions. If we observe that users change spending in the direction of their peers, instead of the direction suggested by the other pieces of information, we can conclude that users indeed react to peer-group information.

We first consider users who lie above their peers in terms of spending but below their average monthly income inflows. In this case, if users reacted to peer spending, we should observe a decline in their spending ratio, and this decline should be proportional to their distance from peers' spending. To the contrary, we would expect a change in the opposite direction if users were reacting to learning that their income was higher than their spending and they could spend more. In the left columns of Table 4, we find evidence consistent with a reaction to peer information: Users decrease their spending, and this decrease is proportional to their distance from their peers.

The second subsample are users who lie above their peers but below the US average in terms of spending. In this case, similar to the first test, if users were reacting to information about the average US consumer, their spending should increase, rather than decrease, after signup. In the right columns of Table 4, instead, we find evidence consistent with a reaction to peer information.

Overall, we conclude that, among the three pieces of information that users see in the same vivid homepage picture, they appear to react to information about peers and not to the other pieces of information.

Another difference in terms of information is that users who overspend relative to their peers see vivid messages (in red) about their spending on the app, whereas under-spenders see less vivid messages (in black). We address this point in section 6, in which we show that even though they all see the same vivid overspending messages, users for whom peer groups are more informative react substantially more than users for whom peer groups are less informative. The mere vividness of overspending messages thus triggers no reaction.

5 Causality? Quasi-Random Assignment to Peer Group Information

The baseline results do not rule out a few concerns about endogeneity in our setting. For instance, the endogenous timing of signup might be correlated with users' intentions to change their spending behavior, irrespective of the information they obtain at signup about their peers. In this case, users might not react because of exposure to peer information but because of pre-determined reasons. Another potential concern, which we have already discussed, is that mean reversion in spending drives our results. Although we proposed direct tests in the baseline setting to tackle the issue of mean reversion, one might still be concerned.

To address these challenges, ideally we would compare the reactions of users who have similar incomes and spend similar amounts before signup, but observe different information about peers at signup. Fortunately, our setting allows for this type of test. We exploit a feature of the design of peer groups that uses sharp income thresholds to assign users to peer groups at signup. In this way, we can use the quasi-random sharp assignment rule to instrument for the amount of peer spending to which otherwise similar users are compared.

Specifically, *Status* matches users with peer groups based on whether users fall above or below a set of annual income thresholds that are pre-specified and of which users are not aware when they sign up. The annual income thresholds are \$35K, \$50K, \$75K, \$100K, and \$150K. Because income is a continuous variable, small differences in income capture similar subscribers. And yet, users who report similar incomes but end up falling below or above an unknown threshold will be matched to different peer groups, and will face different information about peers' spending.

For example, consider a user who reports an annual income of \$99K and one who reports an annual income of \$101K. These two users would be quite similar with respect to their annual income, and we can test directly that they are similar along the set of observables we have, which include demographics and, crucially, the amount of spending in the months before signup. We can thus rule out by construction that a mechanical mean reversion of users' spending drives any results in this setting.

Whereas users around the thresholds are similar, the peer groups to which they are matched, and hence the information about peers' spending, can be substantially different. In our example, the user who reports an income of \$99K receives information about the average spending of peers whose income is between \$75K (just above the immediately lower threshold) and \$99K, whereas the similar user who reports an income of \$101K receives information about the average spending of peers whose income is between \$100K and \$149K (just below the immediately higher threshold).

The user just below the threshold faces a peer group that spends, on average, substantially less than the group of the users just above the threshold. The treatment effect of peers' spending information is thus systematically different for the two users, and, in particular, the user just below the threshold will likely appear as an overspender relative to her peers, whereas the user just above the threshold will likely appear as an underspender relative to her peers.

For this test, we need to restrict the sample to a group of users who are close enough to the thresholds so that they do not differ on observables. Moreover, we need to have a large enough mass 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 the thresholds (which includes the exact threshold value) is higher than the mass of users just below the threshold. In our identification sample, we thus include users with an income up to \$6K below the threshold and users with an income up to \$2K above the threshold.[11]

5.1 Balancing of Observables around Sharp Income Thresholds

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 thresholds differ on observables. The main difference we want to rule out is that spending before signup differs systematically across the thresholds. Detecting no differences would rule out the possibility of a mechanical effect due to mean reversion in spending.

We provide evidence against this concern in Figure 9. For each threshold, we plot the average monthly spending in the three months before signup (in thousands \$) for users around the threshold at intervals of \$500 of annual income. Around each average, we report intervals whose size is one third of a standard deviation above and below the average. We cannot reject the null that users above and below each threshold spent different amounts before signup. The highest threshold (\$150K) is the one for which average pre-spending is more varied across income levels, largely because the sample of users with these yearly incomes is smaller, and hence averages are based on a small number of observations. At the same time, even for this threshold, we do not detect systematic economic or statistical patterns.

We also move on to assess the balancing of other observables we use in the analysis. We report these balancing results in Table 5. We consider the observables that enter the set of controls in our baseline regressions, that is, the dummy for whether the user is a homeowner, the log of users' credit scores, the log of age, and the logs of users' asset and debt balances. For each variable, we report the results for regressing it on a dummy that equals 1 if the user is above the threshold, and 0 otherwise. Each panel of Table 5 refers to one of the sharp thresholds. Overall, we cannot detect economic or statistically significant patterns in the observables that vary systematically above and below the sharp thresholds. 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 either in terms of sign or in terms of economic and statistical significance across thresholds. Overall, the observables appear to be well balanced above and below each threshold. Although, of course, we cannot test directly whether unobservable characteristics are also balanced, the fact that important determinants of spending behavior–including spending by

[11]Later, we run a placebo test that compares users close to rounded income values for which the app does not assign different peer groups and verify that the reactions of these users do not differ.

itself—are well balanced reassures us about the assumption that the peer-group information assignment is quasi random around the sharp income thresholds.

5.2 IV Results

Armed with our identification sample, we use the quasi-random assignment of peer-group information to similar users to design an IV strategy. We use the dummy for whether a user is above the threshold as an instrument for the peer spending the user sees after signup: Users just above the threshold are assigned to a peer group whose income is, on average, substantially higher, and hence will observe higher peer spending, relative to users just below the threshold, who are assigned to a peer group whose income is substantially lower. Therefore, users below the thresholds are more likely to be classified as overspenders with respect to their peers, relative to users below the thresholds. We use this design because peer spending is a continuous variable contrary to the dummy for being overspenders, and hence this design avoids the forbidden regression problem.

We estimate a two-stage least-squares model. The first stage consists of the following specification:

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

where $Peer\ Spending_i$ is the peer spending user i sees at signup, and $Dummy\ Above\_i$ is a dummy variable for whether the users' incomes are below 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, relative to those users below the threshold.

In the second stage, we use 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 consumption in the three months around signup.

Even in the second stage, we predict 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 underspender, the more likely she is to cut her spending.

We first verify the relevance of our instrument by assessing the results from the first-stage regressions in Table 6, where the right columns control directly for individual-level observables. Consistent with the app's sharp assignment rule, users just above the threshold do observe a substantially higher amount of peer spending relative to other users—about three quarters of a standard deviation higher. The first-stage F statistics are 606 and 438 in the specifications with and without controls, respectively, which suggests our instrument is not weak.

Moving on to the second-stage estimates, irrespective of whether we control for observables, the spending ratio of users who see a higher level of peer spending and hence are more likely to be underspenders is higher, consistent with the baseline result in the full sample. The size of the second-stage coefficients of interest range between 7.5 and 11 percentage points.

5.3 Placebo IV Analysis

To corroborate our interpretation of the IV analysis and the exclusion restriction we need to assume for a causal interpretation, we propose a placebo test. We set alternative placebo income thresholds different from the true thresholds that *Status* uses to assign users to peer groups. To determine the placebo thresholds, 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 thresholds are set at round dollar values like the true thresholds. This criterion allows us to rule out that our IV results might be related to systematic differences of the behavior of income rounders and non-rounders. Indeed, rounders and non-rounders might differ in the extent to which they care about their personal finances, peers' spending, and other dimensions. By setting the placebo thresholds to round values, we ensure that, as in the actual IV, we have rounders in one group but not in the other.

Third, we ensure the thresholds allow us to obtain a large-enough placebo IV sample so that statistical power is not a problem for the test. For instance, if we set a placebo threshold at a very high dollar value, such as around $200K, we would have a very small amount of observations in the sample.

These three criteria lead us to select the following values for the placebo thresholds: $45K, $60K, $90K, $110K, and $140K. We construct the placebo IV sample exactly as we did in the IV analysis above, 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 threshold should not predict the dollar amount of peer spending, because users above and below the peer thresholds should be assigned to the same peer group. Moreover, in equation (3), we predict that $\hat{\gamma}$ "blows up;" that is, it becomes large in any direction and becomes statistically insignificant, because the first stage does not go through and the variation we use to instrument the endogenous variable is unrelated to the actual variation of such variables in the data.[12]

We report the results for the placebo IV analysis in Table A.2 of the Online Appendix. Indeed, the first stage does not go through, whether we control for observables or not, and consistently the second stage blows up. If anything, the estimated coefficient when controlling for observables is large and *negative* and with a t-statistic close to zero.

6 Informativeness of Peer Groups and Spending Behavior

As we discussed in section 2.3, our setting is not only unique in terms of separating the role of peer pressure from the information about peer spending, but also in terms of the large heterogeneity in the informativeness of peer groups across a broad set of dimensions. This heterogeneity enables us to further test the conjecture that information about peers changes beliefs about optimal spending and spending behavior. Indeed, we can design heterogeneity tests that assess whether users who are assigned to more informative peer groups, that is, peer groups that look more similar to themselves, do react more to the information about peer spending relative to other users.

On top of digging deeper into the information channel, this analysis provides yet another piece of evidence that mean reversion in spending cannot drive our results. If it did, no reason exists for why mean reversion should vary systematically across users based on whether their peer groups were more or less informative, and hence we should fail to detect any differences in users' reactions based on the informativeness of their peer groups.

Moreover, this analysis allows us to separate the role of information about peer spending from the role of nudges, messages, and alerts about overspending provided by the app. Nudges, alerts, and messages are the same for all users who are categorized as overspenders relative to their peers, irrespective of the extent to which information about peers is informative. If these features drove

[12] In fact, estimating a second stage when the first stage does not go through is meaningless. We only report the results for completeness and to verify our predictions, but clearly if the first stage does not go through an econometrician should stop and not attempt to estimate a meaningless second stage.

our results, instead of peer information, we should observe no heterogeneity in reactions based on the informativeness of peer information. Instead, the information channel predicts that users should react more if the peer-group information is more accurate, irrespective of any other piece of information users see in the app.

To perform this analysis, we consider the dimensions of heterogeneity in peer group informativeness discussed in section 2.3, and we repeat the IV analysis separately for users assigned to more or less informative peer groups.

6.1 Peer-Group Similarity

The first split exploits the fact that, to make peer-group averages well behaved and meaningful, *Status* imposes a minimal number of 5,000 individuals to construct the peer groups. Based on this rule, the app relaxes in some cases one or more categories in the construction of the peer group to guarantee the 5,000 threshold is met.

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 would easily find 5,000 observations in Manhattan with the characteristics of user A, it is likely to miss the same number 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 the location in which they live in. Intuitively, then, if users attach any information value to the peer signal, user A should believe that the signal she gets is more informative, compared to user B.

Based on this intuition, we split our IV sample into users whose characteristics are all matched to the closest possible ranges and users for which at least one characteristic is matched to a broader range, and hence for whom the signal should be less informative. We then estimate the two-stage least-squares specification in equations (2) and (3) separately across these two subsamples and directly report second-stage estimates in the top-left columns of Table 7. Consistent with our conjecture, users for whom the peer spending signal is more precise drive the baseline 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.

6.2 Peer-Group Size

Our second proxy for informativeness is the size of peer groups. The number of peers in the group that users are assigned to is vividly displayed to users, as shown 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. Even if users lacked statistical literacy, it should be clear to them that information based on a large group of individuals should be more reliable than information based on a small group of people.

In the top middle columns of Table 7, we split our sample between users assigned to peer groups with sizes above (High) and below (Low) the median. Consistent with a stronger reaction for more informative peer groups, we see the effect is about twice as large and statistically different from zero for users assigned to a larger peer group, whereas it is smaller in size and does not differ from zero statistically for other users.

6.3 Peer-Group Width

Third, we consider the fact that, again due to the need to guarantee meaningful-sized peer groups, users are assigned to peer groups by categories that have different ranges of values. For example, consider the ranges of values for income. Because the income distribution in the population is heavily skewed—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 dramatically wide group. For such a broad group, average peer spending might be irrelevant.

In the top-right columns of Table 7, we split our sample between mid-income users and high-income users, which include those earning more than \$100,000 a year. Mid-income users, for whom the peer group is more narrowly defined, display a large economic reaction, which is about four times larger than the baseline effect. The estimate is noisy, however—we can only reject the null of no reaction at the 10% level. By contrast, the estimated effect for high-income users, who are assigned to wider peer groups, is very close to zero economically and statistically.

6.4 Vividness of Comparison of Spending Relative to Peers

Figure 4 suggests yet another dimension along which information about the differences between users' and peers' spending is vivid. In each month, users observe their own cumulative spending based on transaction data (blue line) plotted against the average peer spending and the average overall US spending linearly projected throughout the month. If a user signs up for *Status* in the first few days of the month, she will only see a small blue area between her spending and her peers' spending, irrespective of her position relative to her peers. Instead, if the user signs up later in the calendar month, as is the case in Figure 4, the user will observe a very large blue area that captures the difference between her spending and her peers' spending. Based on this variation, we conjecture that users who sign up later in the month, irrespective of which month of the year, face a more vivid comparison and hence might react more.

In the bottom-left columns of Table 7, we confirm this conjecture, because we find that the reaction is about twice as large for users who sign up later in the month, and hence face a vivid comparison to peers, relative to the reaction of users who sign up earlier in the month.

6.5 Access to 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 before signing up for *Status*. 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 a fewer peers, that is, people who look similar to them under most characteristics, whom they observe on a regular basis, and from whom they could infer spending information before signup. 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.1 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, *Status* users from less dense towns should react more than urban users, which is what we document in the bottom middle columns of Table 7.[13]

This split could also further corroborate the notion that users react to learning information about peers instead of peer pressure, because peer pressure might be higher in smaller localities, in which households are likely to know each other well. In contrast, the anonymous apartment buildings and condominiums of large urban conglomerates might not promote social interactions. And, yet, inhabitants of such large buildings would directly see a large part of others' spending by observing daily package deliveries to a single location in the building. Instead, in more rural areas, keeping track of information about all spending, including non-conspicuous spending, for peers scattered in a broader geographic area is likely to be harder.

6.6 Number of Logins

The last dimension we consider is the number of times users are exposed to the signal about peers' spending, which we measure with the number of times a user logs into the app. This dimension is motivated by D'Acunto, Malmendier, Ospina, and Weber (2019), who find that, when forming beliefs, agents who are exposed to the same signal more frequently weigh it more than other agents.

Of course, the decision to log in is fully endogenous to several users' characteristics, and we do not argue that the number of logins represents an exogenous dimension for sorting users. Whether this exposure is driven by endogenous or exogenous motivations, though, is not relevant to our test: In both cases, users who log in more often are exposed to the signal about peers more frequently, which increases the vividness and ability to recall the signal. Users who log in less often, instead, will see the signal fewer times. We thus conjecture that users who are exposed to the peer signal more often might change their spending by a larger amount.

Consistent with this intuition, in the bottom-right columns of Table 7, we find the size of the estimated effect is about twice as large for users who log in at least twice a month (Many) relative to other users (Few).

The fact that we detect stronger results for active users further rules out that any of our findings might be due to unobserved spending by users who stop using the app over time and for whom the app stops tracking spending accurately.

[13]Note the urban and rural subsamples do not sum up to the full regression-discontinuity sample. This discrepancy arises because *Status* 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 just 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.

Overall, our heterogeneity analysis reveals that, not only do users react to peer information in a setting that allows for quasi-random assignment of individuals to peer groups, but also that the informativeness of these peer groups is what determines whether and to what extent users react to peer information.

7 External Validity and Demographic Heterogeneity: Evidence from a Randomized Control Trial

Due to the setting, our results so far cannot speak to two important issues. The first issue 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 not to the general population, because app users in our setting cared enough about obtaining information about their peers that they decided to sign up. Other parts of the population might not be interested in peer information, and providing them with peer information might not change their spending behavior. Moreover, similar to any other income aggregator studied in the literature, such as *Mint.com* or *Meniga*,[14] the app is only meaningful to users if they provide sensitive information such as their linked accounts and Social Security numbers to access credit reports. 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 we do not observe some demographics that we know from earlier research are important to determine spending behavior (D'Acunto, Hoang, and Weber, 2016). 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 effect might be higher or lower across different demographics; that is, unobserved characteristics might interact with peer information to determine heterogeneous effects.

[14] 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).

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 propose a different randomized control trial. We recruited a representative US population through an online platform, Amazon Mechanichal Turk (mTurk), which is increasingly used for experimental research in the US and abroad.[15] We invited respondents to answer a survey, without any specific 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 because they were interested in their own finances, which are the main concerns of our baseline analysis in terms of external validity.

We report the survey questions in the Online Appendix. The main aim of the survey was to elicit respondents' marginal propensity to consume based on the work of Attanasio and Weber (1995) and subsequent literature that is reviewed in Jappelli and Pistaferri (2010), both before and after receiving truthful information about the marginal propensity to consume of representative individuals with similar income levels. Eliciting MPC is yet another way show that mean reversion has no role in our results.

The survey consisted of three parts. In the first part, we elicited respondents' age groups 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 the marginal propensity to consumer (MPC) of their income peers. We then elicited respondents' MPC following Jappelli and Pistaferri (2014) both before and after observing the information about their peers. Moreover, because the survey was run during the COVID-19 pandemic, we asked respondents to provide their MPC based on their conditions at the time of the survey as well as based on their conditions before the COVID-19 pandemic started.

The treatment condition thus depends on the income group of the subject as well as her stated MPC before observing the one of her peers. To avoid any demand effects—the possibility that respondents guessed the aim of our trial and conformed to what they thought was the hypothesis we were testing—we told respondents explicitly on two occasions that we had no hypotheses about whether providing

[15] For instance, see Kuziemko, Norton, Saez, and Stantcheva (2015), DellaVigna and Pope (2017), Bazley, Cronqvist, and Mormann (2017), Lian, Ma, and Wang (2019), D'Acunto, Hoang, Paloviita, and Weber (2020), D'Acunto (2018), and D'Acunto (2019).

information should have changed their MPC in any respect, let alone the direction, and that all answers would be of extreme interest to us as long as they reflected the respondents' true opinions. In this way, we made clear to the subjects that we found a reaction or a non-reaction to peer information equally plausible and interesting, and hence respondents would not have answered based on demand effects. Note that based on the first MPC-elicitation answer, subjects would end up seeing they were reporting a higher or a lower MPC than their peers. Whether subjects ended up above or below the peers was driven by their first answer. We then elicited again their MPC to an unexpected reimbursement equal to one month of their income, both in normal times and during the COVID-19 pandemic. Our survey does not include an additional control group to whom no information is provided. This group would barely help the analysis, because then we would not know if the difference-in-differences coefficient we estimated was due to exposure to peer information or exposure to information about spending in general.

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

The aim of this survey was to test whether a representative sample of the general population would react to information about peers when thinking about their spending behavior, after being randomly provided with such information.

Armed with our survey, we can assess the two questions we proposed above. First, does a non-selected draw of the US population respond to the provision of peer information about spending in a similar manner as the users of our app? Indeed, we find that our random sample of survey respondents reacts strikingly similarly to the app users in our baseline analysis. In Figure 10, 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 8, we report the average change in reported MPC before and after seeing information about the MPC of income peers, for both MPCs in normal times and MPCs at the time the survey was fielded, that is, during the COVID-19 crisis. In both cases, respondents converge to the behavior of income peers, and the effect appears stronger for respondents whose MPC was higher than their income peers', as well as for respondents who were further away from their income peers.

The second question is: Is the spending response to peer information only driven by specific demographics that we cannot identify in *Status*? If so, we could not exclude that the app-based results might be driven by a peculiar (unobserved) demographic composition of users. Reassuringly, in Table 9, we find 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, or individuals who are more or less risk averse. This result reduces concerns related to the fact that certain demographics are unobserved in Status.[16]

7.1 Overall Effect in a Non-selected Population

Finally, we can use our RCT to assess the overall average effect of the intervention—providing consumers with information about their income peers' MPC—on the population. Because of the potential selection of consumers who are more sensitive than others to their peers' spending in the baseline analysis, we cannot extrapolate those results to make claims about the general population, whereas this issue does not arise in our RCT.

The drawback from inferring population-level effects from our RCT is that the MPC is self-reported and not backed up by actual spending data, and hence we are assessing spending plans instead of actual amounts spent. Note, though, that the elicitation of MPC we propose is standard practice in economic research (e.g., see Parker and Souleles (2019); Jappelli and Pistaferri (2014); D'Acunto, Hoang, and Weber (2016)), and we follow this earlier work closely to estimate MPCs in our setting.

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 declines 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 sobering effect, on average, in terms of spending plans. A sobering 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. As we discussed in the opening of the introduction, the lack of wealth accumulation due to the high spending of many US households has attracted interest

[16]Note the baseline results from our RCT suggest potential heterogeneity in the extent of the reaction across demographics, which would be an interesting aspect to investigate in future research.

from policymakers. Our results, which hold true not only in a part of the population that selected into obtaining peer information, but also in a representative, non-selected sample, suggest that the provision of crowdsourced information about peer spending through FinTech apps could act as a form of robo-advising for spending and reduce the MPC of high-spending consumers. To the best of our knowledge, the design and effects of robo-advising tools for spending is very limited, and our results pave the way for additional research in this area (D'Acunto and Rossi (2020); Gargano and Rossi (2020)).

Interestingly, respondents' MPC is *lower* in times of crisis than in normal times, which seems at odds with the standard life-cycle consumption framework. This result could be consistent with precautionary savings motives and with recent research in behavioral macroeconomics, which shows most consumers in a representative population have a stagflationary macroeconomic model in mind and believe agents should save more and spend less during economic crises than during normal times (e.g., see Andre, Pizzinelli, Roth, and Wohlfart (2019) and D'Acunto, Hoang, Paloviita, and Weber (2020)). This result begs for additional research using actual transaction-level spending information to inform theoretical advances in models of consumption and saving choices.

8 Conclusions

We study the effects of providing individuals with crowdsourced information about similar but unknown peers' spending through a FinTech app. Our results are as follows: First, users who spend more than anonymous peers reduce their spending, and users who spend less increase their spending, which suggests users find the signals they receive informative. Moreover, users' reaction is asymmetric—overspenders cut spending more than underspenders increase it. The reaction is larger when the signal about peers is more informative and for users who are further away from their peers in terms of pre-signup spending. We document that mean reversion does not drive our findings, and support a causal interpretation of our results exploiting sharp income thresholds in the definition of peer groups.

We also implement a randomized control trial to support the external validity of our findings. Using a population that did not select into receiving peer information, we find that the average marginal propensity to consume in the economy drops after the provision of peer information, due to consumers' asymmetric reaction. This finding is a first step toward understanding the potential policy applications of our results and especially the design of robo-advising tools targeting the consumption/saving choice (D'Acunto and Rossi (2020)), which is perhaps the most important household finance decision non-expert consumers have to make. Future research should investigates these avenues.

An interesting question for future research is also whether providing information about peer spending has permanent or transitory effects, which would have opposite implications in terms of promoting policies that disseminate information about others' spending behavior.

Moreover, our setting does not allow us to test whether reacting to peers' spending is optimal for users. Whether reacting to peer information is optimal and whether peer spending contains any relevant information for consumers is a completely open question and should be addressed by future research.

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, 4205–30.

Agarwal, S., W. Qian, and X. Zou, 2017, "Thy neighbor's misfortune: Peer effect on consumption," *Available at SSRN 2780764*.

Andre, P., C. Pizzinelli, C. Roth, and J. Wohlfart, 2019, "Subjective models of the macroeconomy: Evidence from experts and a representative sample," .

Argyle, B., T. Nadauld, and C. Palmer, 2020, "Monthly Payment Targeting and the Demand for Maturity," *Review of Financial Studies*.

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, 1121–1157.

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, 2224–2276.

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

Bazley, W., H. Cronqvist, and M. Mormann, 2017, "In the Red: The Effects of Color on Investment Behavior," *Working Paper*.

Bernheim, B. D., J. Skinner, and S. Weinberg, 2001, "What accounts for the variation in retirement wealth among US households?," *American Economic Review*, 91, 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, 1161–1201.

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, 1273–1301.

Chalmers, J., W. T. Johnson, and J. Reuter, 2014, "The effect of pension design on employer costs and employee retirement choices: Evidence from Oregon," *Journal of Public Economics*, 116, 17–34.

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

Chen, H., P. De, Y. J. Hu, and B.-H. Hwang, 2014, "Wisdom of crowds: The value of stock opinions transmitted through social media," *The Review of Financial Studies*, 27, 1367–1403.

Coibion, O., D. Georgarakos, Y. Gorodnichenko, and M. Weber, 2020, "Forward guidance and household expectations," Discussion paper, National Bureau of Economic Research.

Coibion, O., Y. Gorodnichenko, and M. Weber, 2019, "Monetary policy communications and their effects on household inflation expectations," Discussion paper, National Bureau of Economic Research.

Da, Z., and X. Huang, 2020, "Harnessing the wisdom of crowds," *Management Science*, 66, 1847–1867.

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., 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," Discussion paper, National Bureau of Economic Research.

D'Acunto, F., U. Malmendier, J. Ospina, and M. Weber, 2019, "Exposure to daily price changes and inflation expectations," Discussion paper, National Bureau of Economic Research.

D'Acunto, F., U. Malmendier, and M. Weber, 2020, "Gender Roles and the Gender Expectations Gap," Discussion paper, National Bureau of Economic Research.

D'Acunto, F., N. Prabhala, and A. G. Rossi, forthcoming, "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," Discussion paper, National Bureau of Economic Research.

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

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*, pp. 19–06.

DellaVigna, S., and D. Pope, 2017, "Predicting Experimental Results: Who Knows What?," *Journal of Political Economy.*

Duflo, E., and E. Saez, 2003, "The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment," *The Quarterly Journal of Economics*, 118, 815–842.

Gargano, A., and A. G. Rossi, 2018, "Does it pay to pay attention?," *The Review of Financial Studies*, 31, 4595–4649.

Gargano, A., and A. G. Rossi, 2020, "There's an App for That: Goal-Setting and Saving in the FinTech Era," *Available at SSRN.*

Gargano, A., A. G. Rossi, and R. Wermers, 2017, "The freedom of information act and the race toward information acquisition," *The Review of Financial Studies*, 30, 2179–2228.

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, 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," Discussion paper, National Bureau of Economic Research.

Hau, H., Y. Huang, H. Shan, and Z. Sheng, 2019, "How FinTech enters China's credit market," in *AEA Papers and Proceedings*, , vol. 109, pp. 60–64.

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

Jappelli, T., and L. Pistaferri, 2014, "Fiscal policy and MPC heterogeneity," *American Economic Journal: Macroeconomics*, 6, 107–36.

Kalda, A., forthcoming, "Peer Financial Distress and Individual Leverage," *The Review of Financial Studies*.

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

Keys, B. J., and J. Wang, 2019, "Minimum payments and debt paydown in consumer credit cards," *Journal of Financial Economics*, 131, 528–548.

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, 1478–1508.

Leary, J., and J. Wang, 2016, "Liquidity Constraints and Budgeting Mistakes: Evidence from Social Security Recipients," *Working Paper*.

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, 2107–2148.

Lieber, E., and W. Skimmyhorn, 2018, "Peer effects in financial decision-making," *Journal of Public Economics*, 163, 37–59.

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, 205–224.

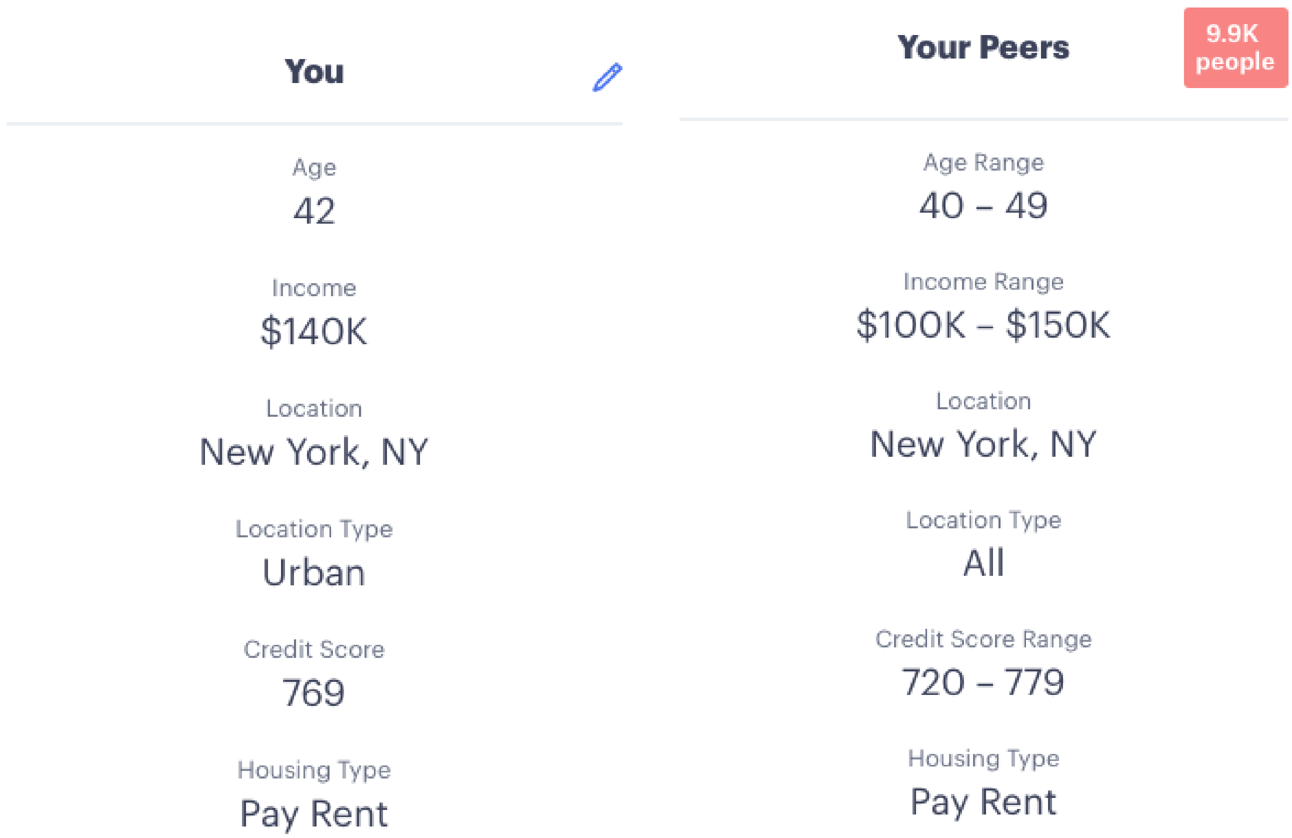
Maturana, G., and J. Nickerson, 2019, "Teachers Teaching Teachers: The Role of Workplace Peer Effects in Financial Decisions," *The Review of Financial Studies*, 32, 3920–3957.

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

Ouimet, P., and G. Tate, 2020, "Learning from coworkers: Peer effects on individual investment decisions," *The Journal of Finance*, 75, 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, 273–90.

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
Peer Group Assignment for a Fictitious Account

This figure shows the graphics 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. Only after being assigned to a peer group, the users can observe the characteristics of the peer group to which they are assigned as in the graphics.

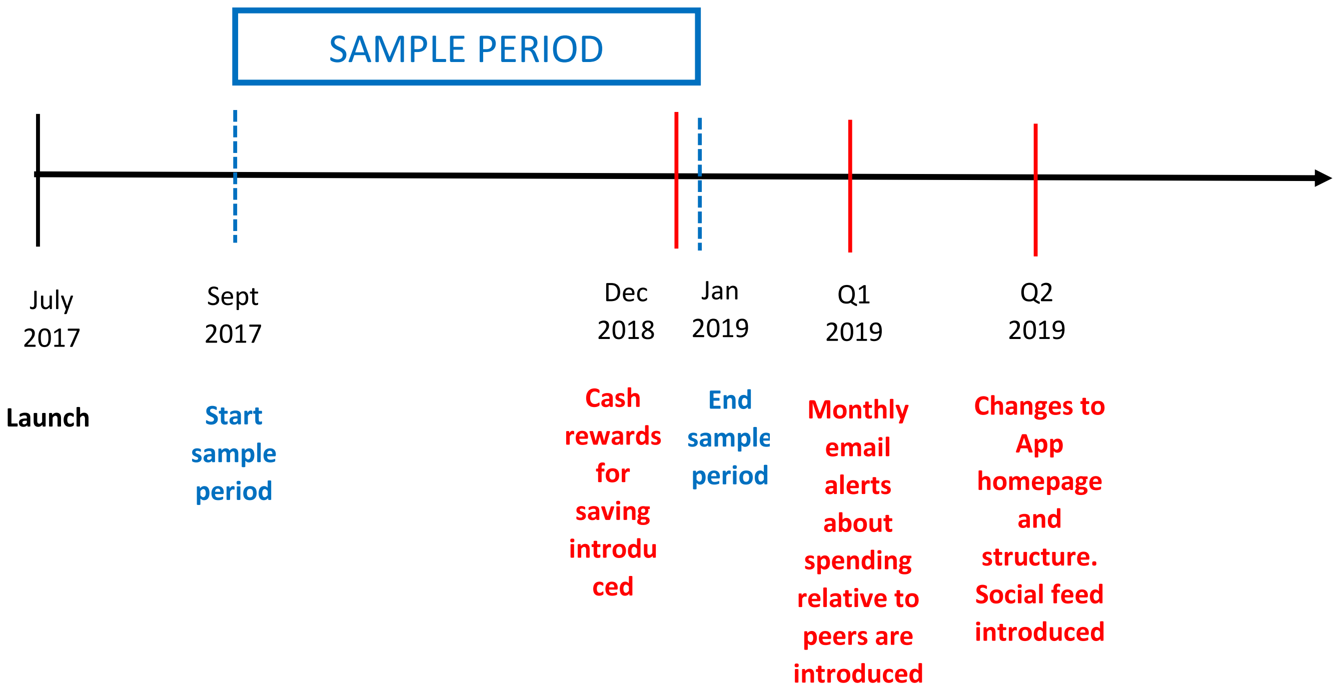


Figure 3
Evolution of the Status App

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. In particular, many of the 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.

Figure 4
Top Part of *Status Money* Home Page During Our Sample Period

This figure shows the top part of the homepage of *Status Money*, i.e. the main image users see when signing in to the app during our sample period. Our analysis is based on the information and graphics in this figure. Note that some of this figure's features, including its position and vividness in the app, have changed after our sample period, as we show in the Online Appendix.

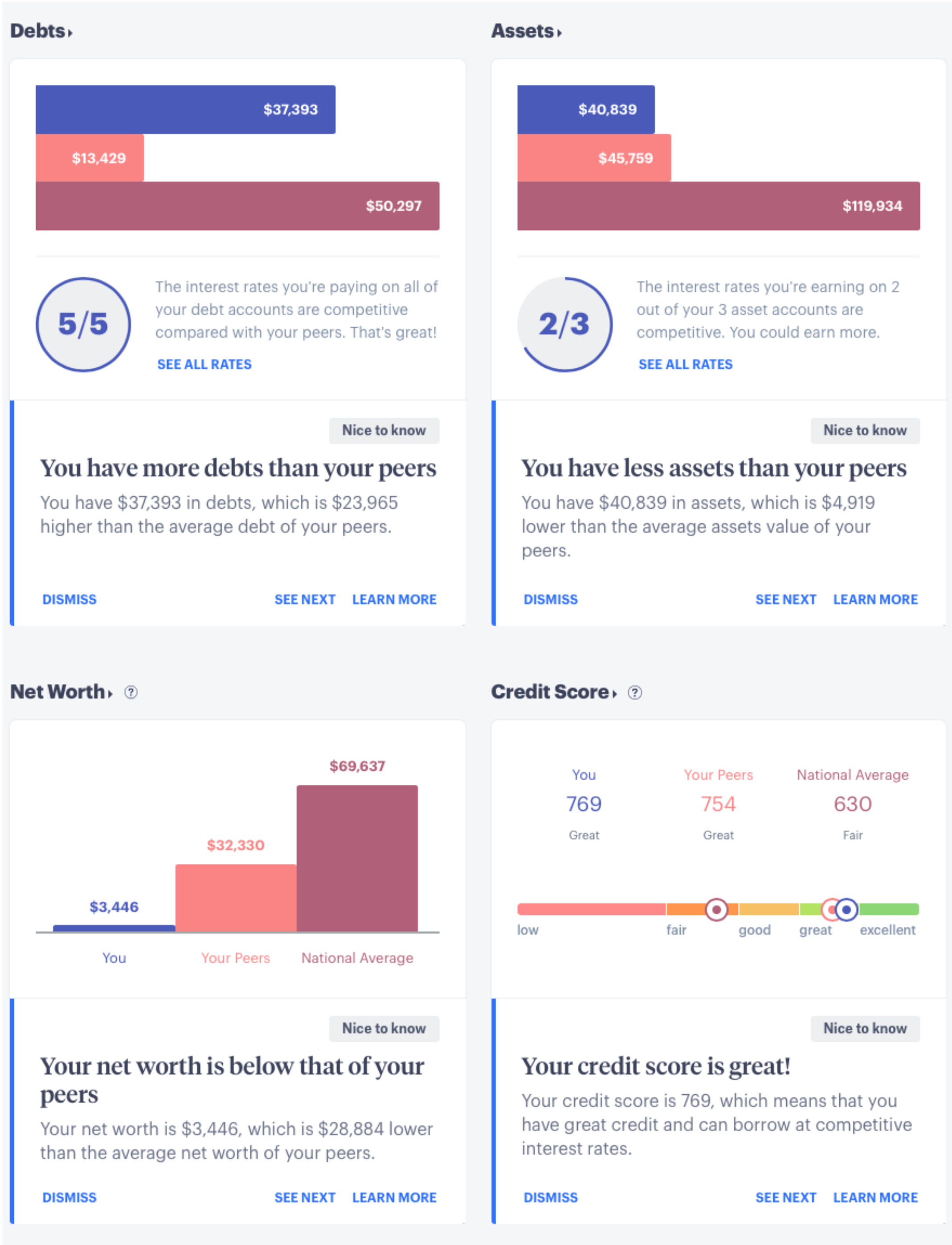


Figure 5
Bottom Part of *Status Money* Home Page During Our Sample Period

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.

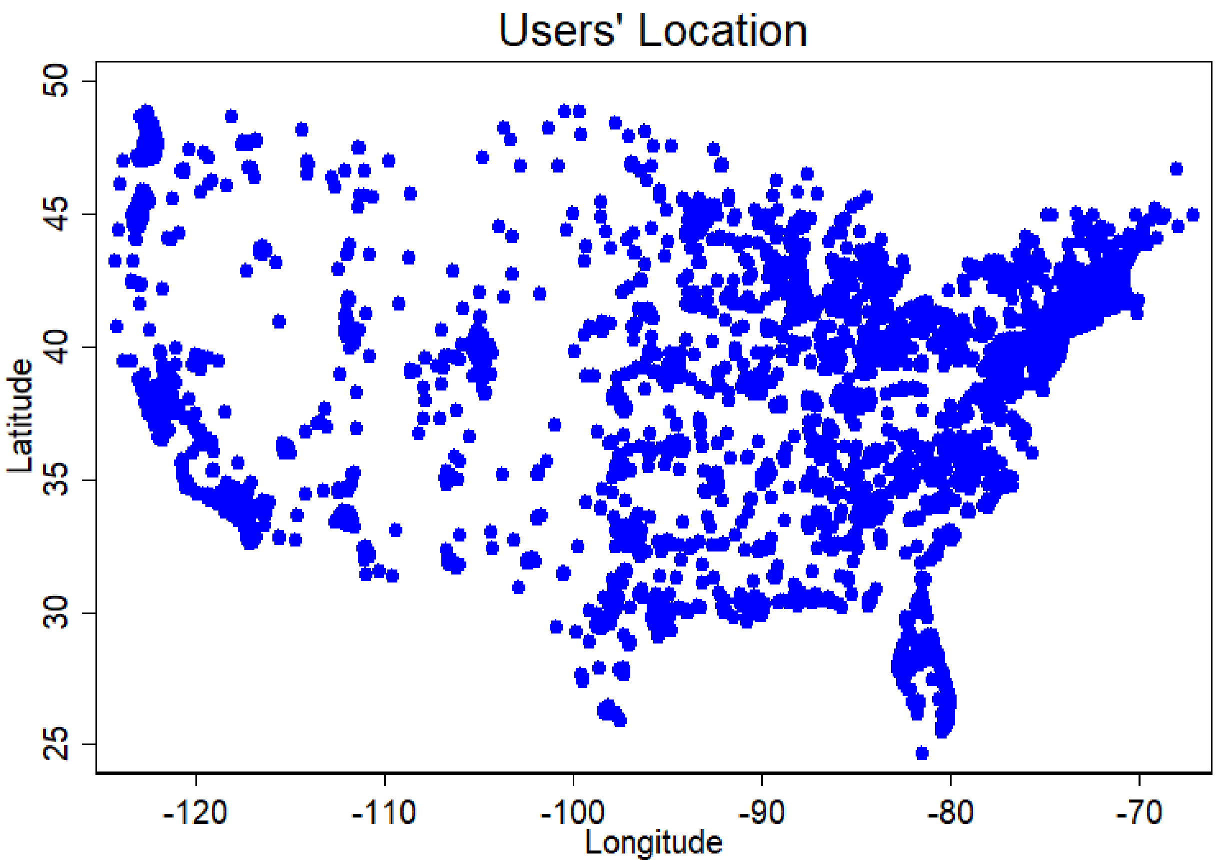


Figure 6
Geographic Location of Users

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.

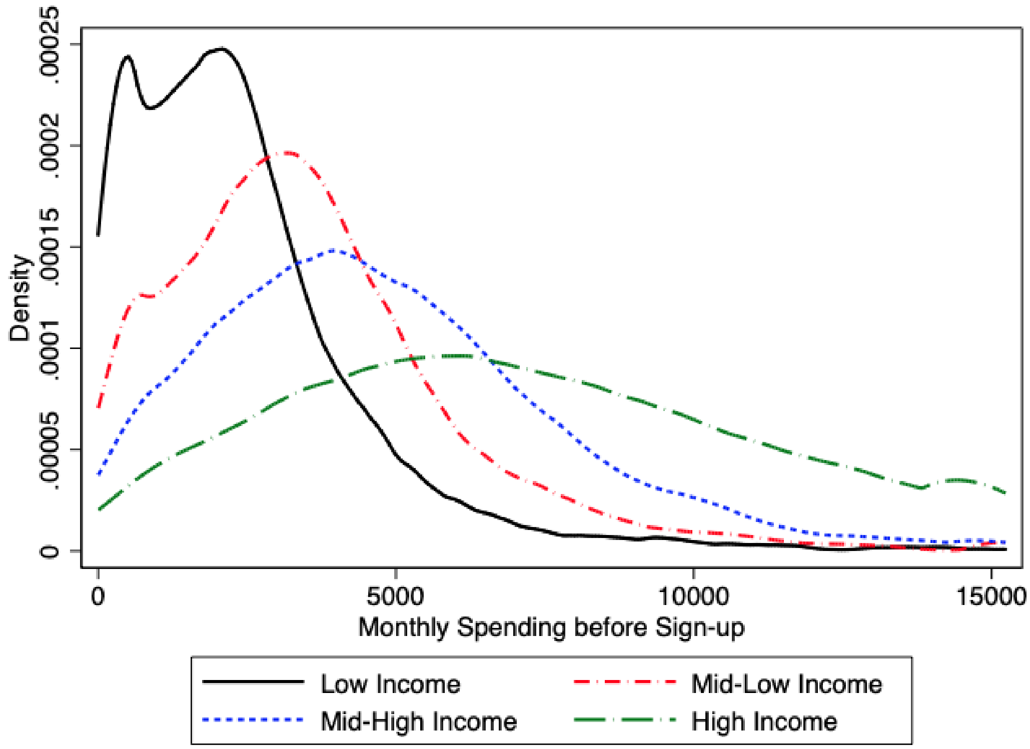
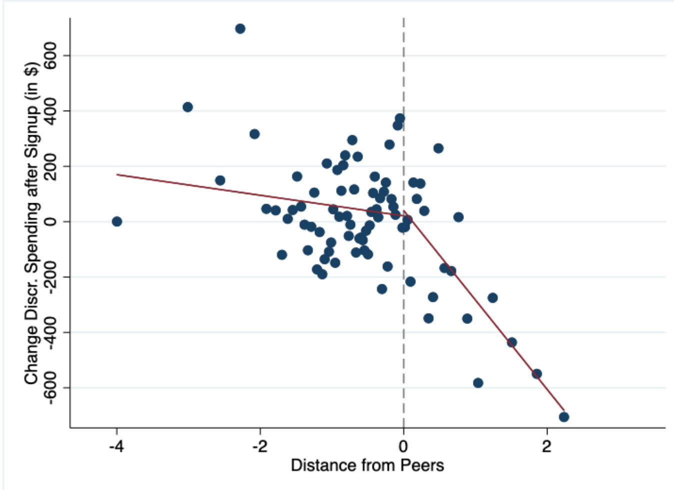
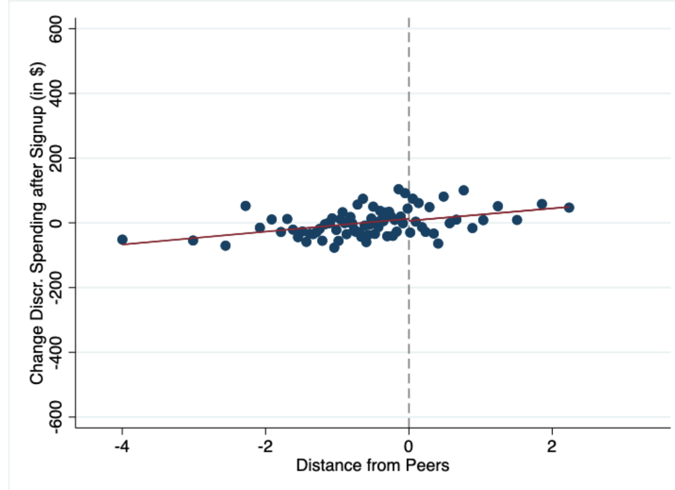


Figure 7
Distribution of Monthly Spending by Income Quartiles

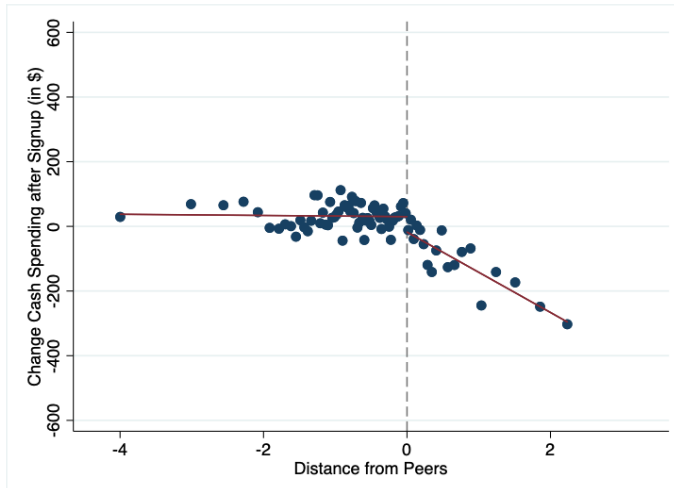
This figure plots the density of the dollar value of monthly spending across 4 quartiles of our sample based on income levels.



(a) Discretionary Spending



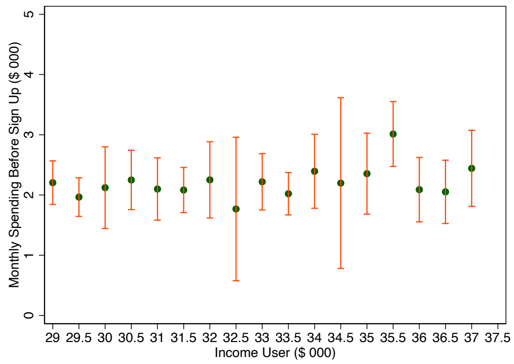
(b) Non-Discretionary Spending



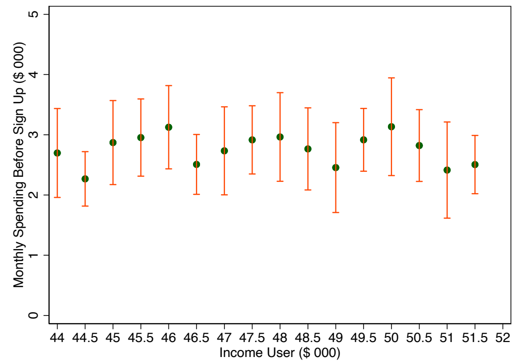
(c) Checking Account Withdrawals

Figure 8
Quality of Expense Categorization—Discretionary vs. Non-discretionary Spending and Cash Withdrawals

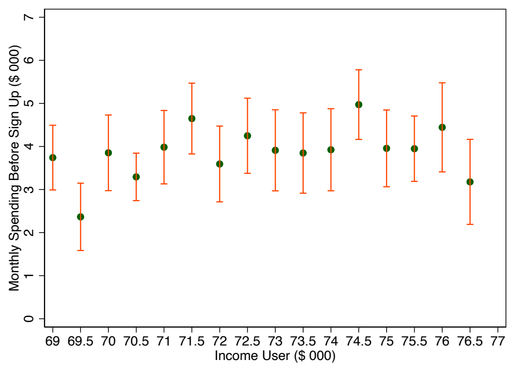
This figure shows binned scatterplots of changes in discretionary, non-discretionary consumption and cash withdrawals after signing up for *Status* and differences in consumption between individuals and their peer group. In all subfigures, the x-axis measures the difference in consumption with respect 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.



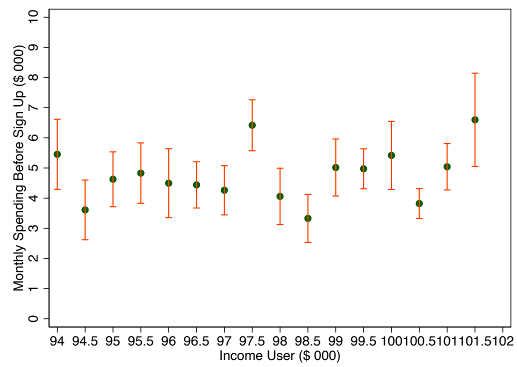
(a) Income Threshold $35,000



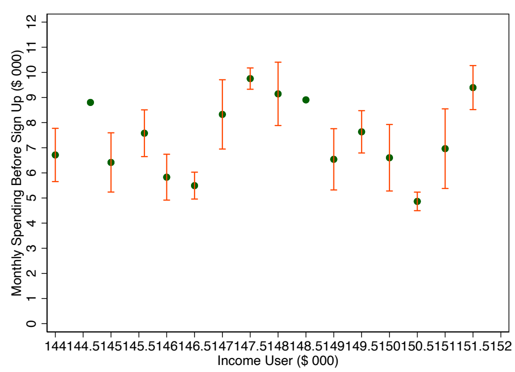
(b) Income Threshold $50,000



(c) Income Threshold $75,000



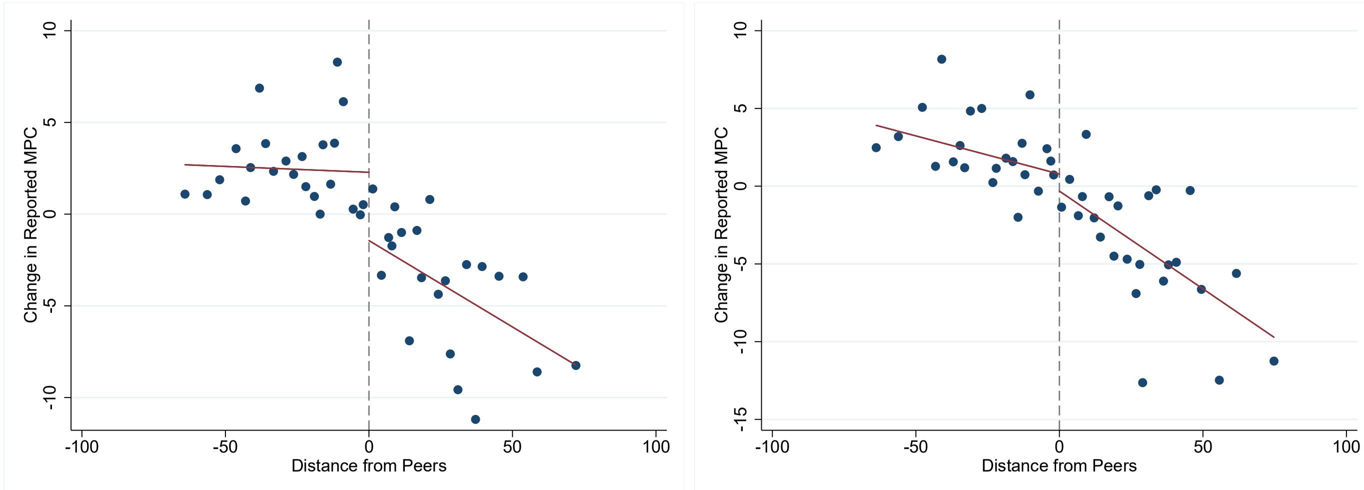
(d) Income Threshold $100,000



(e) Income Threshold $150,000

Figure 9
Average Monthly Spending in Three Months before Sign-up around Income Thresholds

This figure reports the average monthly spending in the three months before signup to *Status* by income around the different income thresholds we use in our identification sample. Bars around each point estimates indicate a third of a standard deviation in either direction.



(a) Current Times (COVID-19 Crisis)

(b) Normal Times

Figure 10

Change in Reported MPC After Peer Information: Randomized Control Trial

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. Baseline Sample		
	Observations	Mean	St. Dev.
Age	21,133	32.01	7.85
Credit Score	19,450	734.36	75.93
Home Ownership	21,133	0.39	0.49
Annual Income ($)	21,047	91,634	62,324
Monthly Spending (Total) ($)	21,133	4,860	4,026

	Panel B. Identification Sample		
	Observations	Mean	St. Dev.
Age	5,629	31.37	7.57
Credit Score	5,236	730.06	75.99
Home Ownership	5,629	0.36	0.48
Annual Income ($) income	5,629	71,916	33,399
Monthly Spending (Total) ($)	5,629	2,806	2,520

This table reports summary statistics of the main variables used in the paper. For each variable, we report the number of observations, the average, and the standard deviation. Panel A reports results for the baseline sample. Panel B reports results for the identification sample, see Section 5 for details.

Table 2. Peer Spending Information and Change in Spending

	Above	Below	Distance Above	Distance Below	Distance Above	Distance Below	Distance Above	Distance Below
Average Change	-0.233***	0.074***						
	(-42.00)	(8.34)						
Distance Peers			-0.103***	-0.086***	-0.104***	-0.096***	-0.039***	-0.083***
			(-11.31)	(-7.03)	(-10.47)	(-7.43)	(-3.54)	(-6.61)
Spend Before							-0.096***	-0.335***
							(-13.11)	(-28.29)
Homeownership					0.002	0.058***	0.037***	0.0823***
					(0.17)	(2.63)	(2.76)	(3.90)
log of Credit Score					-0.040	-0.301***	-0.036	-0.349***
					(-1.46)	(-3.93)	(-1.35)	(-4.73)
log of Age					0.014	-0.152***	0.081***	0.122**
					(0.46)	(-3.02)	(2.62)	(2.48)
log of Asset Balance					-0.001	-0.011**	0.009***	0.019***
					(-0.40)	(-2.40)	(3.05)	(4.25)
log of Debt Balance					0.002	-0.011***	0.007**	0.012***
					(0.55)	(-2.58)	(2.36)	(2.77)
Constant			-0.166***	-0.002	0.048	2.543***	-0.096**	1.511***
			(-20.14)	(-0.11)	(0.23)	(4.88)	(-1.62)	(2.99)
Observations	5,012	15,667	5,012	15,667	4,179	10,688	4,179	10,688

This table reports results for the sensitivity of the ratio of spending post signup to spending pre signup to peer consumption, spending pre sign-up as well as controls. Spending ratios are computed using consumption three months before and after signup. Regression estimates are computed for all users in columns 1 and 2, users with below-peer spending in columns 3, 5, and 7 and and for users with above-peer spending in columns 4, 6, and 8.

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

	2 Months Before		3 Months Before		Avg. Quarter Before	
	Above	Below	Above	Below	Above	Below
Distance Peers	-0.020**	-0.088***	-0.021***	-0.097***	-0.048***	-0.104***
	(-2.53)	(-8.67)	(-2.79)	(-9.78)	(-4.20)	(-9.48)
Spend Before	-0.099***	-0.318***	-0.099***	-0.320***	-0.088***	-0.321***
	(-13.40)	(-26.44)	(-13.72)	(-26.76)	(-10.85)	(-26.88)
Homeownership	0.037***	0.087***	0.037***	0.088***	0.037***	0.091***
	(2.78)	(4.13)	(2.79)	(4.21)	(2.79)	(4.32)
log of Credit Score	-0.034	-0.339***	-0.033	-0.340***	-0.035	-0.337***
	(-1.28)	(-4.60)	(-1.23)	(-4.62)	(-1.31)	(-4.57)
log of Age	0.085***	0.107**	0.083***	0.103**	0.080***	0.099**
	(2.77)	(2.19)	(2.71)	(2.09)	(2.62)	(2.00)
log of Asset Balance	0.008***	0.018***	0.008***	0.018***	0.008**	0.017***
	(2.73)	(4.15)	(2.74)	(4.10)	(2.48)	(3.92)
log of Debt Balance	0.007**	0.011***	0.007**	0.011***	0.007**	0.011***
	(2.57)	(2.68)	(2.57)	(2.66)	(2.42)	(2.62)
Constant	-0.369*	1.509***	-0.372*	1.523***	-0.333*	1.510***
	(-1.83)	(3.00)	(-1.85)	(3.03)	(-1.65)	(3.00)
Observations	4,179	10,688	4,179	10,688	4,179	10,688

This table reports results for the sensitivity of the ratio of spending post signup to spending pre signup to peer consumption, spending pre sign-up as well as controls. Spending ratios are computed using consumption three months before and after signup.

Table 4. Information About Peers or other Information?
Own Income and Average US Spending

	Below Own Income		Below Average US Spending	
	Above Peers	Distance Above Peers	Above Peers	Distance Above Peers
Average Change	-0.202***		-0.068***	
	(-22.72)		(-2.65)	
Distance Peers		-0.122***		-0.423**
		(-4.74)		(-2.23)
Constant		-0.152***		-0.001
		(-11.07)		(-0.04)
Observations	1,885	1,885	530	530

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.

Table 5. Balancing of Observables around Exogenous Income Thresholds

	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)	(0.14)	(-0.03)	(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 individuals above an income threshold. We only include users that earn up to \$6K less than the threshold (above dummy 0) and users up to \$2K more the threshold (above dummy 1). The income thresholds are \$35K in Panel A, \$50K in Panel B, \$75K in Panel C, \$100K in Panel D, and \$150K in Panel E.

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

	First Stage	Second Stage	First Stage	Second Stage
Above Dummy	0.743***		0.738***	
	(24.62)		(20.92)	
Peer Spending		0.111***		0.074**
		(3.08)		(2.13)
Spending Before	0.344***	-0.305***	0.388***	-0.296***
	(23.33)	(-15.63)	(20.70)	(-14.41)
Homeownership			-0.276***	0.111***
			(-6.93)	(3.62)
log of Credit Score			0.078	-0.166***
			(0.85)	(-2.49)
log of Age			0.002	0.095
			(0.02)	(1.39)
log of Asset Balance			0.045***	0.012*
			(5.39)	(1.93)
log of Debt Balance			0.022***	0.008
			(2.66)	(1.30)
Observations	5,629	5,629	4,046	4,046
IV F-stat	606.1		437.6	

This table reports results for a two-stage-least-squares identification strategy that compares users just below and users at or above each of 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 \$6K below the threshold as well as those whose income is at most \$2K above the threshold. We then estimate the following first-stage specification:

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

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. In the second stage, we use 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.

Table 7. IV Analysis: Effects by Informativeness of Peer Groups

	Peer Group Similarity		Number of Peers		Peer Group Width	
	High	Low	High	Low	Narrow	Wide
Peer Spending	0.189***	0.054	0.137***	0.076	0.434*	-0.013
	(3.10)	(1.24)	(2.84)	(1.38)	(1.78)	(-0.33)
Spending Before	-0.325***	-0.293***	-0.316***	-0.291***	-0.441***	-0.143***
	(-9.62)	(-12.42)	(-11.95)	(-10.32)	(-9.69)	(-10.06)
Observations	2,129	3,500	3,377	2,252	4,512	1,117
	Vivid Comparison		Pre Information Density		Log-ins (# Signals)	
	No	Yes	Rural/Low	Urban/High	Many	Few
Peer Spending	0.077	0.152***	0.421***	0.070*	0.138**	0.073
	(1.48)	(3.11)	(2.59)	(1.92)	(2.87)	(1.33)
Spending Before	-0.297***	-0.316***	-0.446***	-0.239***	-0.313***	-0.297***
	(-10.60)	(-12.03)	(-7.09)	(-11.31)	(-11.99)	(-10.16)
Observations	3,027	2,602	1,013	2,600	3,260	2,369

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 \$6K below the threshold as well as those whose income is at most \$2K above the threshold. We then estimate the following first-stage specification:

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

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. In the second stage, we use 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.

Table 8. 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 changes in the reported marginal propensity to consumer (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.

Table 9. 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 regressing the changes in the reported marginal propensity to consume (MPC) by respondents to our randomized control trial before and after observing information about income peers' MPC. Each panel splits the sample by demographics that are unobserved in the *Status* sample.

Online Appendix:

Crowdsourcing Financial Information to Change Spending Behavior

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

Not for Publication

A.1 Additional Figures

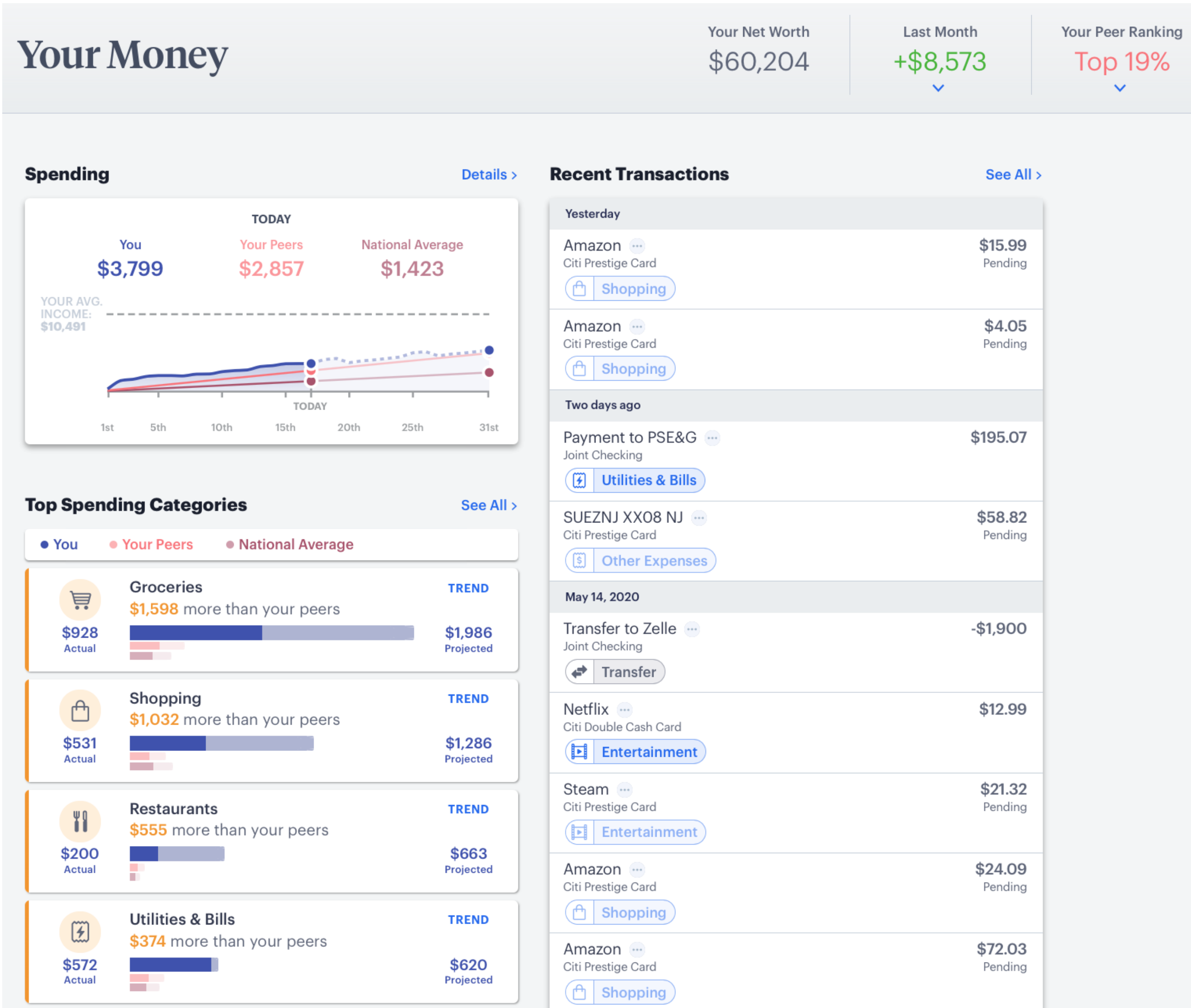


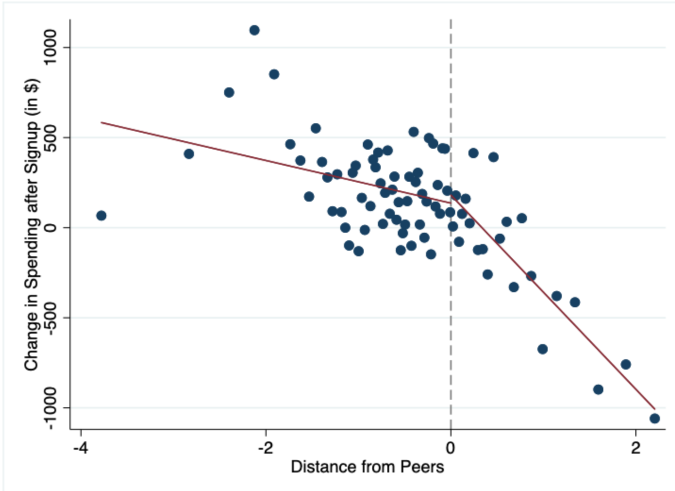
Figure A.1
Top Part of *Status Money* Home Page After our Sample Period

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.

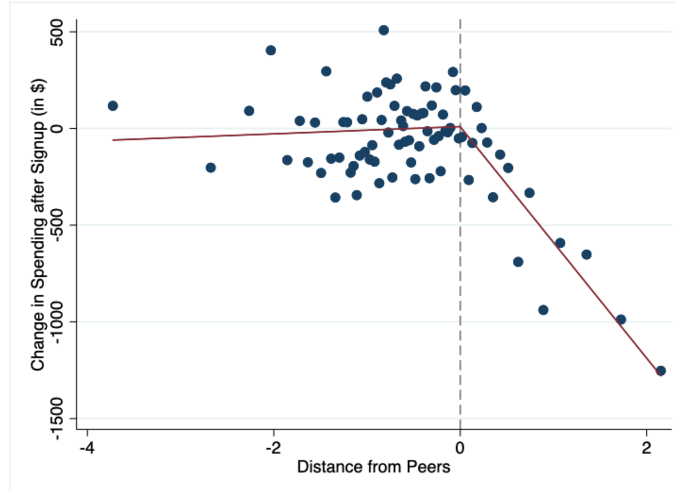
Moe's Southwest Grill	Citi Prestige Card	Restaurants
Amazon Marketplace	Citi Prestige Card	Shopping
Amazon Marketplace	Citi Prestige Card	Shopping
Essen	Savor	Restaurants
Payment to Honda	Joint Checking	Loan Payments
Amazon Marketplace	Citi Prestige Card	Shopping
Netflix	Citi Prestige Card	Entertainment
Essen	Citi Prestige Card	Restaurants
SARAY CUISINE CLIFFSIDE PAR NJ	Citi Prestige Card	Restaurants
Acme Fresh Market	Citi Prestige Card	Groceries
SQ *BOURKE STREET BAKE	Savor	Restaurants
Chipotle Mexican Grill	Savor	Restaurants

Figure A.2
Example of Transactions Categorized by Status Money

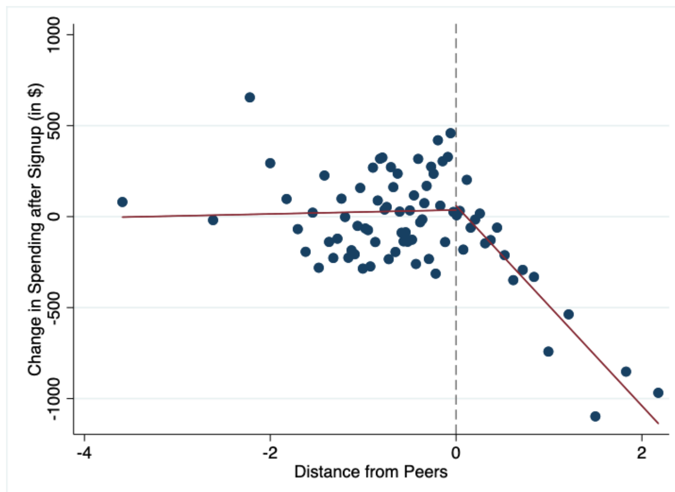
This figure shows an example of transactions categorized by *Status Money*.



(a) At least 2 Accounts Linked



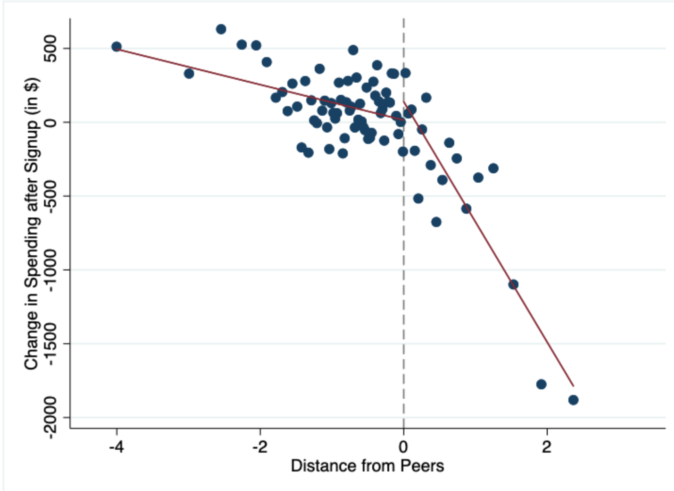
(b) Users Younger than 35



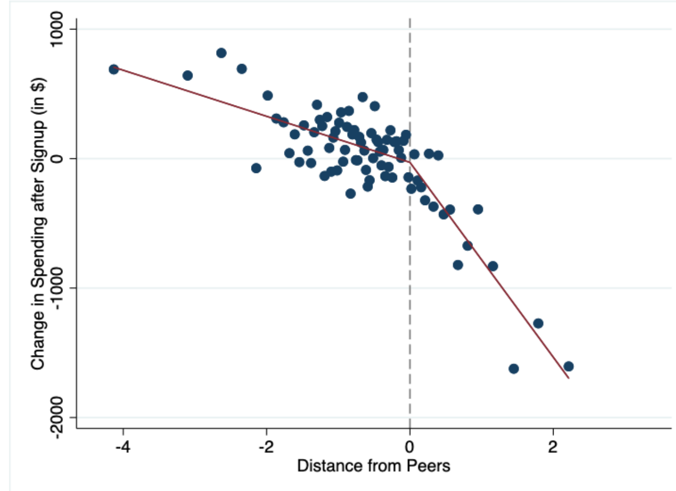
(c) Excluding Over $200K Income

Figure A.3
Raw-Data Change in Spending by Distance from Peers across Sub-populations

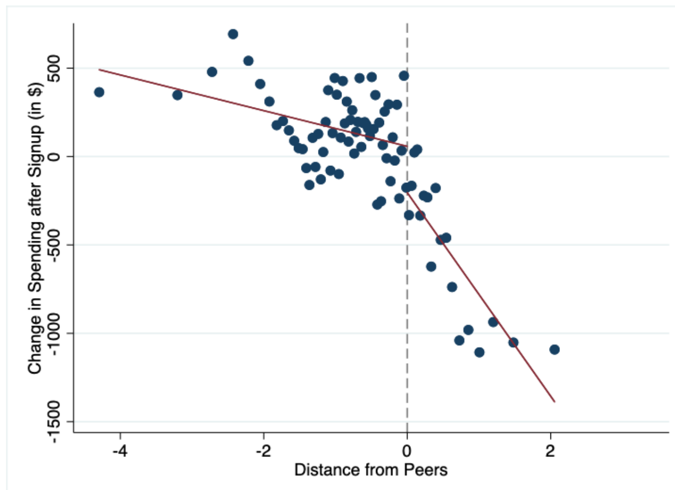
This figure shows binned scatterplots of changes in total consumption after signing up for *Status* and differences in consumption between individuals and their peer group. In all subfigures, the x-axis measures the difference in consumption with respect to peers, normalized by its standard deviation. 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. 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.4
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* 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.

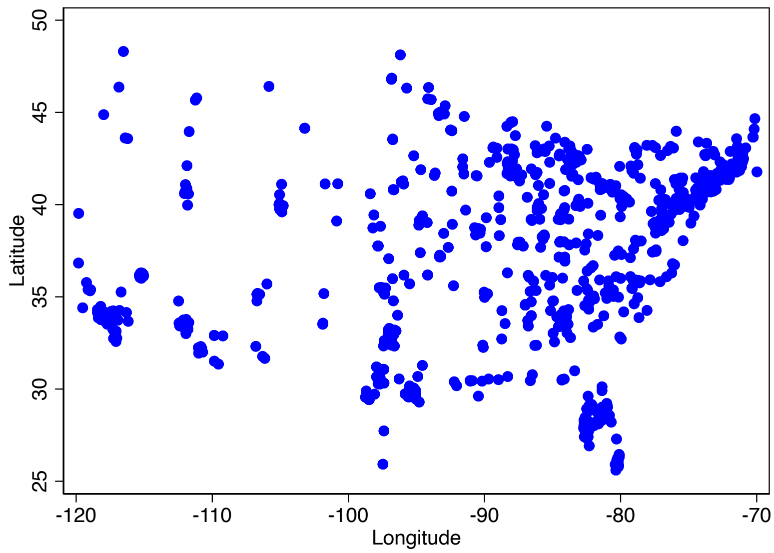


Figure A.5
Location of Respondents to the Randomized Control Trial

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. 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.

Table A.2. Placebo Instrumental-Variable (IV) Analysis

	First Stage	Second Stage	First Stage	Second Stage
Above Dummy	0.078		0.003	
	(0.795)		(0.025)	
Peer Spending		0.942		-29.08
		(0.423)		(-0.0252)
Spending Before	0.120***	-0.566**	0.168***	4.481
	(3.455)	(-2.024)	(4.087)	(0.0231)
Homeownership			0.256***	7.508
			(4.268)	(0.026)
log of Credit Score			0.212**	6.011
			(2.187)	(0.025)
log of Age			-0.487***	-14.140
			(-3.318)	(-0.025)
log of Asset Balance			-0.018	-0.509
			(-1.426)	(-0.025)
log of Debt Balance			0.0185	0.515
			(1.340)	(0.024)
Observations	678	678	477	477
IV F-stat		0.632		0.001

This table reports results for a two-stage-least-squares identification strategy that compares users just below and users at or above the following placebo income thresholds that *Status Money* does not use to define peer groups, that is, \$45K, \$60K, \$90K, \$110K, and \$140K. For each threshold, we only keep users whose income is at most \$6K below the threshold as well as those whose income is at most \$2K above the threshold. We then estimate the following first-stage specification:

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

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. In the second stage, we use 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.

A.2 Survey Questions

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 truly reflects your opinion. 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?*