Crowdsourcing Financial Information to Change Spending Behavior

Francesco D'Acunto
Boston College

Alberto G. Rossi
Georgetown University
& University of Maryland

Michael Weber
University of Chicago
& NBER

Abstract

We document five effects of providing individuals with crowdsourced spending information about their peers (individuals with similar demographics) through a FinTech app. First, users that overspend with respect to peers reduce their spending significantly and while users that underspend keep constant or increase their spending. Second, users' distance from their peers' spending affects the reaction monotonically in both directions. Third, users' reaction is severely asymmetric – spending cuts are three times as large as increases. Fourth, lower income users react more than others. Fifth, discretionary spending drives the reaction in both directions and especially cash withdrawals, which are commonly used for incidental expenses and anonymous transactions. We argue none of Bayesian updating, peer pressure, or the fact that bad news loom more than (equally-sized) good news alone can explain all these facts.

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

Keywords: FinTech, Learning, Beliefs and Expectations, Peer Pressure, Financial Decision-Making, Saving, Consumer Finance.

We are indebted to Majd Maksad and Korash Hernandez (the founders of *Status* Money Inc.) for providing us with invaluable feedback and comments at each step of this project in addition to allowing us to use the company's data. For relevant comments and suggestions, we thank ... All errors are our own.

1 Introduction

Low savings limit the wealth accumulation of US 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; and Lusardi and Mitchell, 2007). Channels that contribute to this phenomenon – whether neoclassical or not – include among others liquidity constraints (Zeldes, 1989; Jappelli and Pagano, 1994), hyperbolic discounting (Laibson, 1997), limited attention (Madrian and Shea, 2001; Carroll, Choi, Laibson, Madrian, and Metrick, 2009), expectations-based reference-dependent preferences (Pagel, 2017), and the lack of financial literacy (Van Rooij, Lusardi, and Alessie, 2012; Lusardi and Mitchell, 2014).

An issue related to financial literacy is that most US households have little information about the income, spending, and saving rates that would guarantee the appropriate wealth accumulation before retirement – they are often financially illiterate and/or have no access to financial advice (Lusardi and Mitchell, 2017). In principle, households could obtain information about saving norms while observing the overall spending of peers (D'Acunto, Malmendier, Ospina, and Weber, 2018). But although consumption is sometimes conspicuous (Charles, Hurst, and Roussanov, 2009), the overall spending of peers is mostly unobserved and hence households can barely learn about the prevailing saving rates of those with similar incomes and demographic characteristics.

If this information friction was material, telling households the spending of peers with similar income and other demographic characteristics might change households' beliefs about the appropriate spending and saving rates. This update would happen irrespective of whether peers' saving rates are optimal or unoptimal as long as households believe the signal they receive is credible and valuable (Gargano and Rossi, 2018; Gargano, Rossi, and Wermers, 2017). Moreover, this information might affect households' beliefs and choice both directly – through learning about others' spending – and indirectly – through peer pressure, that is, the concern of lagging behind with respect to peers (Bursztyn, Ederer, Ferman, and Yuchtman, 2014).

In this paper, we study the effects of providing households with crowdsourced information about their peers' consumption spending through a free-to-use FinTech application (app) called *Status*. Upon subscribing to the app, users provide a set of demographic characteristics, which include their annual income, credit score, age range, homeownership status, location of residence, and location type. Using transaction-based data from a large sample of US consumers, *Status* computes the average monthly spending of consumers with similar characteristics as the users (*peer households*). Moreover, users link their credit, debit, and other financial accounts to the app. Using users' past and present transactions from their own financial accounts, *Status* computes users' own recent average monthly spending. *Status* then produces vivid and easy-to-grasp graphics that compare the evolution of the users' monthly consumption spending with the evolution of the peers' spending. Figure 1 shows one of these graphics *Status* users see on their homepage.[1] These graphics give users simple and immediate feedback on whether their spending is higher, similar, or lower than peers' spending. Displaying this crowdsourced information in an easy-to-understand setting is a crucial feature of *Status*, which aims to avoid the potential ineffectiveness of financial-literacy trainings for non-sophisticated households.

We find being exposed to peers' spending changes users' own spending decisions based on whether users overspend or underspend relative to peers. On average, users that overspend relative to peers reduce their seasonally-adjusted spending by $237 per month around the adoption of the app. Instead, users that underspend increase their seasonally-adjusted consumption spending by $71.

We lever the granular data we access to assess the potential channels and mechanisms that drive this average effect of information about peers, which masks substantial heterogeneity. First, all the users that overconsume relative to peers reduce their monthly spending, whereas all the users that underconsume relative to peers keep constant or increase slightly their monthly consumption spending .

A second robust fact is the distance of users' spending from the peers' average spending

[1] Figure A.1 and Figure 3 show other graphics and Section 2 describes details about the setting.

Figure 1. Graphics comparing Users' and Peers' Spending on *Status*' homepage

affects households' reactions monotonically in both directions – the further away is the user from the peers' consumption spending, the stronger is the convergence of the users to peers' spending. A one-standard-deviation increase in the distance from peers' consumption spending for the overspenders is associated with a drop in monthly consumption spending of about 9.3% in the two months after adoption of the app.

These two results paired with the fact that users converge to the levels of peers' consumption spending both above and below the threshold suggest users find the crowdsourced information *Status* diffuses valuable, relevant, and learn from it. Note that *Status* does not say the average behavior of peers is optimal in any respect. Users might assume the average behavior of peers includes information about optimal spending behavior conditional on demographics and that *Status* lets them harness this "wisdom of the crowds" (e.g., see Galton, 1907; Wolfers and Zitzewitz, 2004; Da and Huang, forthcoming).

The third fact we document is that the reaction to information about peers is severely asymmetric across the positive and negative domains – users that spend more than their peers cut their monthly spending normalized by income by 3% in the two months around adoption, whereas underspenders increase it on average by 1%. This asymmetric sensitivity of users to peers' information based on their consumption spending relative to peers is a robust feature of the data. We also show in a fourth fact that the cut in consumption spending is substantial for overspenders in the lowest quartile of the distribution by income, whereas the size of the change is about 50% smaller for underspenders in the highest quartile by income.

The asymmetries of the reaction of overspenders relative to underspenders, and especially of low-income overspenders, suggests that on top of learning about saving rates users might face pressure when they are compared directly to peers. Receiving bad news about spending relative to peers looms more than receiving same-size good news, which is hard to reconcile with Bayesian updating.

We then move on to assess which spending categories users adjust more after they obtain information about peers' spending. Consistent with the presence of frictions in consumption spending, the whole margin of adjustment comes from discretionary spending relative to non-discretionary spending, which households can barely reduce.[2] Cash withdrawals show a dramatic drop after sign up for households that overspend with respect to their peers relative to food and drink expenses, utilities, or fees and tuitions. Because cash is mainly used for incidental expenses (Bagnall, Bounie, Huynh, Kosse, Schmidt, Schuh, and Stix, 2014) and for transactions consumers want to keep anonymous (Acquisti, Taylor, and Wagman, 2016), the change in spending behavior we document might reduce expenses that are the least likely to provide goods and services to the benefit of the whole household as opposed to the benefit of one member of the household.

The baseline facts we discussed above do not rule out the possibility that users that signed up to *Status* had already decided they would cut or increase their spending based on what they correctly knew or guessed about their peers (for instance, see D'Acunto, Prabhala, and Rossi (forthcoming)). These users might have signed up to *Status* to enjoy other features of the app, such as the income-aggregation function or the possibility to set dynamic targets for consumption and savings. They might have changed their spending irrespective of the information they received about peers.

[2]As we discuss below, non-discretionary spending includes groceries, fees, mortgage payments, and tuitions. Discretionary spending includes outside food and drink spending, clothes, entertainment, travels, and cash withdrawals.

To tackle this endogeneity concern, we propose an identification strategy for the causal effect of providing users with information about peers on users' change in consumption spending. Our identification strategy exploits the fact that *Status* constructs peer groups based on ranges of demographic values. *Status* computes the average monthly consumption based on the transactions of peers whose income falls in the same range as the user's. Because of this feature, two users whose incomes are close to the boundaries of two peer groups, but so that one user falls slightly below the threshold and one user falls at the threshold will be provided with different information about the average peers' monthly spending even if their incomes are almost indistinguishable. Importantly, users do not know the thresholds *Status* uses to construct peer groups and hence they cannot strategically manipulate their position on one side of the discontinuities or the other to avoid receiving negative news about their consumption spending relative to peers.

For an example of the identification design, consider two adjacent yearly income ranges *Status* uses to compute peers' spending are \$25K-\$49K and \$50K-\$75K. Suppose user A declares he/she earns \$49K, whereas user B rounds his/her yearly income to \$50K. Although these reported incomes only differ by \$1K – which is likely to represent the mere tendency of B to round and A to not round, and hence potentially underlying the same exact yearly income for A and B – users A and B will observe substantially different information about their peers' consumption spending. Specifically, user B will observe a peer spending value that is the average of the transactions of US consumers earning between \$25K-\$49K, whereas A will observe a higher peer spending value – the average of the transactions of US consumers earning between \$50K-\$75K. Because users do not know the thresholds *Status* uses to compute the peers, we argue that users that fall around the income thresholds for the peer groups are assigned quasi-randomly to alternative pieces of information about peers.

This strategy confirms our baseline results – users that happen to be assigned to a peer group whose spending is lower cut their consumption more relative to users that are almost identical in terms of income levels but are assigned to a peer group whose average spending is higher.

Note *Status* is marketed as an app that improves saving decisions by providing accessible information about peers' spending but also other services. In particular, *Status* users are not only exposed to information about peers but also to information about the national average spending in the US as well as users' own average monthly income. One might wonder whether the average effect we attribute to reaction to information about peers is at least in part driven by reaction to other types of information users obtain at the time of sign up.

In the last part of the paper, we assess three economic channels through which exposure to information on *Status* might affect users' spending decisions. The first channel – *wisdom of the crowds* – implies users update their beliefs about the optimal spending rate after observing information about peers. For this channel to be relevant, users need to believe the information *Status* gives them is an informative signal about their optimal spending rate irrespective of whether this is true or not. Although this channel can explains some of the facts we documented, it can barely explain the asymmetric reaction of overspenders relative o underspenders.

The second channel we consider is *peer pressure* – individuals might obtain disutility from behaving worse than their peers. In this case, overspending might be perceived as a negative behavior because it reduces users' financial health with respect to peers. This channel can explain the reaction of overspenders but can barely explain why underspenders – who are not behaving worse than their peers – would react at all.

The third channel we consider is *overreaction to negative news.* Under this channel, users learn from peers' spending but negative news about the difference between own and peers' spending looms more than positive news. Although this channel has the potential to explain all our facts, we find it is unlikely to drive our results fully in direct tests in which users react more to information about peers over more negative information about average US consumers or about their overspending with respect to their own income.

In terms of economic channels, we conclude that only a combination of the three channels we consider can explain fully the five facts we document.

Overall, our results suggest that providing households crowdsourced information based on micro data households could have barely accessed on their own allows them to learn about peers' spending/saving choices and affects their own spending/saving choices systematically. Our results appear consistent with a role for both Bayesian learning and peer pressure as the economic channels that might help explain households' reactions. FinTech apps thus can provide a cost-effective and vivid, salient way to transmit financial literacy and financial information to households and affect their choices.

The persistence of the effects of providing information about peers is an aspect further research needs to assess. Within the time frame we observe, which includes about 3 months around the adoption of *Status* for our working sample, we do not detect any dissipation of the effect or any reversal of users' choices after the first reaction. Whether the information *Status* provides will have far-reaching implications for savings throughout the users' working life and up to retirement will require longer time series, but the size of the average reactions at adoption documented in this paper are large in economic magnitude.

2 Institutional Setting

Purpose of the app we study (*Status*)

Status is an app designed to help individuals make more informed decisions in the personal finance space. The app shows users how individuals comparable to them manage their finances, that is, how they spend their money, what interest rates they pay on their loans, and what credit cards they use—among others.

Clients sign-up

To enroll into the app, 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 the credit bureau that returns all of the user's credit-score-related information.[3] Finally, the app asks users to link their checking and savings accounts, their

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

credit card accounts, as well as taxable and non-taxable accounts.

For each user, the app constructs a peer group based on the user's age, income, location, credit score, and housing type. Peer groups are constructed to be as precise as possible subject to the constraint that there should be at least 5,000 individuals in each group. The trade-off is that coarse groups may not be too informative because associated with individuals to whom users do not relate. On the other hand, spending patterns constructed using too few individuals may be too noisy and provide non-credible information. Note that for the sake of testing whether users react to information about peers' spending, whether such information is accurate or inaccurate is not relevant as long as users think the information they obtain contains an informative signal about their optimal spending rate.

In Figure 2, we provide an example of the screenshot *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 of age, 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, who live in New York City, pay rent, and have a credit score that ranges between 720 and 779.

Main features of the app

Once the user is enrolled, the app retrieves automatically information from the users' savings and investment 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.

The main feature of the home page is comparing user's spending to their peers' spending. Figure 1 in the introduction displays the vivid graphics that compare the users' own daily spending based on daily transactions to the projected average daily spending of the peer group and the national US average. The screenshot is taken as of October 30th. On the top, the plot shows the total spending of the client, which turns out to be $17,799, together with the peers consumption, \$8,651, and the national average, \$4,222. The blue line presents the cumulative spending of the client over the course of the month until October 30^{th}. It also presents a forecast of total spending until the end of the month. On the same graph, the light and dark red lines presents the peers and national average cumulative consumption over the month. The app also displays as a grey dotted line the clients average income, \$10,204. As a final piece of information, the app explicitly communicate the client how he/she is doing in terms of consumption for the current month. Note the users' spending is based on their actual daily transactions. 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.

Note this discrepancy in the way users' and peers' data are treated is not relevant to the scope of our inquiry – whether users react to peers' information – unless the difference in frequency and timing of the pieces of information makes certain users believe the information *Status* gives them is not credible. But in this case, we would observe no reaction of users to peers' information irrespective of their distance from peers' spending. If anything, this feature of the app might reduce the average reaction on the side of users.

The bottom of the home page reports more comprehensive statistics regarding the clients' debts, assets, net worth and credit score (see Figure 3). Our fictitious user has a debt of \$37,393, which is compared to peers' debt of \$13,429 and a national average debt of 50,297. On top of this information, the app tells the client that the interest rate he/she is charged is competitive with the national average. The client has \$40,839 in assets, compared to \$45,759 for the peers and \$119,934 for the national average. The interest rate on the asset accounts 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 client's credit score. He/she has a credit score of 769, the peers average 754 and the national average is 630.

3 Data

Status collects and displays large amounts of information from and to their clients. In many cases, some of this information is calculated on the fly when the client requests it and it is not stored on their servers. For this study, we were able to obtain a subset of the information *Status* collects. We categorize and describe the variables we collect below.

Client characteristics at sign-up

Under this header, we have the following variables: *memberid*, a unique client identifier; *joindate*, the date on which the client joined *Status*; *creditscore_member*, the credit score of the client; *income_member*, the income of the client; *housingtype_member*, whether the client owns or rents the house he/she lives in; *age_member*, the age of the client; *citystate_member*, the city and state of the client; and *zip5*, the client's zip-code.

Peer group information

Under this header, we have the following variables: *peergroup_id*, the id of each peer-group; *peergroup_creditscore*, the credit score range of the peer group; *peergroup_income*, the income range of the peer group; *peergroup_housingtype*, whether the peer group is constituted by clients that rent or own their place; *peergroup_agerange*, the age range of the peer group; *peergroup_location*, the location of the peer group; *peergroup_areatype*, whether the peer group is populated by clients in rural or urban areas; *peergroup_size*, the size of the peer group; *peergroup_creditscore_avg*, the average credit score of the peer group; *peergroup_debts_avg*, the average debt of the peer group; *peergroup_assets_avg*, the average assets of the peer group; *peergroup_networth_avg*, the average net worth of the peer group; *peergroup_income_avg*, the average income of the peer group.

Client usage information

Under this header, we have the following variables: *custompeergroupind*, whether the customer constructed one or more custom peer groups; *sessioncount_30*, *sessioncount_60*, *sessioncount_90*, the number of logins during the first, second, and third month after signing up; *trackerind*, whether the customer has an active tracker; *activetrackers*, the number of active trackers; *accountslinked*, the number of accounts linked as of August 21, 2018; *accountslinkedafterjoin_day0*, the number of accounts linked on the day the client signed up; *accountslinkedafterjoin_week1*, *accountslinkedafterjoin_week2*, *accountslinkedafterjoin_week3*, *accountslinkedafterjoin_week4*, the number of account linked during the first, second, third, and fourth week after joining the *Status*; *assetaccountslinked*, the number of asset accounts linked as of August 21, 2018; *assetbalance*, the asset balance as of August 21, 2018, computed across all asset accounts; *debtaccountslinked*, the number of debt accounts linked as of August 21, 2018; *debtbalance*, the debt balance as of August 21, 2018, computed across all debt accounts; *savingsbalance*, the savings balance as of August 21, 2018; *investmentbalance*, the investments balance as of August 21, 2018.

Client spending information

Under this header, we have the following variables: *spend_prev30_all*, *spend_prev60_all*, and *spend _prev90_all*, are the total spending over the first, second, and third month before signing up for *Status*; *spend_post30_all*, *spend_post60_all*, and *spend_post90_all*, are the total spending over the first, second, and third month after signing up for *Status*. The total spending is also broken down by categories, that is: checking account withdrawals, auto and gas, education, entertainment, fees, gifts and charity, groceries, health and medical, home improvement, housing, loans, restaurants shopping, travel, utilities and bills, and other expenses. From these categories, we construct discretionary and non-discretionary spending. The first contains: checking account withdrawals, entertainment, restaurants, shopping, travel, fees, and other expenses. The second contains: groceries, utilities and bills, health and medical, auto and gas, and education.

Peer group spending information

Under this header, we have the same variables contained under "client spending information," but computed at the peer group level.

3.1 Basic Client Characteristics

Table 1 reports the basic characteristics of the clients in our sample. For each variable, we report number of observations, average, and standard deviation. The first three variables are demographic characteristics: *Age*, *Credit Score*, and *Home Ownership*. The average client is 30 years old, with a standard deviation of 7 years, indicating that *Status* users are rather young. The average credit score is 728, higher than the average US credit score of 687. Thirty-eight percent of users are homeowners, which is below the US average in line with the fact our sample is on average younger than the US average consumer.

The average client earns approximately $90,000 per year, with a large standard deviation of $61,000, suggesting that our sample spans individuals with varying levels of income. The majority of the *Status* clients have a positive net worth. The average assets are $42,462, while the average debt—including credit card debt—equals $29,971.

Figure 4 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 $2,200, $3,409, $4,470, and $6,974. Second, the within-group standard deviation of spending increases with income. Higher-income individuals have more varied levels of spending than lower-income individuals, which is consistent with low-income individuals facing spending constraints.

4 Sign up and Spending: Baseline Results

Our first set of analyses tests whether the two pieces of information subscribers receive at sign up—whether they spend more or less compared to their peers, and how different is their spending with respect to peers' spending—have any effects on subscribers' subsequent spending behavior. We first compute the overall spending for each subscriber for the 60 days before sign up and the 60 days after sign up, and measure the change in aggregate spending across the two periods. Because spending is cyclical, we deduct the average change in spending across all users from the change in consumption of each user. We refer to this quantity as seasonally adjusted spending in some cases and simply as spending in other cases.

As reported in Panel A of Table 2, in the raw data we find that the average subscriber that overspends with respect to his peer group reduces his spending after signing up to *Status* by an average of \$474/2=\$237 per month in the first 60 days after sign up. Users that underspend compared to their peers instead increase their monthly spending by \$142/2=\$71.

To allow a more appropriate comparison across subscribers with different levels of income, we normalize the change in aggregate spending by the subscribers' income to make sure systematic differences in the propensity to spend across income levels do not drive any results. The results, reported in Panel B Table 2, suggest that overspenders reduce their spending by 3% of their income, whereas underspenders *increase* their spending by 1% of their income.

Next, we ask whether users' distance from peers' spending affects their reaction in terms of change in spending. To address this question, we first rely on the raw data and plot the average change in spending at the level of groups of users as a function of the groups' distance from peers' spending for both groups of users that overspend and underspend. Figure 5 reports the results of this analysis. Subfigure (a) reports the results for changes in spending, while Subfigure (b) reports the results for changes in spending, normalized by income. Each binned scatterplot divides the 17,500 clients in 100 groups. Figure 5 documents two features of the raw data. First, the distance of each group of users from their peers' spending is monotonically related to users' change in spending – the farther is the group from the peers' spending level, the higher the change in their spending, irrespective of the sign. The second fact is a substantial asymmetric sensitivity of users' change in spending to their distance from peers' spending based on whether the group of users spend more or less than their peers.

As an aside, note that the average subscriber underspends compared to the peers. This detail is likely driven by the fact that peers' consumption is computed in one specific month, July 2017, in which consumer spending in the United States was particularly high. Because our regressions include a constant, the constant captures any systematic difference between the change in spending of all subscribers and peers, and hence this feature of the data does not confound our baseline results.

We repeat the analysis described above more formally by estimating the following set of linear equations by ordinary least squares:

$$\Delta Spending_i = \beta_0 + \beta_1 \ Distance \ from \ Peers_i + \epsilon_i \tag{1}$$

We standardize the distance to peers so that the β_1 coefficient can be interpreted as the association between a standard deviation increase in $Distance \ from \ Peers_i$ and the change in spending after users sign up to *Status*. We estimate this specification separately for users above and below the spending of their peer group.

The results for estimating equation (1), reported in Table 3, show that the distance to peers' spending impacts users' change in spending in both directions. Subscribers far away from the average spending of their peer group are the ones that change their spending by more relative to other users. The relationship between distance and change in spending is monotonic in both directions.

Table 3 also confirms the asymmetric sensitivity to peer consumption based on whether the user over- or underspends before signing up for *Status*. Users that learn they underspend compared to the peers barely change their spending attitude. They increase their consumption by \$183 (Panel A), which corresponds to an income-normalized increase in consumption of only 1% (Panel B). To the contrary, subscribers that learn they over-consume cut their spending by \$1,126 (Panel A) and their income-normalized spending by 9.3% (Panel B), compared to their pre-subscription spending.

Overall, our analysis suggests that subscribers who learn they are overspending compared to their peers cut their monthly spending substantially by an average of 3% of their monthly income, and the cut is proportionally larger the more subscribers overconsume compared to peers. Subscribers that learn they underspend compared to their peers (barely) react to this news by increasing their spending by 1% of their monthly income.

4.1 Multivariate Analysis

The baseline results reported in Table 3 do not control for client characteristics. In Table 4, we repeat the analysis in Panel B of Table 3, but include demographic controls. We estimate:

$$\Delta Spending_i = \beta_0 + \beta_1 \ Distance\ from\ Peers_i + \boldsymbol{\gamma}' \boldsymbol{x}_i + \epsilon_i, \tag{2}$$

where the vector of controls $\boldsymbol{x}_i$ contains: *asset balance*; *income*; *home ownership*; *credit score*; *age* and *age-squared*; and *debt balance*. The coefficient estimates on $Distance\ from\ Peers_i$ remain largely unchanged relative to the univariate counterpart. The below-peer consumption coefficient changes from -1.01 to -1.20. Whereas the first is statistically significant at the 1% level, the second is only significant at the 5% level. The above-peer consumption coefficient changes from -9.34 to -11.95—both statistically significant at the 1% level. Among the controls, the only regressor significant at the 5% level across all specifications is Asset Balance, which suggests that the higher is the amount of assets available to users, the more users increase their spending after signing up for *Status*.

4.2 Heterogeneous Effects Across Income Levels

After having tested for the baseline effects of peer spending information on subscribers' spending decisions, we move on to assess the potential heterogeneity of the effects across client characteristics. Take income level as an example. One could think of arguments suggesting that the effects might both be stronger and weaker at lower levels of income. On the one hand, individuals with lower income might react more to overspending because they have fewer resources to hire financial advisors, and hence the information about peers might be more useful to them. On the other hand, lower income individuals might have less discretionary spending than others, making it hard for them to change their spending in the short run irrespective of the information they receive regarding their peers.

Figure 6 reports the results for estimating the baseline regression of the change of normalized spending over income on indicator variables for whether the subscriber overspends with respect to his/her peers. Each Subfigure reports the results for estimating the coefficients separately across four quartiles of income. Lower-income subscribers react more when they learn they overspend relative to higher-income subscribers that learn they overspend. While all the figures may appear similar at first sight, the range of the y-axis is much larger in Subfigure (a) and (b) (from -40% to 20%) that report results for the first and second income quartile, respectively. The y-axis in Subfigure (c) ranges only from -20% to 10%, while the one for the top income quartile (Subfigure (d)) ranges only from -10% to 10%.

To test more formally whether the sensitivity of spending differences to the distance from peer spending changes systematically across income groups, we estimate the following linear regression by ordinary least squares:

$$\Delta Spending_i = \beta_0 + \beta_1 \ Distance_i + \sum_{j=1}^{3} \delta_j \ Distance_i \times Income_{i,j} + \boldsymbol{\gamma}' \boldsymbol{x}_i + \epsilon_i, \tag{3}$$

where Δ *Spending_i* is the change in consumption of individual i after signing up for *Status*; $Distance_i$ is the difference in consumption between individual i and the average consumption of his/her peer group at the time of sign-up; and $\boldsymbol{x}_i$ is a vector of control variables. The vector of control variables contains: *Asset Balance*, the total asset quartile dummy of the client at the time of sign-up; *Income*, the income quartile dummy; *Credit Score*, the credit score quartile dummy of the client at the time of sign-up; *Debt Balance*, the debt balance quartile dummy at the time of sign-up; *Age* and Age^2, the age and squared age of the client; *Home Ownership*, an indicator variable of whether the client is a home-owner. We report the estimated coefficients for $Distance_i$ and the interaction between $Distance_i$ and the quartile dummies of the control variables. In all cases, the base case is the fourth quartile.

The estimates computed across all customers show that the ones in the higher income quartile react the least. The coefficient on the distance to peers is -2.28 and significant at the 1% level. The coefficient estimate on the interaction between consumption with respect to peers and income decreases monotonically with income. It equals -2.73, for the third income quartile, -4.66 for the second income quartile, and -6.32 for the first income quartile—all significant at the 5% level. As a result, the sensitivity to peer consumption equal $-6.32-2.28 = -8.60$ for the lowest income quartile, $-4.66 - 2.28 = -6.94$ for the second income quartile, and $-2.73 - 2.28 = -5.01$ for the third income quartile.

For users below peer consumption at sign-up, none of the interactions are significant, indicating that there is very little heterogeneity in the response across various income groups among underspenders.

For the users above peer consumption at sign-up, on the other hand, we find the coefficient on distance is economically large – -5.20 relative to the omitted category (top income quartile) and significant at the 5% level. The coefficient on the interaction between consumption distance and an indicator for the lowest income quartile is also economically large – -29.59 significant at the 1% level. Also, the estimate for the second income group is economically large, -7.26 significant at the 1% level. This indicates that the two lowest income quartiles have a higher sensitivity to excess spending relative to wealthier individuals.

4.3 Heterogeneous Effects Across Spending Categories

The results computed so far are estimated using clients' total spending. We now exploit the richness of the categorization of transactions into consumption categories we observe in the data.

As a first pass, we categorize spending into discretionary and non-discretionary spending, as described in Section 3. Intuitively, we would expect that most of the users' reaction in terms of change in spending involves discretionary spending, because users can barely reduce non-discretionary spending and might have no reason to increase it either.

We re-estimate the baseline results separately for the two types of consumption. The results, reported in Figure 7, suggest that as conjectures the vast majority of spending changes are related to changes in discretionary spending. As shown in Subfigure (a), overspending users cut their discretionary spending substantially more relative to underspenders. Subfigure (b) shows instead that individuals barely react at all in terms of non-discretionary spending.

The regression line is flat both above and below zero, indicating that investors do not adjust their non-discretionary consumption. Also, note that Subfigure (a) and (b) have very different y-axes. The one for discretionary spending ranges from -15% to 10%, while the one for non-discretionary spending ranges only from -1% to 2%.

While many of the individual spending categories do not display much of a reaction – some categories are noisy – at least two categories display intriguing results. The first is checking account withdrawals. As shown in Subfigure (a) of Figure 8, checking withdrawals respond dramatically to information about peer spending in both directions. This phenomenon might occur for a number of reasons. Cash is mainly used for incidental expenses (Bagnall, Bounie, Huynh, Kosse, Schmidt, Schuh, and Stix, 2014) and for transactions consumers want to keep anonymous (Acquisti, Taylor, and Wagman, 2016). The latter group might include both legal and illegal entertainment expenses. One interpretation of this result might be that individuals limit their spending on vices once they discover they overspend relative to their peers, although the data at hand do not allow us to ultimately pin down how users employed the cash they withdrew before signing up to *Status*.

The second spending category we consider is the amount spent to service loans and credit card debt, reported in Subfigure (b) of Figure 8. Individuals seem more reluctant to take out loans and might cut on their borrowing through credit cards when they find they are overspending relative to their peers.

5 Reaction to Information? Estimating the Kink's Location

A limitation of the results reported so far is that we imposed the threshold between those reacting positively and those reacting negatively to information is set at the point of no distance from the average peers' consumption. But if individuals were not basing their reaction only on the value of peers' consumption *Status* shows them, the actual threshold might fall at a value different from zero. Moreover, our results so far do not allow testing whether the regression slope coefficients are statistically different below and above the threshold.

To address these two concerns, we estimate the location of the threshold non-parametrically using two complementary approaches. The first approach builds on Hansen (1996, 2000). It estimates a threshold model with unknown threshold. To build intuition, consider the case of one regressor. The threshold regression estimates the optimal threshold for a linear model that has different intercept and slope estimates below and above the threshold. Hansen (1996) also proposes a test for whether the coefficient estimates below and above the threshold are statistically different from each other.

For the second approach, we follow Hansen (2017) and estimate a regression kink model with unknown threshold. This model is similar to the one described above, but does not allow for discontinuities. The approach is thus similar to estimating a linear spline model that has a single endogenously determined node. Hansen (2017) also develops the asymptotic theory to make statistical inference about the threshold.

Threshold Regression Results

We estimate the threshold regression model on the full set of 17,673 observations and report the results in Panel A of Table 6. The first two columns report the linear regression on the full sample. Columns 2 and 3 (4 and 5) repeat the estimates below (above) the endogenously determined threshold.

The threshold is very precisely estimated to be 0.235, with a 95% confidence interval of [0.233; 0.237]. The heteroskedasticity-consistent Lagrange multiplier test for a threshold developed by Hansen (1996) rejects the null of no threshold with a p-value of 0.00.

The regression estimates below and above the threshold are very similar to the ones reported in Panel B of Table 3. The coefficient equal -1.01 (significant at the 1% level) for the customers below peer consumption. The coefficient is instead -11.09 (significant at the 1% level) for those above peer consumption. To give a visual representation of the results, we present a binned scatterplot of the threshold regression estimates in Subfigure (a) of Figure 9.

Kink Regression Results

Panel B of Table 6 reports the results for the regression with endogenous kink.The threshold is estimated at 0.546, with a 95% confidence interval of [0.34; 0.77] and the null of no-threshold is rejected with a p-value of 0.00. The constant is not statistically different from zero.

The coefficient on consumption difference equals on -0.726 below the threshold and is statistically different from zero. The coefficient above the threshold is instead 15 times larger (in absolute value) as the coefficient equals -11.197. This result indicates, once again, that investors that over-consume are much more responsive to peer-group consumption information, compared to individuals that under-consume, with respect to their peers.

6 Is the Effect of Information About Peers Causal?

In addition to the baseline results reported so far, we now present the estimates from an identification strategy that tests whether the effects we uncover are causal. This concern is relevant, because subscribers might decide to sign up to *Status* Money only after they have already realized they are overspending. In this case, *Status* Money could be used by subscriber as an app that allows tracking one's aggregate spending simply by consolidating all the spending accounts. If the latter interpretation is true, it might be that overspending subscribers are completely uninterested in the information regarding peers, and they might start to cut their spending after sign up merely because they had already decided to do so before subscribing.

To tackle this potential issue, we move on to analyze a set of "identification subsamples," that is, subsamples of subscribers for which the potential external motives to cut spending on top of peer pressure are identical. For this reason, any systematic difference in the change in spending across subscribers in the identification samples cannot be attributed to external motives and should be attributed to the causal effect of peer spending information. To construct our identification samples, we exploit a feature of the design of peer groups on *Status* Money that allows for a regression discontinuity identification design (RDD).

The intuition behind the design is that subscribers' income is a continuous variable, and small differences in income capture very similar subscribers. For instance, if a subscriber reports an annual income of $99K and another subscriber an annual income of $100K, the two subscribers are very similar. At the same time, though, the design of peer groups follows discontinuous thresholds based on subscribers' income. For instance, one threshold is set between $75K and $99K and the adjacent threshold between $100K and $150K. Based on this design, subscribers that report an income of $99K will receive information about the average spending of peers whose income is between $75K and $99K, whereas very similar subscribers that report an income of $100K will receive information about the average spending of peers whose income is between $100K and $150K.

Although the two subscribers are very similar, one of them faces a peer group that spends on average substantially less than the other, and hence the extent of the information treatment will be larger for the $99K income subscriber than for the $100K income subscriber.

We extend this intuition to subscribers just below and at each of the income thresholds *Status* Money uses to define peer groups, that is, $35K, $50K, $65K, $75K, $100K, and $150K. For each threshold, we only keep the clients that are at the lower threshold of the group as well as those clients that are in the lower income group, but are within $3K of the threshold. For example, taking the $100K as an example, we only keep those that declare $100K in annual income as well as those with an income between $97K and $100K. We then undertake a Two-Stage-Least-Squares strategy. In the baseline strategy, we estimate the following first-stage specification:

$$Peer\ Spending\_i = \alpha + \beta\ Dummy\ Above\_i + \epsilon_i, \tag{4}$$

where $Peer\ Spending\_i$ is the peer-spending value for client i, $Dummy\ Above\_i$ is a dummy variable for whether the income is exactly equal to the lower-bound of a threshold. In the second stage, we use the instrumented $Peer\ Spending\_i$ in Equation 4 as the main covariate in the following specification:

$$\Delta Spending_i = \alpha + \beta\ \widehat{Peer\ Spending\_i} + \epsilon_i, \tag{5}$$

where $\Delta Spending_i$ is the change in consumption before and after signing up.

The results—reported in Panel A of Table 7—show that the instrument in the first step is not weak, as the t-statistic associated with *Above_dummy* exceeds 18 across all specifications. The second stage results reported in Panel B show that the causal effect of a higher threshold is positive and significant across all specifications. The t-statistics are always greater than 2.5 and stable across specifications. Economically, the coefficients range from 0.77 (associated with the $5K threshold) to 0.98 (associated with the $3K threshold), indicating that a unit increase in peer-group consumption causes an increase in the clients' consumption between 77 and 98 cents.

7 Understanding the Mechanisms: Learning, Peer Pressure, Overreacting to Negative News

In this section, we discuss the economic channels that might help explain the facts we have documented so far. As we discussed in the introduction, three non-mutually-exclusive channels might contribute to the results.

First is a neoclassical channel – Bayesian updating. Users might believe crowdsourced information about peers' spending contains valuable information regarding the optimal spending rate and might update their beliefs accordingly. Even if any individual peer might not be optimizing their spending based on the users' own characteristics, users might think the average spending of a large group of peers provides a valuable signal. We label this channel *wisdom of the crowds* (for instance, see Da and Huang (forthcoming)). This channel does not involve any non-standard assumptions about users' preferences or beliefs and could explain both the convergence of users' spending to peers' spending as well as the monotonic relationship between the distance of users from their peers and the size of the reaction – convergence requires a stronger reaction the farther away is users' spending from the spending of their peers.

At the same time, the *wisdom of the crowds* channel can barely explain the asymmetry of the reaction based on whether users overspend or underspend relative to their peers. Under the *wisdom of the crowds* interpretation, the reactions of users should be similar in absolute value and symmetric with respect to the kink – the point of zero distance from peers' consumption – whereas we observe a substantially stronger reaction by users that overspend relative to users that underspend. Thus, the *wisdom of the crowds* channel cannot fully explain all our results.

Note that one could consider a non-Bayesian alternative of this channel – *conformism*. Under conformism, individuals obtain utility from mixing with the crowd and reducing their idiosyncrasies relative to their peers. But in this case we would need to assume that conforming to peers from a worse starting point looms more to individuals than conforming to peers from a better starting point in order to explain the asymmetric reaction around the kink.

The second channel we consider is *peer pressure*. By *peer pressure* we mean individuals dislike to perform worse than their peers. In the context of spending, if users were told they overspent relative to peers they might want to amend this behavior and cut on their spending because they obtain utility from perceiving their financial health is not worse than that of their peers. Note that the version of *peer pressure* we propose can help explain the stronger reaction by users that overspend relating to peers, but is unlikely to explain the (slight) convergence of underspenders to their peers' level of spending. Underspenders perform better than their peers in terms of financial health and hence if *peer pressure* was the only channel at play they would not change their behavior after they sign up to *Status*.

The third channel we consider is *overreaction to negative news*. This channel is a modification of the *wisdom of the crowds* channel that adds a non-Bayesian assumption regarding individuals' reaction to learning from information to account for our results in both the overspending and underspending domains. *Overreaction to negative news* suggests that individuals learn from the information we provide them as if peers' spending is a valuable signal but negative news loom more to them than equally-sized positive news. This channel would predict that both overspenders and underspenders react to obtaining information about their peers, but overspenders react more than underpsenders at the same distance from their peers. In principle, this channel could explain all our baseline facts.

Disentangling the three channels above in field data, which include no randomized exposure to different pieces and types of news, is challenging. We propose a set of tests and arguments to assess the potential role of one or more of the channels and their relative magnitude.

First of all, recall that Bayesian learning seems the only plausible channel to explain the reaction of underspenders. We could thus conjecture that the size of the reaction we document in the underspending domain is the effect of Bayesian learning and the convergence to peers' spending Bayesian learning predicts. At the same time, the *wisdom of the crowds* channel predicts a symmetric reaction around the kink for overspenders and underspenders. We could thus use the size of the reaction in the underspending domain to obtain a lower bound for the size of the reaction of overspenders due to non-Bayesian channels. This lower bound is the difference between the size of the reaction we document and the size of the reaction of underspenders – Bayesian learning can barely explain any stronger reaction to the underspenders' one. Although we do not have a structural model to interpret the magnitudes of the reactions in our paper, Panel B of Table 2 documents the absolute value of the normalized change of users' monthly spending is 3 times as large for overspenders than for underspenders. Under our conjecture, this result would suggest that non-Bayesian channels might explain most of overspenders' reaction.

We propose a set of direct tests aiming at disentangling the two non-Bayesian channels we propose – *peer pressure* and *overreaction to negative news*. These tests exploit a feature of *Status* we have not discussed so far. As Figure 1 in the introduction shows, *Status* users do not only observe information about their own spending and the spending of peers on their homepage, but also information about the (i) average spending of all US households and (ii) their own average monthly income.

Under the *peer pressure* channel, we should find that overspending users' reaction in terms of reducing their spending should be most sensitive to the distance of their spending from their peer group. The reaction should be less sensitive to the distance between overspending users and the average US household or users' own average monthly income. This prediction stems from the fact that reacting to overspending with respect to one's own income has nothing to do with comparing oneself with peers. Moreover, the information about peers is explicitly labeled as such and *Status* is marketed as providing crowdsourced and tailored information about one's own peers based on similar demographic characteristics. Users should thus interpret this piece of information as more representative of peers' spending than the information about the average US household.

Under the *overreaction to negative news* channel, instead, users should react most to the worst piece of news they obtain from *Status*, that is, the information that is farthest away from their spending between peers' spending, average US households, and average income.

Across the four panels of Table 8, we regress overspending users' change in spending on the distance of their pre-sign up spending from 4 different points – peers' spending (Panel A), the average US household's spending (Panel B), users' average income (Panel C), and the maximum distance among these three (Panel D). Across columns, we start with the results for the full sample and exclude alternatively the top decile, quintile, or tercile of the sample to ensure none of our results is driven by outliers or extreme reactions. Across the board and for each subsample, the coefficients attached to the distance of users' spending to peers' spending are systematically larger than any of the other coefficients. In particular, the coefficients on the distance from peers are about 3 times as large as those on the distance of the average US household and about twice as large as those on the average users' income and the maximum distance across any three values.

8 Conclusions

We document five effects of providing individuals with crowdsourced information about their peers' spending through a FinTech app. First, all the users that overconsume with respect to peers reduce their spending and all the users that underconsume keep constant or increase their spending. Second, users' distance from their peers' spending affects the reaction monotonically in both directions. We interpret these facts as consistent with convergence after learning about peers' spending. Third, users' reaction is severely asymmetric – overconsumers cut spending substantially more than underconsumers increase it. Fourth, the reaction is substantially larger for the less wealthy. We argue these two results are not consistent with Bayesian updating but might be driven by peer pressure or the fact that bad news loom more than (equally-sized) good news. Fifth, discretionary spending drives the reaction in both directions and especially cash withdrawals, commonly used for incidental expenses and transactions for which individuals want to maintain anonymity. Users thus cut the potentially unnecessary expenses to the benefit of their whole household.

References

Acquisti, A., C. Taylor, and L. Wagman, 2016, "The economics of privacy," *Journal of Economic Literature*, 54, 442–92.

Bagnall, J., D. Bounie, K. Huynh, A. Kosse, T. Schmidt, S. Schuh, and H. Stix, 2014, "Consumer cash usage: a cross-country comparison with payment diary survey data," .

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

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.

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.

Carroll, G. D., J. J. Choi, D. Laibson, B. C. Madrian, and A. Metrick, 2009, "Optimal defaults and active decisions," *The quarterly journal of economics*, 124, 1639–1674.

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

Da, Z., and X. Huang, forthcoming, "Harnessing the wisdom of crowds," *Management Science*.

D'Acunto, F., U. Malmendier, J. Ospina, and M. Weber, 2018, "Shopping, Inflation Expectations, and Household Behavior," *Working Paper*.

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

Galton, F., 1907, "Vox populi (The wisdom of crowds)," *Nature*, 75, 450–451.

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

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.

Hansen, B. E., 1996, "Inference when a nuisance parameter is not identified under the null hypothesis," *Econometrica: Journal of the econometric society*, pp. 413–430.

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

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

Jappelli, T., and M. Pagano, 1994, "Saving, growth, and liquidity constraints," *The Quarterly Journal of Economics*, 109, 83–109.

Laibson, D., 1997, "Golden Eggs and Hyperbolic Discounting*," *The Quarterly Journal of Economics*, 112, 443–478.

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.

Lusardi, A., and O. S. Mitchell, 2014, "The economic importance of financial literacy: Theory and evidence," *Journal of economic literature*, 52, 5–44.

Lusardi, A., and O. S. Mitchell, 2017, "How ordinary consumers make complex economic decisions: financial Literacy and retirement readiness," *Quarterly Journal of Finance*, 7, 1750008.

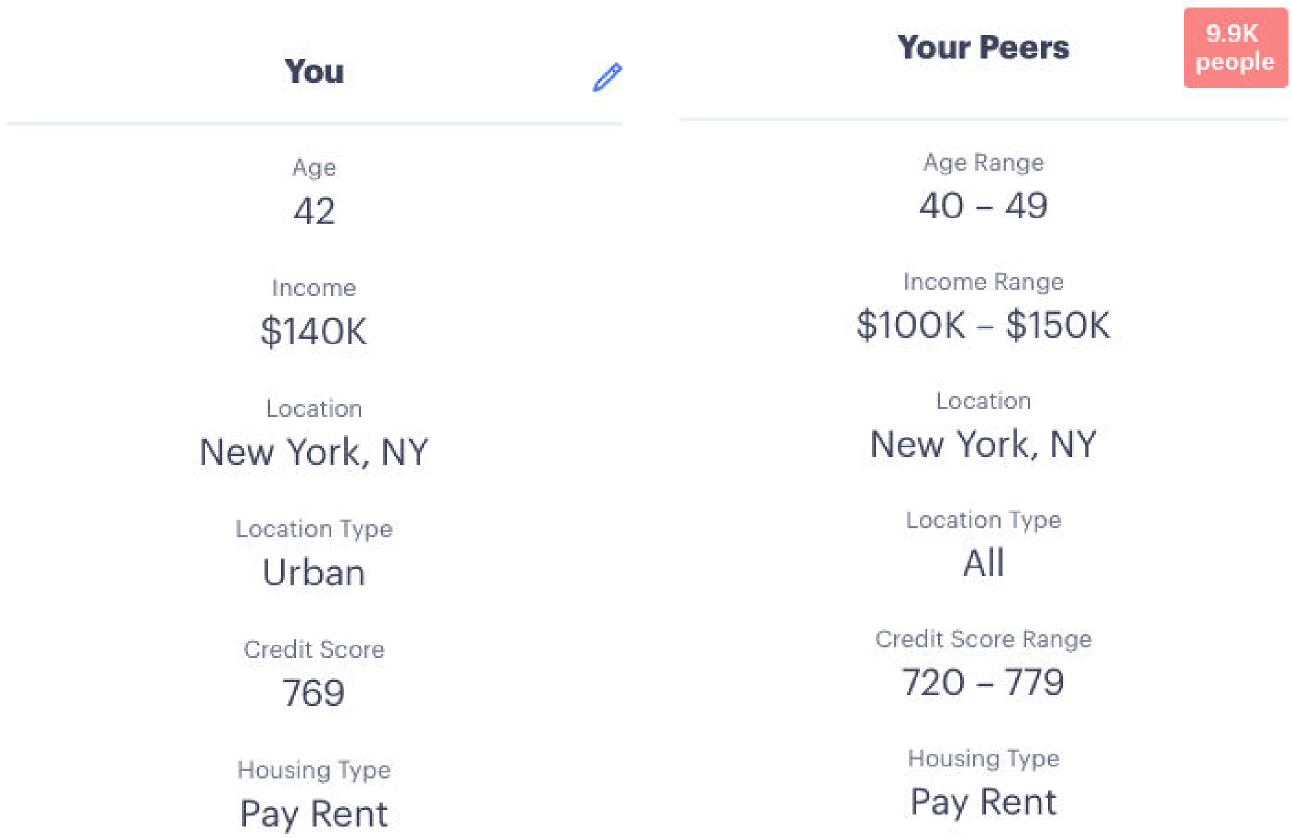
Madrian, B. C., and D. F. Shea, 2001, "The power of suggestion: Inertia in 401 (k) participation and savings behavior," *The Quarterly journal of economics*, 116, 1149–1187.

Pagel, M., 2017, "Expectations-Based Reference-Dependent Life-Cycle Consumption," *The Review of Economic Studies*, 84, 885–934.

Van Rooij, M. C., A. Lusardi, and R. J. Alessie, 2012, "Financial literacy, retirement planning and household wealth," *The Economic Journal*, 122, 449–478.

Wolfers, J., and E. Zitzewitz, 2004, "Prediction markets," *Journal of economic perspectives*, 18, 107–126.

Zeldes, S. P., 1989, "Consumption and liquidity constraints: an empirical investigation," *Journal of political economy*, 97, 305–346.



(a) Client Profile

(b) Peer Group Information

Figure 2
Peer Group for a Sample Account

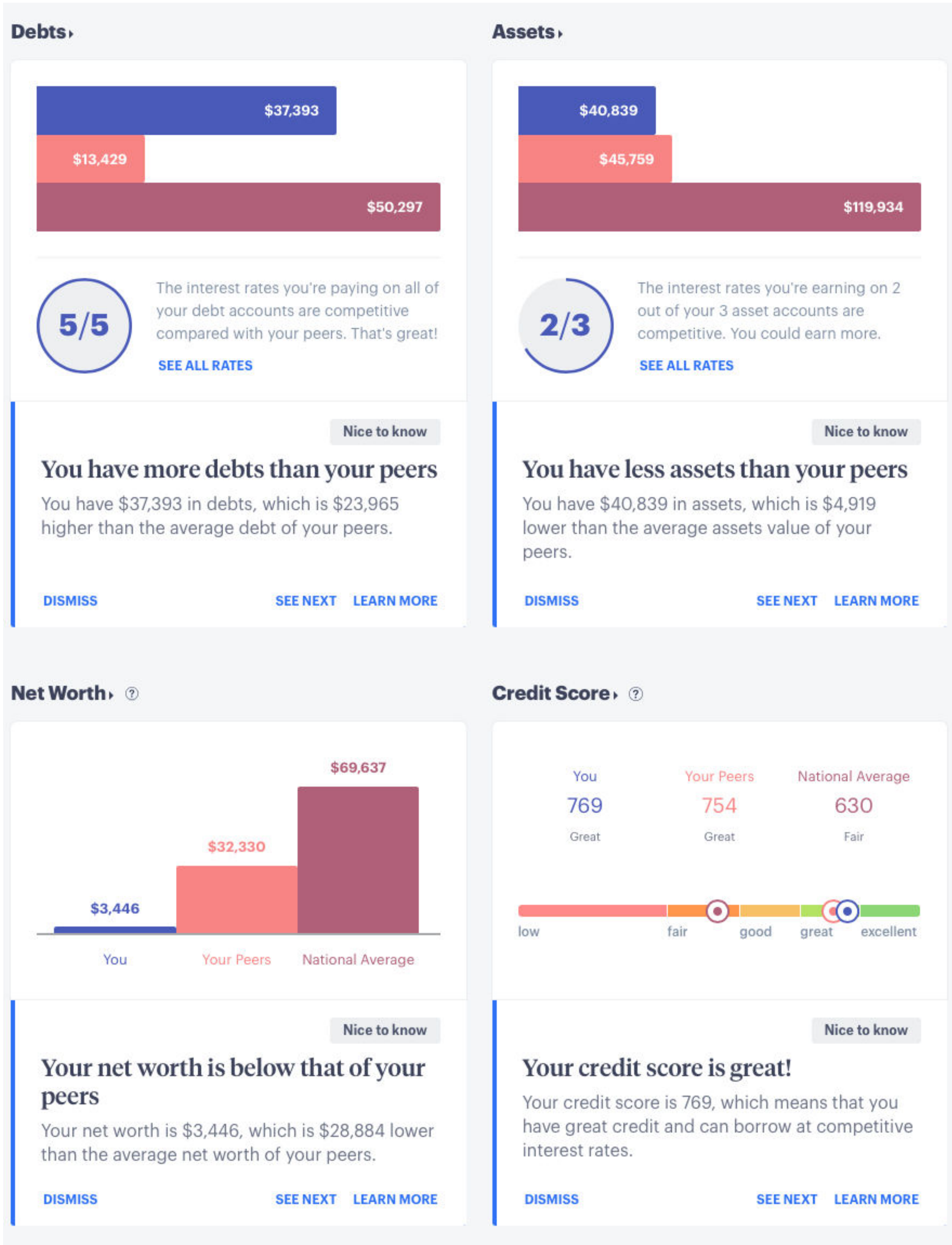


Figure 3
Status Home Page

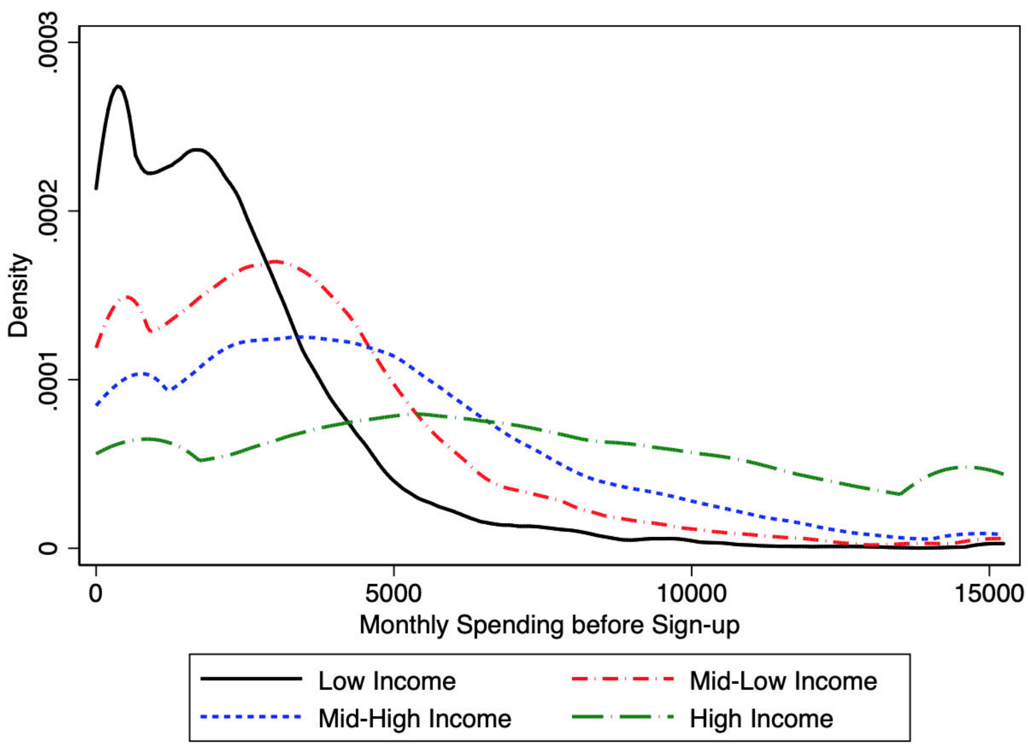
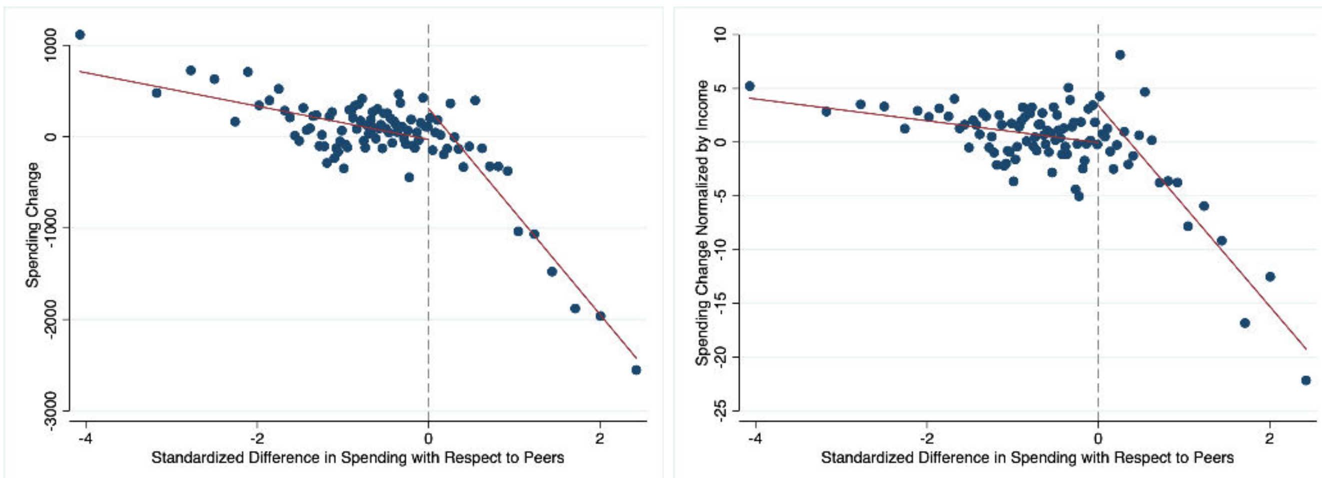


Figure 4
Distribution of Monthly Spending by Income Quartiles

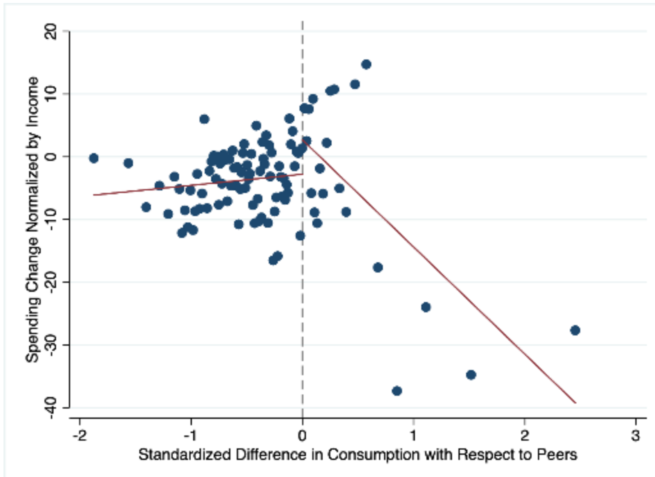


(a) Changes in Spending

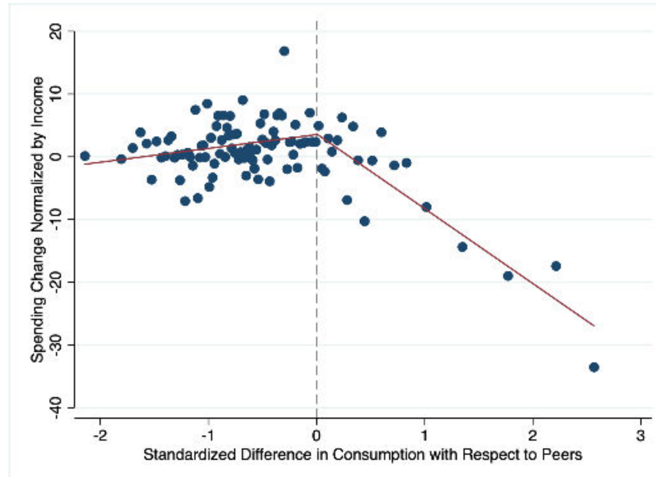
(b) Changes in Spending Normalized by Income

Figure 5
Consumption With Respect to Peers and Changes in Consumption

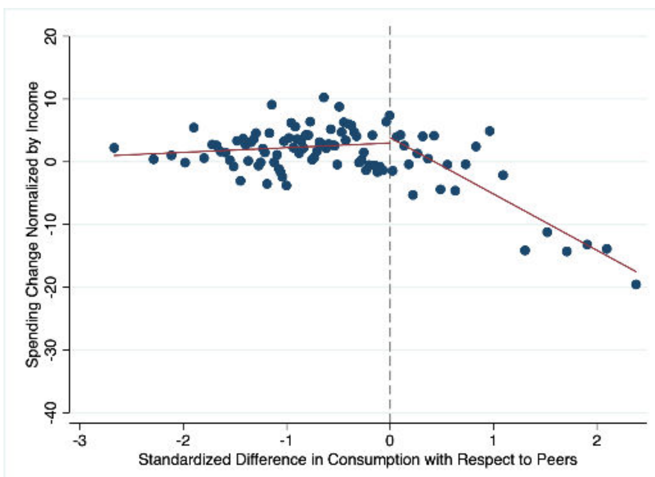
This figure shows binned scatterplots of changes in overall consumption after signing up for *Status* and differences in consumption between individuals and their peer group at the time of sign-up. The x-axis measures the difference in consumption with respect to peers, normalized by its standard deviation. The y-axis in Subfigure (a) reports results for dollar changes in spending, computed using two months before and after investors's sign-up. The y-axis in Subfigure (b) normalizes the changes in consumption by income. The binned scatterplot divides the 17,500 clients in 100 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.



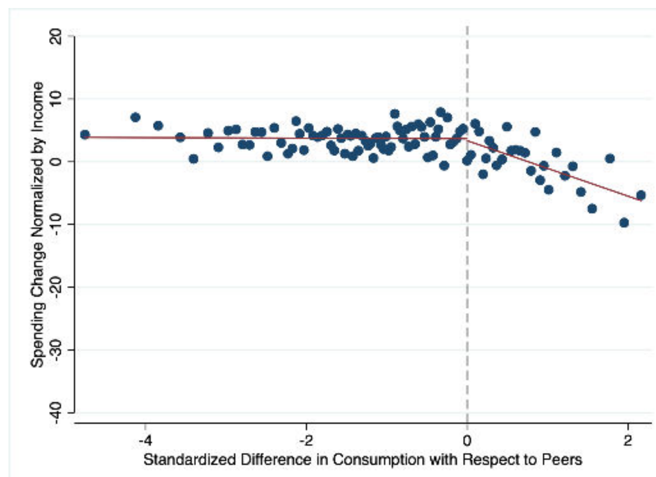
(a) First Income Quartile



(b) Second Income Quartile



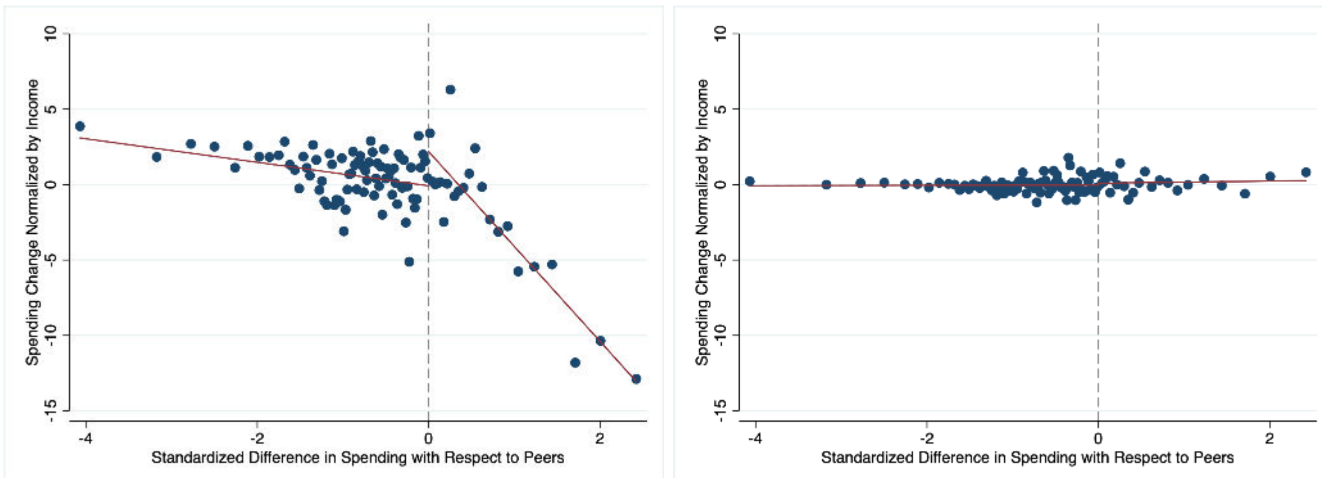
(c) Third Income Quartile



(d) Fourth Income Quartile

Figure 6
Consumption With Respect to Peers and Changes in Consumption—by Income Quartiles

This figure shows binned scatterplot of changes in overall 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 normalized by income, computed using two months before and after investors's sign-up. Each subfigure reports the results for an income quartile and the binned scatterplot divides the clients in each income quartile in 100 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.

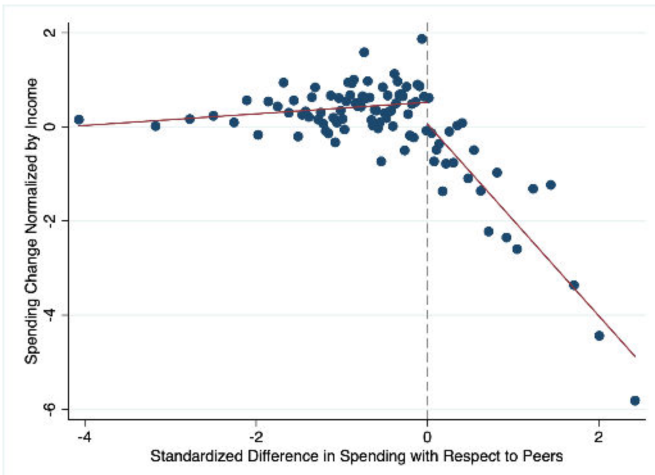


(a) Discretionary Consumption

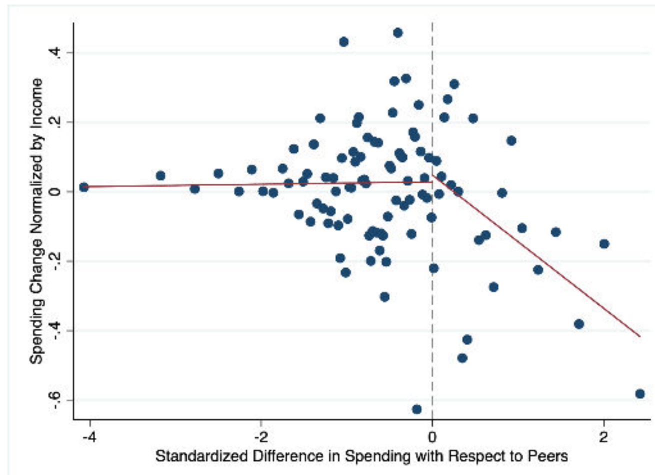
(b) Non-Discretionary Consumption

Figure 7
Consumption With Respect to Peers and Changes in Consumption—Discretionary and Non-Discretionary Consumption

This figure shows binned scatterplot of changes in discretionary and non-discretionary 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 normalized by income, computed using two months before and after investors's sign up. Subfigure (a) reports the results for discretionary consumption. Subfigure (b) the ones for non-discretionary consumption. Each binned scatterplot divides the 17,500 clients in 100 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.



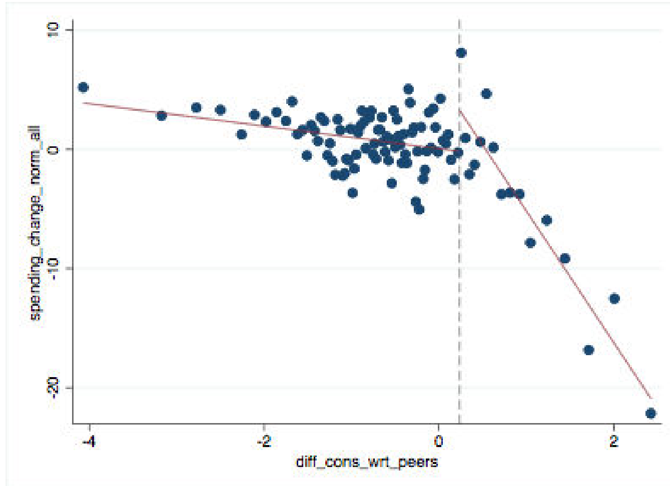
(a) Checking Withdrawals



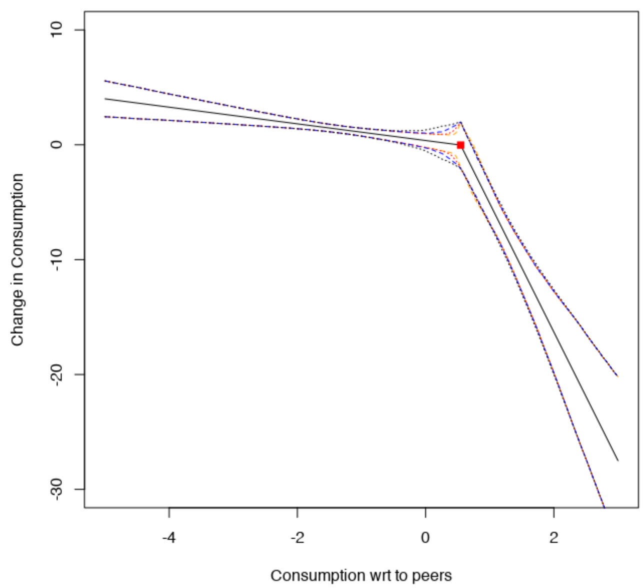
(b) Loans

Figure 8
Consumption With Respect to Peers and Changes in Consumption—Withdrawal from Checking Accounts and Loans

This figure shows binned scatterplot of changes in checking account withdrawals and changes in the loans taken out by clients 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 normalized by income, computed using two months before and after investors's sign-up. Subfigure (a) reports the results for checking account withdrawals. Subfigure (b) the ones for loans taken out by the client. Each binned scatterplot divides the 17,500 clients in 100 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) Threshold Regression with Unknown Threshold



(b) Kink Regression with Unknown Threshold

Figure 9
Consumption With Respect to Peers and Changes in Consumption—Endogenous Threshold Models

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

Table 1. Summary Statistics

	Observations	Mean	St. Dev.
Age	17,673	30	7
Credit Score	16,335	728	84
Home Ownership	17,676	0.38	0.49
Annual Income ($)	17,598	90,055	61,796
Assets ($)	15,325	42,462	68,066
Debts ($)	12,332	29,971	64,637
Monthly Spending (Total) ($)	17,676	4,334	4,073
Monthly Spending (Discretionary) ($)	17,676	2,772	2,906
Monthly Spending (Non-Discretionary) ($)	17,676	680	679
Monthly Spending (Other) ($)	17,676	882	1,475

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.

Table 2. Spending Changes After Signing up for ***Status***

Panel A. Dollar Value Changes in Consumption

	Below Peers		Above Peers	
	Coeff.	*t*-stat	Coeff.	*t*-stat
ΔSpending	142.24***	(5.84)	-474.01***	(-7.81)
Observations	13,596		4,080	

Panel B. Consumption Changes Scaled by Income

	Below Peers		Above Peers	
	Coeff.	*t*-stat	Coeff.	*t*-stat
ΔSpending	0.924***	(4.27)	-3.079***	(-5.25)
Observations	13,596		4,080	

This table presents results for changes in consumption after signing up for *Status*. Panel A reports results for dollar changes in consumption, while Panel B scales the changes in consumption by income. Within each panel, changes in consumption are computed for clients with below-peer consumption in columns 1 and 2; and for clients with consumption above peers in columns 3 and 4. Consumption changes are computed using a 60-day window after and before signing up for *Status*. To account for cyclicality in monthly consumption, we deduct from the change in consumption of each client the average change in consumption across all the clients that sign-up in the same month. As a result, the change in consumption across all clients is equal to zero.

Table 3. Sensitivity to Peer Consumption and Spending Changes

Panel A. Dollar Value Changes in Consumption

	Below Peers		Above Peers	
	Coeff.	t-stat	Coeff.	t-stat
Difference from Peers	-182.6***	(-5.56)	-1,126.4***	(-12.40)
Constant	-30.5	(-0.77)	307.4***	(3.54)
Observations	13,596		4,077	

Panel B. Consumption Changes Scaled by Income

	Below Peers		Above Peers	
	Value	t-stat	Value	t-stat
Difference from Peers	-1.01***	(-3.48)	-9.34***	(-10.57)
Constant	-0.03	(-0.10)	3.41***	(4.03)
Observations	13,596		4,077	

This table reports results for the sensitivity of spending changes to peer consumption. We estimate the following simple linear regression by ordinary least squares:

$$\Delta\_Cons_i = \beta_0 + \beta_1 \; Distance\_from\_Peers_i + \epsilon_i$$

where Δ_Cons_i is the change in consumption of individual i after signing up for *Status*, while $Distance\_from\_Peers_i$ is the difference in consumption between individual i and the average consumption of his/her peer group at the time of sign-up. Consumption changes are computed using a 60-day window after and before signing up for *Status*. To account for cyclicality in monthly consumption, we deduct from the change in consumption of each client the average change in consumption across all the clients that sign-up in the same month. $Distance\_from\_Peers_i$ is standardized so that the coefficient estimates represent the relation between spending changes and a standard deviation increase in $Distance\_from\_Peers_i$. Panel A reports results for dollar changes in consumption, while Panel B scales the changes in consumption by income. Within each panel, regression estimates are computed for clients with below-peer consumption in columns 1 and 2; and for clients with above-peer consumption in columns 3 and 4.

Table 4. Sensitivity to Peer Consumption and Spending Changes, Controlling for Client Characteristics

	All		Below Peers		Above Peers	
	Value	t-stat	Value	t-stat	Value	t-stat
Distance	-4.24***	(-10.85)	-1.20**	(-2.14)	-11.95***	(-5.87)
Asset Balance	0.97***	(3.39)	0.87***	(2.73)	2.38**	(2.42)
Income	-0.40	(-0.40)	-0.77	(-0.63)	3.00	(1.21)
Home Ownership	1.95*	(1.94)	2.63**	(2.15)	-0.20	(-0.07)
Credit Score	0.00	(0.29)	-0.00	(-0.17)	-0.01	(-0.28)
Age	0.04	(0.09)	-0.45	(-1.02)	1.19	(0.95)
Age^2	-0.00	(-0.04)	0.01	(1.13)	-0.01	(-0.88)
Debt Balance	0.42**	(2.46)	0.42**	(2.30)	0.25	(0.46)
Constant	-11.52	(-0.88)	6.79	(0.50)	-70.59*	(-1.82)
Observations	9,597		6,826		2,771	

This table reports results for the sensitivity of spending changes to peer consumption. We estimate the following simple linear regression by ordinary least squares:

$$\Delta\_Cons_i = \beta_0 + \beta_1 \; Distance\_from\_Peers_i + \boldsymbol{\gamma}'\boldsymbol{x}_i + \epsilon_i$$

where Δ_Cons_i is the change in consumption of individual i after signing up for *Status*; $Distance\_from\_Peers_i$ is the difference in consumption between individual i and the average consumption of his/her peer group at the time of sign-up; and $\boldsymbol{x}_i$ is a vector of control variables. Consumption changes are computed using a 60-day window after and before signing up for *Status* and are scaled by income. To account for cyclicality in monthly consumption, we deduct from the change in consumption of each client the average change in consumption across all the clients that sign-up in the same month. $Distance\_from\_Peers_i$ is standardized so that the coefficient estimates represent the relation between spending changes and a standard deviation increase in $Distance\_from\_Peers_i$. The vector of control variables contains: *Asset Balance*, the total assets of the client at the time of sign-up; *Income*, the income of the client; *Home Ownership*, an indicator variable of whether the client is a home-owner; *Credit Score*, the credit score of the client at the time of sign-up; *Age* and Age^2, the age and squared age of the client; *DebtBalance*, the debt balance at the time of sign-up. Within each panel, regression estimates are computed across all clients in columns 1 and 2; for clients with below-peer consumption in columns 3 and 4; and for clients with above-peer consumption in columns 5 and 6.

Table 5. Sensitivity to Peer Consumption and Spending Changes with Client Characteristics Interactions

	All		Below Peers		Above Peers	
	Value	t-stat	Value	t-stat	Value	t-stat
Distance	-2.282***	(-7.19)	-0.884*	(-1.89)	-5.196**	(-2.56)
Distance × Income_1	-6.319**	(-2.20)	2.239	(0.72)	-29.591***	(-2.69)
Distance × Income_2	-4.662***	(-4.25)	0.500	(0.32)	-7.257*	(-1.65)
Distance × Income_3	-2.735***	(-3.99)	-0.094	(-0.10)	-5.379	(-1.64)
Constant	-4.742	(-0.65)	6.186	(0.87)	-22.891	(-0.88)
Other Controls	✓		✓		✓	
Observations	12,256		9,247		3,009	

This table reports results for the sensitivity of spending changes to peer consumption. We estimate the following simple linear regression by ordinary least squares:

$$\Delta\_Cons_i = \beta_0 + \beta_1\ Distance_i + \sum_{j=1}^{3} \delta_j\ Distance_i \times Income_{i,j} + \boldsymbol{\gamma}'\boldsymbol{x}_i + \epsilon_i$$

where Δ_Cons_i is the change in consumption of individual i after signing up for *Status*; $Distance_i$ is the difference in consumption between individual i and the average consumption of his/her peer group at the time of sign-up; and $\boldsymbol{x}_i$ is a vector of control variables. Consumption changes are computed using a 60-day window after and before signing up for *Status* and are scaled by income. To account for cyclicality in monthly consumption, we deduct from the change in consumption of each client the average change in consumption across all the clients that sign-up in the same month. $Distance\_from\_Peers_i$ is standardized so that the coefficient estimates represent the relation between spending changes and a standard deviation increase in $Distance\_from\_Peers_i$. The vector of control variables contains: *Asset Balance*, the total asset quartile dummy of the client at the time of sign-up; *Income*, the income quartile dummy; *Credit Score*, the credit score quartile dummy of the client at the time of sign-up; *DebtBalance*, the debt balance quartile dummy at the time of sign-up. *Age* and Age^2, the age and squared age of the client; *Home Ownership*, an indicator variable of whether the client is a home-owner. We report the estimated coefficient estimates for $Distance_i$ and the interaction between $Distance_i$ and the quartile dummies of the control variables. In all cases, the base case is the fourth quartile. Within each panel, regression estimates are computed across all clients in columns 1 and 2; for clients with below-peer consumption in columns 3 and 4; and for clients with above-peer consumption in columns 5 and 6.

Table 6. Sensitivity to Peer Consumption and Spending Changes. Results from Endogenous Threshold models

Panel A. Threshold Regression Results

	All		Below Threshold		Above Threshold	
	Value	t-stat	Value	t-stat	Value	t-stat
Distance from Peers	-2.52***	(-11.79)	-1.01***	(-3.96)	-11.09***	(-7.81)
Constant	-1.43***	(-5.82)	0.07	(-0.19)	5.94***	(4.27)
Observations	17,673		14,846		2,827	

Threshold Estimate = 0.235; Confidence Interval = [0.233, 0.237]

Hansen (1996) Lagrange Multiplier for threshold: p-value = 0.00

Panel B. Kink Regression Results

	Coeff.	t-stat	Low CI	High CI
Constant	-0.026	-0.05	-0.95	0.89
Below Threshold	-0.726***	-3.02	-1.13	-0.32
Above Threshold	-11.197***	-5.80	-15.15	-7.24

Threshold Estimate = 0.546; Confidence Interval = [0.34, 0.77]

Hansen (2015) Wald test for threshold: p-value = 0.00

This Table reports results for endogenous threshold regressions estimating the sensitivity of spending changes to peer consumption. In panel A, we report the results for the threshold regressions of Hansen (2000). The procedure automatically selects the optimal threshold and estimates unconstrained linear regression estimates below and above the threshold. In addition to the regression coefficient estimates, we report results for the threshold estimates, the confidence interval for the threshold, and the p-value of the Hansen (1996) Lagrange Multiplier test for the presence of a threshold. Panel B reports the results for the regression kink model with unknown threshold proposed in Hansen (2017). The procedure automatically selects the optimal threshold and estimates a piecewise linear regression model that is continuous at the threshold. In addition to the parameter estimates, we report results for the threshold estimates, the confidence interval for the threshold, and the p-value of the Hansen (2017) Wald test for the presence of a threshold.

Table 7. Sensitivity to Peer Consumption and Spending Changes. Results from Two-Stage-Least-Squares Estimates

Panel A. First-Stage Estimates

	\$3K Thresh	**\$4K Thresh**	**\$5K Thresh**
Above_dummy	2068.9***	2087.4***	2064.3***
	(18.67)	(22.77)	(24.39)
Constant	5145.1***	5126.6***	5149.7***
	(56.31)	(73.73)	(84.04)

Panel B. Second-Stage Estimates

	\$3K Thresh	**\$4K Thresh**	**\$5K Thresh**
$\widehat{Peer\_Spending\_i}$	0.984**	0.825***	0.769***
	(2.53)	(2.64)	(2.66)
Constant	-6473.8**	-5323.6***	-4921.1***
	(-2.51)	(-2.66)	(-2.70)

This table reports results for a two-stage-least-squares identification strategy that considers clients just below and at each of the income thresholds *Status* Money uses to define peer groups, that is, \$35K, \$50K, \$65K, \$75K, \$100K, and \$150K. For each threshold, we only keep the clients that are at the lower threshold of the group as well as those clients that are in the lower income group, but are within \$3K of the threshold. Taking the \$100K threshold as an example, we only keep those that declare \$100K in annual income as well as those with an income between \$97K and \$100K. 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 client i and $Dummy\_Above\_i$ is a dummy variable for whether the income is exactly equal to the threshold value. The results for the first-stage are reported in Panel A. In the second stage, we use the instrumented $Peer\_Spending\_i$ of the first stage as the main covariate in the following specification:

$$\Delta\_Cons_i = \alpha + \beta\ \widehat{Peer\_Spending\_i} + \epsilon_i,$$

where Δ_Cons_i is the change in consumption before and after signing up for *Status*. The results for the second-stage are reported in Panel B. Within each panel, we report results for the specification that uses a \$3K threshold as well as two additional specifications that use \$4K and \$5K thresholds, respectively.

Table 8. Interpretation: Peer Pressure vs Overreaction to Negative News

	Panel A. Distance from Peer Spending							
	Full Sample		Exc. Top Decile		Exc. Top Quintile		Exc. Top Tercile	
	Value	t-stat	Value	t-stat	Value	t-stat	Value	t-stat
Dist. from Peers	-9.342***	(-10.57)	-8.512***	(-12.48)	-9.343***	(-13.83)	-10.693***	(-15.41)
Observations	4,077		3,669		3,261		2,718	
	Panel B. Distance from Average US Spending							
	Full Sample		Exc. Top Decile		Exc. Top Quintile		Exc. Top Tercile	
	Value	t-stat	Value	t-stat	Value	t-stat	Value	t-stat
Dist. from Avg US	-3.652***	(-5.49)	-0.874	(-1.53)	-0.502	(-0.86)	-3.065***	(-4.78)
Observations	3,652		3,287		2,922		2,435	
	Panel C. Distance from Average Monthly Income							
	Full Sample		Exc. Top Decile		Exc. Top Quintile		Exc. Top Tercile	
	Value	t-stat	Value	t-stat	Value	t-stat	Value	t-stat
Dist. from Avg Income	-6.456***	(-12.32)	-6.415***	(-17.10)	-7.141***	(-19.92)	-8.464***	(-23.61)
Observations	8,086		7,282		6,473		5,388	
	Panel D. Maximum Distance from Peer Spending, US Average Spending, and Average Monthly Income							
	Full Sample		Exc. Top Decile		Exc. Top Quintile		Exc. Top Tercile	
	Value	t-stat	Value	t-stat	Value	t-stat	Value	t-stat
Maximum Distance	-4.478***	(-10.22)	-3.362***	(-10.54)	-3.338***	(-10.76)	-5.587***	(-17.22)
Observations	5,151		4,636		4,121		3,434	

This table reports results on the economic channels driving the effects we document in the paper. Across the four panels, we regress overspending users' change in spending on the distance of their pre-sign up spending from 4 different points – peers' spending (Panel A), the average US household's spending (Panel B), users' average income (Panel C), and the maximum distance among these three (Panel D). Within each panel, we report across columns the results for the full sample, as well as results that exclude the top decile, quintile and tercile of observations.

Online Appendix:

Crowdsourcing Financial Information to Change Spending Behavior

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

Not for Publication

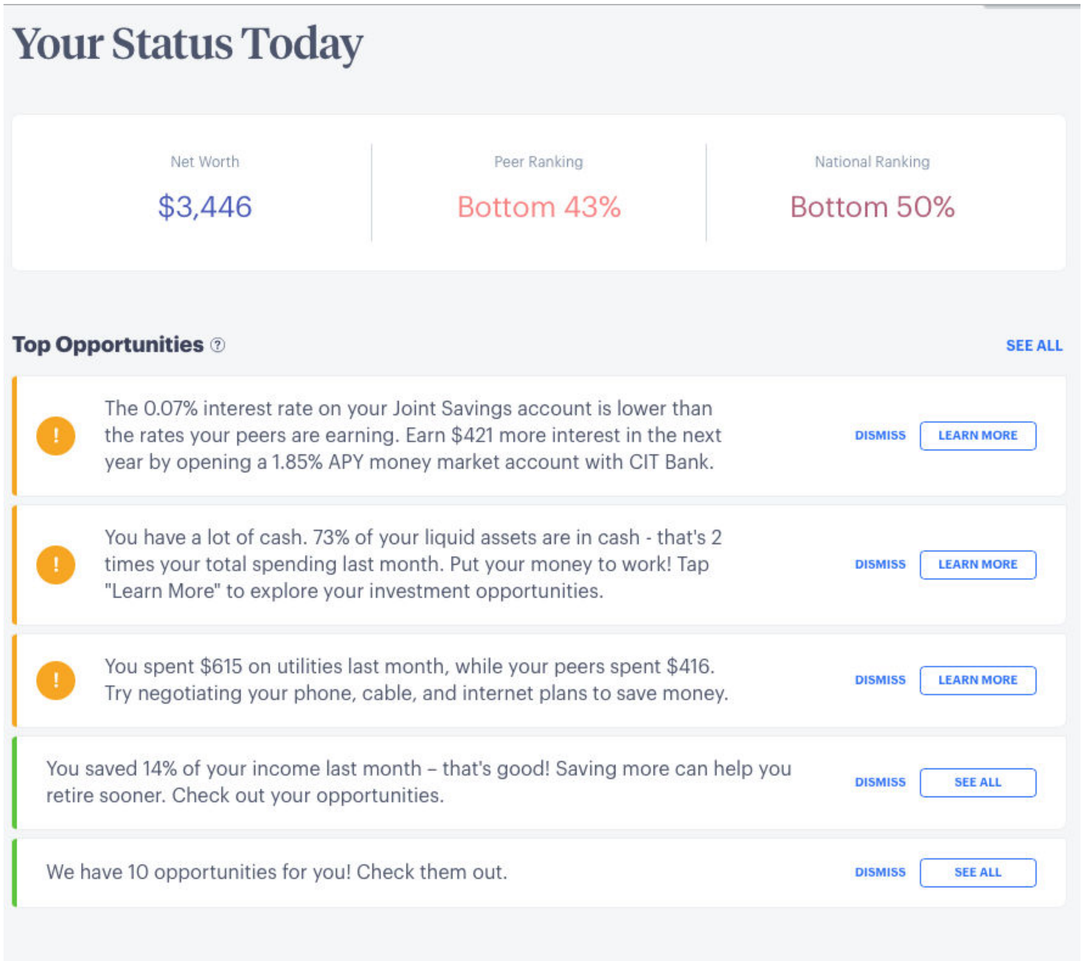


Figure A.1. ***Status*** **Home Page**